Effectiveness of exercise therapy: A best-evidence summary of systematic reviews

Nynke Smidt, Henrica CW de Vet, Lex M Bouter and Joost Dekker

for the Exercise Therapy Group[a]

The purpose of this project was to summarise the available evidence on the effectiveness of exercise therapy for patients with disorders of the musculoskeletal, nervous, respiratory, and cardiovascular systems. Systematic reviews were identified by means of a comprehensive search strategy in 11 bibliographic databases (08/2002), in combination with reference tracking. Reviews that included (i) at least one randomised controlled trial investigating the effectiveness of exercise therapy, (ii) clinically relevant outcome measures, and (iii) full text written in English, German or Dutch, were selected by two reviewers. Thirteen independent and blinded reviewers participated in the selection, quality assessment and data-extraction of the systematic reviews. Conclusions about the effectiveness of exercise therapy were based on the results presented in reasonable or good quality systematic reviews (quality score ≥ 60 out of 100 points). A total of 104 systematic reviews were selected, 45 of which were of reasonable or good quality. Exercise therapy is effective for patients with knee osteoarthritis, sub-acute (6 to 12 weeks) and chronic (≥ 12 weeks) low back pain, cystic fibrosis, chronic obstructive pulmonary disease, and intermittent claudication. Furthermore, there are indications that exercise therapy is effective for patients with ankylosing spondylitis, hip osteoarthritis, Parkinson's disease, and for patients who have suffered a stroke. There is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with neck pain, shoulder pain, repetitive strain injury, rheumatoid arthritis, asthma, and bronchiectasis. Exercise therapy is not effective for patients with acute low back pain. It is concluded that exercise therapy is effective for a wide range of chronic disorders. **[Smidt N, de Vet HCW, Bouter LM and Dekker J (2005): Effectiveness of exercise therapy: A best-evidence summary of systematic reviews. *Australian Journal of Physiotherapy* 51: 71–85]**

Key words: Exercise Therapy; Exercise Movement Techniques; Meta-analysis; Physical Therapy Techniques

Introduction

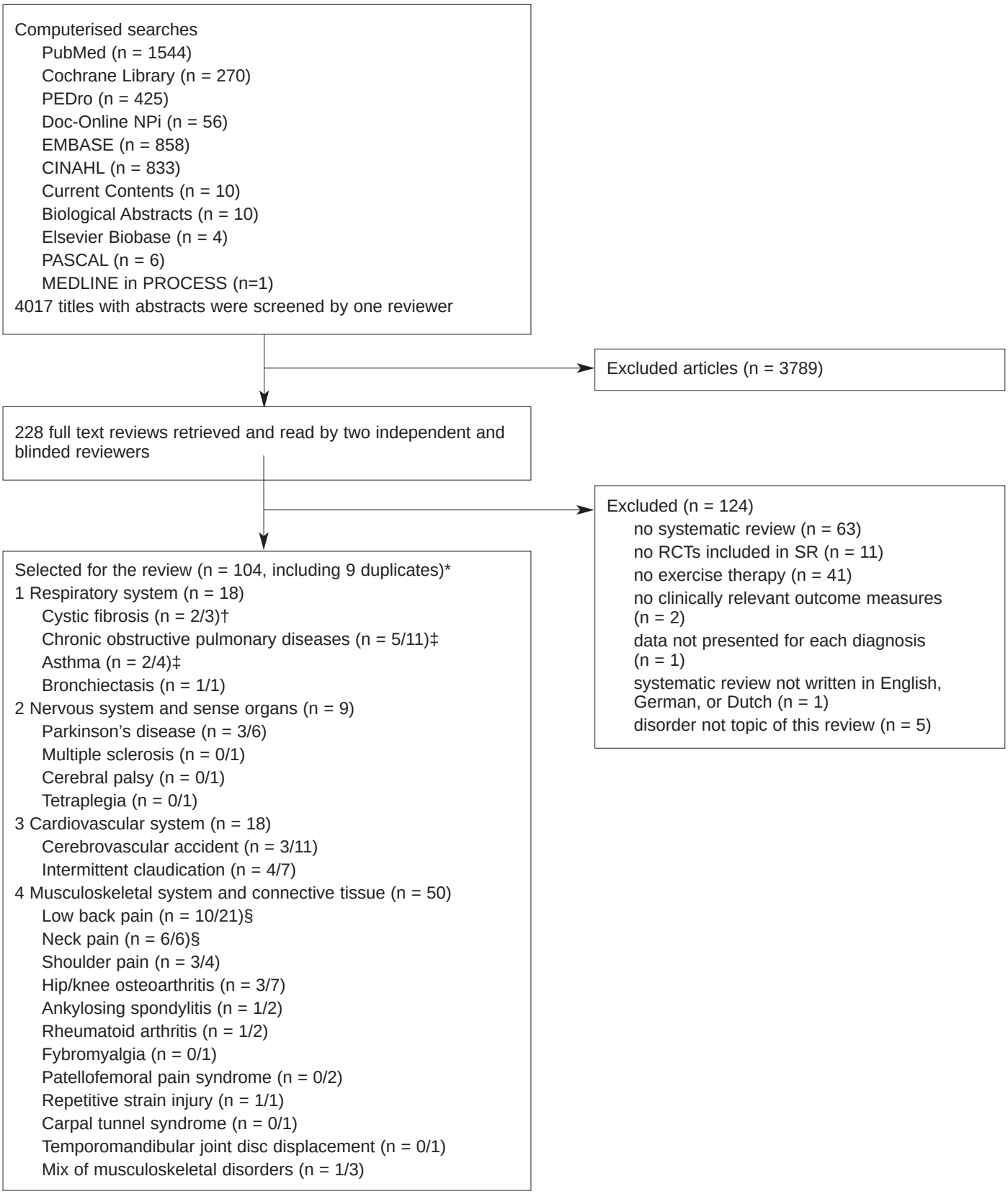
Exercise therapy is a regular component in the management of various (chronic) disorders, such as musculoskeletal, neurological, cardiovascular, and respiratory disorders (Chartered Society of Physiotherapy 2001, ACCP/AACVPR Pulmonary Rehabilitation Guidelines Panel, American College of Chest Physicians, American Association of Cardiovascular and Pulmonary Rehabilitation 1997, Gordon et al 2004, Pina et al 2003, Woolf et al 2004). Exercise therapy involves the prescription of muscular contraction and bodily movement ultimately to improve the overall function of the individual and to help meet the demands of daily living (Tan and Horn 1998).

There is no up-to-date overview of the effectiveness of exercise therapy compared with no treatment or an alternative treatment (Beckerman et al 1993a, Beckerman et al 1993b, Bouter et al 1992, Herbert et al 2001). Such an overview will help: care providers to choose the most appropriate treatment option; policy makers in making decisions concerning health care; and research agencies in setting priorities in the field of physiotherapy. Our objective was therefore to assess and summarise the available evidence on the effects of exercise therapy in a best-evidence summary of systematic reviews.

Method

Searching One reviewer (NS) searched computerised bibliographical databases (MEDLINE 01/1966–03/2002, PEDro 03/2002, CINAHL 01/1990–07/2002, EMBASE 01/1990–08/2002, Cochrane Library Issue 3 2002, Current Contents 01/1999–07/2002, Biological Abstracts 01/1999–07/2002, Elsevier Biobase 01/1999–07/2002, PASCAL 01/1999–07/2002, MEDLINE in PROCESS 01/1999–07/2002, and DocOnline (NPI) 10/1988–03/2002), using an approach based on the comprehensive search strategy outlined by Hunt and McKibbon (1997). The following specific subject (MeSH) headings and free text words were used to identify reviews of exercise therapy: pain, physical education and training, physical fitness, relaxation, physical endurance, physical therapy, exercise, motion therapy, and physiotherapy. In addition, references from retrieved reviews were screened.

Selection We included systematic reviews that met the following criteria: (i) the full text of the systematic review is published and it is based on a transparent and reproducible protocol (at least reporting on inclusion criteria, search date(s), and database(s)); (ii) at least one randomised controlled trial is included in the review; (iii) exercise therapy is compared with no treatment, other conservative types of treatment (e.g. steroid injections), surgery, or some other type of exercise therapy (e.g. flexion versus extension exercises); (iv) at least one clinically relevant outcome measure is included (e.g. pain, activities of daily living (ADL), walking distance, return to work) is included; (v) the results and conclusions are presented separately for each diagnosis; (vi) reviews are written in English, German or Dutch; (vii) the focus is on patients with disorders of the following: the musculoskeletal system and connective tissue, the nervous system and sense organs, the respiratory system, and the cardiovascular system (excluding coronary heart diseases),



*Duplicates came from Green et al (1998, 2002), Koes et al (1991a, 1991b), Lacasse et al (1997a, 1997b), Ram et al (2000, 2002), van Baar et al (1998a, 1999, 2001), van den Ende et al (1998, 2002), and van Tulder et al (1999, 2000a, 2002b). †Two of the 3 systematic reviews on the effectiveness of exercise therapy for cystic fibrosis had quality scores of at least 60 points out of 100. ‡In one review (<60 points) exercise therapy was investigated for both asthma and COPD therefore the sum of the individual sub-categories add up to more than the total category; §In one review (≥ 60 points) exercise therapy was investigated for both neck and low back pain therefore the sum of the individual sub-categories add up to more than the total category.

Figure 1. Selection of systematic reviews.

according to the International Classification of Diseases (ICD-10) are the topic of this summary (Anonymous 1992).

To determine whether a review should be included, the abstracts of all identified articles were read by one reviewer (NS). If there was any doubt, the full article was retrieved and read by two reviewers, independently. The articles were blinded for authors, journal, acknowledgements, and year of publication by a research assistant who was not involved in this study in any other way (KJ, see acknowledgements). Disagreements between reviewers about the final selection of the articles were discussed and resolved in a consensus meeting.

Quality assessment The quality of the systematic reviews was assessed according to the list of criteria developed by Assendelft et al (1995). This list consists of criteria for the selection of studies (30 points), assessment of the methodological quality of randomised controlled trials (20 points), description of the interventions (15 points), data presentation (20 points), and evaluation (15 points) (see Appendix I). The maximum quality score is 100 points. A total of 13 independent, blinded reviewers (see authors' affiliations) participated in the final selection and assessment of the quality of the systematic reviews. One reviewer (NS) assessed all systematic reviews and 12 other reviewers (MEB, SMAB, AH, SHJK, GK, TL, RPSP, MR, CT, CBT, APV, DAWMW) each evaluated a selection of the included reviews. Disagreements were discussed and resolved in a consensus meeting. If consensus could not be reached, a third reviewer (RWJGO) made the final decision.

The systematic reviews were categorised according their quality score: good quality (≥ 80 points), reasonable quality (60–79 points), moderate quality (40–59 points), poor quality (20–39 points), and very poor quality (< 20 points). Our conclusions regarding the effectiveness of exercise therapy are based on the results of reasonable quality (60–79 points) or good quality (≥ 80 points) systematic reviews (De Vet et al 2001).

Data extraction An overview of each systematic review (≥ 60 points) was made, including the research question(s) and details of all the randomised controlled trials investigating exercise therapy included in the systematic review (interventions in the experimental and control group, methodological quality, sample size (statistical power), outcome measures, timing of outcome assessment, and effectiveness of the exercise therapy (statistical significance)).

The conclusions reported in each systematic review were discussed with a panel of experts in the field of physiotherapy, general practice, rehabilitation medicine, and epidemiology (JHA, RAB, JD, PJMH, RABO, ST, HCWV). For each systematic review, categorisation of the conclusions was based on the following two research questions:

A What is the effectiveness of exercise therapy, compared to no treatment, a placebo, or a wait-and-see policy?

B What is the effectiveness of exercise therapy, compared to other treatments (e.g. steroid injections)? Is one specific type of exercise therapy more effective than others?

The following are all the possible conclusions that could be drawn for Question A:

- Exercise therapy is effective, compared to no treatment, placebo, or a wait-and-see policy (positive).
- Exercise therapy is not effective, compared to no treatment, placebo, or a wait-and-see policy (negative).
- Exercise therapy is less effective than no treatment, placebo, or a wait-and-see policy (harmful).
- There is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to no treatment, placebo, or a wait-and-see policy (insufficient evidence).
- There is insufficient evidence, but there are indications to support the effectiveness of exercise therapy, compared to no treatment, placebo, or a wait-and-see policy (insufficient evidence but indications).

The following are all the possible conclusions that could be drawn for Question B:

- Exercise therapy is effective, compared to other treatments (positive).
- Exercise therapy is equally effective, compared to other treatments (equal).
- Exercise therapy is less effective, compared to other treatments (negative).
- There is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to other treatments (insufficient evidence).
- There is insufficient evidence, but there are indications to support the effectiveness of exercise therapy, compared to other treatments (insufficient evidence but indications).

If the panel felt that the conclusions were not sufficiently justified by the data presented in the systematic review at issue, the conclusions reported in the systematic review were not endorsed, and the panel drew its own conclusions about the effectiveness of exercise therapy. In such cases, the panel's conclusions were based on randomised controlled trials that were of good methodological quality (≥ 50% of the quality score reported in the systematic review) with large sample sizes (smallest group n ≥ 50).

For each disorder, the panel's final conclusions with regard to the effectiveness of exercise therapy were based on the conclusions of all available systematic reviews. If the conclusions of the systematic reviews were conflicting, the sources of discordance among the conclusions of systematic reviews were explored (Jadad et al 1997). The panel based its final conclusions on the most complete systematic review, using the decision tool described by Jadad et al (1997).

Results

Selection of studies The results of our search strategy are presented in a flow chart (Fig. 1). Out of a total of 4017 abstracts, 228 reviews were considered to be potentially eligible for our best-evidence summary. Reviewing the full text resulted in the inclusion of 104 systematic reviews, including nine duplicates. The systematic reviews have been marked with an asterisk in the reference list.

Quality assessment The overall inter-rater agreement for the quality assessment was 86% (Cohen's Kappa 0.73). Most of the disagreements were caused by differences in

Table 1. Results and conclusions of systematic reviews (quality score ≥ 60 points) on the effectiveness of exercise therapy (n = 45)[a]

Systematic review[b]	Disease	Score[c]	No. RCTs		Quality[d]		Conclusions SR		Dissent[e]	Conclusions Panel	
			A	B	A	B	A	B		A	B
Thomas et al (1995)	Cystic fibrosis	66	3	7	2	5	+	?	No	+	?
Bradley & Moran (2002)	Cystic fibrosis	75	3	4	0	0	?	?	No	?	?
Ram et al (2000, 2002)	Asthma	70	8	NA	0	NA	?	NA	No	?	NA
Holloway & Ram (2002)	Asthma	64	3	3	1	0	?	?	No	?	?
Bradley et al (2002)	Bronchiectasis	64	2	1	0	0	?	?	No	?	?
Smith et al (1992)	COPD, asthma and bronchitis	65	12	3	10	2	-	?	Yes	?	?
Lacasse et al (1996)	COPD	70	14	NA	14	NA	+	NA	No	+	NA
Lacasse et al (1997a, 1997b)	COPD	62	6	18	0	2	+	?	No	+	?
Cambach et al (1999)	COPD	64	12	2	2	0	+	?	No	+	?
Lacasse et al (2002)	COPD	71	23	NA	6	NA	+	?	No	+	?
De Goede et al (2001)	Parkinson's dis.	65	5	3	2	2	+	?	Yes	? (ind)	?
Deane et al (2002b)	Parkinson's dis.	64	NA	7	NA	5	NA	?	No	NA	?
Deane et al (2002a)	Parkinson's dis.	66	11	NA	5	NA	?	?	No	? (ind)	?
Kwakkel et al (1997)	CVA	61	3	5	0	0	?	?	No	?	?
van der Lee et al (2001)	CVA	79	1	12	1	9	?	? (ind)	No	? (ind)	? (ind)
Snels et al (2002)	CVA	62	NA	2	NA	1	NA	?	No	?	?
Brandsma et al (1998)	Inter. claudication	69	5	5	4	5	?	?	No	? (ind)	?
Robeer et al (1998)	Inter. claudication	76	6	6	5	6	+	+	Yes	? (ind)	?
Girolami et al (1999)	Inter. claudication	66	6	NA	6	NA	+	NA	No	+	NA
Leng et al (2002)	Inter. claudication	74	6	6	3	2	+	?	No	+	?
van der Heijden et al (1995)[f]	Neck & back pain	73	1	0	0	0	?	?	No	?	?
Aker et al (1996)	Neck pain	69	7	6	6	6	?	?	No	?	?
Hurwitz et al (1996)	Cervical spine dis.	72	1	4	0	2	?	?	No	?	?
Kjellman et al (1999)	Neck pain	63	4	6	1	1	?	?	No	?	?
Philadelphia Panel (2001b)	Neck pain	71	3	NA	0	NA	+	NA	Yes	?	?
Gross et al (2002)	Neck disorders	76	0	1	0	1	?	?	No	?	?
van der Heijden et al (1997)	Shoulder disorders	66	2	4	1	1	?	?	No	?	?
Green et al (1998, 2002)	Shoulder disorders	75	0	3	0	0	?	?	No	?	?
Philadelphia Panel (2001c)	Shoulder pain	66	1	NA	0	NA	?	NA	No	?	NA
Konijnenberg et al (2001)	RSI	75	2	4	1	2	?	?	No	?	?
Dagfinrud & Hagen (2002)	Ankyl. spondylitis	84	1	2	1	1	?	?	No	? (ind)	?
van Baar et al (1998a, 1999, 2001)	Hip and knee osteoarthritis	87	8	4	4	3	+[g]	?	No	+[g]	?
Philadelphia Panel (2001a)	Knee pain	72	6	NA	3	NA	+	NA	No	+	NA
Fransen et al (2002)	Knee osteoarthritis	78	11	3	7	2	+	?	No	+	?
van den Ende et al (1998, 2002)	Rheumatoid arth.	74	4	2	3	2	?	?	No	?	?
Koes et al (1991a, 1991b)	Low back pain (not specified)	66	5	14	3	1	?	?	No	?	?
van der Heijden et al (1995)[f]	Neck & back pain	73	0	2	0	0	?	?	No	?	?
Scheer et al (1995)	Low back pain (< 4 weeks)	63	5	5	4	4	?	?	No	?	?
van Tulder et al (1997)	Low back pain (≤ 6 weeks)	69	5	8	2	0	-	?	No	-	?
	Low back pain (> 12 weeks)	69	6	16	2	4	+	?	No	+	?
Hilde & Bo (1998)	Low back pain (> 4 weeks)	69	5	8	1	4	?	?	No	?	?
van Tulder et al (1999, 2000a, 2002b)	Low back pain (≤ 12 weeks)	83	3	11	2	3	-	?	No	-	?

	Low back pain (> 12 weeks)	83	8	19	3	10	+	±	No	+	±
	Low back pain (mixed group)	83	1	4	1	0	?	?	No	?	?
van Tulder et al (2000b)	Low back pain (> 12 weeks)[h]	81	11	13	3	3	+	?	No	+	?
Guzman et al (2001)	Low back pain (> 12 weeks)[i]	72	1	9	0	4	?	?	No	?	?
Philadelphia Panel (2001d)	Low back pain (< 4 weeks)	72	4	NA	2	NA	+	NA	Yes	?	NA
	Low back pain (4–12 weeks)	72	3	NA	1	NA	+	NA	No	+	NA
	Low back pain (> 12 weeks)	72	8	NA	2	NA	+	NA	No	+	NA
van Tulder et al (2002a)	Low back pain (≤ 12 weeks)[j]	88	3	2	1	1	?	?	No	?	?
	Low back pain (> 12 weeks)[j]	88	6	5	1	1	?	?	No	?	?
	Low back pain (mixed group)[j]	88	1	2	0	0	?	?	No	?	?
Beckerman et al (1993a)	Musculo. dis.	62	NA	NA	NA	NA	NA	NA	No	NA	NA

[a]Because there are nine duplicates, the number of systematic reviews presented in this table is 45. Duplicates came from van Tulder et al (1999, 2000a, 2002b), Koes et al (1991a, 1991b), Green et al (1998, 2002), van Baar et al (1998a, 1999, 2001), van den Ende et al (1998, 2002), Ram et al (2000, 2002) and Lacasse et al (1997a, 1997b). [b]The systematic reviews are ranked in order of publication (for each disorder), equally ranked reviews are ordered alphabetically. [c]Total quality score of the systematic review; the quality score is calculated as the sum of all items. [d]Number of randomised controlled trials of high quality based on the methodological quality presented in the systematic review; RCTs with at least 50% of the maximum quality score were regarded as 'high quality'. [e]Disagreement between the conclusions in the systematic review and the conclusions of the panel. [f]The systematic review of van der Heijden (1995) investigated exercise therapy for patients with low back pain and neck pain and is therefore presented twice in this table. [g]Conclusions were drawn regarding the effectiveness of exercise therapy for patients with knee osteoarthritis; only one large (smallest group > 50) RCT of methodological good quality (≥ 50% quality scores) investigated the effectiveness of exercise therapy for hip osteoarthritis and found positive results on pain, observed disabilities, and patients, global assessment. [h]RCTs investigated the effectiveness of cognitive behavioural therapy (exercise therapy is included). [i]RCTs investigated the effectiveness of multidisciplinary biopsychosocial rehabilitation (exercise therapy is included). [j]RCTs investigated the effectiveness of back schools (exercise therapy is included).

A, these columns contain data relating to the effectiveness of exercise therapy compared to no treatment, a placebo or a wait-and-see policy. Ankyl. Spondylitis = ankylosing spondylitis. B, these columns contain data relating to the effectiveness of exercise therapy compared to another treatment. ?(ind) = insufficient evidence to support the effectiveness of exercise therapy, but there are indications to support the effectiveness of exercise therapy. Cervical spine dis. = cervical spine disorders. Inter. Claudication = intermittent claudication. Musculo. dis. = musculoskeletal disorders. NA = not applicable (was not investigated in the review). Parkinson's dis. = Parkinson's disease. RCT = randomised controlled trial. Rheumatoid arth. = rheumatoid arthritis. RSI = repetitive strain injury. + = Exercise therapy is effective. ? = Insufficient evidence to support or refute the effectiveness of exercise therapy. ± = Exercise therapy is equally effective compared to other treatments. – = Exercise therapy is not effective compared to no treatment.

interpretation when discussing the power of the randomised controlled trials (see Appendix I, item L) and the heterogeneity of randomised controlled trials and outcomes (items N1, N2, N3, N4).

The mean (standard deviation) quality score of 95 systematic reviews (excluding the duplicate reviews) was 56 (17), ranging from 17 to 88 points (see Appendix II in the addenda at the AJP website, www.physiotherapy.asn.au/AJP). The most prevalent flaws were associated with the assessment of the methodological quality of the individual randomised controlled trials in the systematic review (items D1, D2, D5, D6, F, G), the data presentation (items J1, J2, J3, J4, L) and the evaluation of the results (items N1, N3, N4). There were 45 systematic reviews with a quality score of at least 60 points. These reviews investigated the effectiveness of exercise therapy for cystic fibrosis, chronic obstructive pulmonary disease (COPD), asthma, bronchiectasis, Parkinson's disease, cerebrovascular accident (CVA), intermittent claudication, osteoarthritis, ankylosing spondylitis, rheumatoid arthritis, repetitive strain injury (RSI), neck pain, shoulder pain, and low back pain. Systematic reviews investigating the effectiveness of exercise therapy for patients with fibromyalgia, patellofemoral pain syndrome, carpal tunnel syndrome, temporomandibular joint displacement, multiple sclerosis, and cerebral palsy had low quality scores (< 60 points) (Baker and Tickle-Degnen 2001, Crossley et al 2001, Dodd et al 2002, Feuerstein et al 1999, Kropmans et al 1999, Rossy et al 1999, Stiller and Huff 1999, Zomerdijk et al 1998). Consequently, these disorders will not be discussed.

For each systematic review (≥ 60 points), the quality score, the total number of randomised controlled trials, the number of high quality randomised controlled trials, the conclusions reported in the review, and the final conclusions of the panel

are presented in Table 1. In five cases the panel disagreed with the authors of the systematic review with regard to the conclusions. These disagreements were mainly caused by inadequate reporting of the results of the randomised controlled trials in the systematic review (Philadelphia Panel 2001d, Robeer et al 1998) or because the conclusions were based on both randomised controlled trials and controlled clinical trials (De Goede et al 2001, Philadelphia Panel 2001b). In one systematic review the overall conclusions were drawn for a very heterogeneous patient population, namely patients with COPD, asthma, and bronchitis (Smith et al 1992).

Characteristics of the systematic review Details of each systematic review (≥ 60 points), including the research question(s), information on randomised controlled trials, the conclusions of the authors, and the final conclusions of the panel are presented in the Appendix III (see addenda at the AJP website, www.physiotherapy.asn.au/AJP).

Cystic fibrosis Three systematic reviews investigated the effectiveness of exercise therapy for patients with cystic fibrosis (Boyd et al 1994, Bradley and Moran 2002, Thomas et al 1995). Based on the results of two reasonable quality systematic reviews, we concluded that exercise therapy in addition to percussion, vibration, and postural drainage, has beneficial effects on FEV_1 (Forced Expiration Volume within one second) (Bradley and Moran 2002, Thomas et al 1995). The exercise therapy consisted of aerobic exercises (e.g. swimming), strength training exercises, and inspiratory muscle training. It is unclear whether exercise therapy is also effective for outcome measures such as quality of life. There is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to no treatment (no randomised controlled trials available), or compared to treatment consisting of percussion, vibration, and postural drainage, or other treatments for patients with cystic fibrosis.

Asthma Four systematic reviews investigated the effectiveness of exercise therapy for patients with asthma (Ernst 2000, Gosselink and Wagenaar 1993a, Gosselink and Wagenaar 1993b, Holloway and Ram 2002, Ram et al 2002, Ram et al 2000). Based on the results of two reasonable quality systematic reviews (Holloway and Ram 2002, Ram et al 2002, Ram et al 2000), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for children and adults with asthma, compared to no treatment or other conservative treatments.

Bronchiectasis With regard to bronchiectasis, we found only one reasonable quality systematic review (Bradley et al 2002). Due to the strict selection criteria applied in this systematic review, only two randomised controlled trials with poor quality reporting (abstract only) were included. Based on the results of this review, we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with bronchiectasis.

Chronic obstructive pulmonary disease (COPD) During the period 1992–2002, 11 systematic reviews on the effectiveness of exercise therapy for COPD were published (Bekkering et al 1998, Cambach et al 1999, Chavannes and Vollenberg 2002, Devine and Peacy 1996, Gosselink and Wagenaar 1993a, Gosselink and Wagenaar 1993b, Lacasse et al 1996, Lacasse et al 1997a, Lacasse et al 1997b, Lacasse et al 2002, Lötters et al 2002, Ries et al 1997, Smith et al 1992). Based on the results of five reasonable quality systematic reviews (Cambach et al 1999, Lacasse et al 1996, Lacasse et al 1997a, Lacasse et al 1997b, Lacasse et al 2002, Smith et al 1992) we concluded that exercise therapy, consisting of aerobic exercises (e.g. walking, cycling) and strengthening exercises, is effective in improving the maximum and functional exercise capacity and quality of life of patients with COPD. Exercise therapy in a supervised program is probably more effective than exercise therapy in an unsupervised program, which showed no beneficial effects, compared to no treatment (Lacasse et al 2002). However, there were no randomised controlled trials included in the systematic reviews that directly compared the effectiveness of supervised exercise therapy to unsupervised exercise therapy. There is insufficient evidence to support or refute the effectiveness of a specific type of exercise therapy. There is also insufficient evidence to draw conclusions with regard to the effectiveness of exercise therapy, compared to other conservative treatments.

Parkinson's disease Six systematic reviews investigated the effectiveness of exercise therapy for patients with Parkinson's disease (Deane et al 2002a, Deane et al 2002b, Deane et al 2002c, De Goede et al 2001, Nieuwboer et al 1994, Rubinstein et al 2002). Based on the results of three reasonable quality systematic reviews (Deane et al 2002a, Deane et al 2002b, De Goede et al 2001), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with Parkinson's disease. There are indications that intensive exercise therapy, consisting of general mobility activities focussing on balance, posture, walking, range of motion, fine motor dexterity, and functional exercises has positive effects on the activities of daily living for patients with Parkinson's disease. However, this is based on randomised controlled trials with poor methodological quality or randomised controlled trials with small sample sizes (Comella et al 1994, Gauthier et al 1987, Patti et al 1996).

Cerebrovascular accident (CVA) Eleven systematic reviews investigated the effectiveness of exercise therapy in patients who had suffered a stroke (CVA) (de Bie et al 1995, Hiraoka 2001, Kwakkel et al 1997, Langhorne et al 1996, Ottenbacher and Jannell 1993, Pedro-Cuesta et al 1992, Pomeroy and Tallis 2000, Schoppink et al 1996, Snels et al 2002, van der Lee 2001, van der Lee et al 2001). Based on the results of three reasonable quality systematic reviews (Kwakkel et al 1997, Snels et al 2002, van der Lee et al 2001), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients who had suffered a stroke or for patients with hemiplegic shoulder pain, compared to no treatment or other conservative treatments. There are indications that (time-) intensive exercise therapy has more positive effects on the activities of daily living in patients who had suffered a stroke than less intensive exercise therapy. The exercise therapy consisted of neuromuscular facilitation and functional exercises, focusing on training of toilet transfers, rising from a sitting position, and walking. However, this was based on randomised controlled trials with poor methodological quality (Peacock et al 1972, Sivenius et al 1985, Smith et al 1981, Werner and Kessler 1996). More research is needed to confirm these results.

Intermittent claudication Seven systematic reviews investigated the effectiveness of exercise therapy in patients with intermittent claudication (Brandsma et al 1998, Gardner and Poehlman 1995, Girolami et al 1999, Leng et al 2002, Neill 1999, Radack and Wyderski 1990, Robeer et al 1998).

Based on the results of four reasonable quality systematic reviews (Brandsma et al 1998, Girolami et al 1999, Leng et al 2002, Robeer et al 1998), we concluded that exercise therapy is effective for patients with intermittent claudication, compared to no treatment. Exercise therapy consisted of (treadmill) training in walking, and lower limb strengthening exercises (e.g. stair climbing). The patients were encouraged to continue with daily walking exercises at home until they felt moderate pain. There are also indications that exercise therapy is more effective in improving maximal walking time than angioplasty (Creasy et al 1990) (Weighted Mean Difference (WMD) = 3.30 minutes; 95% CI 2.21 to 4.39) or antiplatelet therapy (Mannarino et al 1991) (WMD = 1.06 minutes; 95% CI 0.15 to 1.97), and there are indications that exercise therapy is equally as effective as surgery (Lundgren et al 1989) (WMD = -1.66 minutes; 95% CI -4.55 to 1.23). However, this was based on randomised controlled trials that either had small sample sizes or the methodological quality was not described in the review (unclear) (Creasy et al 1990, Lundgren et al 1989, Mannarino et al 1991). No conclusions can be drawn with regard to the effectiveness of a specific type of exercise therapy for patients with intermittent claudication.

Osteoarthritis Seven systematic reviews investigated the effectiveness of exercise therapy for patients with knee or hip osteoarthritis (Fransen et al 2002, McCarthy and Oldham 1999, Pendleton et al 2000, Petrella 2000, Philadelphia Panel 2001a, Puett and Griffin 1994, van Baar et al 1998a, Van Baar et al 1999, van Baar et al 2001). Based on the results of three reasonable or good quality systematic reviews, we concluded that exercise therapy, consisting of strengthening, stretching, and functional exercises, is effective for patients with knee osteoarthritis, compared to no treatment (Fransen et al 2002, Philadelphia Panel 2001a, van Baar et al 1998a, van Baar et al 1999, van Baar et al 2001). There are indications that exercise therapy (e.g. strengthening and stretching exercises, functional training, and ADL instruction) is effective for patients with hip osteoarthritis. However, this is based on one large randomised controlled trial with good methodological quality (van Baar et al 1998b). There is insufficient evidence to support or refute the effectiveness of a specific type of exercise therapy (individual, group therapy, or hydrotherapy) for patients with knee or hip osteoarthritis.

Ankylosing spondylitis Two systematic reviews investigated the effectiveness of exercise therapy for patients with ankylosing spondylitis (Ammer 1997, Dagfinrud and Hagen 2002). Based on one good quality systematic review (Dagfinrud and Hagen 2002), we concluded that there are indications to support the effectiveness of exercise therapy, compared to no treatment for patients with ankylosing spondylitis. The exercise therapy consisted of functional exercises and exercises to improve mobility, strength, and endurance, using normal movement patterns and proprioceptive neuromuscular facilitation. The patients received disease education and were encouraged to continue their exercises daily at home. However, this was based on only one small good quality randomised controlled trial (Kraag et al 1990). No conclusions can be drawn with regard to the effectiveness of exercise therapy, compared to other types of exercise therapy or other treatments.

Rheumatoid arthritis Two systematic reviews investigated the effectiveness of exercise therapy for patients with rheumatoid arthritis (Augustinus et al 2000, van den Ende et al 1998, van den Ende et al 2002). Based on one reasonable quality systematic review, we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with rheumatoid arthritis (van den Ende et al 1998, van den Ende et al 2002).

Repetitive strain injury With regard to repetitive strain injury, we found only one reasonable quality systematic review (Konijnenberg et al 2001). We concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with repetitive strain injury.

Neck pain We found six systematic reviews investigating the effectiveness of exercise therapy for patients with non-specific neck pain (Aker et al 1996, Gross et al 2002, Hurwitz et al 1996, Kjellman et al 1999, Philadelphia Panel 2001b, van der Heijden et al 1995). Based on the results of these six reasonable quality systematic reviews, we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to no treatment or other conservative treatments, for patients with (non-specific) neck pain.

Shoulder pain During the period 1997–2002, four systematic reviews on the effectiveness of exercise therapy for shoulder pain were published (Green et al 1998, Green et al 2002, Johansson et al 2002, Philadelphia Panel 2001c, van der Heijden et al 1997). Based on the results of three reasonable quality systematic reviews (Green et al 1998, Green et al 2002, Philadelphia Panel 2001c, van der Heijden et al 1997), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with shoulder pain or shoulder complaints.

Low back pain A total of 21 systematic reviews, published between January 1985 and July 2002, investigated the effectiveness of exercise therapy (also exercise therapy including cognitive behavioural therapy, back school, multidisciplinary rehabilitation) in patients with low back pain (Elders et al 2000, Di Fabio 1995, Faas 1996, Guzman et al 2001, Hilde and Bo 1998, Koes et al 1991a, Koes et al 1991b, Koes et al 1995, Maier-Riehle and Härter 2001, Ottenbacher and Di Fabio 1985, Philadelphia Panel 2001d, Scheer et al 1995, Scheer et al 1997, Smith et al 2002, van Duijvenbode 1996, van Duijvenbode 1999, van Tulder et al 1997, van Tulder 1999, van Tulder et al 2000a, van Tulder et al 2000b, van Tulder et al 2002a, van Tulder et al 2002b, Weinhardt et al 2001). Ten systematic reviews had reasonable or good scores for quality (Guzman et al 2001, Hilde and Bo 1998, Koes et al 1991a, Koes et al 1991b, Philadelphia Panel 2001d, Scheer et al 1995, van Tulder et al 1997, van Tulder 1999, van Tulder et al 2000a, van Tulder et al 2000b, van Tulder et al 2002a, van Tulder et al 2002b).

For patients with acute low back pain (< 6 weeks) there is no difference in the effectiveness of exercise therapy (e.g. stretching, strengthening, extension/flexion exercises), compared to no treatment, care provided by a general practitioner, or manipulations (high velocity techniques). For patients with sub-acute (6 to 12 weeks) and chronic (> 12 weeks) low back pain, we concluded that exercise therapy is effective compared to no treatment. The exercise therapy consisted of aerobic exercises (e.g. walking, jogging), and intensive strengthening exercises for the abdomen and trunk muscles. Exercise therapy (e.g. aerobic exercises, progressive muscle relaxation) in combination with cognitive behavioural therapy is also more effective than no treatment for patients with chronic low back pain.

For patients with chronic (> 12 weeks) low back pain, exercise therapy (e.g. strengthening exercises) is more effective than continued care provided by a general practitioner, and equally as effective as conventional physiotherapy (e.g. traction, massage, ultrasound, mobilisation exercises, hot and cold packs). There is insufficient evidence to support or refute the effectiveness of a particular type of exercise therapy for patients with sub-acute or chronic low back pain. There are indications that intensive multidisciplinary bio-psychosocial rehabilitation with functional restoration (including intensive aerobic exercises, stretching exercises, and muscle relaxation therapy) is more effective than physical training plus back school for patients with chronic low back pain. However, this was based on only one good quality randomised controlled trial with a short and long-term follow-up (Bendix et al 1995).

There are indications that exercise therapy, consisting of abdominal strengthening exercises, in addition to back school, is effective for patients with chronic low back pain, compared to back school without exercise therapy. However, this was also based on only one randomised controlled trial with good methodological quality (Klaber-Moffett et al 1986). There is insufficient evidence to support or refute the effectiveness of cognitive behavioural therapy plus exercise therapy compared to other conservative treatments for patients with chronic low back pain. There is also insufficient evidence to draw conclusions with regard to the (in)effectiveness of back schools for patients with acute, sub-acute or chronic low back pain.

Discussion

Exercise therapy is effective for patients with knee osteoarthritis, sub-acute and chronic low back pain, cystic fibrosis, COPD, and intermittent claudication. Furthermore, there are indications that exercise therapy is effective for patients with ankylosing spondylitis, hip osteoarthritis, and Parkinson's disease, and also for patients who have suffered a stroke. We concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with neck pain, shoulder pain, RSI, rheumatoid arthritis, asthma, and bronchiectasis. Exercise therapy is not effective for patients with acute low back pain. Based on the available literature, we found no evidence that exercise therapy is harmful or that it provoked harmful side effects. However, systematic reviews provide little information on the safety aspects of exercise therapy. This is mainly due to the inadequate reporting of adverse effects in randomised controlled trials (Ernst and Pittler 2001).

For certain diseases (fibromyalgia, patellofemoral pain syndrome, carpal tunnel syndrome, temporomandibular joint displacement, multiple sclerosis, cerebral palsy), only systematic reviews with low scores for quality (< 60 points) were available (Baker and Tickle-Degnen 2001, Crossley et al 2001, Dodd et al 2002, Feuerstein et al 1999, Kropmans et al 1999, Rossy et al 1999, Stiller and Huff 1999, Zomerdijk et al 1998). For these disorders, we recommend that systematic reviews be carried out using methods that accord to the current state of knowledge (Egger et al 2001).

Although a number of systematic reviews were of reasonable or good quality, there was still insufficient evidence to draw firm conclusions with regard to the (in)effectiveness of exercise therapy for neck pain, shoulder pain, repetitive strain injury, rheumatoid arthritis, asthma, and bronchiectasis. This was mainly due to the contradictory results, the poor methodological quality of the randomised controlled trials, inadequate reporting, small sample sizes, and the large variation in outcome measures and study populations. We recommend that searches be conducted for new published, large randomised controlled trials of good quality (since the last search date of the most recent systematic review of reasonable or good quality) on the effectiveness of exercise therapy for the following disorders: neck pain, shoulder pain, RSI, rheumatoid arthritis, asthma, and bronchiectasis. If no new randomised controlled trials have been published, or the retrieved randomised controlled trials are of poor methodological quality, we recommend that a new, large randomised controlled trial with good methodological quality be carried out.

We found indications to support the effectiveness of exercise therapy for patients with ankylosing spondylitis, hip osteoarthritis, Parkinson's disease and patients who had suffered a stroke, but more randomised controlled trials are needed to confirm these results.

With regard to the disorders for which exercise therapy appeared to be effective, it still remains to be determined whether exercise therapy should be included in a supervised or an unsupervised program, and whether exercise at home is sufficient or referral should be made to a physiotherapist. There is also insufficient evidence to support or refute the effectiveness of specific types of exercise therapy for almost all disorders. More research is also needed to investigate how the short-term effectiveness of exercise therapy can be maintained in the long-term. Programs or methods with which care-providers could encourage the compliance of patients with home exercises and motivate them to continue their exercises in the future would be very useful.

This best-evidence summary of systematic reviews has a number of limitations. First, different weights were applied to the five quality criteria, including the selection of studies, methodological quality assessment of the randomised controlled trials, description of the intervention, data presentation, and evaluation. Total quality scores were calculated by summing up the weights of all quality items. The advantage of using an overall quality score is its simplicity, but methodologically it is debatable. If equal weights were applied to each quality item, the division of systematic reviews into good, reasonable, moderate, poor, and very poor quality would be quite similar, and the final conclusions with regard to the effectiveness of exercise therapy would still be the same.

Second, the choice of the cut-off point for reasonable or good quality was arbitrary. The quality of the reporting of the results of systematic reviews with low scores for quality (< 60 points) was often too poor to draw conclusions with regard to the effectiveness of exercise therapy. If, for example, the cut-off point was set at 50 points, another 11 reviews would have been included. However, our conclusions with regard to the effectiveness of exercise therapy for the disorders discussed in this review would remain the same (data not shown). We could only draw new conclusions with regard to the effectiveness of exercise therapy for patellofemoral pain syndrome.

Third, our conclusions were based on statistically significant differences, rather than clinically relevant differences. Unfortunately, based on the results presented in the

systematic reviews, it was not possible to calculate effect sizes. Therefore, the clinically relevant differences were not taken into account in our conclusions.

Finally, a few systematic reviews on the same topic reported conflicting conclusions. However, based on the guidelines developed by Jadad et al (1997), explaining differences in research questions, assessment of the quality of randomised controlled trials, number of randomised controlled trials, and statistical methods for data-analysis, the panel succeeded in drawing clear conclusions.

In conclusion, exercise therapy has been shown to be effective for a wide range of (chronic) disorders.

Footnotes

[a]Contributing authors (members of the Exercise Therapy Group):

Johan H Arendzen, Leiden University Medical Center, Leiden, The Netherlands

Rob A de Bie, University of Maastricht, The Netherlands

Sita MA Bierma-Zeinstra, Erasmus Medical Center, Rotterdam, The Netherlands

Lex M Bouter, VU University Medical Center, The Netherlands

Henrica CW de Vet, VU University Medical Center, Amsterdam, The Netherlands

Joost Dekker, VU University Medical Center, Amsterdam The Netherlands (chair)

Paul JM Helders, University Medical Center Utrecht, The Netherlands

Samyra HJ Keus, Leiden University Medical Center, Leiden, The Netherlands

Gert Kwakkel, VU University Medical Center, Amsterdam, The Netherlands

Ton Lenssen, University of Maastricht, The Netherlands

Rob AB Oostendorp, University Medical Center St. Radboud, Nijmegen, The Netherlands

Raymond WJG Ostelo, VU University Medical Center, Amsterdam, The Netherlands

Max Reijman, Erasmus Medical Center, Rotterdam, The Netherlands

Nynke Smidt, VU University Medical Center, Amsterdam, The Netherlands

Caroline B Terwee, VU University Medical Center, Amsterdam, The Netherlands

Carlo Theunissen, University of Maastricht, The Netherlands

Siep Thomas, Erasmus Medical Center, Rotterdam, The Netherlands

Margriet E van Baar, Erasmus Medical Center, Rotterdam, The Netherlands

Alex van 't Hul, VU University Medical Center, Amsterdam, The Netherlands

Roland PS van Peppen, VU University Medical Center, Amsterdam, The Netherlands

Arianne P Verhagen, Erasmus Medical Center, Rotterdam, The Netherlands

Daniëlle AWM van der Windt, VU University Medical Center, Amsterdam, The Netherlands

Acknowledgements This article was based on the work of the 'Panel on Exercise Therapy' of the Dutch Health Council. The authors would like to thank Karin Johnson for blinding the articles.

Correspondence Nynke Smidt, Institute for Research in Extramural Medicine, VU University Medical Center, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands. Email: <n.smidt@vumc.nl>

References

Systematic reviews have been marked with an asterisk.

ACCP/AACVPR Pulmonary Rehabilitation Guidelines Panel; American College of Chest Physicians; American Association of Cardiovascular and Pulmonary Rehabilitation (1997): Pulmonary rehabilitation: Joint ACCP/AACVPR evidence-based guidelines. *Chest* 112: 1363–1396.

*Aker PD, Gross AR, Goldsmith CH and Peloso P (1996): Conservative management of mechanical neck pain: Systematic overview and meta-analysis. *BMJ* 313: 1291–1296.

*Ammer K (1997): Physiotherapy in seronegative spondylarthropathies—a systematic review. *European Journal of Physical Medicine and Rehabilitation* 7: 114–119.

Anonymous (1992): *International Statistical Classification of Diseases and Related Health Problems (ICD-10)*, 1989 Revision. Geneva: World Health Organization.

Assendelft WJ, Koes BW, Knipschild PG and Bouter LM (1995): The relationship between methodological quality and conclusions in reviews of spinal manipulation. *Journal of the American Medical Association* 274: 1942–1948.

*Augustinus L, Noordzij D, Verkerk K and De Visser D (2000): *Systematische Review Naar Actieve Bewegingsvormen Bij Reumatoide Arthritis.* Rotterdam: Hogeschool Rotterdam (in Dutch).

*Baker NA and Tickle-Degnen L (2001): The effectiveness of physical, psychological and functional interventions in treating clients with multiple sclerosis: a meta-analysis. *American Journal of Occupational Therapy* 55: 324–331.

*Beckerman H, Bouter LM, van der Heijden GJMG, De Bie RA and Koes BS (1993a): Efficacy of physiotherapy for musculoskeletal disorders: What can we learn from research? *British Journal of General Practice* 43: 73–77.

*Beckerman H, Bouter LM, van der Heijden GJMG, De Bie RA and Koes BS (1993b): The efficacy of physiotherapy for musculoskeletal disorders – Overview of the current state of knowledge. *European Journal of Physical Medicine and Rehabilitation* 3: 236–241.

*Bekkering GE, Cambach W, Hendriks HJM, Chadwick-Straver RVM, Gosselink R, Jongmans M, Paterson WJ, van der Schans CP, Verhoef-de Wijk MCE and Decramer M (1998): Revalidatie bij patiënten met COPD: een systematische literatuurstudie. *Nederlands Tijdschrift voor Fysiotherapie* 108: 120–127 (in Dutch).

Bendix AF, Bendix T, Ostenfeld S, Bush E and Andersen A (1995): Active treatment programs for patients with chronic low back pain: A prospective, randomised, observer-blinded study. *European Spine Journal* 4: 148–152.

*Bouter LM, Beckerman H, van der Heijden GM, Koes BW and De Bie RA (1992): [Effectiveness of physiotherapy; a summary of 9 meta-analyses.] *Nederlands Tijdschrift voor Geneeskunde* 136: 1058–1061 (in Dutch).

*Boyd S, Brooks D, Agnew-Coughlin J and Ashwell J (1994): Evaluation of the literature on the effectiveness of physical therapy modalities in the management of children with cystic fibrosis. *Pediatric Physical Therapy* 6: 70–74.

*Bradley J and Moran F (2002): Physical training for cystic fibrosis. The Cochrane Library, Issue 2. Oxford: Update Software.

*Bradley J, Moran F and Greenstone M (2002): Physical training for bronchiectasis. The Cochrane Library, Issue 3. Oxford: Update Software.

*Brandsma JW, Robeer BG, van den Heuvel S, Smit B, Wittens CHA and Oostendorp RAB (1998): The effect of exercises on walking distance of patients with intermittent claudication: A study of randomized clinical trials. *Physical Therapy* 78: 278-288

*Cambach W, Wagenaar RC, Koelman TW, van Keimpema T and Kemper HCG (1999): The long-term effects of pulmonary rehabilitation in patients with asthma and chronic obstructive pulmonary disease: A research synthesis. *Archives of Physical Medicine and Rehabilitation* 80: 103–111.

Chartered Society of Physiotherapy (2001): Neurology, Parkinson's disease, multiple sclerosis and severe traumatic brain injury. *Physiotherapy Effectiveness Bulletin* 3: 1–11.

*Chavannes NH and Vollenberg JJH (2002): Effects of physical activity in mild to moderate COPD: A systematic review. *British Journal of General Practice* 52: 574–578.

Comella CL, Stebbins GT, Brown-Toms N and Goetz CG (1994): Physical therapy and Parkinson's Disease: A controlled clinical trial. *Neurology* 44: 376–378.

Creasy TS, McMillan PJ, Fletcher EWL, Collin J and Morris PJ (1990): Is percutaneous transluminal angioplasty better than exercise for claudication? Preliminary results from a prospective randomised trial. *European Journal of Vascular Surgery* 4: 135–140.

*Crossley K, Bennell K, Green S and McConnell J (2001): A systematic review of physical interventions for patellofemoral pain syndrome. *Clinical Journal of Sport Medicine* 11: 103–110.

*Dagfinrud H and Hagen K (2002): Physiotherapy interventions for ankylosing spondylitis. The Cochrane Library, Issue 2. Oxford: Update Software.

*Deane KHO, Ellis-Hill C, Clarke CE, Playford ED and Ben-Shlomo Y (2002a): Physiotherapy for Parkinson's disease: A comparison of techniques. The Cochrane Library, Issue 2. Oxford: Update Software.

*Deane KHO, Jones D, Playford ED, Ben-Shlomo Y and Clarke CE (2002b): Physiotherapy versus placebo or no intervention in Parkinson's disease. The Cochrane Library, Issue 2. Oxford: Software Update.

*Deane KHO, Ellis-Hill C, Jones D, Whurr R, Ben-Shlomo Y, Playford D and Clarke CE (2002c): Systematic review of paramedical therapies for Parkinson's disease. *Movement Disorders* 17: 984–991.

*de Bie RA, Schoppink EM and Kool JP (1995): Effectiviteit van oefentherapie bij CVA-patiënten: een meta-analyse van de literatuur. *Jaarboek Fysiotherapie*, pp 197–212 (in Dutch).

*De Goede CJT, Keus SHJ, Kwakkel G and Wagenaar RC (2001): The effects of physical therapy in Parkinson's disease: A research synthesis. *Archives of Physical Medicine and Rehabilitation* 82: 509–515.

De Vet HCW, Kroese MEAL, Scholten RJPM and Bouter LM (2001): A method for research programming in the field of evidence-based medicine. *International Journal of Technology Assessment in Health Care* 17: 433–441.

*Devine EC and Pearcy J (1996): Meta-analysis of the effects of psychoeducational care in adults with chronic obstructive pulmonary disease. *Patient Education and Counseling* 29: 167–178.

*Di Fabio RP (1995): Efficacy of comprehensive rehabilitation programs and back school for patients with low back pain meta-analysis. *Physical Therapy* 75: 865–878.

*Dodd KJ, Taylor NF and Damiano DL (2002): A systematic review of the effectiveness of strength-training programs for people with cerebral palsy. *Archives of Physical Medicine and Rehabilitation* 83: 1157–1164.

Egger M, Smith GD and Altman DG (2001): *Systematic Reviews in Health Care: Meta-Analysis in Context* (2nd Ed). London: BMJ Publishing Group.

*Elders LAM, van der Beek AJ and Burdorf A (2000): Return to work after sickness absence due to back disorders—a systematic review on intervention strategies. *International Archives of Occupational and Environmental Health* 73: 339–348

*Ernst E (2000): Breathing techniques—adjunctive treatment modalities for asthma? A systematic review. *European Respiratory Journal* 15: 969-972.

Ernst E and Pittler MH (2001): Assessment of therapeutic safety in systematic reviews: Literature review. *BMJ* 323: 546.

*Faas A (1996): Exercises: Which ones are worth trying, for which patients, and when? *Spine* 21: 2874–2879.

*Feuerstein M, Burrell LM, Miller VI, Lincoln A, Huang GD and Berger R (1999): Clinical management of carpal tunnel syndrome: A 12-year review of outcomes. *American Journal of Industrial Medicine* 35: 232–245.

*Fransen M, McConnell S and Bell M (2002): Therapeutic exercise for people with osteoarthritis of the hip or knee. A systematic review. *Journal of Rheumatology* 29: 1737–1745.

*Gardner AW and Poehlman ET (1995): Exercise rehabilitation programs for the treatment of claudication pain. A meta-analysis. *Journal of the American Medical Association* 27: 975–980.

Gauthier L, Dalziel S and Gauthier S (1987): The benefits of group occupational therapy for patients with Parkinson's disease. *American Journal of Occupational Therapy* 41: 360–365.

*Girolami B, Bernardi E, Prins MH, Cate JW ten, Hettiarachchi R, Prandoni P, Girolami A and Büller HR (1999): Treatment of intermittent claudication with physical training, smoking cessation, pentoxifylline, or nafronyl. *Archives of Internal Medicine* 159: 337–345.

Gordon NF, Gulanick M, Costa F, Fletcher G, Franklin BA, Roth EJ and Shephard T (2004): Physical activity and exercise recommendations for stroke survivors. *Circulation* 109: 2031–2041.

*Gosselink HAAM and Wagenaar RC (1993a): Efficacy of breathing exercises in chronic obstructive pulmonary disease and asthma. A meta-analysis of the literature (part 1). *Journal of Rehabilitation Sciences* 6: 66–79.

*Gosselink HAAM and Wagenaar CR (1993b): Efficacy of breathing exercises in chronic obstructive pulmonary disease and asthma. A meta-analysis of literature (part 2). *Journal of Rehabilitation Sciences* 6: 105–113.

*Green S, Buchbinder R, Glazier R and Forbes A (2002): Interventions for shoulder pain. The Cochrane Library, Issue 2. Oxford: Update Software.

*Green S, Buchbinder R, Glazier R and Forbes A (1998): Systematic review of randomised controlled trials of interventions for painful shoulder: Selection criteria, outcome assessment, and efficacy. *BMJ* 316: 354–360.

*Gross AR, Aker PD, Goldsmith CH and Peloso P (2002): Physical medicine modalities for mechanical neck disorders. The Cochrane Library, Issue 1. Oxford: Update Software.

*Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E and Bombardier C (2001): Multidisciplinary rehabilitation for chronic low back pain: Systematic review. *BMJ* 322: 1511–1516.

Herbert RD, Maher CG, Moseley AM and Sherrington C (2001): Effective physiotherapy. *BMJ* 323: 788-790.

*Hilde G and Bo K (1998): Effect of exercise in the treatment of chronic low back pain: A systematic review, emphasising type and dose of exercise. *Physical Therapy Review* 3: 107–117.

*Hiraoka K (2001): Rehabilitation effort to improve upper extremity function in post-stroke patients: A meta-analysis. *Journal of Physical Therapy and Science* 13: 5–9.

*Holloway E and Ram FSF (2002): Breathing exercises for asthma. The Cochrane Library, Issue 1. Oxford: Update Software.

Hunt DL and McKibbon KA (1997): Locating and appraising

systematic reviews. *Annals of Internal Medicine* 126: 532–538.

*Hurwitz EL, Aker PD, Adams AH, Meeker WC and Shekelle PG (1996): Manipulation and mobilization of the cervical spine. A systematic review of the literature. *Spine* 21: 1746–1760.

Jadad RA, Cook DJ and Browman GP (1997): A guide to interpreting discordant systematic reviews. *Canadian Medical Association Journal* 156: 1411–1416.

*Johansson K, Öberg B, Adolfsson L and Foldevi M (2002): A combination of systematic review and clinicians' beliefs in interventions for subacromial pain. *British Journal of General Practice* 52: 145–152.

*Kjellman GV, Skargren EI and Öberg BE (1999): A critical analysis of randomised clinical trials on neck pain and treatment efficacy. A review of the literature. *Scandinavian Journal of Rehabilitation Medicine* 31: 139–152.

Klaber-Moffett JA, Chase SM, Portek I and Ennis JR (1986): A controlled prospective study to evaluate the effectiveness of a back school in the relief of chronic low-back pain. *Spine* 11: 120–122.

*Koes BW, Bouter LM, Beckerman H, van der Heijden GJMG and Knipschild PG (1991a): Oefentherapie bij rugklachten. Een geblindeerd literatuuronderzoek. *Nederlands Tijdschrift voor Fysiotherapie* 101: 229–234 (in Dutch).

*Koes BW, Bouter LM, Beckerman H, van der Heijden GJMG and Knipschild PG (1991b): Physiotherapy exercises and back pain: A blinded review. *BMJ* 302:1572–6.

*Koes BW, Bouter LM and van der Heijden GJMG (1995): Methodological quality of randomized clinical trials on treatment efficacy in low back pain. *Spine* 20: 228–235.

*Konijnenberg HS, De Wilde NS, Gerritsen AAM, van Tulder MW and De Vet HCW (2001): Conservative treatment for repetitive strain injury. *Scandinavian Journal of Work, Environment and Health* 27: 299–310.

Kraag G, Stokes B, Groh J, Helewa A and Goldsmith C (1990): The effects of comprehensive home physiotherapy and supervision on patients with ankylosing spondylitis—A randomised controlled trial. *Journal of Rheumatology* 17: 228–233.

*Kropmans ThJB, Dijkstra PU, Stegenda B and De Bont LGM (1999): Therapeutic outcome assessment in permanent temporomandibular joint disc displacement. *Journal of Oral Rehabilitation* 26: 357–363.

*Kwakkel G, Wagenaar RC, Koelman TW, Lankhorst GJ and Koetsier JC (1997): Effects of intensity of rehabilitation after stroke. A research synthesis. *Stroke* 28: 1550–1556.

*Lacasse Y, Wong E, Guyatt GH, King D, Cook DJ and Goldstein RS (1996): Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease. *Lancet* 348: 1114–1119.

*Lacasse Y, Guyatt GH and Goldstein RS (1997a): The components of a respiratory rehabilitation program. A systematic review. *Chest* 111: 1077–1088.

*Lacasse Y, Goldstein RS and Guyatt GH (1997b): Respiratory rehabilitation in chronic obstructive pulmonary disease: Summary of a systematic overview of the literature. *Reviews in Clinical Gerontology* 7: 327–347.

*Lacasse Y, Brosseau L, Milne S, Martin S, Wong E, Guyatt GH and Goldstein RS (2002): Pulmonary rehabilitation for chronic obstructive pulmonary disease. The Cochrane Library, Issue 3. Oxford: Update Software.

*Langhorne P, Wagenaar R and Partridge C (1996): Physiotherapy after stroke: More is better? *Physiotherapy Research International* 1: 75–87.

*Leng GC, Fowler B and Ernst E (2002): Exercise for intermittent claudication. The Cochrane Library, Issue 1. Oxford: Update Software.

*Lötters F, van Tol B, Kwakkel G and Gosselink R (2002): Effects of controlled inspiratory muscle training in patients with COPD: A meta-analysis. *European Respiratory Journal* 20: 570–576.

Lundgren F, Dahlhof AG, Lundholm K, Schersten T and Volkman R (1989): Intermittent claudication—surgical reconstruction or physical training? A prospective randomised trial of treatment efficiency. *Annals of Surgery* 209: 346–355.

*Maier-Riehle B and Härter M (2001): The effects of back schools—A meta-analysis. *International Journal of Rehabilitation Research* 24: 199–206.

Mannarino E, Pasqualini L, Innocente S, Scricciolo V, Rignanese A and Ciuffetti G (1991): Physical training and antiplatelet treatment in stage II peripheral arterial occlusive disease: Alone or combined? *Angiology* 42: 513–521.

*McCarthy CJ and Oldham JA (1999): The effectiveness of exercise in the treatment of osteoarthritic knees: A critical review. *Physical Therapy Reviews* 4: 241–250.

*Neill R (1999): A meta-analysis of the treatment of intermittent claudication. *Journal of Family Practice* 48: 418–419.

*Nieuwboer A, De Weerdt W, Nuyens G, Hantson L and Feys H (1994): Een literatuurstudie naar de effecten van fysiotherapie bij de ziekte van Parkinson. *Nederlands Tijdschrift voor Fysiotherapie* 5: 122–128 (in Dutch).

*Ottenbacher K and Di Fabio RP (1985): Efficacy of spinal manipulation/mobilization therapy. A meta-analysis. *Spine* 10: 833–837.

*Ottenbacher KJ and Jannell S (1993): The results of clinical trials in stroke rehabilitation research. *Archives of Neurology* 50: 37–44.

Patti F, Reggio A, Nicoletti F, Sellaroli T, Deinite G and Nicoletti F (1996): Effects of rehabilitation therapy in Parkinson's disability and functional independence. *Journal of Neurologic Rehabilitation* 10: 223–231.

Peacock PD, Riley CHP, Lampton TD, Raffel SS and Walker JS (1972): The Birmingham Stroke Epidemiology and Rehabilitation Study. In: Stewart GT (Ed.): *Trends in Epidemiology*. Springfield, Ill: Charles C Thomas, pp. 231–345.

*Pedro-Cuesta de J, Widen-Holmqvist L and Bach-y Rita P (1992): Evaluation of stroke rehabilitation by randomised controlled studies: A review. *Acta Neurologica Scandinavica* 86: 433–439.

*Pendleton A, Arden N, Dougados M, Doherty M, Bannwarth B, Bijlsma JWJ, Cluzeau F, Cooper C, Dieppe PA, Günther K-P, Hauselmann HJ, Herrero-Beaumont G, Kaklamanis PM, Leeb B, Lequesne M, Lohmander S, Mazieres B, Mola E-M, Pavelka K, Serni U, Swoboda B, Verbruggen AA, Weseloh G and Zimmermannn-Gorska I (2000): EULAR recommendations for the management of knee osteoarthritis: Report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Annals of the Rheumatic Diseases* 59: 936–944.

*Petrella RJ (2000): Is exercise effective treatment for osteoarthritis of the knee? *British Journal of Sports Medicine* 34: 326–331.

*Philadelphia Panel (2001a): Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for knee pain. *Physical Therapy* 81: 1675–1700.

*Philadelphia Panel (2001b): Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for neck pain. *Physical Therapy* 81: 1701–1717.

*Philadelphia Panel (2001c): Philadelphia panel evidence-based clinical practice guidelines on selected rehabilitation interventions for shoulder pain. *Physical Therapy* 81: 1719–1730.

*Philadelphia Panel (2001d): Philadelphia panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. *Physical Therapy* 81: 1641–1674.

Pina IL, Apstein CS, Balady GJ, Belardinelli R, Chaitman BR, Duscha BD, Fletcher BJ, Fleg JL, Myers JN and Sullivan MJ (2003): Exercise and heart failure. A statement from the American Association Committee on Exercise, Rehabilitation,

and Prevention. *Circulation* 107: 1210–1225.

*Pomeroy VM and Tallis RC (2000): Physical therapy to improve movement performance and functional ability poststroke. Part 1. Existing evidence. *Clinical Gerontology* 10: 261–290.

*Puett DW and Griffin MR (1994): Published trials of non-medicinal and noninvasive therapies for hip and knee osteoarthritis. *Annals of Internal Medicine* 121: 133–140.

*Radack K and Wyderski RJ (1990): Conservative management of intermittent claudication. *Annals of Internal Medicine* 113: 135–146.

*Ram FSF, Robinson SM and Black PN (2002): Physical training for asthma. The Cochrane Library, Issue 1. Oxford: Update Software.

*Ram FSF, Robinson SM and Black PN (2000): Effects of physical training in asthma: A systematic review. *British Journal of Sports Medicine* 43: 162–167.

*Ries AL, Carlin BW, Carrieri-Kohlman V, Casaburi R and Celli BR (1997): Pulmonary rehabilitation. Joint ACCP/AACVPR evidence-based guidelines. *Chest* 112: 1363–1396.

*Robeer GG, Brandsma JW, Heuvel SP van den, Smit B, Oostendorp RAB and Wittens CHA (1998): Exercise therapy for intermittent claudication; a review of the quality of randomised clinical trials and evaluation of predictive factors. *European Journal of Vascular and Endovascular Surgery* 15: 36–43.

*Rossy LA, Buckelew MA, Buckelew SP, Dorr N, Hagglund KJ, Thayer JF, McIntosh MJ, Hewett JE and Johnson JC (1999): A meta-analysis of fibromyalgia treatment interventions. *Annals of Behavioral Medicine* 21: 180–191.

*Rubinstein TC. Giladi N and Hausdorff JM (2002): The power of cueing to circumvent dopamine deficits: A review of physical therapy treatment of gait disturbances in Parkinson's disease. *Movement Disorders* 17: 1148–1160.

*Scheer SJ, Radack KL and O'Brian DR (1995): Randomized controlled trials in industrial low back pain relating to return to work. Part 1. Acute interventions. *Archives of Physical Medicine and Rehabilitation* 76: 966–973.

*Scheer SJ, Watanabe TK and Radack KL (1997): Randomized controlled trials in industrial low back pain. Part 3. Subacute/chronic pain interventions. *Archives of Physical Medicine and Rehabilitation* 78: 414–423.

*Schoppink EM, Kool JP and De Bie RA (1996): Oefentherapie bij CVA-patiënten. Een meta-analyse. *Nederlands Tijdschrift voor Fysiotherapie* 2: 41–51 (in Dutch).

Sivenius J, Pyorala K, Heinonen OP, Salonen JT and Riekkinen P (1985): The significance of intensity of rehabilitation of stroke: A controlled trial. *Stroke* 16: 928–931.

*Smith K, Cook D, Guyatt GH, Madhavan J and Oxman AD (1992): Respiratory muscle training in chronic airflow limitation: A meta-analysis. *American Review of Respiratory Disease* 145: 533–539.

Smith DS, Goldberg E, Ashburn A, Kinsella G, Sheikh K, Brennan PJ, Meade TW, Zutshi DW, Perry JD and Reeback JS (1981): Remedial therapy after stroke: A randomised controlled trial. *BMJ* 282: 5–9.

*Smith D, McMurray N and Disler P (2002): Early intervention for acute back injury: Can we finally develop an evidence-based approach? *Clinical Rehabilitation* 16: 1–11.

*Snels IAK, Dekker JHM, van der Lee JH, Lankhorst GJ, Beckerman H and Bouter LM (2002): Treating patients with hemiplegic shoulder pain. *American Journal of Physical Medicine & Rehabilitation* 81: 150–160.

*Stiller K and Huff N (1999): Respiratory muscle training for tetraplegic patients: A literature review. *Australian Journal of Physiotherapy* 45: 291–299.

Tan JC and Horn SE (1998): Practical Manual of Physical Medicine and Rehabilitation: Diagnostics, Therapeutics, and Basic Problems. St Louis: Mosby, pp. 156–177.

*Thomas J, Cook DJ and Brooks D (1995): Chest physical therapy management of patients with cystic fibrosis. A meta-analysis. *American Journal of Respiratory and Critical Care Medicine* 151: 846–850.

*van Baar ME, Assendelft WJJ and Dekker J (2001): Effectiviteit van oefentherapie als behandeling voor patiënten met artrose van heup of knie: Een systematisch literatuuroverzicht van gerandomiseerde klinische trials. *Nederlands Tijdschrift voor Fysiotherapie* 111: 56–66 (in Dutch).

*van Baar ME, Assendelft WJJ, Dekker J, Oostendorp RAB and Bijlsma JWJ (1999): Effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: A systematic review of randomized clinical trials. *Arthritis and Rheumatism* 42: 1361–1369.

*van Baar ME, Assendelft WJJ, Dekker J, Oostendorp RAB and Bijlsma JWJ (1998a): Exercise therapy is effective in patients with osteoarthritis of hip or knee: A systematic review of randomised clinical trials. In: Effectiveness of exercise therapy in osteoarthritis of hip or knee (Thesis). Utrecht: University Utrecht, pp. 107–131.

van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Lemmens JA and Bijlsma JW (1998b): The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: A randomized clinical trial. *Journal of Rheumatology* 25: 2432–2439.

*van den Ende CHM, Vliet Vlieland TPM, Munneke M and Hazes JMW (2002): Dynamic exercise therapy for rheumatoid arthritis. The Cochrane Library, Issue 1. Oxford: Update Software.

*van den Ende CHM, Vliet Vlieland TPM, Munneke M and Hazes JMW (1998): Dynamic exercise therapy in rheumatoid arthritis: A systematic review. *British Journal of Rheumatology* 37: 677–687.

*van der Heijden GJMG, Beurskens AJHM, Koes BW, Assendelft WJJ, De Vet HCW and Bouter LM (1995): The efficacy of traction for back and neck pain: A systematic, blinded review of randomized clinical trial methods. *Physical Therapy* 75: 93–104.

*van der Heijden GJMG, van der Windt DAWM and De Winter AF (1997): Physiotherapy for patients with soft tissue shoulder disorders: A systematic review of randomised clinical trials. *BMJ* 315: 25–30.

*van der Lee JH (2001): Constraint-induced therapy for stroke: More of the same or something completely different? *Current Opinion in Neurology* 14: 741–744.

*van der Lee JH, Snels IAK, Beckerman H and Lankhorst GJ (2001): Exercise therapy for arm function in stroke patients: A systematic review of randomized controlled trials. *Clinical Rehabilitation* 15: 20–31.

*van Duijvenbode I (1999): Oefentherapeutische behandeling helpt bij (sub)chronische lage rugpijn: Een meta-analyse. *Tijdschrift voor Oefentherapie Mensendieck* 2: 34–40 (in Dutch).

*van Duijvenbode ICD (1996): Het effect van therapeutische oefeningen bij lage rugpijn: Een meta-analyse (Thesis). Amsterdam: Faculteit der Bewegingswetenschappen VU Amsterdam (in Dutch).

*van Tulder MW (1999): Oefentherapie bij acute en chronische lage rugpijn: Een systematisch literatuuroverzicht. *Tijdschrift voor Oefentherapie Mensendieck* 4: 12–19 (in Dutch).

*van Tulder MW, Koes BW and Bouter LM (1997): Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine* 22: 2128–2156.

*van Tulder MW, Esmail R, Bombardier C and Koes BW (2002a): Back schools for non-specific low back pain. The Cochrane Library, Issue 1. Oxford: Update Software.

*van Tulder MW, Malmivaara A, Esmail R and Koes BW (2002b): Exercise therapy for low back pain. The Cochrane Library, Issue 1. Oxford: Update Software.

*van Tulder MW, Malmivaara A, Esmail R and Koes B (2000a):

Exercise therapy for low back pain. A systematic review within the framework of the Cochrane collaboration back review group. *Spine* 25: 2784–2796.

*van Tulder MW, Ostelo R, Vlaeyen JW, Linton SJ, Morley SJ and Assendelft WJ (2000b): Behavioral treatment for chronic low back pain. A systematic review within the framework of the Cochrane back review group. *Spine* 25: 2688–2699.

*Weinhardt C, Heller KD and Weh L (2001): Konservative Therapie des chronischen Rückenschmerzes: Spezifisches Krafttraining der Rückenmuskulatur oder Steigerung der allgemeinen körperlichen Fitness. *Zeitschrift fur Orthopadie und ihre Grenzgebiete* 139: 490–495 (in German).

Werner RA and Kessler S (1996): Effectiveness of an intensive outpatient rehabilitation program for postacute stroke patients. *American Journal of Physical Medicine and Rehabilitation* 75: 114–120.

Woolf AD, Zeidler H, Haglund U, Carr AJ, Chaussade S, Cucinotta D, Veale DJ and Martin-Mola E (2004): Musculoskeletal pain in Europe; its impact and a comparison of population and medical perceptions of treatment in eight European countries. *Annals of the Rheumatic Diseases* 63: 342–347.

*Zomerdijk TE, Beetsma AF, Dekker J, van Wijck R and van Horn JR (1998). Conservatieve behandeling van het patellofemoraal pijnsyndroom—Een systematisch literatuuronderzoek. *Nederlands Tijdschrift voor Fysiotherapie* 108: 95–102 (Dutch).

Appendix 1. Criteria for the assessment of the quality of the systematic reviews.

Criteria			Maximal points
Study selection			*(30)*
A	Description of inclusion and exclusion criteria of the systematic review		
	1	Study setting(s) included (i.e. industry, general practice, hospital)	2
	2	Interventions type(s) included	2
	3	Outcome type(s) included (i.e. pain, general improvement, disability questionnaire)	2
	4	Years covered	2
	5	Language(s) covered	2
B	Search strategy		
	1	Established bibliographic database included (Medline (or PubMed), and at least one other database)	5
	2	Additional efforts to locate non-indexed randomised clinical trials (RCTs) (e.g. reference tracking, correspondence with experts, manual search of non-indexed journals)	5
C	Emphasis on RCTs: RCTs only, or results or RCTs discussed separately from other study designs		10
Methodological Quality Assessment			*(20)*
D	Assessment (of the validity) of RCTs included that is explicit (reproducible by readers of the review) regarding:		
	1	Similarity of treatment groups at baseline (prognostic factors)	2
	2	Similarity of treatment characteristics (co-interventions)	2
	3	Adequacy of treatment of missing values (dropouts, loss to follow-up)	2
	4	Blinding of outcome assessment	2
	5	Relevance of outcome measures	2
	6	Adequacy of statistical analysis (i.e. intention-to-treat analysis)	2
E	Number of reviewers (at least two independent reviewers)		4
F	Blinding of reviewer(s): (blinded for source of article: journal, year of the trial, publication, institute)		2
G	Agreement of reviewer(s): reported (quantitatively in percentage agreement or Kappa statistics) and acceptable (cut-off Kappa statistics > 0.60, where Kappa statistics is not reported look at percentage agreement, which should be at least 80%). In the event of reviewer, use of an assessment list with established reliability.		2
Intervention			*(15)*
H	Description of (index) intervention(s) (exercises) per RCT		
	1	Description of therapeutic exercise (i.e. strength, endurance and cardiovascular fitness, mobility and flexibility, stability, relaxation, coordination, balance, and functional skills)	3
	2	Profession or training of care provider	1
	3	Treatment frequency or number of treatments	2
	4	Duration of treatment period	2
I	Description of control intervention(s): per RCT		
	1	Type (e.g. conservative treatments, wait-and-see policy, surgery)	3
	2	Treatment frequency or number of treatments	2
	3	Duration of treatment period	2
Data Presentation			*(20)*
J	Outcome presentation (for the most important (clinical relevant) outcome measures)		
	1	The original data of the main outcome(s) are presented separately per RCT per group	5
	2	Presentation of the mean difference (effect size, standardised mean differences, weighted mean differences) or ratio of outcome(s) (relative risk, risk difference, odds ratio) between intervention group(s) and control group(s)	3
	3	Presence of confidence interval (i.e. 95% CI) or standard deviation (SD) per RCT	3
	4	Graphic presentation of the most important outcome(s) (indicating outliers and distribution) per RCT (presentation of a tree plot, meta-analysis)	3

K		Adequate summary of research findings: statistical pooling of the most important outcome(s); discussion of the reason why pooling is not indicated or warranted; or pooling of the subset considered to be valid and similar enough	3
L		Discussion of the power of negative RCTs	
	1	Calculation (quantitative) of the power of each RCT	3
		or	
	2	Narrative elaboration (qualitative) on the power of each negative RCT	2
		or	
	3	Overall narrative elaboration on the power of the negative RCTs (i.e. remarks about small sample sizes)	1
Evaluation			*(15)*
M		Overall conclusion regarding the aggregated level of available RCTs on the effectiveness of the (index) intervention presented	5
N		Discussion of heterogeneity of RCTs and outcomes	
	1	Identification of relevant subgroups (e.g. age, study setting, disease classification) with explicit motivation	4
	2	Discussion of variety of treatment modalities in the intervention groups (i.e. high dose exercises)	2
	3	Discussion of variety of treatment modalities in control groups (placebo, existing modality)	2
	4	Discussion of relationship between methodological quality of RCTs and outcome	2
Total			100

Teixeira et al. International Journal of Behavioral Nutrition and Physical Activity 2012, 9:78
http://www.ijbnpa.org/content/9/1/78

INTERNATIONAL JOURNAL OF BEHAVIORAL NUTRITION AND PHYSICAL ACTIVITY

REVIEW Open Access

Exercise, physical activity, and self-determination theory: A systematic review

Pedro J Teixeira[1*], Eliana V Carraça[1], David Markland[2], Marlene N Silva[1] and Richard M Ryan[3]

Abstract

Background: Motivation is a critical factor in supporting sustained exercise, which in turn is associated with important health outcomes. Accordingly, research on exercise motivation from the perspective of self-determination theory (SDT) has grown considerably in recent years. Previous reviews have been mostly narrative and theoretical. Aiming at a more comprehensive review of empirical data, this article examines the empirical literature on the relations between key SDT-based constructs and exercise and physical activity behavioral outcomes.

Methods: This systematic review includes 66 empirical studies published up to June 2011, including experimental, cross-sectional, and prospective studies that have measured exercise causality orientations, autonomy/need support and need satisfaction, exercise motives (or goal contents), and exercise self-regulations and motivation. We also studied SDT-based interventions aimed at increasing exercise behavior. In all studies, actual or self-reported exercise/physical activity, including attendance, was analyzed as the dependent variable. Findings are summarized based on quantitative analysis of the evidence.

Results: The results show consistent support for a positive relation between more autonomous forms of motivation and exercise, with a trend towards identified regulation predicting initial/short-term adoption more strongly than intrinsic motivation, and intrinsic motivation being more predictive of long-term exercise adherence. The literature is also consistent in that competence satisfaction and more intrinsic motives positively predict exercise participation across a range of samples and settings. Mixed evidence was found concerning the role of other types of motives (e.g., health/fitness and body-related), and also the specific nature and consequences of introjected regulation. The majority of studies have employed descriptive (i.e., non-experimental) designs but similar results are found across cross-sectional, prospective, and experimental designs.

Conclusion: Overall, the literature provides good evidence for the value of SDT in understanding exercise behavior, demonstrating the importance of autonomous (identified and intrinsic) regulations in fostering physical activity. Nevertheless, there remain some inconsistencies and mixed evidence with regard to the relations between specific SDT constructs and exercise. Particular limitations concerning the different associations explored in the literature are discussed in the context of refining the application of SDT to exercise and physical activity promotion, and integrating these with avenues for future research.

Introduction

Physical activity and exercise, when undertaken regularly, are highly beneficial for health, and for physical and psychological well-being [e.g., [1]. Yet, only a minority of adults in modern societies reports engaging in physical exercise at a level compatible with most public health guidelines [2]. For instance, 2009 data indicate that, on a typical week, 60% of adults in Europe engaged in *no* physical exercise or sports [3]. In the US, less than 50% of adults are considered regularly physically active [4] while in Canada new accelerometer data shows that only 15% of adults meet national physical activity recommendations [5]. Such findings suggest that many people lack sufficient motivation to participate in the 150 minutes of moderately intense exercise or physical activity[a] per week recommended [6]. Indeed, approximately 40% of Europeans agree with the statement: "Being physically active does not really

* Correspondence: pteixeira@fmh.utl.pt
[1]Faculty of Human Kinetics, Technical University of Lisbon, Estrada da Costa, 1495-688, Cruz Quebrada, Portugal
Full list of author information is available at the end of the article

interest me – I would rather do other things with my spare time" [3].

Lack of motivation can broadly be explained by two orders of factors. First, as highlighted in the previous statistic, people may not be sufficiently interested in exercise, or value its outcomes enough to make it a priority in their lives [7]. Many individuals experience competing demands on their time from educational, career, and family obligations, possibly at the expense of time and resources that could be invested in exercising regularly. Second, some people may not feel sufficiently competent at physical activities, feeling either not physically fit enough or skilled enough to exercise, or they may have health limitations that present a barrier to activity [8]. Whether it be low interest or low perceived competence, the physical activity participation data indicate that many people are either unmotivated (or *amotivated*), having no intention to be more physically active, or are insufficiently motivated in the face of other interests or demands on their time.

In addition to those who are unmotivated, another source of short-lived persistence in exercise behaviors comes from people who do express personal motivation to exercise regularly, yet initiate exercise behaviors with little follow through. Specifically, a significant percentage of people may exercise because of *controlled* motivations, where participation in activities like going to the gym or running regularly is based on a feeling of "having to" rather than truly "wanting to" participate [7]. Controlled forms of motivation, which by definition are not *autonomous* (i.e., they lack volition), are predominant when the activity is perceived primarily as a means to an end and are typically associated with motives or goals such as improving appearance or receiving a tangible reward [9]. One hypothesis then is that the stability of one's motivation is at least partially dependent on some of its qualitative features, particularly the degree of perceived autonomy or of an *internal perceived locus of causality* [10]. That is, the level of reflective self-endorsement and willingness associated with a behavior or class of behaviors should be associated with greater persistence. An utilitarian approach to exercise (and to exercise motivation), such as might be prevalent in fitness clubs or other settings where exercise is externally prescribed, could thus be partially responsible for the high dropout rate observed in exercise studies [e.g., [11]. In fact, the pervasiveness of social and medical pressures toward weight loss, combined with externally prescriptive methods may be ill-suited to promote sustained increases in population physical activity levels.

In sum, large numbers of individuals are either unmotivated or not sufficiently motivated to be physically active, or are motivated by types of externally-driven motivation that may not lead to sustained activity. This highlights the need to look more closely at goals and self-regulatory features associated with regular participation in exercise and physical activity. Self-determination theory (SDT) is uniquely placed among theories of human motivation to examine the differential effects of qualitatively different types of motivation that can underlie behavior [12]. Originating from a humanistic perspective, hence fundamentally centered on the fulfillment of needs, self-actualization, and the realization of human potential, SDT is a comprehensive and evolving macro-theory of human personality and motivated behavior [12]. In what follows we will briefly describe key concepts formulated within SDT (and tested in SDT empirical studies) that are more relevant to physical activity and exercise, all of which will be implicated in our empirical review.

First, SDT distinguishes between intrinsic and extrinsic types of motivation regulating one's behavior. *Intrinsic motivation* is defined as doing an activity because of its inherent satisfactions. When intrinsically motivated the person experiences feelings of enjoyment, the exercise of their skills, personal accomplishment, and excitement [13]. To different degrees, recreational sport and exercise can certainly be performed for the associated enjoyment or for the challenge of participating in an activity. In contrast to intrinsic motivation, *extrinsic motivation* refers to doing an activity for instrumental reasons, or to obtain some outcome separable from the activity *per se*. For example, when a person engages in an activity to gain a tangible or social reward or to avoid disapproval, they are extrinsically motivated. SDT, however, conceptualizes qualitatively different types of extrinsic motivation, that themselves differ in terms of their relative autonomy. Some extrinsic motives are relatively heteronomous, representing what in SDT are described as *controlled* forms of motivation. For example, externally regulated behaviors are those performed to comply with externally administered reward and punishment contingencies. Also controlled are extrinsic motivations based on introjected regulation, where behavior is driven by self-approval. Controlled forms of extrinsic motivation are expected within SDT to sometimes regulate (or motivate) short-term behavior, but not to sustain maintenance over time [14]. Yet not all extrinsic motives are controlled. When a person does an activity not because it is inherently fun or satisfying (intrinsic motivation), but rather because it is of personal value and utility, it can represent a more autonomous form of behavioral regulation. Specifically in SDT, identified and integrated forms of behavioral regulation are defined as those in which one's actions are self-endorsed because they are personally valued. Examples include exercising because one values its outcomes and desires to maintain good health [7]. Thus, in SDT, these different forms of motivation are conceptualized as lying

along a continuum from non-autonomous to completely autonomous forms of behavioral regulation.

Third, SDT introduces the concept of *basic psychological needs* as central to understanding both the satisfactions and supports necessary for high quality, autonomous forms of motivation. Specifically SDT argues that there are basic psychological needs for autonomy, competence, and relatedness, all of which are conceived as essential and universal nutriments to psychological health and the development of internal motivation. Satisfaction of these basic needs results in increased feelings of vitality and well-being [15]. Like any other activity, engaging in sports and exercise can be more or less conducive to having one's psychological needs realized [16]. For example, experiences of competence vary upon success or failure at challenging physical tasks or as a function of feedback from, for example, a fitness professional. Perceptions of personal connection (relatedness) with others (e.g., fellow members of a fitness class or weight loss program) can vary greatly as a function of the interpersonal environment. Feelings of autonomy (versus feeling controlled) differ as a function of communication styles in exercise settings. According to SDT, in fact, need fulfillment in any context is closely associated with the characteristics of that social milieu, that is, whether important others support the needs for autonomy (e.g., take the perspective of the client/patient, support their choices, minimize pressure), relatedness (e.g., create an empathetic and positive environment, show unconditional regard), and competence (e.g., limit negative feedback, provide optimally challenging tasks). The concept of *need support* is thus thought to largely explain individual differences in the development and enactment of motivation across the lifespan [12]. Consequently, the design of health behavior change interventions that enhance satisfaction of participants' basic needs is a matter of much interest in SDT studies, including in the area of exercise and physical activity [17,18].

More recently, *goal contents* have also been explored from an SDT perspective in relation to a range of behaviors, including exercise [e.g., [19,20]. It should be noted that most authors have referred to goal contents in exercise contexts as *motives*, or more specifically *participation motives* [e.g., [64,79]. Operationally both terms are identical and we will use them interchangeably herein. Whereas intrinsic motivation and the various forms of extrinsic motivation represent the regulatory processes underlying a behavior, motives or goal contents are the outcomes that individuals are pursuing by engaging in the behavior [12]. Goal contents are differentiated according to the extent to which their pursuit is likely to satisfy basic psychological needs. Specifically, SDT distinguishes *intrinsic goals* (e.g., seeking affiliation, personal growth, or health) as those thought to be more closely related to the fulfillment of basic psychological needs, from *extrinsic goals* (e.g., seeking power and influence, wealth, or social recognition) that are thought to be associated with "substitute needs" which are neither universal nor truly essential to well-being and personal development. Factor analytic studies have borne out this theoretical distinction, and a number of studies have shown the predicted differential consequences of intrinsic versus extrinsic goal importance [21,22]. Within the domain of exercise and physical activity, extrinsic goals (e.g., when exercise is performed primarily to improve appearance) or intrinsic goals (e.g., to challenge oneself or to improve/preserve health and well-being) can clearly be distinguished. It should be noted that different goals or motives towards a given activity often naturally co-exist in the same person, some being more intrinsic, some less. Similar to what occurs with motivational regulations (which can have more or less autonomous elements, see more below), it is the relative preponderance of certain types of motives versus others which is thought to determine more or less desirable outcomes [e.g., [19,20].

Finally, SDT also proposes that people have dispositional tendencies, named *causality orientations* [14] which describe the way they preferentially orient towards their environments, resulting in characteristic motivational and behavioral patterns. Although some people may be more inclined to seek out and follow their internal indicators of preference in choosing their course of action, others may more naturally tend to align with external directives and norms, while still others may reveal to be generally amotivated, more passive, and unresponsive to either internal or external events that could energize their actions [12]. Although this topic has not been explored at length in previous research, these orientations can manifest themselves (and be measured) in exercise and physical activity contexts and the *Exercise Causality Orientation Scale* has been developed to measure individual differences in orientations around exercise [9].

Previous review papers of the topic of SDT and physical activity have primarily focused on describing the rationale for the application of this particular theoretical framework to physical activity behaviors, reviewing illustrative studies [7,23,24]. Meanwhile, the SDT-related exercise empirical research base has grown considerably in recent years, warranting a more comprehensive and systematic review of empirical data. Systematic reviews and meta-analyses of empirical studies provide the highest level of evidence for the appraisal and synthesis of findings from scientific studies. Accordingly, the present review includes 66 empirical studies published up to June 2011 that assessed relations between SDT-based constructs or interventions and exercise outcomes. We included experimental and cross-sectional studies that have measured

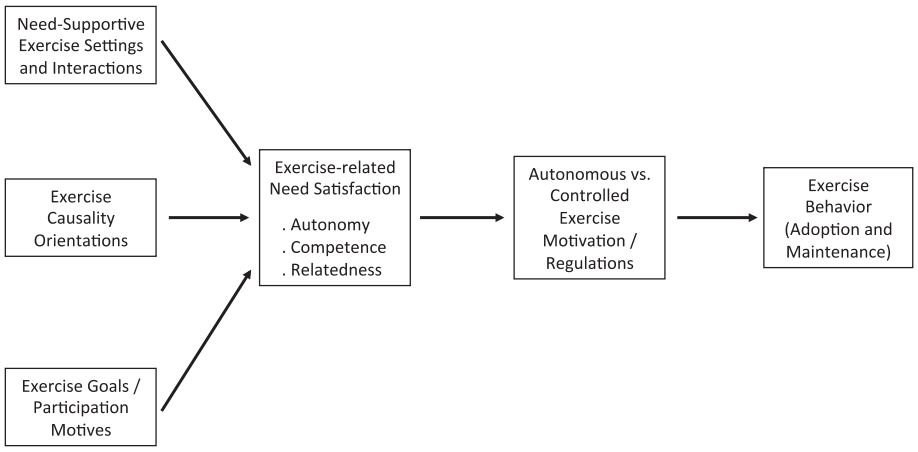


Figure 1 General SDT process model for exercise behavior. Adapted from the general health process model (Ref Ryan et al., Europ Health Psych, 2009), this graph includes the 5 groups of variables analyzed in this review as exercise predictors and their expected relationships (in a simplified version). Although this review only covers direct relationships between each class of variables (e.g., need satisfaction in exercise) and exercise behaviors, since few articles have simultaneously tested various steps of this model, the SDT model for exercise assumes that a sizable share of variance of exercise associated with SDT variables may be explained via indirect or mediating mechanisms, as depicted. See Discussion for more details.

exercise causality orientations, autonomy/need support and need satisfaction, exercise motives or goals, and exercise self-regulations and motivation. We also studied SDT-based interventions as predictors of exercise behavioral outcomes. Figure 1 depicts a general model of SDT and exercise, where its major constructs and theoretical links are highlighted.

Methods

Data sources and procedure

This review is limited to articles written in English and published in peer-reviewed journals covering adult samples. Research on autonomy and exercise in adolescents and children (typically based in school and physical education) was excluded, as well as studies with competitive athletic samples. Both are specific settings and were considered distinct from leisure-time or health-related exercise participation in adults, the focus of this review. The review includes both cross-sectional and longitudinal studies, investigating clinical and/or general population samples, and using diverse quantitative methodological approaches. A systematic literature search of studies published between 1960 and June 2011 was undertaken on the computerized psychological and sport databases PsycINFO and SportDiscus. The following strategy was used: TX (autonomous motivation OR autonomous regulation OR intrinsic motivation OR controlled regulation OR autonomy OR self-determination OR treatment regulations OR goals OR motives OR basic needs OR autonomy-supportive climate) AND TX (physical activity OR exercise OR exercise behavior OR leisure-time physical activity) Limiters were: Scholarly (peer-reviewed) journals; English Language; Adulthood (> 18 yr); Specific subjects: exercise OR motivation OR self-determination. This search yielded 660 articles. Abstracts were read and, of those, all potentially relevant full manuscripts were retrieved (n = 73). At this stage, studies were excluded which did not include either SDT variables or physical activity variables (accounting for most of the excluded studies), that used non-adult samples, and that reported achievement/performance outcomes related to PE classes. Next, reference lists of retrieved articles, previous review articles on the topic, and books were also reviewed, and manual searches were conducted in the databases and journals for authors who regularly publish in this area. This search yielded 11 additional manuscripts, totaling 84 potentially relevant manuscripts. Next, manuscripts were read and the following inclusion criteria used to select the final set of manuscripts: inclusion of non-athletic samples; outcomes included exercise/physical activity behaviors; reported direct associations between self-determination variables and physical activity outcomes. A total of 66 studies fulfilled all inclusion criteria and thus were included in this review. Of these, ten were experimental, eleven prospective, forty-two cross-sectional, and three used mixed designs.

Studies were initially coded with a bibliography number, but independent samples (*K*) were considered as the unit of analysis in the current review since a few studies used the same sample while other studies reported analyses on multiple samples. Data tables (Table 1) were constructed and encompassed sample characteristics of study populations, motivational predictors of exercise behavior, instruments of assessment, exercise-related outcomes, research designs, and statistical methods used to test the associations.

Table 1 Description of reviewed studies

Reference	Design	Sample			Measures	Significant Predictors	Outcomes	Analysis/Observations
		Size (%F)	Features	Location				
I. Exercise self-regulations and related measures								
ThØgersen-Ntoumani & Ntoumanis, 2006 [52]	Cross-sectional	375 (51)	Exercisers (Mean 38.7 yr)	UK	Exercise self-regulations (BREQ) + amotivation (AMS)	*MV*: IM (+) [a], ID (+) [a,b], INTR (+) [a]; EXT (–) [a,b], AMOT (–) [a]	Exercise stages of change [a]; Exercise relapses (fewer) [b]	Multivariate logistic regressions, adjusting for sex and age; Manovas
Rose et al., 2005 [56]	Cross-sectional	184 (55)	Healthy adults (17–60 yr)	UK	Exercise self-regulations (BREQ)	*MV*: IM (+) [a], ID (+), INTR (+), EXT (–)	Exercise stages of change	Discriminant function analysis (IM was redundant); Manovas [a]
Ingledew et al., 2009 [79]	Cross-sectional	251 (52)	University Students (Mean 19.5 yr)	UK	Exercise self-regulations (BREQ-2)	*MV*: IM (+), ID (+), INTR (n.s.), EXT (n.s.)	Self-reported exercise (measure analogous to LTEQ)	Partial Least Squares Analysis (PLS); Mediation analysis
Edmunds et al., 2006 [44]	Cross-sectional	369 (52)	Healthy individuals (Mean 31.9 yr)	UK	Exercise self-regulations (BREQ)	*MV*: IM (n.s.), ID (+), INTR (+), EXT (–) *BIV*: IM (+), ID (+), INTR (+), EXT (n.s.)	Self-reported exercise (total and strenuous PA; LTEQ)	Multiple regressions; Mediation analysis. No associations with mild/ moderately intense PA.
Wilson et al., 2006 [85]	Cross-sectional	139 (64)	Undergraduate students (Mean 19.5 yr)	Canada	Exercise extrinsic self-regulations (BREQ) and Integrated Regulation scale (INTEG)	*MV*: INTEG (+), ID (+), INTR (+), EXT (n.s.) *BIV*: INTEG (+), ID (n.s.), INTR (n.s.), EXT (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; Multiple regression analysis
McDonough et al., 2007 [50]	Cross-sectional	558 (72)			Exercise self-regulations (BREQ)	*MV*: RAI (+) *BIV*: RAI (+), IM (n.s.), ID (+), INTR (n.s.), EXT (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; SEM; Mediation analysis. Only RAI was tested in multivariate analysis.
Daley & Duda, 2006 [55]	Cross-sectional	409 (61)	Undergraduate students (19.9 yr)	UK	Exercise self-regulations (BREQ-2)	*MV*: IM (+), ID (++), INTR (+); EXT (– M); AMOT (– F)	Exercise stages of change; Physical activity status (from inactive to active)	Discriminant function analysis
Wilson et al., 2004 [45]	Cross-sectional	276 (64)	Undergraduate students (20.5 yr)	Canada	Exercise self-regulations (BREQ-2)	*MV*: IM (n.s.); ID (+), INTR (+ F; - M), EXT (n.s.), AMOT (n.s.) *BIV*: IM (+); ID (+), INTR (+ F), EXT (n.s.), AMOT (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; Multiple regressions analysis
Markland, 2009 [9]	Cross-sectional	102 F	Healthy individuals (Mean 29.2 yr)	UK	Exercise self-regulations (BREQ-2)	*MV*: IM (+), ID (+), AMOT (n.s.) *BIV*: IM (+), ID (+), INTR (+), EXT (n.s.), AMOT (–)	Self-reported exercise (LTEQ)	Bivariate correlations; Multiple regression/ mediation (Preacher & Hayes): INTR and EXT not analyzed.

Table 1 Description of reviewed studies *(Continued)*

Ingledew & Markland, 2008 [46]	Cross-sectional	252 (48)	Office workers (Mean 40 yr)	UK	Exercise self-regulations (BREQ-2)	*MV*: IM (n.s.), ID (+), INTR (n.s.), EXT (–) *BIV*: IM (+), ID (+), INTR (n.s.), EXT (–)	Self-reported exercise (measure analogous to LTEQ)	Bivariate correlations; SEM
Peddle et al., 2008 [43]	Cross-sectional	413 (46)	Colorectal cancer survivors (Mean 60 yr)	Canada	Exercise self-regulations (BREQ-2)	*MV*: IM (n.s.), ID (+), INTR (+), EXT (n.s.), AMOT (n.s.) *BIV*: IM (+), ID (+), INTR (+), EXT (n.s.), AMOT (–)	Self-reported exercise (LTEQ)	Bivariate correlations; Path analysis
Landry & Solmon, 2004 [86]	Cross-sectional	105 F	African-American (Mean 56 yr)	USA	Exercise self-regulations (BREQ)	*MV*: IM (+), ID (+), INTR (–), EXT (n.s.) *BIV*: RAI (+); IM (+), ID (+), INTR (n.s.), EXT (n.s.)	Exercise stages of change; exercise categories	Anovas; Discriminant function analysis
Milne et al., 2008 [87]	Cross-sectional	558 F	Breast cancer survivors (Mean 59 yr)	Australia	Exercise self-regulations (BREQ-2)	*MV*: IM (+), ID (+), INTR (n.s.), EXT (n.s.), AMOT (n.s.) *BIV*: IM (+), ID (+), INTR (n.s.), EXT (–), AMOT (–)	Self-reported exercise (LTEQ); exercise categories (meeting vs. not meeting guidelines)	Anovas; Hierarchical regression analysis
Mullan & Markland, 1997 [57]	Cross-sectional	314 (49.7)	Healthy individuals (Mean 35–40 yr)	UK	Exercise self-regulations (BREQ)	*MV*: IM (+), ID (+), INTR (n.s.), EXT (n.s.) *BIV*: RAI (+)	Exercise stages of change	Anova (RAI was analyzed); Discriminant function analysis;
Lewis & Sutton, 2011 [48]	Cross-sectional	100 (50)	95% undergraduates, members of a university gym; age not specified	UK	Exercise self-regulations (BREQ-2)	*MV*: IM (+); ID (n.s.), INTR (n.s.), EXT (–), AMOT (n.s.) *BIV*: IM (+); ID (+), INTR (+), EXT (–), AMOT (–)	Exercise frequency	Bivariate correlations; Multiple regression analysis
Markland & Tobin, 2010 [88]	Cross-sectional	133 F	Exercise referral scheme clients (Mean 54.5 yr)	UK	Exercise self-regulations (BREQ-2)	*BIV*: IM (+), ID (+), INTR (n.s.), EXT (n.s.), AMOT (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations
Wilson et al., 2002 [49]	Cross-sectional	500 (81)	Aerobic exercisers (Mean 34 yr)	Canada	Exercise self-regulations (BREQ)	*BIV*: IM (+), ID (+), INTR (+), EXT (–)	Self-reported exercise (LTEQ)	Bivariate correlations. Differences between PA intensities.
Sebire et al., 2009 [19]	Cross-sectional	410 (71)	Exercisers (Mean 41.4 yr)	UK	Exercise self-regulations (BREQ)	*MV*: RAI (+) *BIV*: RAI (+)	Self-reported exercise (LTEQ)	Bivariate correlations; Hierarchical regression analysis
Brickell & Chatzisarantis, 2007 [42]	Cross-sectional	252 (61)	College students (Mean 23.2 yr)	Canada	Exercise self-regulations (BREQ)	*MV*: IM (n.s.), ID (+), INTR (n.s.), EXT (n.s) *BIV*: IM (+), ID (+), INTR (+), EXT (n.s)	Self-reported exercise (LTEQ)	Multiple regression analysis
Edmunds et al., 2006 [51]	Cross-sectional	339 (53)	Symptomatic vs asymptomatic for exercise dependence (Mean 32.1 yr)	UK	Exercise self-regulations (BREQ) and Integrated Regulation scale (INTEG)	*MV*: Symptomatic: INTR (+ tendency); Asymptomatic: ID (+). Remaining variables not significant.	Self-reported exercise (total and strenuous PA; LTEQ)	Multiple regressions. No associations with moderately intense PA.

Table 1 Description of reviewed studies *(Continued)*

Moreno et al., 2007 [89]	Cross-sectional	561 (53)	Healthy adults (Mean 31.8 yr)	Spain	Exercise self-regulations (BREQ-2)	*MV*: IM (n.s.), ID (–), INTR (n.s.), EXT (–), AMOT (–)	Exercise duration (0-45 min vs. 45-60 min vs. > 60 min)	Manovas
Hall et al., 2010 [90]	Cross-sectional	470 (54)	Adults (Mean 44.9 yr)	Canada	Exercise self-regulations (BREQ-2); Self-reported exercise (LTEQ)	*BIV*: IM (+), ID (+), INTR (+), EXT (n.s.), AMOT (–)	Exercise status (active vs. inactive)	Anovas
Standage et al., 2008 [91]	Cross-sectional	52 (50)	University students (Mean 22 yr)	UK	Exercise self-regulations; Autonomous and controlled motivations (BREQ)	*MV*: AutMot (+), CtMot (n.s.) BIV: IM (+), ID (+), INTR (n.s.), EXT (n.s), AutMot (+), CtMot (n.s.)	Accelerometry	Bivariate correlations; Sequential regression analysis
Duncan et al., 2010 [41]	Cross-sectional	1079 (57)	Regular exercisers (Mean 24.2 yr)	Canada	Exercise self-regulations (BREQ-2) + Integrated reg. scale	*MV*: IM (n.s.), INTEG (+), ID (+)*, INTR (n.s.), EXT (n.s), AMOT (n.s) *BIV*: IM (+), INTEG (+), ID (+), INTR (+), EXT (– F)*, AMOT (–)	* PA frequency; PA intensity; PA duration (LTEQ)	Bivariate correlations; Multiple regression analysis
Sorensen et al. 2006 [54]	cross-sectional	109 (59)	Psychiatric patients (Mean age group 31–49 yr)	Norway	Exercise regulations (based on BREQ)	*MV*: IM (+), ID (n.s.), INTR (n.s.), EXT (n.s.) *BIV*: IM (+), ID (n.s.), INTR (n.s.), EXT (–)	Self-reported exercise level	Bivariate correlations; Logistic regressions
Puente & Anshel, 2010 [77]	Cross-sectional	238 (57)	College students (Mean 20.4 yr)	USA	Exercise self-regulations (SRQ-E)	*MV*: RAI (+) *BIV*: RAI (+)	Exercise frequency	Bivariate correlations; SEM
Halvary et al., 2009 [76]	Cross-sectional	190 (44)	Healthy adults (Mean 21.8 yr)	Norway	Autonomous motivation (SRQ)	*MV*: AutMot (+) *BIV*: AutMot (+)	Exercise frequency and duration	Bivariate correlations; SEM; Mediation analysis
Wilson et al., 2006 [29]	Cross-sectional	220; 220 (56)	Cancer survivors (Mean 60–64 yr) vs non-cancer (Mean 50 yr)	Canada	Autonomous and controlled motivation (TSRQ-PA)	*MV*: AutMot (+), CtMot (–) in both samples *BIV*: AutMot (+), CtMot (n.s.) in both samples	Self-reported exercise (min/wk of MVPA)	Bivariate correlations; Multiple regression analysis
Hurkmans et al., 2010 [92]	Cross-sectional	271 (66)	Patients with Rheumatoid Arthritis (Mean 62 yr)	Netherlands	Exercise self-regulations (TSRQ-PA). Adated RAI.	*MV*: RAI (+) *BIV*: RAI (+)	Self-reported exercise (SQUASH)	Bivariate correlations; Multiple regression analysis
Lutz et al., 2008 [93]	Cross-sectional	535 (60)	University students (Mean 20 yr)	USA	Exercise self-regulations (EMS). Adapted RAI.	*MV*: RAI (+) *BIV*: RAI (+)	Self-reported exercise (LTEQ)	Bivariate correlation; Preacher & Hayes mediation analysis

Table 1 Description of reviewed studies *(Continued)*

Wininger, 2007 [28]	Cross-sectional	143; 58 (76)	Undergraduates (Mean 21–22 yr)	USA	Exercise self-regulations (EMS)	*MV* *: IM (+), INTEG (+), ID (+), INTR (+), EXT (n.s.), AMOT (–) *BIV* **: IM experience sensations (+), INTEG (n.s.), ID (n.s.), INTR (n.s.), EXT (n.s.), AMOT (–)	* Exercise stages of change; ** Distance walked on treadmill	Bivariate correlations; Manovas
Craike, M., 2008 [47]	Cross-sectional	248 (53)	Healthy adults (Mean 48 yr)	Australia	Exercise self-regulations (based on BREQ and EMS)	*MV*: IM (+), ID (n.s.), INTR (n.s.), EXT (–)	Self-reported LTPA	SEM
Tsorbatzoudis et al., 2006 [94]	Cross-sectional	257 (55)	Healthy adults (Mean 31 yr)	Greece	Exercise self-regulations (SMS)	*MV*: IM (+), ID (+), INTR (+), EXT (–), AMOT (–)	Exercise frequency (from the least to the most frequent)	Multivariate analysis of variance; multiple regressions
Chatzisarantis & Biddle, 1998 [95]	Cross-sectional	102 (50)	University employees (Mean 40 yr)	UK	Behavioral regulations for PA (SMS adaptation)	*MV*: Autonomous group (vs controlled) based on RAI scores (+)	Self-reported exercise (LTEQ)	SEM
Matsumoto & Takenaka, 2004 [96]	Cross-sectional	486 (53)	Healthy individuals (Mean 45 yr)	Japan	Exercise self-regulations (SDMS); profiles of self-determination	*BIV*: IM (+), ID (+), INTR (+), EXT (n.s.) AMOT (–); Self-determined profile (+)	Exercise stages of change	Bivariate correlations and cluster analysis
McNeill et al., 2006 [97]	Cross-sectional	910 (80)	Healthy individuals (Mean 33 yr)	USA	Intrinsic and extrinsic motivations (MPA)	*MV*: Intrinsic motivation (+); Extrinsic motivation for social pressure	Self-reported exercise (minutes of walking, and MVPA)	SEM. Indirectly through self-efficacy.
Russell & Bray, 2009 [98]	Cross-sectional and prospective (6 + 6wk)	68 (13)	Cardiac rehabilitation outpatients (Mean 64.9 yr)	Canada	Exercise self-regulations (BREQ-2)	*MV*: RAI (+) *BIV*: RAI (+)	Self-reported exercise (7Day-PAR)	Bivariate correlations; Multiple regression analysis
Russell & Bray, 2010 [99]	Cross-sectional and Observational (14wk)	53 M	Exercise cardiac rehabilitation patients (Mean 62.8 yr)	Canada	Exercise self-regulations (SRQ-E)	*MV*: AutMot (+) *BIV*: AutMot (+), CtMot (n.s.)	Exercise frequency; duration (+); volume (+) – 7Day-PAR	Bivariate correlations; Hierarchical regression analysis
Fortier et al., 2009 [100]	Prospective (6mo)	149 F	Healthy adults (Mean 51.8 yr)	Canada	Exercise self-regulations (TSRQ-adapted)	*MV*: AutMot (n.s.) *BIV*: AutMot (n.s.), CtMot (n.s.)	Duration, Frequency, and Energy Expenditure (CHAMPS)	Bivariate correlations; Mediation/regression analysis
Rodgers et al., 2010 [31]	Prospective	1572 (60)	Initiate vs. long-term exercisers (Mean 22–51 yr)	Canada	Exercise self-regulations (BREQ)	*MV*: IM (+), ID (+), INTR (n.s.), EXT (–) overtime for initiates, but < to regular exercisers	Self-reported exercise (LTEQ); Initiate vs. long-term exercisers	Manovas. Total N from 6 samples: initiates (60, 134, 38, 84), regular exercisers (202, 1054)
Barbeau et al., 2009 [101]	Prospective (1mo)	118 (65)	Healthy adults (Mean 19 yr)	Canada	Exercise self-regulations (BREQ-2)	*MV*: AutMot (+), CtMot (n.s.) *BIV*: AutMot (+), CtMot (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; Path analysis
Hagger et al., 2006 [35]	Prospective (4wk)	261 (64)	University students (Mean 24.9 yr)	UK	Relative autonomy index (based on PLOC scale)	*BIV*: RAI (+)	Self-reported exercise (frequency)	Bivariate correlations; SEM

Table 1 Description of reviewed studies *(Continued)*

Hagger et al., 2006 [34]	Prospective (4 wk)	261 (64)	Exercise sample of university students (Mean 24.9 yr)	UK	Relative autonomy index (based on PLOC Scale)	*BIV*: RAI (+)	Self-reported exercise (frequency)	Bivariate correlations
Kwan et al., 2011[53]	Prospective (4 wk)	104 (58)	Undergraduate students; active (Mean 18.2 yr)	USA	Exercise self-regulations (BREQ-2)	*BIV*: IM (+), ID (n.s.), INTR (n.s.), EXT (n.s.), AMOT (n.s.), RAI (n.s)	Self-reported exercise (online diary)	Bivariate correlations
Edmunds et al., 2007 [38]	Prospective (uncontrolled intervention) (3mo)	49 (84)	Overweight/Obese patients (Mean BMI: 38.8; Mean 45 yr) on an exercise scheme	UK	Exercise self-regulations (BREQ-2); Integrated regulation subscale (EMS)	*MV*: IM (n.s.), INTEG (+), ID (–)*, INTR (+)*, EXT (n.s.) *BIV*: ID (+), INTR (–)	Self-reported exercise (LTEQ);	Bivariate correlations; Multilevel regression analysis.* ID and INTR multivariate outcomes resulted from net suppression; thus, not considered by the authors.
Wilson et al., 2003 [58]	Experimental (12wk)	53 (83)	Adults (Mean 41.8 yr; BMI: 19.9 ± 3.0 kg/m^2)	Canada	Exercise self-regulations (BREQ)	*MV*: IM (+), ID (+) *BIV*: IM (+), ID (+), INTR (n.s.), EXT (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; Multiple regression analysis. IM and ID increased from pre- to post-exercise program
Sweet et al., 2009 [102]	Experimental (12mo)	234 (38)	Inactive with type 2 diabetes (Mean 53 yr) on an exercise program	Canada	Exercise self-regulations (BREQ)	*MV*: AutMot (+) *BIV*: AutMot (+)	Amount of PA (kcal/month)	Bivariate correlations; Regression/Mediation analysis
Fortier et al., 2011 [36]	Experimental (13wk); RCT	120 (69)	Inactive patients (Mean 47.3 yr): intensive vs. brief PA intervention	Canada	Exercise self-regulations (BREQ-2)	*BIV*: IM, ID, INTR, EXT, and RAI were not significant predictors	Self-reported exercise (LTEQ)	Bivariate correlations
Fortier et al., 2007 [17]	Experimental (13wk); RCT	120 (69)	Autonomy supportive vs brief PA counseling (Mean 47.3 yr)	Canada	Treatment self-regulations (TSRQ-PA)	*MV*: AutMot (+) *BIV*: AutMot (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; Path/Mediation analysis
Levy & Cardinal, 2004 [40]	Experimental (2mo); RCT	185 (68)	Adults (Mean 46.8 yr); SDT-based mail intervention vs. controls	USA	Exercise self-regulations (EMS)	*MV*: IM, INTEG, ID, INTR, EXT, and AMOT were not significant predictors	Self-reported exercise (LTEQ)	Manovas with repeated measures
Mildestvedt et al., 2008 [68]	Experimental (4wk); RCT	176 (22)	Cardiac rehabilitation patients (Mean 56 yr): SDT-based vs standard rehab treatment	Norway	Autonomous and controlled motivations (TSRQ)	*BIV*: AutMot (+); CtMot (n.s.)	Self-reported exercise (composite score); exercise intensity	ANOVAs with repeated measures
Silva et al., 2010 [33]	Experimental (12mo); RCT	239 F	OW/Obese women (Mean 38 yr); SDT-treatment vs controls	Portugal	Exercise self-regulations (SRQ-E)	MV: IM (+)*, ID (n.s.), INTR (n.s.), EXT (n.s.) BIV: IM (+), ID (+), INTR (+), EXT (n.s.)	Self-reported exercise: MVPA * (7-day PAR); lifestyle PA index	Bivariate correlations; PLS analysis; Mediation analysis
Silva et al., 2010 [32]	Experimental (1 yr + 2y follow-up); RCT	221 F	OW/Obese women (Mean 38 yr); SDT-treatment vs controls	Portugal	Exercise self-regulations (SRQ-E) at 1 yr and 2 yr	MV: AutMot 2 yr (+), INTR 2 yr (n.s.), EXT 2 yr (n.s.) BIV: AutMot 1 and 2 yr (+), INTR 2 yr (+), EXT 2 yr (n.s.)	2-yr self-reported exercise: MVPA (7-day PAR)	Bivariate correlations; PLS analysis; Mediation analysis

Table 1 Description of reviewed studies *(Continued)*

II. Exercise-related psychological need satisfaction

Puente & Anshel, 2010 [77]	Cross-sectional	238 (57)	College students (Mean 20.4 yr)	USA	Basic Psychological Needs Scale (BPNS); Perceived Competence Scale (PCS)	*MV*: Competence (+) *BIV*: Autonomy (n.s.), Competence (+)	Exercise frequency	Bivariate correlations; SEM; Relatedness not measured.
Edmunds et al., 2006 [44]	Cross-sectional	369 (52)	Healthy individuals (Mean 31.9 yr)	UK	Psychological need satisfaction (BNSWS adapted)	*MV*: Autonomy (n.s.), Competence (+), Relatedness (n.s.) *BIV*: Autonomy (+), Competence (+), Relatedness (+)	Self-reported exercise (total and strenuous PA; LTEQ)	Bivariate correlations; Regression analysis; mediation analysis
Edmunds et al., 2006 [51]	Cross-sectional	339 (53)	Symptomatic vs asymptomatic for exercise dependence (Mean 32.1 yr)	UK	Psychological need satisfaction (BNSWS adapted)	*BIV*: Autonomy (n.s.), Competence (+), Relatedness (n.s.)	Self-reported exercise (total and strenuous PA; LTEQ)	Bivariate correlations. No associations with mild/moderately intense PA
Peddle et al., 2008 [43]	Cross-sectional	413 (46)	Colorectal cancer survivors (Mean 60 yr)	Canada	Psychological need satisfaction (PNSE)	*BIV*: Autonomy (+), Competence (+), Relatedness (+)	Self-reported exercise (LTEQ)	Bivariate correlations
McDonough et al., 2007 [50]	Cross-sectional	558 (72)	Recreational dragon boat paddlers (Mean 45 yr)	Canada	Exercise need satisfaction (PNSE)	*MV*: Autonomy (–), Competence (+) *BIV*: Autonomy (n.s.), Competence (+), Relatedness (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations; SEM
Sebire et al., 2009 [19]	Cross-sectional	410 (71)	Exercisers (Mean 41.4 yr)	UK	Exercise need satisfaction (PNSE)	*BIV*: Exercise need satisfaction (composite score) (+)	Self-reported exercise (LTEQ)	Bivariate correlations
Milne et al., 2008 [87]	Cross-sectional	558 F	Breast cancer survivors (Mean 59 yr)	Australia	Perceived competence (PCS)	*MV*: Competence (+) *BIV*: Competence (+)	Self-reported exercise (LTEQ); Exercise categories (meeting vs. not meeting guidelines)	Anovas; Hierarchical regression analysis
Halvary et al., 2009 [76]	Cross-sectional	190 (44)	Healthy adults (Mean 21.8 yr)	Norway	Perceived competence (PCS)	*MV*: Competence (n.s.) *BIV*: Competence (+)	Exercise frequency and duration	Bivariate correlations; SEM/Mediation analysis
Vlachopoulos & Michailidou, 2006 [103]	Cross-sectional	508 (50)	Greek adults (Mean 30 yr)	Greece	Psychological needs satisfaction in exercise (BPNES)	*MV*: Autonomy (n.s.), Competence (+); Relatedness (n.s.)	Exercise frequency	SEM
Markland & Tobin, 2010 [88]	Cross-sectional	133 F	Exercise referral scheme clients	UK	Autonomy need (LCE); Perceived Competence (IMI); Relatedness (8-item scale)	*BIV*: Autonomy (+), Competence (+), Relatedness (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations

Table 1 Description of reviewed studies *(Continued)*

Russell & Bray, 2009 [98]	Cross-sectional and prospective (6 + 6wk)	68 (13)	Cardiac rehabilitation outpatients (Mean 64.9 yr)	Canada	Exercise need satisfaction (PNSE)	*BIV*: Autonomy (n.s.), Competence (+)*, Relatedness (n.s.)	Self-reported exercise (7Day-PAR) at 3wk and 6wk* follow-up	Bivariate correlations
Barbeau et al., 2009 [101]	Prospective (1mo)	118 (65)	Healthy adults (Mean 19 yr)	Canada	Exercise need satisfaction (PNSES)	*BIV*: Autonomy (+), Competence (+), Relatedness (+)	Self-reported exercise (LTEQ)	Bivariate correlations
Hagger et al., 2006 [34]	Prospective (4 wk)	261 (64)	Exercise sample of university students (Mean 24.9 yr)	UK	Psychological need satisfaction	*BIV*: Psychological need satisfaction (composite score) (+)	Self-reported exercise (frequency).	Bivariate correlations
Edmunds et al., 2007 [38]	Prospective (uncontrolled intervention) (3mo)	49 (84)	OW/Obese patients (BMI: 38.75; Mean 45 yr)	UK	Psychological need satisfaction (PNSS)	*MV*: Autonomy (n.s.), Competence (n.s.); Relatedness (n.s.)	Self-reported exercise (LTEQ); (Increase in relatedness overtime)	Multilevel regression analysis; Paired T-tests
Fortier et al., 2007 [17]	Experimental (13 wk); RCT	120 (69)	Healthy adults (Mean 47.3 yr)	Canada	Perceived Competence (PCES)	*MV*: Competence (+)	Self-reported exercise (LTEQ)	Path analysis; Mediation analysis
Levy & Cardinal, 2004 [40]	Experimental (2mo); RCT	185 (68)	Adults (Mean 46.8 yr); SDT-based mail intervention vs. controls	USA	Perceived autonomy satisfaction (LCE)	*MV*: Autonomy (+ F), Competence (n.s.), Relatedness (n.s.)	Self-reported exercise (LTEQ)	Manovas with repeated measures
Silva et al., 2010 [33]	Experimental (12mo); RCT	239 F	OW/Obese women (Mean BMI: 31.5; Mean 38 y); SDT-based weight loss treatment vs controls	Portugal	Perceived autonomy satisfaction (LCE); Competence (IMI)	*BIV*: Autonomy (+), Competence (+)	Self-reported exercise: MVPA (7-day PAR); lifestyle PA index	Bivariate correlations
III. Exercise motives and related measures								
Ingledew et al., 2009 [79]	Cross-sectional	251 (52)	University Students (Mean 19.5 yr)	UK	Exercise motives (EMI-2)	*MV*: Intrinsic motives: Stress management (+), Affiliation (+), Challenge (+); Extrinsic: Health/ fitness (+); body-related (n.s.)	Self-reported exercise (measure analogous to LTEQ)	Partial Least Squares Analysis (PLS); Mediation analysis
Ingledew & Markland, 2008 [46]	Cross-sectional	252 (48)	Office workers (Mean 40 yr)	UK	Exercise motives (EMI-2)	*BIV*: Intrinsic motives (n.s.), Extrinsic motives: health/ fitness (+) and body-related (–)	Self-reported exercise (measure analogous to LTEQ)	Bivariate correlations
Frederick & Ryan, 1993 [59]	Cross-sectional	376 (64)	Healthy individuals (Mean 39 yr)	USA	Exercise motives (MPAM)	Intrinsic motives: interest/ enjoyment (+); competence (+); Extrinsic motives: body-related (+)	Self-reported exercise (levels, intensity)	Differences between PA categories; correlations and Manovas
Frederick et al., 1996 [104]	Cross-sectional	118 (68)	College students (Mean 22 yr)	USA	Exercise motives (MPAM-r)	*MV*: Extrinsic: body-related (+ M) *BIV*: Intrinsic motives (+ F), Extrinsic: body-related (+ M)	Self-reported exercise: frequency, volume	Bivariate correlations; Multiple regression analysis
Buckworth et al., 2007 [30] a	Cross-sectional	184;220 (60)	University students (Mean 18–22 yr)	USA	Exercise motives (EMI and IMI; total and subscales)	Intrinsic motives (except choice) (+); Extrinsic motives (except tangible rewards) (+)	Exercise stages of change	Anovas and profile analysis
Sebire et al., 2009 [19]	Cross-sectional	400 (73)	Exercisers (Mean 41.4 yr)	UK	Exercise goal content (GCEQ)	*MV*: Intrinsic motives (+) *BIV*: Intrinsic motives (+)	Self-reported exercise (LTEQ)	Bivariate correlations; Hierarchical regression analysis

Table 1 Description of reviewed studies *(Continued)*

Segar et al., 2006 [64]	Cross-sectional	59 F	Healthy adults (Mean 45.6 yr)	USA	Body and non-body shape motives for exercise (via inductive, qualitative methods)	*BIV*: Body motives (–); non-body shape motives (+).	Self-reported exercise (LTEQ)	Hierarchical regression analysis
Sit et al., 2008 [105]	Cross-sectional	360 F	Chinese adults (30–59 yr)	China	Exercise motives (MPAM-r)	*MV*: Intrinsic motives : competence/challenge (+), interest/enjoyment (+); Extrinsic: fitness/health (+); appearance (n.s.)	Exercise stages of change	Manovas
Davey et al., 2009 [106]	Cross-sectional	134 (66)	Employees (estimated mean age between 25–44 yr)	New Zeland	Exercise motives (based on MPAM-r and SMS)	*MV*: Intrinsic motives: enjoyment (+), competence/ challenge (+); Extrinsic: appearance (–); Fitness (n.s.)	Total number of steps in 3wk	Multiple regression analysis
Segar et al., 2008 [65]	Prospective	156 F	Healthy women (Mean 49.3 yr)	USA	Extrinsic and Intrinsic goals (based on a list of goals and on cluster analysis)	*MV*: Intrinsic goals (+); Extrinsic goals: weight maintenance/toning (–); health benefits (–)	Self-reported exercise (LTEQ)	Linear mixed model
Ingledew et al., 1998 [107]	Prospective (3mo)	425 (34)	Government employees (Mean 40 yr)	UK	Exercise motives (EMI-2)	*MV*: Intrinsic motives: enjoyment (+); Extrinsic: body-related (+ action; - maintenance); health pressures (+ preparation; - action/ maintenance)	Exercise stages of change	Discriminant function analysis
Ryan et al., 1997 [27] a	Prospective (10wk)	40 (80)	University students and employees (Mean 21 yr)	USA	Exercise motives (MPAM)	*MV*: Intrinsic motives: enjoyment (+), competence (+); body-related motives (n.s.)	Reduced dropout and attendance to exercise classes	Manovas and multiple regressions
Ryan et al., 1997 [27] b	Prospective (10wk)	155 (57)	New fitness center members (Mean 19.5 yr)	USA	Exercise motives (MPAM-R)	*MV*: Intrinsic motives: enjoyment (+), competence (+), social interactions (+); Extrinsic motives: fitness (n.s.), appearance (n.s.)	Attendance to and duration of exercise workout	Manovas and multiple regressions
Buckworth et al., 2007 [30] b	Experimental (10wk)	142 (66)	College Students (Mean 21.3 yr)	USA	Exercise motives (EMI and IMI);	*BIV*: Intrinsic motives: effort/competence (+) and interest/enjoyment (+); Extrinsic motives: appearance (+) *	Exercise patterns (from stable inactive to stable active); Activity vs. Lecture (no activity) Classes *	Anovas with repeated measures.
IV. Perceived need support								
Peddle et al., 2008 [43]	Cross-sectional	413 (46)	Colorectal cancer survivors (Mean 60 yr)	Canada	Perceived need support (PAS, based on HCCQ-short)	*BIV*: Need support (+)	Self-reported exercise (LTEQ)	Bivariate correlations
Milne et al., 2008 [87]	Cross-sectional	558 F	Breast cancer survivors (Mean 59 yr)	Australia	Perceived need support (mHCCQ)	*MV*: Need support (+)*BIV*: Need support (+)	Self-reported exercise (LTEQ); exercise categories (meeting vs. not meeting guidelines)	Anovas; Hierarchical regression analysis

Table 1 Description of reviewed studies *(Continued)*

Hurkmans et al., 2010 [92]	Cross-sectional	271 (66)	Patients with Rheumatoid Arthritis (Mean 62 yr)	Netherlands	Perceived need support (HCCQ-mod)	*MV*: Need support (n.s.) *BIV*: Need support (n.s.)	Self-reported PA (SQUASH)	Bivariate correlations; Multiple regression analysis
Halvary et al., 2009 [76]	Cross-sectional	190 (44)	Healthy adults (Mean 21.8 yr)	Norway	Perceived need support (SCQ based on HCCQ)	*BIV*: Need support (+)	Exercise frequency and duration	Bivariate correlations
Markland & Tobin, 2010 [88]	Cross-sectional	133 F	Exercise referral scheme clients	UK	Need support (15-item scale)	*BIV*: Need support (n.s.)	Self-reported exercise (LTEQ)	Bivariate correlations
Puente & Anshel, 2010 [77]	Cross-sectional	238 (57)	College students (Mean 20.4 yr)	USA	Exercise need support (SCQ)	*BIV*: Need support (+)	Exercise frequency	Bivariate correlations
Russel & Bray, 2010 [99]	Cross-sectional and prospective (14wk)	53 M	Exercise cardiac rehabilitation patients (Mean 62.8 yr)	Canada	Perceived need support (HCCQ-short)	*MV*: Need support (n.s.) *BIV*: Need support (+)	Exercise frequency; duration (+); volume – 7Day-PAR	Bivariate correlations; Hierarchical regression analysis
Levy et al., 2008 [108]	Prospective (8-10wk)	70 (37)	Injured exercisers in rehabilitation (Mean 33 yr; 69% recreational)	UK	Perceived need support (HCCQ)	*MV*: Need support (+) [a, c] *BIV*: Need support (+) [a, c]	Exercise adherence: [a] clinical, [b] home-based; [c] attendance	Bivariate correlations; Manovas
Edmunds et al., 2007 [38]	Uncontrolled Prospective (3mo)	49 (84)	OW/Obese patients (BMI: 38.75; Mean 45 yr) on an exercise scheme	UK	Perceived need support (HCCQ)	*MV*: Need support (n.s.)	Self-reported exercise (LTEQ);	Multilevel regression analysis
Fortier et al., 2007 [17]	Experimental (13 wk); RCT	120 (69)	Autonomy supportive vs. brief PA counseling (Mean 47.3 yr)	Canada	Perceived need support (HCCQ)	*BIV*: Need support	Self-reported exercise (LTEQ)	Bivariate correlations
Mildestvedt et al., 2008 [68]	Experimental (4wk); RCT	176 (22)	Cardiac rehabilitation patients (Mean 56 yr): autonomy supportive vs. standard rehab	Norway	Perceived need support (mHCCQ)	*MV*: Need support (n.s.)	Self-reported exercise (composite score); exercise intensity	Manovas with repeated measures
Silva et al., 2010 [33]	Experimental (12mo); RCT	239 F	OW/Obese women (Mean BMI: 31.5; Mean 38 y): SDT-based WL treatment vs. controls	Portugal	Perceived need support (HCCQ)	*MV*: Need support (+) *BIV*: Need support (+)	Self-reported exercise: MVPA (7-day PAR); lifestyle PA index	Bivariate correlations; PLS/mediation analysis
Silva et al., 2010 [32]	Experimental (1 yr + 2y follow-up); RCT	221 F	OW/Obese women (Mean BMI: 31.5; Mean 38 y): SDT-based WL treatment vs. controls	Portugal	Perceived need support (HCCQ)	*BIV*: Need support (+)	Self-reported exercise: MVPA (7-day PAR)	Bivariate correlations

Table 1 Description of reviewed studies *(Continued)*

V. Exercise Causality Orientations									
Rose et al., 2005 [56]	Cross-sectional	375 (51)	Volunteers (17–60 yr)	UK	Exercise causality orientations (ECOS)	*MV*: Autonomy O. (+), Controlling O. (– F), and Impersonal O. (–)		Exercise stages of change	Discriminant function analysis. Gender differences
Kwan et al., 2011[53]	Prospective (4 wk)	104 (58)	Undergraduate students; active (Mean 18.2 yr)	USA	Exercise causality orientations (ECOS)	*BIV*: Autonomy O. (+), Controlling O. (–), and Impersonal O. (n.s.)		Self-reported exercise (online diary)	Bivariate correlations
VI. SDT-based Interventions and other SDT-related measures									
Edmunds et al., 2008 [39]	Experimental (10wk)	55 F	Exercisers (Mean 21 yr)	UK	Exercise self-regulations (BREQ-2); Need support (PESS); Basic needs (PNSS); Exercise attendance	Groups: SDT-based exercise classes vs. traditional exercise classes	Higher perceived need support, autonomy and relatedness needs; Competence (+), INTRO (+) and amotivation (–) overtime for both groups	Higher exercise attendance	Multilevel regression analysis
Fortier et al., 2007 [17]	Experimental (13wk); RCT	120 (69)	Healthy adults (Mean 47.3 yr)	Canada	Exercise self-regulations (TSRQ-PA); Perceived Competence (PCES); Need Support (HCCQ); Self-reported exercise (LTEQ)	Groups: autonomy supportive vs. brief PA counseling	Higher perceived need support, autonomous motivation overtime	Higher reported exercise overtime	Ancovas
Fortier et al., 2011 [36]	Experimental (13wk); RCT	120 (69)	Inactive primary care patients (Mean 47.3 yr): intensive vs. brief PA counseling intervention	Canada	Exercise self-regulations (BREQ-2); Self-reported exercise (LTEQ)	Groups: autonomy supportive - intensive vs. brief PA counseling	Higher perceived need support, autonomous motivation overtime	Higher reported exercise overtime	Ancovas
Mildestvedt et al., 2008 [68]	Experimental (4wk); RCT	176 (22)	Cardiac rehabilitation patients (Mean 56 yr): autonomy supportive vs. standard rehab	Norway	Exercise self-regulations (TSRQ); Perceived need support (mHCCQ); Self-reported exercise	Groups: autonomy supportive vs. standard rehab	No significant differences	No significant differences	Anovas with repeated measures
Levy & Cardinal, 2004 [40]	Experimental (2mo); RCT	185 (68)	Adults (Mean 46.8 yr); SDT-based mail intervention vs. controls	USA	Exercise self-regulations (EMS); Perceptions of autonomy (LCE); Competence (PSPP); Self-reported exercise (LTEQ)	Groups: SDT-based mail vs. controls	Women only: increase in perception of autonomy	Women only: increase self-reported exercise	Anovas with repeated measures

Table 1 Description of reviewed studies *(Continued)*

Silva et al., 2010 [18]	Experimental (12mo); RCT	239 F	OW/Obese women (Mean BMI: 31.5; Mean 38 y); RCT	Portugal	Exercise self-regulations (SRQ-E); Need support (HCCQ); Perceived autonomy (LCE); Self-reported exercise (MVPA, lifestyle, steps)	Groups: SDT-based weight loss treatment vs. controls	Higher need supportive climate, autonomy satisfaction, IM, IDENT, INTRO	Higher reported exercise (all measures)	Effect sizes; T-tests
Silva et al., 2011 [32]	Experimental (1 yr + 2y follow-up); RCT	221 F	OW/Obese women (Mean BMI: 31.5; Mean 38 y); RCT	Portugal	Exercise self-regulations (SRQ-E) at 1 yr and 2 yr; Need support (HCCQ); Self-reported exercise (MVPA)	Groups: SDT-based weight loss treatment vs. controls	Higher 2-yr EXT, INTRO and autonomous regulations	Higher 2-yr reported exercise	Effect sizes; T-tests

Legend: F, female; M, male ; BIV, uni/bivariate associations; MV, multivariate associations; IM, intrinsic motivation; INTEG, integrated regulation; ID, identified regulation; INTR, introjected regulation; EXT, external regulation; AMOT, amotivation; RAI, relative autonomy index; AutMot, autonomous motivations; CtMot, controlled motivations; Autonomy O., autonomy orientation; Controlling O., controlling orientation; Impersonal O., impersonal orientation; (+), positive association; (-), negative association; (n.s.), not significant. Superscript letters are used to signal associations between specific predictors and outcomes (check the 'significant predictors' and 'outcomes' columns when applied). (*) is used when specific comments need to be made (check the 'observations' column on those cases).

Organization of SDT predictors

Studies were generally organized based on the self-determination theory process model, depicted in Figure 1. The goal of the present manuscript was not to test this model *per se*, which would involve a considerably larger analysis. Instead, we focused exclusively on relations between each of these categories of variables and exercise outcomes (described below). Results concerning exercise self-regulations are listed first, followed by findings reporting the association between psychological needs satisfaction and exercise behavioral outcomes. Next, results concerning the measures of exercise motives/goals are reported, followed by findings regarding the association between perceived need support and exercise. Exercise causality orientation studies are listed last. In addition, we also identified interventions based on SDT and analyzed their effects on exercise outcomes.

Exercise-related outcomes

Exercise behavior was evaluated through self-reported measures (e.g., *7-day Physical Activity Recall* (PAR) [25], *Godin Leisure-Time Exercise Questionnaire* (LTEQ) [26]) in a total of 55 independent samples (78%). Three studies (representing 4 original samples) used accelerometry or pedometry to measure physical activity (6%). Measures of stages of change for exercise participation were employed in 13 samples (18%). A few other indicators were also used in some cases (8%), namely exercise attendance, number of exercise relapses, and exercise dropout.

Data coding and analyses

Summary tables were created based on the analysis of the available data (Tables 2 and 3). Sample characteristics (i.e., sample size, age, gender) were summarized using a tallying system and resulted in total counts (see Table 2). The percentage of independent samples presenting each characteristic from the total number of samples was also included. A summary of the evidence for each SDT-based construct was determined through a calculation of the percentage of independent samples supporting each association, based on whether the association was statistically significant or not (see Table 3). In all studies, significance level was set at 0.05. The measures of association varied across the studies' statistical methods, as indicated in the column "observations" in Table 1, including correlation and multiple regression coefficients, *t*-test or ANOVA group differences (e.g., between active and inactive groups), discriminant function coefficients, and structural equation model path coefficients, among others. Because many studies included bivariate associations (or direct paths in structural models) and also multivariate associations (in regression or in structural models), these were analyzed separately (see Table 2). A sum code was built for each motivational construct based on the following classification system: Positive (++) for percentage *K* ≥75% and (+) for percentage *K* between 50-75% showing positive associations in both bivariate and multivariate tests; 0/+ or 0/- when the evidence was split between no association (0) and either positive or negative associations, respectively; and (?) for other results indicating inconsistent findings or indeterminate results due to a small number of studies available).

Table 2 Summary of samples characteristics

Characteristics	Samples *K* (%)
Sample size	
< 100	13 (18.0)
100-300	38 (52.8)
300-500	12 (16.7)
≥ 500	9 (12.5)
Gender	
Women only	11 (15.3)
Men only	1 (1.4)
Men and Women – Combined	46 (63.9)
Men and Women – Separately	14 (19.4)
Location	
Western countries	70 (97.2)
Non-western countries	2 (2.8)
Mean age, years	
≤24	21 (29.2)
25-44	28 (38.8)
45-64	22 (30.6)
≥ 65	1 (1.4)
Design	
Cross-Sectional	45 (62.5)
Longitudinal – Observational	16 (22.2)
Longitudinal – Experimental	9 (12.5)
Mixed Method	2 (2.8)
Exercise Data Collection	
Self-reported Exercise	56 (77.8)
Exercise Stages of Change	13 (18.1)
Accelerometry/pedometry	4 (5.6)
Other*	6 (8.3)
Total *K*	72

Note: *Exercise relapses, weekly attendance, exercise adherence (home; clinical), exercise dropout.

Results

Characteristics of studies and samples

The 66 located studies comprised a total of 72 independent samples. The number of samples was higher than the total number of studies because some studies

Table 3 Summary of associations between SDT predictors and exercise-related outcomes

			% *K* Supporting associations			
Predictors	**# of *Studies***	***K***	**+**	**-**	**0**	**Sum code**
Exercise Regulations/Motivations						
Intrinsic motivation	26 (22)	37 (24)	62 (92)	0 (0)	38 (8)	**+**
Integrated regulation	6 (3)	8 (4)	62 (75)	0 (0)	38 (25)	**+**
Identified regulation	27 (24)	38 (26)	74 (85)	2 (0)	24 (15)	**+**
Introjected regulation	26 (25)	37 (27)	30 (52)	5 (4)	65 (44)	**0/+**
External regulation	26 (24)	37 (26)	0 (0)	43 (23)	57 (77)	**0/-**
Amotivation	10 (11)	14 (13)	0 (0)	36 (69)	64(31)	**0/-**
Relative autonomy (e.g., RAI)	8 (13)	8 (12)	88 (83)	0 (0)	12 (17)	**++**
Autonomous regulations	10 (10)	11 (11)	91 (82)	0 (0)	9 (18)	**++**
Controlled regulations	4 (6)	5 (7)	0 (0)	60 (0)	40 (100)	**0/-**
Need-Supportive Climate	6 (11)	6 (11)	50 (73)	0 (0)	50 (27)	**+**
Psychological Needs in Exercise						
Autonomy	4 (9)	5 (10)	20 (50)	20 (0)	60 (50)	**0/+**
Competence	8 (12)	9 (13)	56 (92)	0 (0)	44 (8)	**+**
Relatedness	4 (7)	4 (8)	0 (38)	0 (0)	100 (62)	**0**
Composite score*	0 (2)	0 (2)	0 (100)	0 (0)	0 (0)	**?**
Exercise Motives/Goals						
Intrinsic	7 (5)	8 (8)	100 (75)	0 (0)	0 (25)	**++**
Health/fitness	6 (1)	6 (1)	33 (100*)	33 (0)	33 (0)	**?**
Body-related	7 (5)	8 (8)	25 (63)	25 (12)	50 (25)	**0/+**
Exercise Causality Orientations						
Autonomy*	1 (1)	2 (1)	100 (100)	0 (0)	0 (0)	**?**
Controlling*	1 (1)	2 (1)	0 (0)	50 (100)	50 (0)	**?**
Impersonal*	1 (1)	2 (1)	0 (0)	100 (0)	0 (100)	**?**

Legend: Results derived from multivariate analyses and uni/bivariate analyses (in parenthesis) are presented. *K*, number of samples. Positive (++) was used for percentage *K* ≥75% and (+) for percentage *K* between 50-75% for both bivariate and multivariate associations; 0/+ or 0/- when the evidence was split between no association (0) and either positive or negative associations, respectively; (?) for other results indicating inconsistent findings or indeterminate results (i.e., when only a small number of studies were available, marked with *).

analyzed data originating from more than one sample (two samples: [27], [28], [29]; three samples: [30]; six samples: [31]). On the other hand, 7 studies were published using data from three original samples ([18,33,32]; [35,34]; [17,36]). A summary of the demographic characteristics of participants and samples is reported in Table 2. Samples tended to be mixed gender and included a range of populations (e.g., healthy individuals, chronic disease patients, overweight/obese individuals, exercisers), predominantly from Western cultures (97%), and mainly aged between 25–65 years-old.

From the studies eligible for this review, 53 ($K = 57$) analyzed associations between self-regulations and exercise behavioral outcomes, 17 studies ($K = 17$) investigated the relations between basic psychological needs and exercise, 12 studies ($K = 15$) tested the associations between motives and exercise, and 13 studies ($K = 12$) included measures of perceived need support and evaluated its predictive effect on exercise-related outcomes (see Table 3). Seven intervention studies, corresponding to 6 actual interventions, were identified. It should be noted that relations reported in the intervention studies were also analyzed in the other sections (e.g., regulations, need support, etc.)

Motivational predictors of exercise-related outcomes

Exercise behavioral regulations. A total of 57 samples (53 studies) analyzed associations between regulations and exercise behavior. Of these, 37 were used in cross-sectional designs, 10 in prospective designs, 7 in experimental studies, and 2 in mixed designs. Regulations were assessed with different instruments (53% with the *Behavioural Regulation in Exercise Questionnaire* (BREQ) and with Markland and Tobin's revised version (BREQ-2) [37] and reported results in several ways: Relative autonomy was evaluated as a composite score (e.g., the *Relative Autonomy Index* (RAI), by which individual regulations are weighted and summed to give an index of the extent

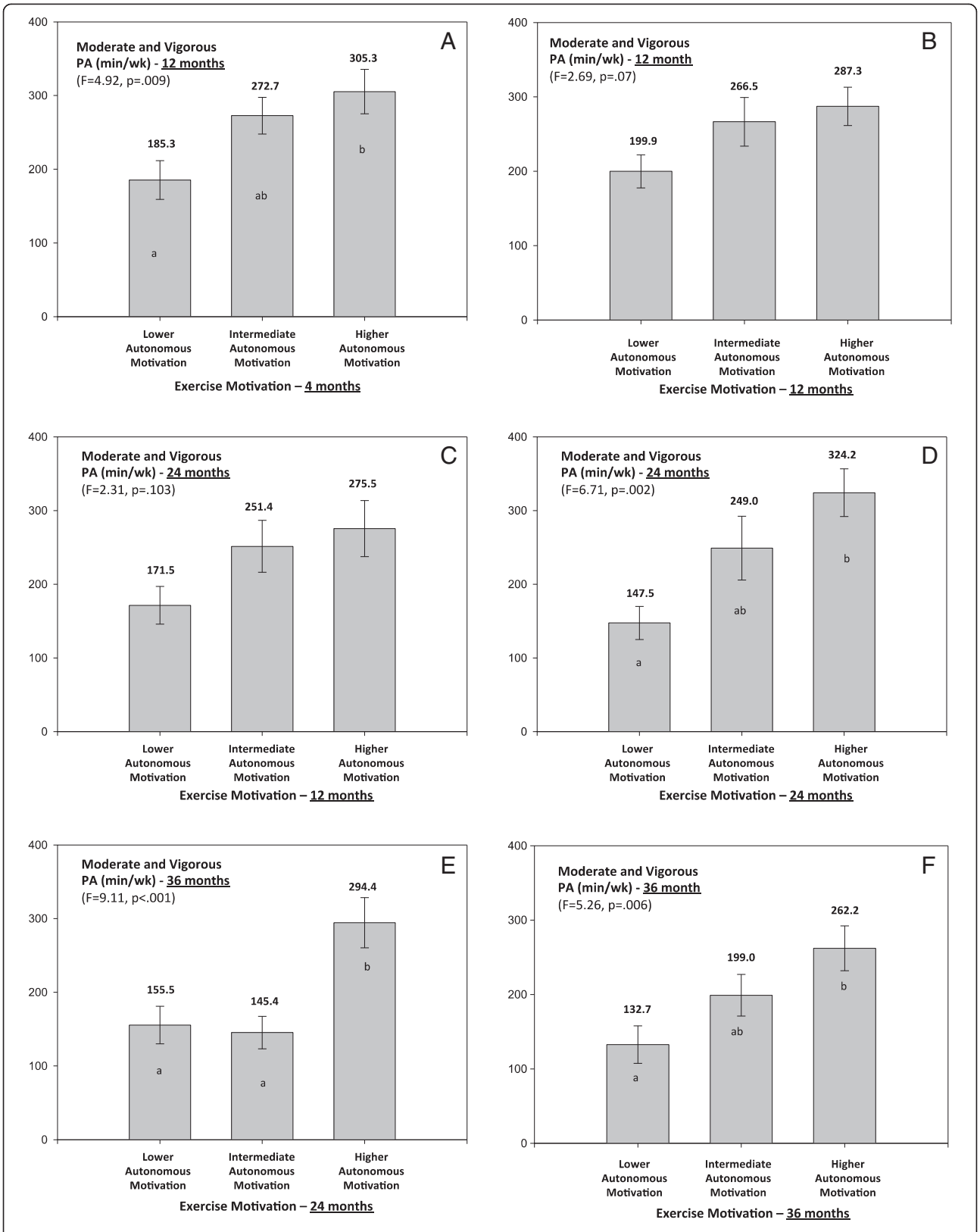


Figure 2 (See legend on next page.)

(See figure on previous page.)
Figure 2 Title. Self-reported minutes of moderate and vigorous exercise per week as a function of exercise autonomous motivation. Analysis includes 141 participants of the PESO trial [67] and data reports to variables assessed at 12 months (intervention end), 24 months (1 year follow-up with no contact) and 36 months (2-year follow-up). The time-point values in exercise and motivational variables at each assessment period were used (not change). Values used for tertile-split groups of autonomous motivation were calculated including all subjects (intervention and control groups collapsed), adjusting for experimental group membership. Autonomous motivation includes the identified regulation and intrinsic motivation subscales of the *Exercise Self-Regulation Questionnaire* [84]. Self-reported exercise was assessed with the 7-day Physical Activity Recall interview [25] and quantifies moderate and vigorous structured physical activity (METs > 3) performed in the previous week (or typical of the previous month if previous week was atypical, see reference 27 for more details). Panels **B, D,** and **F** show cross-sectional associations (variables assessed at the same time point) and panels **A, C,** and **E** show "prospective" associations (motivation assessed one year earlier than exercise). F for one-way ANOVA with letters in bar indicating multiple comparisons with Bonferroni post-hoc tests (different letters indicate different means, $p < .05$).

to which a person's behavior is more or less autonomously regulated) in 23% of the cases (none of which were experimental designs); autonomous and controlled regulations were grouped and analyzed as two higher-level types of regulation in 21% and 14% of the cases, respectively. All major forms of regulation were assessed and discriminated in 71% of the cases.

Nearly all studies using measures of relative autonomy (8 of 9 *K*) reported positive associations with exercise behavior while studies investigating autonomous and controlled forms of regulation ($K = 11$ and $K = 5$, respectively) also found consistent, positive associations favoring autonomous regulations as a predictor of exercise outcomes (82/91%, depending on whether bivariate or multivariate analysis is used). On the other hand, 3 independent samples (60%) showed negative associations in multivariate models for non self-determined regulations, all others (40%) showing no association. In bivariate analyses, results for controlled regulations unanimously showed no association. Results were similar across different study designs, suggesting consistent positive effects of autonomous regulations on exercise behavior, and either negative or null effects associated with controlled regulations. In one study with longer-term follow-up measurements, prospective associations between regulations and exercise behavior were reported [33] (see also Figure 2). The authors found that both 12 and 24-month autonomous regulations, but not controlled regulations, mediated the effects of a SDT-based intervention on self-reported exercise at 24 months [32].

Specific results concerning the separate autonomous types of motivation showed positive associations between identified regulation and exercise behavior in 28 samples (74%) in multivariate analyses and 22 samples (85%) in bivariate analyses. The only exception was a study by Moreno et al. where the mean value for identified regulation was lower in a group reporting 60+ min of exercise than among those who exercised less than 60 min (presumably each day; no details are provided). Of note also are the mixed results found by Edmunds et al. (2007) displaying negative associations for identified regulations in a multilevel model, but positive cross-sectional associations at each of the 3 times points. The authors indicated that the multilevel results "should be ignored as they are a consequence of net suppression" [38]; pg.737]. In 3 studies that analyzed identified regulations [36,40,39], no significant association emerged. Regarding intrinsic motivation, positive associations with exercise behavior were reported in 23 or 22 independent samples (62% or 92%), in multivariate or bivariate analyses respectively. No study reported negative associations and results were consistent independent of study design. Few studies have tested the role of integrated regulation, but it appears to positively predict exercise behavior. Of 8 samples analyzed, 62-75% found positive associations with physical activity, with increased consistency found in bivariate analyses.

In an attempt to further clarify which single self-determined type of motivation is more closely related with behavior outcomes, a comparative analysis between identified and intrinsic motivation findings was undertaken. Twenty-five studies ($K = 31$) reported significant associations for both variables, of which $12\,K$ were derived from multivariate analysis, $5\,K$ from correlational analysis, and $4\,K$ from both types of analysis. Seven studies ($K = 7$) found associations for identified regulation in multivariate analysis, but only bivariate associations for intrinsic motivation [44,45-43,42,41]. Three studies/samples showed the converse [48,47,33], reporting associations for intrinsic motivation in multivariate analysis and only correlational bivariate associations for identified regulation. It should be noted that no study tested whether the differences between the association coefficients (for identified regulation vs. intrinsic motivation) with exercise were significant. Wilson et al. (2002) investigated bivariate predictors of different physical activity intensities [49] and found that at mild intensities, associations were significant only for identified regulation; for moderately intense and strenuous exercise, both identified regulation and intrinsic motivation were significant predictors. Three additional studies/samples showed significant associations only for identified regulation [50,51,38]. In another study ($K = 1$) this regulation was the only variable predicting fewer exercise relapses

[52]. On the other hand, two studies found significant associations only for intrinsic motivation [54,53].

For integrated regulation, only 6 studies ($K = 8$) were available. Comparing results for integrated versus identified regulations no differences were found in the patterns of association for all but one study [85] where there was a significant bivariate association with exercise for integrated but not identified regulation. Comparing results between integrated regulation and intrinsic motivation, two studies show integrated regulation, but not intrinsic motivation, as a significant predictor of exercise in multivariate models [41,38] whereas in a different study the opposite trend was observed using bivariate associations [28].

All studies measuring stages of change for exercise participation ($K = 7$) showed that autonomous regulations increased across stages, being the highest in the action/maintenance stages. However, only one study formally tested differences for regulations' means across stages of change [52]. They found that for identified regulation there was a progressive increase from preparation to action to maintenance stage (ANOVA $F = 25.1$, $p < 0.001$) whereas for intrinsic motivation, maintenance had significantly higher means than both preparation and action stages ($F = 27.5$, $p < 0.001$). Five of these studies used the BREQ/BREQ-2 and 4 of these used discriminant function analysis. In these 4 studies, identified regulation loaded slightly stronger than intrinsic motivation on the primary discriminant functions distinguishing across stages of change. Authors tended to conclude that identified regulation played a more important role in exercise adherence when the full range of stages of change is considered. Finally, in a study examining change in behavioral regulations among exercise initiates, Rodgers et al. showed that both identified and intrinsic motivation increased overtime and that, compared to regular exercisers, initiates' levels of identified and intrinsic motivations remained below regular exercisers' levels even after 6 months of physical activity [31]. Authors also concluded that identified motivation appeared to increase faster than intrinsic motivation in these early stages of exercise adoption [31].

Results from multivariate analysis concerning the controlled types of motivation showed negative associations between external regulation and exercise behavior in 16 independent samples (43%). The remaining samples (57%) showed no associations. The trend for the absence of an association between external regulation and exercise was more apparent in bivariate analysis (77%). Regarding external regulation across stages of change, results show that external regulation generally decreases across stages, being higher in the preparation/action stages than in the maintenance stage. Furthermore, when comparing genders, results suggest that among males external regulation is negatively associated with exercise in the latter stages of change (i.e., maintenance) whereas among female there is no association at this stage.

Regarding introjected regulation, multivariate analysis showed positive associations with physical activity in 11 independent samples (30%), 1 study ($K = 2$) found negative associations (5%) and all others showed no association (65%). Bivariate results pointed in a similar direction, but showed more positive associations (52%). Despite the positive associations with exercise behaviors, the strength of association for introjected regulation appears to be lower compared to self-determined types of motivation, as reported in several studies [e.g., [55,49]. A closer look into the way introjected regulation predicts exercise participation over time shows mixed findings. Rodgers et al. (2010) studied initiate exercisers and found significant, but small, increases in introjection overtime, noting that these changes occurred mainly in the early stages of exercise participation [31]. Increases in introjected regulation were also observed across stages of change in 5 of 7 independent samples, although these were only significant in one case [e.g., [52]. In contrast, Silva and colleagues showed that although introjected regulation was cross-sectionally associated with exercise at 12- and 24-month time points, 12-month regulation did not prospectively predict (nor did it mediate) 24-month exercise outcomes [33,32].

A possible gender effect might be relevant to understand these mixed findings regarding introjected regulations. In effect, a closer examination of all the studies that explored gender differences with respect to the association between exercise regulations and behavior suggests that introjected regulation may be more positively associated with exercise among females, whereas among males the association is negative or zero [e.g., [45,41]. Within the studies examining differences across stages, results suggest that introjection is relevant for both genders in the action stage, but that in the maintenance stage it is more relevant for women than for men [56,55]. It should be noted that only two studies reported associations for men: one showed a positive association in the action stage and negative in the maintenance stage [55] and another study showed a tendency towards a positive association in the action/maintenance stage [57]. For studies with mixed samples and not reporting gender differences (the majority) the associations are mixed. Experimental studies confirm this pattern of mixed results, some showing increases in introjected regulation over the course of an exercise program [e.g., [39] and some showing no significant changes [e.g., [58]. One notes that null or unreliable results from introjection are theoretically expected within SDT, in which introjection is seen as an unstable basis for motivation without positive long-term utility.

Regarding amotivation, 5 independent samples (36%) showed negative associations with exercise outcomes in multivariate analysis; the remaining studies ($K = 9$) showed no associations. Correlational analysis showed negative associations in 9 samples (69%) and no association in 4 samples (31%).

Need satisfaction. A total of 17 samples/studies were used to analyze the associations between basic psychological needs and exercise behavior. Ten samples were evaluated in cross-sectional designs, 3 within prospective studies and 3 in experimental designs. One study used mixed methods (cross-sectional and prospective). Different instruments were used to assess basic needs, a fact that does not facilitate the comparison of results between studies. *The Psychological Need Satisfaction for Exercise Scale* [16] was adopted in 24% of the cases and was the most frequently used measure. Competence was assessed in 14 (82%) independent samples, autonomy in 11 (65%) samples, and relatedness in 9 (53%) independent samples. An examination of the specific multivariate results for each basic need showed that perceived competence was positively associated with physical activity in 56% of the independent samples, while the remaining samples showed no association (44%). The pattern of association was much clearer in correlational analysis with 12 samples (92%) reporting positive associations. Regarding autonomy need satisfaction, findings were mixed and generally ranged from no association (60% in multivariate analysis) to moderate positive or negative associations (20% for each). Nevertheless, positive correlations were reported in 5 studies/samples (50%) using bivariate analysis. Regarding relatedness, multivariate results consistently reported an absence of associations with exercise behavior ($K = 4$, 100%). Correlations showed a similar pattern, even though a general trend towards a positive association with exercise behavior was identified (38%). No negative associations with exercise outcomes were observed for the perceived fulfillment of any of the 3 needs. A composite score was created to assess overall exercise psychological need satisfaction in 2 (of 17) samples; positive associations with exercise behavior were reported in both cases.

Exercise motives. A total of 12 studies ($K = 15$) investigated the associations between motives (or goal contents) and exercise behavior. Of these studies, 8 were cross-sectional, 3 prospective, and 1 used a mixed design (cross-sectional and experimental). Regarding the instruments used to measure exercise motives, there is some inconsistency: the *Motives for Physical Activity Measure* (MPAM) or MPAM revised/adapted versions [59,27] of it were used in 6 independent samples (40%), 3 samples (20%) measured exercise motives using the *Exercise Motivations Inventory - 2* (EMI-2) [60], and in other 3 samples (20%) the *Intrinsic Motivation Inventory* (IMI) [61] was employed to evaluate intrinsic motives and the *Extrinsic Motivation Inventory* (Lee's EMI) [62] to measure extrinsic motives. Sebire and colleagues (2009) [19] used the recently developed *Goal Content for Exercise Questionnaire* [63] while Segar and colleagues used an inductive, qualitative method to assess exercise motives in one study [64], and performed a cluster analysis to identify homogeneous groups of goals, intrinsic and extrinsic, in another study [65].

Multivariate results showed that intrinsic motives (e.g., challenge, affiliation, enjoyment) were positively associated with exercise behavior in all samples ($K = 8$, 100%). A similar trend was observed in correlations (75%). Regarding body-related motives, multivariate findings were mixed regardless of the statistical analysis performed: in multivariate analysis, 25% of the samples showed positive associations and 25% reported negative associations; in correlational analysis, a general trend towards a positive association was identified (63%). The pattern of association was less clear for health/fitness motives with 33% showing positive associations, 33% showing negative associations, and other 33% not finding any association. There was only one study/sample performing correlational analysis to explore the links between health motives and exercise [46]; positive associations were reported. As expected from theory, controlled motives (social recognition, appearance/weight) did not predict, or negatively predicted, exercise participation [46].

Perceived need support. Environments perceived as more need-supportive were positively associated with increased levels of self-reported physical activity in 3 (of 6) independent samples tested with multivariate analysis (50%). This increased to 73% ($K = 8$) in correlational analysis. The remaining studies/samples showed no association. In the majority (67%) of independent samples perceived need support was assessed using the Health Care Climate Questionnaire [66].

SDT-based Interventions. To date, only a few interventions have been designed to promote exercise-related behaviors by specifically increasing personal autonomy in the form of exercise autonomous self-regulation in adults [e.g., [17,40,68,39,67,69]. Some of these trials are still ongoing and all have been conducted in Western cultures. Of 7 interventions (with available data), 6 (86%) found significant differences favoring the SDT-based intervention group for perceived autonomy support, need satisfaction, and autonomous and introjected regulations for exercise, as well as greater self-reported exercise. In addition, one of these interventions found gender differences, reporting significant increases in perceived autonomy support and self-reported exercise only for women [40]. In contrast, there was one study in a clinical setting that did not find significant differences in perceived autonomy support and exercise behavior

between autonomy support group and controls [68]. The authors argued that their additional individual SDT-based 4-week intervention, added to standard cardiac rehabilitation, might have been too limited (i.e., an insufficient number of sessions) to achieve significant between-group differences.

Edmunds and colleagues tested a SDT-based intervention in an exercise setting, examining the effect of an autonomy-supportive teaching style on female exercisers' psychological needs, motivational regulations, and exercise behaviors during a 10-wk exercise program [39]. They found that the intervention increased autonomous self-regulation, need satisfaction, and attendance [39]. Although not a randomized controlled trial, results were similar to those obtained in several RCTs. For instance, Fortier et al. [17] tested an autonomy-promoting counseling protocol for promoting physical activity in sedentary primary care patients in a 13-week RCT. Results showed that the intervention was successful in changing autonomous self-regulation to reach activity goals (vs. a brief counseling protocol) and that higher autonomous regulation for exercise mid-intervention predicted higher levels of physical activity at the end of the intervention in the intervention group. The longest RCT to date to evaluate autonomy support, need satisfaction, motivation, and exercise behaviors was implemented in 239 overweight women, through 30 weekly group sessions for about 1 year, with a 2-year follow-up [67]. A few features of this study clearly distinguish it from the remaining intervention studies reviewed (see Table 2, table VI): larger sample, considerably longer intervention and follow-up assessments up to 3 years, and the use of mediation analysis to predict long-term changes in physical activity. Results showed that the intervention was perceived as need-supportive, it increased perceptions of competence and autonomy for exercise, increased autonomous regulations (and to a lesser degree introjected regulation, but not external regulation), and increased exercise behavior [18]. Exercise level was clearly associated with level of autonomous motivation for all subjects, both concurrently and prospectively, as depicted in Figure 2. Only autonomous regulations were found to mediate the intervention effect on exercise in the long-term [33,32].

Discussion

The aim of this review was to examine the empirical literature on the relations between SDT-based constructs and exercise and physical activity. The review demonstrates the recent growth in the application of this theory to the study of exercise and physical activity motivation, with 53 of the 66 papers identified being published in the last five years. The theory has been applied to a wide range of physical activity contexts including recreational exercise, weight loss programs and clinical populations, and across a range of ages. The majority of studies employed cross-sectional designs but comparable results are found across cross-sectional, prospective, and experimental designs.

Behavioral regulation and exercise

The vast majority of studies included an examination of the relations between behavioral regulation and exercise behavior. Of these, most included some or all of the individual regulations specified within SDT whereas others have collapsed autonomous and controlled forms of regulation into summary scales or adopted the RAI. The results show consistent support for a positive relation between more autonomous forms of motivation and exercise behavior, whether single regulation, summary measures, or the RAI are used. Intervention studies are also clearly supportive as are studies examining the endorsement of different forms of behavioral regulation across the stages of change, consistently showing that more self-determined regulations distinguish between individuals in the later stages from those in the early stages.

When considering the more autonomous forms of behavioral regulation separately, positive associations for identified regulation are found slightly more consistently in comparison to intrinsic motivation in multivariate analyses, whereas intrinsic motivation is somewhat more consistently predictive of exercise behavior in bivariate analyses. A similar trend was found for integrated regulation versus intrinsic motivation, but based on much fewer studies. This could be interpreted as suggesting that, *independent of* other regulatory motives, identified regulation (or integrated regulation) is the single best correlate of exercise. This notwithstanding, the SDT continuum of motivation [10] suggests that regulations that are more closely located in the continuum of autonomy specified by SDT (such as identified and integrated regulation, and intrinsic motivation) are expected to share some degree of variance, highlighting the theoretical expectation that regulatory factors are often simultaneously operative. This renders the question of which sub-type of autonomous motivation is more important in explaining and promoting exercise behaviors difficult to solve. Nonetheless, a number of authors have discussed this issue, attempting to explain results "favoring" either identified or intrinsic motivation. For example, Mullan et al. [57] argued that intrinsic motivation alone is unlikely to sustain long-term regular engagement in exercise, given all the organization and commitment it entails. Edmunds et al. [44] suggested that because sustaining a physically active lifestyle presumably requires a high degree of effort, often for mundane or repetitive activities, regulation by identification with the outcomes may be more important than exercising for fun and enjoyment, or to challenge oneself. Finally, Koestner and

Losier (2002) proposed that in behavioral domains that require engagement in a range of different activities that vary in their intrinsic appeal, internalization of the value of the outcomes of the activities is likely to lead to greater persistence than being intrinsically motivated [70]. Clearly exercise is one such behavioral domain.

Because health promotion campaigns typically market exercise more in terms of health-related outcomes than in terms of its intrinsic value, the primary source of self-determined motivation among active individuals might derive from a valuing of these outcomes, even if they also find exercise intrinsically enjoyable [55]. Conversely, in contexts where enjoyment in and genuine interest for exercise is emphasized over the outcomes, one might expect intrinsic motivation to be more salient to individuals. In support of this, in Silva et al.'s intervention that explicitly emphasized enjoyment, mastery and challenge rather than the outcomes of exercise, intrinsic motivation was a more consistent predictor than identified regulation of moderate and vigorous exercise [33]. Clearer definitions of the nature of the exercise behaviors under investigation (type, intensity, volume, duration, time in the same activity), which may vary within and among studies, and their potential appeal to the individual may shed additional light onto this issue. Some types of physical activity may be inherently intrinsically motivating for many individuals, especially when they involve self-chosen optimal challenges that can help people enjoy the sense of autonomy and mastery, factors that underpin intrinsic motivation.

As Daley and Duda [55] point out, most of the research showing a stronger effect for identified regulation has been cross-sectional and a few studies, including experimental studies lasting for several years, have shown intrinsic motivation to be critical for longer-term engagement [44,32]. Furthermore, a major limitation in interpreting findings concerning a benefit for either identified regulation or intrinsic motivation is that where associations for both have been found, authors have not conducted statistical tests to determine the unique effects of each type of regulation, nor whether the larger effect is in fact statistically significant. Given also the lack of longitudinal or experimental studies to determine whether differential benefits for the two types of regulation might emerge over time, it would be advisable for the time being to recommend fostering *both* identification and intrinsic motivation in order to promote optimal behavioral outcomes. Both of these autonomous forms of motivation share common antecedents in terms of support for autonomy and competence. Identification could be specifically promoted by emphasizing the personal instrumental value of exercising with regard to health, optimal functioning, and quality of life. At the same time, intrinsic motivation could be promoted by emphasizing fun, skill improvement, personal accomplishment, and excitement while exercising. Furthermore, the focus should be not only on the amount of exercise performed, or long-term adherence *per se*, but also on the enhanced well-being and vitality associated with exercise. Indeed, intrinsic motivation has been shown to be not only related to persistence at a task but also with psychological health and improved well-being [15].

The results for more controlled forms of regulation are mixed. No studies have found a positive association for controlled motivation at the summary level of analysis, nor for external regulation at the individual regulation level. However, while a substantial number of studies found a negative association, the majority found no association. There is a trend for external regulation to be negatively associated with exercise in the later stages of change among males, but no association among females, suggesting that more active males might respond more negatively to social pressures to exercise.

Concerning introjected regulation specifically, results are split between positive and null relations with exercise, with a clear predominance of the latter in multivariate analyses. This internally controlling form of regulation is generally theorized to be associated with more maladaptive outcomes such as negative affect, feelings of guilt, and lowered self-esteem [12]. People who feel internally pressured to exercise are likely to experience some degree of guilt or shame if they do not exercise, and the potential to enjoy it and experience the positive well-being consequences of this behavior will be decreased. Furthermore, research examining the motivating forces behind exercise dependence, which is considered to be maladaptive, has found introjected regulation to be the strongest predictor of this type of dependence [51]. Nonetheless, the periodic finding of a positive relation between introjection and adaptive behavioral outcomes in both exercise and other behavioral domains has been attributed to the partial internalization of external pressures from, for example, health promotion messages [52] or parental expectations [71].

When energized primarily by introjected motives, exercise participation may occur at some cost to psychological health, a factor most exercise adherence studies have not quantified. By contrast, recent evidence in overweight women showed that a summary measure of controlled exercise regulation (including introjected and external regulation items) was unrelated to psychological well-being, although controlled motivation to participate in obesity treatment predicted lower quality of life and self-esteem, and higher state anxiety [72]. A more refined analysis of introjected forms of motivation, breaking it into an approach-orientated motivation (to seek positive feelings such as self-aggrandizement and pride) and an avoidance-oriented motivation (to avoid

negative feelings such as shame, guilt, and anxiety) could help clarify the role of introjected regulation on psychological and possibly also on behavioral outcomes [20]. Introjected avoidance regulation has been shown to yield more negative psychological correlates, including less engagement in school or poorer sports performance than introjected approach regulation [73]. The former was also more strongly associated with identified regulation than the latter. To our knowledge, studies have not yet addressed the differential association of these subtypes of introjected regulation with exercise behavior adoption or persistence.

The studies reviewed here also show a trend for an increase in introjection over time in the longitudinal or experimental studies, or across stages of change. However, observed (or assumed) increases in introjection with time do not necessarily mean that this variable explains or mediates increases in exercise. For instance, introjection has been found to be significantly associated with exercise when both were measured at the same time point, but not prospectively [32], suggesting that regulation by introjection may not lead to sustained exercise behavior. Furthermore, and despite observed increases in introjected regulation as a result of an SDT-based intervention [18], only autonomous motivation was predictive of long-term moderate and vigorous exercise in mediation analysis [32]. Unfortunately, there is only one study [32] reporting such long-term prospective associations between experimentally-induced changes in motivation and exercise behavior.

Our analysis of the relation between introjection and exercise for those studies reporting associations separately for males and females provides some evidence for a gender effect. Where such effects occur, introjection appears to be more positively associated with exercise among women, whereas among men there is a negative association or no association, especially in the maintenance stage of change. Some studies also report no differences. Given the pervasive societal and media pressures on women to have a slim and toned physique [74], this is perhaps not surprising. In the majority of studies, gender differences are not reported, making it difficult to draw firm conclusions but the trends we observe here for both introjection and external regulation suggest that future research would do well to consider possible gender differences rather than assuming no such differences and collapsing data across gender.

Finally, with regard to behavioral regulations and exercise, unsurprisingly no studies found a positive association between amotivation and exercise. The remaining studies showed either a predominance of null findings (nearly 70% in multivariate analyses) or negative associations (64% in bivariate analyses). Closer examination of these studies shows a trend for a sample effect. In all five studies showing no association the samples comprised either non-exercisers or a mixture of non-exercisers and exercisers, while the majority of studies showing negative associations comprised regular exercisers. Furthermore, it is noteworthy that fewer studies have assessed amotivation in comparison to those assessing the other regulations. This is understandable given that amotivation refers to the absence of both intrinsic and extrinsic motivation and represents a complete lack of self-determination and volition with respect to the target behavior [12]. Therefore one would expect to rarely see highly amotivated individuals in exercise settings. Additionally, different authors have put forth the hypothesis that individuals could also be autonomously motivated to *not* participate in exercise upon consideration, perhaps even when they can perceive some value in the behavior [7,20]. In some respect, they would be "autonomously amotivated" towards exercising. To the extent this would occur, it might also confound the association between amotivation and exercise, since these individuals might not score high on typical amotivation items such as "I don't see the point in exercising" and "I think that exercising is a waste of time", despite behind sedentary. It should also be noted that, empirically, it is difficult to distinguish amotivation from a lack of controlled or autonomous regulation [46]. Hence, including amotivation along with controlled and autonomous regulation in the same model might introduce a confound and could help explain the absence of associations in multivariate analyses.

Need satisfaction and exercise

Rather less attention has been paid to examining the associations between satisfaction of psychological needs and exercise than for behavioral regulations. The use of different instruments to assess basic need satisfaction (both domain-general and domain-specific measures), differences in the number of needs assessed, and their combined or separate analyses do not facilitate easy comparison of results across studies. Generally, competence satisfaction has been the most frequently assessed need and the literature shows consistent support for a positive association with exercise. In this review, twice as many studies reported bivariate associations between need satisfaction and exercise, compared to multivariate analyses. In bivariate analyses, no studies report a negative association between autonomy and exercise and the remaining results are split equally between positive and null associations whereas multivariate results are more mixed. Results for relatedness satisfaction are also mixed in bivariate analyses, although again no studies found a negative association with exercise. The exercise context might explain a lack of association for relatedness satisfaction. In some contexts, engaging in solitary exercise

being the most obvious, the need for relatedness might simply not be an issue. Inconsistency in the measures used to assess the needs, and therefore their operational definitions, and a lack of applicability of particular scales to different exercise contexts might be concealing positive associations for autonomy.

In interpreting the results for need satisfaction and exercise, it is important to note that only direct effects of need satisfaction on exercise (whether from bivariate or multivariate association or direct paths in structural models) were considered in the present review, a fact that does not consider their indirect effects. In fact, theorizing within SDT stresses that the internalization of behavioral regulations is fostered by the satisfaction of basic psychological needs, and thus autonomous regulations would mediate associations between need satisfaction and behavioral outcomes. In current interpretations of mediation analysis, a significant association between an independent and a dependent variable is not a necessary condition for the possible occurrence of significant indirect (i.e., mediated) effects between them [75]. This highlights the importance of conducting more sophisticated analyses, such as path analysis or structural equation modeling, to clarify the mediating role of need satisfaction in the development of self-determined motivation. Indeed, going beyond the simple direct associations between behavioral regulations or need satisfaction and exercise (which are the main focus of this review), it is important to note that several studies have tested one or more parts of SDT's proposed motivational sequence(s) for physical activity behaviors (see Figure 1). Relations from perceived autonomy support to exercise behavior, via psychological needs and regulatory styles have been tested (in part or all) in several studies and in general these confirm the proposed sequences [17,44,43,77,76,38,33]. In one case this was tested with a longitudinal randomized controlled trial using structural equation modeling [33,32], which empirically supported the motivational sequence proposed by SDT (i.e., need-supportive health care climate -> need satisfaction -> autonomous exercise regulation -> exercise behaviors).

Participation motives and exercise

Following some early work in the 1990s, there has been a resurgence of research in recent years on the role of exercise participation motives or goal contents. The rationale for this is that some motives (e.g., affiliation, skill development) are more intrinsically-oriented and likely to be experienced as autonomous whereas others (e.g., body-related motives such as weight or appearance management) are more extrinsic and likely to be experienced as internally controlling. Studies show a consistent positive association between more intrinsic motives and exercise. Findings for fitness/health and body-related motives are mixed. For fitness/health, although no studies found a negative association, an absence of association is more frequently found than positive associations. This might reflect different ways in which fitness/health motives have been operationalized. Health/fitness motives can reflect health pressures or threats (e.g., medical advice) or be associated with drives for thinness or an attractive image. Yet health and fitness motives can also reflect more positive concerns such as general health promotion, increasing physical strength for performing daily activities, reducing pain (e.g. lower back pain or discomfort in joints), or feeling more energy and vitality. Thus, conceptually, being concerned about health or fitness *per se* cannot be easily defined as either intrinsic or extrinsic, as it depends on what the motive means to the individual [78].

Similarly, results for body-related motives results are also mixed, despite a preponderance of both positive and null findings, relative to negative associations. For a more in-depth understanding of the relation between participation motives and exercise, the characteristics of exercise participation (e.g. type, intensity, total volume) and type of sample need to be taken into account. For example, Frederick and Ryan (1993) compared individuals whose primary physical activity was a sport with individuals whose primary physical activity was a non-sport fitness activity [59]. The sport participants had higher interest/enjoyment and competence motives whereas the fitness participants had higher body-related motives. Furthermore, the apparent positive (at least in the short term) role of these motives on exercise may then be mediated by the development of introjected regulation. Ingledew et al. [79,46] found that body related motives were associated with introjections and a recent study [41] found that introjected regulation predicted exercise intensity among females.

It is important to note, as Markland and Ingledew pointed out [46], that holding controlled motivations is not necessarily problematic, motivationally speaking, as long as self-determined regulations are also held. It has been suggested [20], for example, that a person may strive for a physically appealing body (an "extrinsic" motive) because her partner praises her good looks (controlled motivation) and at the same time she may personally value a fit appearance (autonomous motivation). Thus, although intrinsic goals tend to be pursued for autonomous reasons and extrinsic goals tend to be pursued for controlled reasons [81], the content of, and reasons for pursuing aspirations can be empirically crossed. Therefore, exercise promotion programs should take care not to explicitly or implicitly denigrate appearance/weight motive or any other motive for exercising, which may lead individuals to perceive that their autonomy is threatened, with consequent defiance and dropout [46]. Instead, acknowledging

the validity of individuals' motives in a need-supportive context may ultimately promote movement away from controlled regulations toward more autonomous commitments to be active.

Experimental studies

It is encouraging to see that in more recent years researchers have turned their attention to experimental studies evaluating interventions based on SDT principles. However, all but one were shorter than 3 months in duration and involved a small amount of contact time with the participants, in some cases amounting to approximately 2–3 in-person sessions. The remaining contacts were performed via telephone [e.g., [17,68,69], and one of these interventions relied solely on email booster messages to promote self-determined motivation and behavior change [40]. By contrast, one intervention provided substantially more contact time, (thirty 2-hour group sessions for about 1 year [18,67]). Not surprisingly, intensity, depth, and strategies used to promote personal autonomy and the development of intrinsic motivation for exercise also varied among these interventions. Some interventions were limited to strategies such as encouraging participants to make their own choices, providing information, setting realistic goals, and/or encouraging participants to seek and find forms of social support [e.g., [17,40]. Others included a more comprehensive set of strategies, more fully embracing SDT propositions [18,39,67] including providing a clear rationale for behavior change, acknowledging ambivalence and internal conflict, providing a menu of options, minimizing controlling influences (e.g., use of pressure, demands, and extrinsic rewards), and promoting competence through optimal challenge and giving informative feedback [18,33,32]. In sum, existing interventions are limited in number and highly varied. Longer and more comprehensive longitudinal interventions are needed, especially those which work toward the development of autonomous motivation, allow more time for changes in motivational and behavioral processes to take place, and assess whether those changes (and associations) persist in the long-term.

Conclusions

Overall, this review provides good evidence for the value of SDT in understanding and promoting exercise behavior. The clearest finding of this review concerns the beneficial role of developing autonomous self-regulation, be it predominantly via autonomous forms of extrinsic regulation (i.e., identified and integrated regulation) or enhanced intrinsic motivation. The present literature is consistent in showing that all forms of autonomous regulation predict exercise participation across a range of samples and settings. There is also increasing evidence that a motivational profile marked by high autonomous motivation is important to sustain exercise behaviors over time, although the pool of studies supporting this inference is limited. Longer-term studies and follow-ups will be especially important in evaluating the relative efficacy of identified versus intrinsic regulations in exercise maintenance. For the moment, evidence is consistent with the hypothesis that reporting well-internalized extrinsic regulations, such as personally valuing certain *outcomes* of exercise, is a particularly important factor for initial adoption (when cognitive factors such as rationally weighing pros and cons may be decisive but experiential knowledge of exercise may be limited). Conversely, there is some indication that a predominance of intrinsic motivation (i.e., valuing the actual *experience* of exercise) is especially important for longer-term exercise participation. It is also important to highlight the strong covariance between identified/integrated regulations and intrinsic motivation, especially since these different forms of autonomous motivation share some common antecedents that would be applied in intervention settings.

We suspect future studies may come to identify significant moderating factors for the role of specific regulations on exercise adherence, such as age, gender, previous health conditions, or social norms and social desirability. For instance, current public campaigns against obesity may have enhanced the perceived *utility* of exercise for weight control and health (as a preventive or treatment "medicine"), inadvertently minimizing experiential rewards of exercise such as social interaction, expression of personal skills and abilities, self-development, or pure enjoyment. The experiential qualities of exercise were highlighted as a critical factor for adherence in a recent review of mediators of physical activity behavior change [82]. On this note, it is perhaps no coincidence that in the current public health dialogue about "exercise as medicine", physical activities not typically associated with the term "exercise" such as playing sports, dancing, or outdoor exploration activities are rarely mentioned. From a public health/exercise promotion perspective, this could be a limiting factor if such activities, rich in their intrinsic appeal although less likely to be monitored and supervised, are not considered viable options in professionals' exercise prescriptions or as targets of public policy promotions. Again, future research with long-term outcomes and also exploring predictors of different forms of exercise should help elucidate these issues.

Two additional conclusions can be derived from the present review. One is that having more intrinsic participation motives or goals associated with exercise, such as affiliation and social engagement, challenge, and skill development, is clearly associated with greater exercise participation. Since these motives are associated with intrinsic motivation [22,34], it may be especially important

that health professionals are trained in distinguishing the "signs" of intrinsic (vs. extrinsic) motives in their patients and promoting them at every opportunity, aiming at long-term exercise maintenance. The other is that reporting increased perceived competence for exercise is also positively predictive of more adaptive exercise behavioral outcomes. Together, the previous findings have important implications for practice. It serves as evidence-based support for health professionals to strive not only to provide sufficient structure and optimal challenge to promote feelings of mastery and competence in their clients and patients, but also to encourage professionals to actively explore with the people they counsel reasons to be physically active that go beyond the most common motives such as improved body shape and attractiveness. Finally, as we discussed previously, the consequences of health and fitness-related motives, including weight loss, are perhaps more complex and likely moderated by other motivational aspects.

Limitations in the collective body of work are worthy of consideration as they bear on avenues for future research. A major limitation concerns the heterogeneity of the samples in the majority of studies. Heterogeneity within samples with regard to such factors as age, gender, weight or body composition, and fitness status may be contributing to variability across studies. While general motivational patterns are likely to remain constant (e.g., autonomous motivation being more likely to promote long-term exercise adherence), there may be much to learn by examining motivational profiles that are specific to different demographic groups or to individuals at different stages of change for exercise. For instance, a recent study [63] highlights the existence of different patterns of motivation between long-term exercisers versus beginners. Similarly, more enduring individual differences could be explored. Only one study has examined the relations between exercise causality orientations and exercise, and none have explored general causality orientations, despite the fact that such individual difference measures have been shown to predict adaptive outcomes in other health-related contexts [e.g., [108]. Finally, SDT has a history of strong experimental work on motivational factors but experimental work in the exercise domain itself could be expanded to better examine the causal mechanisms and process aspects of motivation for physical activity. Cross-sectional research is now abundant, and generally supportive, but it needs to be complemented with more applied intervention and translational studies that adequately model, implement, and evaluate key hypotheses about why and how individuals adopt and sustain more physically active lifestyles.

The methodology used in this review may also limit its conclusions. First, unpublished studies, evidence from grey literature, and data from non-English publications were not included. Although this is a frequent occurrence in scientific systematic review papers, it may provide an incomplete account of all studies in this area. Second, the way in which results from each study were classified and quantified (see Table 3) is somewhat arbitrary and subject to criticism and various interpretations. Third, as stated before, the decision to only evaluate direct paths is also inherently limiting considering that the distal effects of some variables on behavior is thought to be mediated by other intermediate variables. Unfortunately, few studies are available to assess these more complete causal paths. Finally, our definition of "behavioral variable" to describe the outcome of choice, lumping together self-report and direct measures of behavior, and also attendance and stages of change is clearly not without reproach. Although we felt this was the best decision considering the relative paucity of studies for various measures, future studies might want to be more specific and/or selective in their outcomes of choice.

In sum, it is clear that the exercise domain has provided fertile ground for testing SDT's precepts. While testing and developing theory is a worthwhile activity in its own right, the real significance of SDT will be realized if it can be employed to actually make a positive difference in peoples' lives. In this regard, the growing evidence for the utility of SDT-based interventions for promoting the adoption and maintenance of exercise is a significant advance. Future studies would do well to include biological markers of successful exercise-related outcomes such as increased fitness and reductions in disease risk factors. Similarly, studies that include markers of psychological well-being and mental health, such as self-esteem, vitality, and symptoms of anxiety and depression symptomatology would also be useful, given that according to SDT only autonomously regulated behaviors can translate into enhanced psychological wellness. Extending SDT´s applicability beyond behavioral engagement in exercise to actual improvements in health and well-being would thus be another important step for SDT research to influence health care policy and delivery.

Endnotes

[a]Exercise outcomes covered in this review include what is normally termed "exercise" (purposeful and formalized leisure-time physical activity, often with the goal of improving fitness or health) but also, in a few cases, less structured forms of exercise (e.g., walking minutes), energy expenditure measures, and accelerometry data (which cannot distinguish between different forms of activity). Although the term "physical activity" would aptly cover the entire range of outcomes in this review, "exercise" is a more specific term to what the large majority of studies measured, with the use of instruments such as the Godin Leisure Time Exercise Questionnaires (LTEQ, used in 55

independent samples [77.5%]). For this reason, we will use the two terms indiscriminately in this review.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
PJT conceived this manuscript and led the writing team. EVC conducted the study search, summarized the quantitative review, and drafted the Results section. DM made substantial contributions to the Discussion section. DM, RMR, EVC, and MNS revised the entire manuscript and made important contributions in various sections. All authors read and approved the final version of the manuscript.

Acknowledgments
None.

Author details
[1]Faculty of Human Kinetics, Technical University of Lisbon, Estrada da Costa, 1495-688, Cruz Quebrada, Portugal. [2]School of Sport, Health and Exercise Sciences, Bangor University, Bangor, UK. [3]Department of Clinical and Social Sciences in Psychology, University of Rochester, Rochester, NY, USA.

Received: 19 December 2011 Accepted: 22 June 2012
Published: 22 June 2012

References
1. Physical Activity Guidelines Advisory Committee: *Physical activity guidelines advisory committee report*. Washington DC: U.S. Department of Health and Human Services; 2008.
2. Sisson SB, Katzmarzik PT: **International prevalence of physical activity in youth and adults.** *Obesity Reviews* 2008, **9:**606–614.
3. **Eurobarometer: Sport and Physical Activity.** *Special Eurobarometer* 2010, **334/Wave 72.3**.
4. **Prevalence of Regular Physical Activity Among Adults—United States, 2001 and 2005.** *MMWR* 2007, **56:**1209–1212.
5. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS: **Physical activity of Canadian Adults: Accelerometer data from the 2007 to 2009 Canadian Health Measures Survey.** *Statistics Canada, Health Reports* 2011 **22, Catalogue no. 82-003-XPE**.
6. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I, Nieman DC, Swain DP: **Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise.** *Medicine & Science in Sports & Exercise* 2011, **43:**1334–1359.
7. Ryan RM, Williams GC, Patrick H, Deci EL: **Self-determination theory and physical activity: The dynamics of motivation in development and wellness.** *Hellenic Journal of Psychology* 2009, **6:**107–124.
8. Korkiakangas EE, Alahuhta MA, Laitinen JH: **Barriers to regular exercise among adults at high risk or diagnosed with type 2 diabetes: a systematic review.** *Health Promot Int* 2009, **24:**416–427.
9. Markland D: **The mediating role of behavioural regulations in the relationship between perceived body size discrepancies and physical activity among adult women.** *Hellenic Journal of Psychology* 2009, **6:**169–182.
10. Ryan RM, Deci EL: **Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.** *American Psychologist* 2000, **55:**68–78.
11. Hall A: **Tonic for Gyms?** *Recreation* 2008:24–26.
12. Deci EL, Ryan RM: **The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior.** *Psychological Inquiry* 2000, **11:**227–268.
13. Deci EL: *Intrinsic Motivation*. New York: Plenum; 1975.
14. Deci E, Ryan R: *Intrinsic motivation and self-determination in human behavior*. New York: Plenum; **1985:**113–148.
15. Ryan RM, Deci EL: **Overview of self-determination theory: An organismic dialectical perspective.** In *Handbook of Self-Determination Research*. Edited by Deci EL, Ryan RM. Rochester, NY: Rochester University Press; **2002:** 3–33
16. Wilson PM, Rogers WT, Rodgers WM, Wild TC: **The psychological need satisfaction in exercise scale.** *Journal of Sport and Exercise Psychology* 2006, **28:**231–251.
17. Fortier MS, Sweet SN, O'Sullivan TL, Williams GC: **A self-determination process model of physical activity adoption in the context of a randomized controlled trial.** *Psychology of Sport and Exercise* 2007, **8:**741–757.
18. Silva MN, Vieira PN, Coutinho SR, Minderico CS, Matos MG, Sardinha LB, Teixeira PJ: **Using self-determination theory to promote physical activity and weight control: a randomized controlled trial in women.** *Journal of Behavioral Medicine* 2010, **33:**110–122.
19. Sebire S, Standage M, Vansteenkiste M: **Examining goal content in the exercise domain: Intrinsic versus extrinsic goals and cognitive, affective, and behavioural outcomes, and psychological need satisfaction.** *Journal of Sport and Exercise Psychology* 2009, **31:**189–210.
20. Vansteenkiste M, Niemiec CP, Soenens B: **The development of the five mini-theories of self-determination theory: an historical overview, emerging trends, and future directions.** In *The Decade Ahead: Theoretical Perspectives on Motivation and Achievement (Advances in Motivation and Achievement, Volume 16)*. Edited by Urdan TC, Karabenick SA: Emerald Group Publishing Limited; 2010: 105–165
21. Grouzet FM, Kasser T, Ahuvia A, Dols JM, Kim Y, Lau S, Ryan RM, Saunders S, Schmuck P, Sheldon KM: **The structure of goals across 15 cultures.** *Journal of Personality and Social Psychology* 2005, **89:**800–816.
22. Kasser T, Ryan RM: **Further examining the American Dream: Differential correlates of intrinsic and extrinsic goals.** *Personality and Social Psychology Bulletin* 1996, **22:**80–87.
23. Hagger MS, Chatzisarantis NL: **Self-determination theory and the psychology of exercise.** *International Review of Sport and Exercise Psychology* 2008, **1:**79–103.
24. Wilson P, Mack D, Grattan K: **Understanding Motivation for Exercise: A Self-Determination Theory Perspective.** *Canadian Psychology* 2008, **49:**250–256.
25. Blair SN, Haskell WL, Ho P, Paffenbager RS Jr: **Vranizan KM, Farquhar JW, Wood PD: Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments.** *American Journal of Epidemiology* 1985, **122:**794–804.
26. Godin G, Shepherd R: **A simple method to assess exercise behavior in the community.** *Canadian Journal of Applied Sport Science* 1985, **10:**141–146.
27. Ryan RM, Frederick CM, Lepes D, Rubio N, Sheldon KM: **Intrinsic motivation and exercise adherence.** *International Journal of Sport Psychology* 1997, **28:**335–354.
28. Wininger SR: **Self-determination theory and exercise behavior: An examination of the psychometric properties of the exercise motivation scale.** *Journal of Applied Sport Psychology* 2007, **19:**471–486.
29. Wilson PM, Blanchard CM, Nehl E, Baker F: **Predicting physical activity and outcome expectations in cancer survivors: An application of self-determination theory.** *Psycho-Oncology* 2006, **15:**567–578.
30. Buckworth J, Lee RE, Regan G, Schneider LK, DiClemente CC: **Decomposing intrinsic and extrinsic motivation for exercise: Application to stages of motivational readiness.** *Psychology of Sport and Exercise* 2007, **8:**441–461.
31. Rodgers WM, Hall CR, Duncan LR, Pearson E, Milne MI: **Becoming a regular exerciser: Examining change in behavioural regulations among exercise initiates.** *Psychology of Sports and exercise* 2010, **11:**378–386.
32. Silva MN, Markland D, Carraça EV, Vieira PN, Coutinho SR, Minderico CS, Matos MG, Sardinha LB, Teixeira PJ: **Exercise Autonomous Motivation Predicts 3-Year Weight Loss in Women.** *Medicine and Sciences in Sports and Exercise* 2011, **43:**728–737.
33. Silva MN, Markland DM, Vieira PN, Coutinho SR, Carraça EV, Palmeira AL, Minderico CS, Matos MG, Sardinha LB, Teixeira PJ: **Helping Overweight Women Become More Active: Need Support and Motivational Regulations for Different Forms of Physical Activity.** *Psychology of Sports and exercise* 2010, **11:**591–601.
34. Hagger MS, Chatzisarantis NL, Harris J: **From psychological need satisfaction to intentional behavior: Testing a motivational sequence in two behavioral contexts.** *Personality and Social Psychology Bulletin* 2006, **32:**131–148.
35. Hagger MS, Chatzisarantis NL, Harris J: **The process by which relative autonomous motivation affects intentional behavior: Comparing effects across diet and exercise behaviors.** *Motivation and Emotion* 2006, **30:**307–321.
36. Fortier MS, Wiseman E, Sweet SN, O'Sullivan TL, Blanchard CM, Sigal RJ, Hogg W: **A moderated mediation of motivation on physical activity in**

the context of the PAC randomized control trial. *Psychology of Sport and exercise* 2011, **12:**71–78.
37. Markland D, Tobin V: **A modification of the behavioural regulation in exercise questionnaire to include an assessment of amotivation.** *Journal of Sport and Exercise Psychology* 2004, **26:**191–196.
38. Edmunds J, Ntoumanis N, Duda J: **Adherence and well-being in overweight and obese patients referred to an exercise on prescription scheme: A Self-Determination Theory perspective.** *Psychology of Sport and Exercise* 2007, **8:**722–740.
39. Edmunds J, Ntoumanis N, Duda J: **Testing a self-determination theory-based teaching style intervention in the exercise domain.** *European Journal of Social Psychology* 2008, **38:**375–388.
40. Levy SS, Cardinal BJ: **Effects of a self-determination theory-based mail-mediated intervention on adults' exercise behavior.** *American Journal of Health Promotion* 2004, **18:**345–349.
41. Duncan LR, Hall CR, Wilson PM, O J: **Exercise motivation: A cross-sectional analysis examining its relationships with frequence, intensity, and duration of exercise.** *International Journal of Behavioral Nutrition and Physical Activity* 2010, **7 (doi:** 10.1186/1479-5868-7-7**)**.
42. Brickell T, Chatzisarantis N: **Using self-determination theory to examine the motivational correlates and predictive utility of spontaneous exercise implementation intentions.** *Psychology of Sport and exercise* 2007, **8:**758–770.
43. Peddle CJ: **R.C. P, T.C. W, J. A., K.S. C: Medical, demographic, and psychosocial correlates of exercise in colorectal cancer survivors: an application of self-determination theory.** *Supportive Care in Cancer* 2008, **16:**9–17.
44. Edmunds J, Ntoumanis N, Duda J: **A Test of Self-Determination Theory in the Exercise Domain.** *Journal of Applied Social Psychology* 2006, **36:**2240–2265.
45. Wilson PM, Rodgers WM, Fraser SN, Murray TC: **The relationship between exercise regulations and motivational consequences.** *Research Quarterly for Exercise and Sport* 2004, **75:**81–91.
46. Ingledew DK, Markland D: **The role of motives in exercise participation.** *Psychology and Health* 2008, **23:**807–828.
47. Craike M: **Application of self-determination theory to study determinants of regular participation in leisure-time physical activity.** *World Leisure Journal* 2008, **50:**58–70.
48. Lewis M, Sutton A: **Understanding exercise behaviour: Examining the intereaction of exercise motivation and personality in predicting exercise frequency.** *Journal of Sport Behavior* 2011, **34:**82–97.
49. Wilson PM, Rodgers WM, Fraser SN: **Examining the Psychometric Properties of the Behavioral Regulation for Exercise Questionnaire.** *Measurement in Physical Education and Exercise Science* 2002, **6:**1–21.
50. McDonough MH, Crocker PR: **Testing self-determined motivation as a mediator of the relationship between psychological needs and affective and behavioral outcomes.** *Journal of Sport and exercise Psychology* 2007, **29:**645–663.
51. Edmunds J, Ntoumanis N, Duda JL: **Examining exercise dependence symptomatology from a self-determination perspective.** *Journal of Health Psychology* 2006, **11:**887–903.
52. Thögersen-Ntoumani C, Ntoumanis N: **The role of self-determined motivation in the understanding of exercise-related behaviours, cognitions and physical self-evaluations.** *Journal of Sports Sciences* 2006, **24:**393–404.
53. Kwan BM: **Caldwell Hooper AE, Magnan RE, Bryan AD: A longitudinal diary study of the effects of causality orientations on exercise-related affect.** *Self and Identity* 2011, **10:**363–374.
54. Sorensen M: **Motivation for physical activity of psychiatric patients when physical activity was offered as part of treatment.** *Scandinavian Journal of Medicine Sciences in Sports* 2006, **16:**391–398.
55. Daley AJ, Duda J: **Self-determination, stage of readiness to change for exercise, and frequency of physical activity in young people.** *European Journal of Sport Sciences* 2006, **6:**231–243.
56. Rose EA, Parfitt G, Williams S: **Exercise causality orientations, behavioural regulation for exercise and stage of change for exercise: exploring their relationships.** *Psychology of Sport and Exercise* 2005, **6:**399–414.
57. Mullan E, Markland D: **Variations in Self-Determination Across the Stages of Change for Exercise in Adults.** *Motivation and Emotion* 1997, **21:**349–362.
58. Wilson PM, Rodgers WM, Blanchard CM, Gessell J: **The Relationship Between Psychological Needs, Self-Determined Motivation, Exercise Attitudes, and Physical Fitness.** *Journal of Applied Social Psychology* 2003, **33:**2373–2392.
59. Frederick CM, Ryan RM: **Differences in motivation for sport and exercise and their relations with participation and mental health.** *Journal of Sport Behavior* 1993, **16:**124–146.
60. Markland D, Ingledew D: **The measurement of Exercise Motives: factorial validity and Invariance across gender of a revised Exercise Motivation Inventory.** *Brit J Health Psych* 1997, **2:**361–376.
61. McAuley E, Duncan T, Tammen VV: **Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis.** *Res Q Exerc Sport* 1989, **60:**48–58.
62. Lee RE, DiClemente CC: **Extrinsic and intrinsic motivation: Which is important for exercise?** *Medicine and Science in Sports and Exercise* 2001, **33:**S112.
63. Sebire S, Standage M, Vansteenkiste M: **Development and validation of the Goal Content for Exercise Questionnaire.** *Journal of Sport and Exercise Psychology* 2008, **30:**353–378.
64. Segar ML, Spruijt-Metz D, Nolen-Hoeksema S: **Go figure? Body-shape motives are associated with decreased physical activity participation among midlife women.** *Sex Roles* 2006, **54:**175–187.
65. Segar ML, Eccles JS, Richardson CR: **Type of physical activity goal influences participation in healthy midlife women.** *Women's Health Issues* 2008, **18:**281–291.
66. Wilson P, Mack D, Grattan K: **Understanding Motivation for Exercise: A Self-Determination Theory Perspective.** *Canadian Psychology* 2008, **49**.
67. Silva MN, Markland D, Minderico CS, Vieira PN, Castro MM, Coutinho SR, Santos TC, Matos MG, Sardinha LB, Teixeira PJ: **A randomized controlled trial to evaluate self-determination theory for exercise adherence and weight control: rationale and intervention description.** *BMC Public Health* 2008, **8:**234–246.
68. Mildestvedt T, Meland E, Eide GE: **How important are individual counseling, expectancy beliefs and autonomy for the maintenance of exercise after cardiac rehabilitation?** *Scandinavian Journal of Public Health* 2008, **36:**832–840.
69. Jolly K, Duda J, Daley AJ, Eves FF, Mutrie N, Ntoumanis N, Rouse PC, Lodhia R, Williams GC: **Evaluation of a standard provision versus an autonomy promotive exercise referral programme: Rationaly and study design.** *BMC Public Health* 2009, **9:**176.
70. Koestner R, Losier GF: **Distinguishing three ways of being internally motivated: a closer look at introjection, identification, and intrinsic motivation.** In *Handbook of self-determination research.* Edited by Deci EL, Ryan RM. Rochester, NY: University of Rochester Press; 2002: 101–121
71. Vallerand RJ, Fortier MS, Guay F: **Self-determination and persistence in a real-life setting: toward a motivational model of high school dropout.** *Journal of Personality and Social Psychology* 1997, **72:**1161–1176.
72. Vieira PN, Mata J, Silva MN, Coutinho SR, Santos TC, Minderico CS, Sardinha LB, Teixeira PJ: **Predictors of psychological well-being during behavioral obesity treatment in women** *Journal of Obesity (doi:101155/2011/936153)* 2011.
73. Assor A, Vansteenkiste M, Kaplan A: **Identified Versus Introjected Approach and Introjected Avoidance Motivations in School and in Sports: The Limited Benefits of Self-Worth Strivings.** *Journal of Educational Psychology* 2009, **101:**482–497.
74. Cusumano DL, Thompson JK: **Body image and body shape ideals in magazines: exposure, awareness, and internalization.** *Sex Roles* 1997, **37:**701–721.
75. Shrout PE, Bolger N: **Mediation in experimental and nonexperimental studies: New procedures and recommendations.** *Psychological Methods* 2002, **7:**422–445.
76. Halvary H, Ulstad SA, Bagoien TE, Skjesol K: **Autonomy support and its links to physical activity and competitive performance: Mediations through motivation, competence, action orientation and harmonious passion, and the moderator role of autonomy support by perceived competence.** *Scandinavian Journal of Educational Research* 2009, **53:**533–555.
77. Puente R, Anshel MH: **Exercisers' perceptions of their fitness instructor's interacting style, perceived competence, and autonomy as a function of self-determined regulation to exercise, enjoyment, affect, and exercise frequency.** *Scandinavian Journal of Psychology* 2010, **51:**38–45.
78. Markland D, Ingledew DK: **Exercise Participation Motives: A Self-Determination Theory Perspective.** In *Intrinsic Motivation and Self-Determination in Exercise and Sport.* Edited by Hagger MS, Chatzisarantis NL: United States of America: Human Kinetics; 2007: 23–35

79. Ingledew DK, Markland D, Ferguson E: **Three levels of exercise motivation.** *Applied Psychology: Health and Well-Being* 2009, **1:**336–355.
80. Cash TF, Novy PL, Grant JR: **Why do women exercise? Factor analysis and further validation of the Reasons for Exercise Inventory.** *Perceptual and Motor Skills* 1994, **78:**539–544.
81. Sheldon KM, Ryan RM, Deci EL, Kasser T: **The independent effects of goal contents and motives on well-being: It's both what you pursue and why you pursue it.** *Personality and Social Psychology Bulletin* 2004, **30:**475–486.
82. Rhodes RE, Pfaeffli LA: **Review Mediators of physical activity behaviour change among adult non-clinical populations: a review update.** *International Journal of Behavioral Nutrition and Physical Activity* 2010, **7:**37–48.
83. Williams GC, Grow VM, Freedman ZR, Ryan R, Deci E: **Motivational predictors of weight loss and weight-loss maintenance.** *Journal Personal Social Psychology* 1996, **70:**115–126.
84. Ryan R, Connell J: **Perceived locus of causality and internalization: Examining reasons for acting in two domains.** *Journal of Personality and social Psychology* 1989, **57:**749–761.
85. Wilson PM, Rodgers WM, Loitz CC, Scime G: **"It's who I am. . .really!": The importance of integrated regulation in exercise contexts.** *Journal of Applied Biobehavioral Research* 2006, **11:**79–104.
86. Landry JB, Solmon M: **African American women's self-determination across the stages of change for exercise.** *Journal of Sport and Exercise Psychology* 2004, **26:**457–469.
87. Milne HM: **K.E. W, A. G., S. G, K.S. C: Self-determination theory and physical activity among breast cancer survivors.** *Journal of Sport and Exercise Psychology* 2008, **30:**23–38.
88. Markland D, Tobin VJ: **Need support and behavioural regulations for exercise among exercise referral scheme clients: The mediating role of psychological need satisfaction.** *Psychology of Sport and Exercise* 2010, **11:**91–99.
89. Moreno JA, Cervelló EM, Martínez A: **Measuring self-determination motivation in a physical fitness setting: validation of the Behavioral Regulation in Exercise Questionnaire-2 (BREQ-2) in a Spanish sample.** *The Journal of Sport Medicine and Physical Fitness* 2007, **47:**366–378.
90. Hall CR, Rodgers WM, Wilson PM, Norman P: **Imagery use and self-determined motivations in a community sample of exercisers and non-exercisers.** *Journal of Applied Social Psychology* 2010, **40:**135–152.
91. Standage M, Sebire S, Loney T: **Does exercise motivation predict engagement in objectively assessed bouts of moderate-intensity exercise?: A self-determination theory perspective.** *Journal of Sport and exercise Psychology* 2008, **30:**337–352.
92. Hurkmans EJ, Maes S, de Gucht V, Knittle K, Peeters AJ, Ronday HK: **Vliet Vlieland T: Motivation as a determinant of physical activity in patients with rheumatoid arthritis.** *Arthritis Care & Research* 2010, **62:**371–377.
93. Lutz RS, Karoly P, Okun MA: **The why and the how of goal pursuit: Self-determination goal process cognition, and participation in physical exercise.** *Psychology of Sport and Exercise* 2008, **9:**559–575.
94. Tsorbatzoudis H, Alexandris K, Zahariadis P, Grouios G: **Examining the relationship between recreational sport participation and intrinsic and extrinsic motivation and amotivation.** *Perceptual and Motor Skills* 2006, **103:**363–374.
95. Chatzisarantis NL, Biddle SJ: **Functional significance of psychological variables that are included in the Theory of Planned Behaviour: A Self-Determination Theory approach to the study of attitudes, subjective norms, perceptions of control, and intentions.** *European Journal of Social Psychology* 1998, **28:**303–322.
96. Matsumoto H, Takenaka K: **Motivational profiles and stages of exercise behavior change.** *International Journal of Sport and Health Science* 2004, **2:**89–96.
97. McNeill LH, Wyrwick KW, Brownson RC, Clarck EM, Kreuter MW: **Individual, social environmental, and physical environmental influences in physical activity among black and white adults: A structural equation analysis.** *Annals of Behavioral Medicine* 2006, **31:**36–44.
98. Russell KL, Bray SR: **Self-determined motivation predicts independent, home-based exercise following cardiac rehabilitation.** *Rehabilitation Psychology* 2009, **54:**150–156.
99. Russell KL, Bray SR: **Promoting self-determined motivation for exercise in cardiac rehabilitation: the role of autonomy support.** *Rehabilitation Psychology* 2010, **55:**74–80.
100. Fortier MS, Kowal J, Lemyre L, Orpana HM: **Intentions and actual physical activity behavior change in a community-based sample of middle-aged women:Contributions from the theory of planned behavior and self-determination theory.** *International Journal of Sport and Exercise Psychology* 2009, **9:**46–67.
101. Barbeau A, Sweet SN, Fortier MS: **A path-analytic model of self-determination theory in a physical activity context.** *Journal of Applied Biobehavioral Research* 2009, **14:**103–118.
102. Sweet SN, Fortier MS, Guérin E, Tulloch H, Sigal RJ, Kenny GP, Reid RD: **Understanding physical activity in adults with type 2 diabetes after completing an exercise intervention trial: A mediation model of self-efficacy and autonomous motivation.** *Psychol Health Med* 2009, **14:**419–429.
103. Vlachopoulos SP, Michailidou S: **Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: The basic psychological needs in exercise scale.** *Measurement in Physical Education and Exercise Science* 2006, **10:**179–201.
104. Frederick CM, Morrison C, Manning T: **Motivation to participate, exercise affect, and outcome behaviors toward physical activity.** *Perceptual and Motor Skills* 1996, **82:**691–701.
105. Sit CH, Kerr JH, Wong IT: **Motives for and barriers to physical activity participation in middle-aged Chinese women.** *Psychology of Sport and exercise* 2008, **9:**266–283.
106. Davey J, Fitzpatrick M, Garland R, Kilgour M: **Adult participation motives: Empirical evidence from a workplace exercise program.** *European Sport Management Quarterly* 2009, **9:**141–162.
107. Ingledew DK, Markland D, Medley AR: **Exercise Motives and Stages of Change.** *Journal of Health Psychology* 1998, **3:**477–489.
108. Levy AR, Polman RC, Borkoles E: **Examining the relationship between perceived autonomy support and age in the context of rehabilitation adherence in sport.** *Rehabilitation Psychology* 2008, **53:**224–230.

doi:10.1186/1479-5868-9-78
Cite this article as: Teixeira *et al.*: Exercise, physical activity, and self-determination theory: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity* 2012 **9**:78.

Hill et al. BMC Geriatrics (2018) 18:3
DOI 10.1186/s12877-017-0683-1

BMC Geriatrics

RESEARCH ARTICLE Open Access

What works in falls prevention in Asia: a systematic review and meta-analysis of randomized controlled trials

Keith D. Hill[1*], Plaiwan Suttanon[2], Sang-I Lin[3], William W.N. Tsang[4], Asmidawati Ashari[5,1], Tengku Aizan Abd Hamid[5], Kaela Farrier[1] and Elissa Burton[1]

Abstract

Background: There is strong research evidence for falls prevention among older people in the community setting, although most is from Western countries. Differences between countries (eg sunlight exposure, diet, environment, exercise preferences) may influence the success of implementing falls prevention approaches in Asian countries that have been shown to be effective elsewhere in the world. The aim of this review is to evaluate the scope and effectiveness of falls prevention randomized controlled trials (RCTs) from the Asian region.

Method: RCTs investigating falls prevention interventions conducted in Asian countries from (i) the most recent (2012) Cochrane community setting falls prevention review, and (ii) subsequent published RCTs meeting the same criteria were identified, classified and grouped according to the ProFANE intervention classification. Characteristics of included trials were extracted from both the Cochrane review and original publications. Where ≥2 studies investigated an intervention type in the Asian region, a meta-analysis was performed.

Results: Fifteen of 159 RCTs in the Cochrane review were conducted in the Asian region (9%), and a further 11 recent RCTs conducted in Asia were identified (total 26 Asian studies: median 160 participants, mean age:75.1, female:71.9%). Exercise (15 RCTs) and home assessment/modification ($n = 2$) were the only single interventions with ≥2 RCTs. Intervention types with ≥1 effective RCT in reducing fall outcomes were exercise (6 effective), home modification (1 effective), and medication (vitamin D) (1 effective). One multiple and one multifactorial intervention also had positive falls outcomes. Meta-analysis of exercise interventions identified significant benefit (number of fallers: Odds Ratio 0.43 [0.34,0.53]; number of falls: 0.35 [0.21,0.57]; and number of fallers injured: 0.50 [0.35,0.71]); but multifactorial interventions did not reach significance (number of fallers OR = 0.57 [0.23,1.44]).

Conclusion: There is a small but growing research base of falls prevention RCTs from Asian countries, with exercise approaches being most researched and effective. For other interventions shown to be effective elsewhere, consideration of local issues is required to ensure that research and programs implemented in these countries are effective, and relevant to the local context, people, and health system. There is also a need for further high quality, appropriately powered falls prevention trials in Asian countries.

Keywords: Falls prevention, Effectiveness, Asia, Elderly, Community

* Correspondence: Keith.Hill@Curtin.edu.au
[1]School of Physiotherapy and Exercise Science, Curtin University, GPO Box U1987, Perth, WA 6845, Australia
Full list of author information is available at the end of the article

Background

Falls are recognized as a major cause of death and growing burden of disease world-wide [1, 2]. There has been strong growth in the available research evidence from randomized controlled trials investigating interventions to reduce falls among older people living in the community, with the most recent (2012) Cochrane review reporting 159 studies with 79,193 participants [3]. However, data from several developed countries indicate that despite this high volume of quality research, that key national indicators such as rates of falls related hospitalizations continue to increase (for example, in Australia, trend data from 1999 to 2013 indicate an average 2% increase in age-standardized fall related hospitalizations per year) [4–6]. Importantly, if rates of fall related hospitalizations remain stable or increase, then with aging populations this means substantial growth in actual case numbers of people being hospitalized due to falls (see Fig. 8.3 in [4]).

Population aging is occurring at differing rates between countries [7]. Japan is the "oldest" country in the world, with 26% of its population aged over 65 years [8]. Many developed countries already have more than 15% of their population aged greater than 65 years, [8] and are expected to exceed 20% by 2050. In contrast, many developing countries in Asia such as Malaysia, Thailand, Indonesia, Cambodia, Vietnam and the Philippines have less than 10% of their population aged more than 65 years [9, 10]. However, Asia is home to 60% of the world's population, [8] and has the fastest aging population of any region in the world [11]. There is a need for a strong preventive approach to minimize the risk of falls and associated injuries as countries' aging populations grow.

The available research provides mixed evidence about the magnitude of the problem of falls between countries. While it is widely accepted from prospective, large scale representative samples in a number of Western countries that between 30 and 39% of the population aged greater than 65 will experience one or more falls in a 12 month period, [12–14] the proportion of older people reporting falls in the limited number of studies in Asian countries has generally been lower and considerably varied (14–34%, median 18%) [15–17]. Limitations to the design of most of these studies in Asian countries include utilizing retrospective data collection (recall of falls in the preceding 12 months), and lack of representative sampling. There may also be some cultural factors influencing whether or not a fall will be reported by an older person, [18] which may contribute to varying and under-reporting of falls outcomes.

Factors contributing to falls risk may also vary between countries and cultures [18, 19]. Differences in intrinsic factors may include sunlight exposure, diet, stature, exercise patterns and preferences, and knowledge and attitudes towards ageing and falls prevention. For example, in a recent qualitative study of older people in Thailand, falls prevention was not noted as a perceived benefit of exercise, and some family and cultural values were considered to be potential barriers to older people becoming involved in exercise [20]. Another qualitative study in India reported that older people often considered falls were random events, and not considered a health priority [21]. A study in China reported fatalistic perceptions about falls being common among older people, low levels of knowledge about falls prevention interventions were evident, and falls were often hidden from family and doctors, and were not often discussed openly [22]. Differences in extrinsic factors also exist, including the home and outdoor environment, and footwear [17, 21, 23]. Type of housing and flooring surfaces vary substantially across Asia, and outdoor environments such as footpaths are often poorly maintained or non-existent. A study in China highlighted environmental factors such as adequate lighting on stairs and adequate step width as protective of falls (and therefore poor lighting and narrow step width as risk factors for falls) [17]. Additionally, there can be differences in health services and systems, and engagement of older people in these. Focus is often on treatment of acute health conditions, and less on preventative care.

The growing evidence highlighting different rates of falls and potential falls risk factors between countries raises the question of whether interventions that have been shown to be effective in one country may need some tailoring if being introduced into countries that have considerable diversity to the country where the intervention was shown to be effective. One recent scoping review has considered this issue in the context of falls prevention studies that have been conducted in South East Asian countries [23]. This review reported some unique aspects of socioeconomic, geographical and cultural differences of South East Asian countries, and identified limited quality research investigating falls prevention interventions in this region. However, the majority of the studies reported were non-randomized trials, studies were limited to only South East Asian countries, and there was no quality assessment of the included studies. The objective of our systematic review paper was to review the *randomized controlled trial* research evidence conducted among older people living in the community across the Asian region to identify (1) the type and number of falls prevention interventions shown to be effective in Asian populations; and (2) gaps for future research investigating falls prevention interventions in Asian countries.

Methods

The 2012 community setting falls prevention Cochrane review was used as the basis for this review as it is the most recent and extensive systematic review and meta-analysis available on this topic [3]. Studies in the Cochrane review were categorized as to whether the study was undertaken in an (i) Asian, or (ii) non-Asian country. Additionally, a systematic review of the falls prevention literature published

since the 2012 Cochrane review was also conducted, to identify more recent falls prevention RCTs conducted in the Asian region. This update component was guided by the Additional file 1: PRISMA checklist to ensure the results are reported systematically [24].

Intervention types were classified according to the Prevention of Falls Network Europe (ProFaNE) classification as (i) single interventions, (ii) multiple interventions (two or more single interventions, applied to all participants), and (iii) multifactorial interventions (two or more interventions, targeted to an individual's risk factor profile, often based on a falls risk assessment process – different participants receive a different mix of interventions) [25]. Single interventions were further classified as described in the ProFANE classification as (a) exercise, (b) medication, (c) psychological, (d) environmental/assistive technology, (e) interventions to increase participant knowledge, (f) surgical interventions (eg cataract surgery, cardiac pacemaker surgery), (g) interventions to address incontinence, and (h) fluid or nutritional therapy. This intervention classification system was used in the 2012 Cochrane review, [3] and was also used to classify the more recent RCTs.

Eligibility criteria for RCTs published since 2012 Cochrane review

The additional studies included in this review met the following eligibility criteria: peer reviewed articles published from January 2012 – November 2016; studies undertaken in Asia; randomized controlled trials; people aged 60 years and over – at least 50% of the sample; living in the community, and reporting at least one falls outcome.

Information sources for additional RCTs

Studies were identified from six databases: Medline (Proquest), CINAHL, PubMed, PsycINFO, SPORTDiscus and Scopus for the time period described above. Reference lists from the included articles were also scanned. Only papers in English were included, no unpublished data, conference proceedings, books, poster abstracts or theses were included.

Search strategy for additional RCTs

The search strategy used a mix of keywords which could be identified in the title and/or abstract. The search strategy included fall* [Title/abstract] AND communit* [Title/abstract] AND RCT [Title/abstract] OR randomi* controlled trial [Title/abstract]. There were differences between the databases for language and syntax, and where this occurred only the abstract was searched.

Study selection for additional RCTs

Study selection was a three-stage process, stage one involved one author (KF) screening all of the additional identified papers based on their titles; stage two involved screening abstracts according to the eligibility criteria; and stage three involved two independent researchers (KF and EB) screening full articles. Where disagreements occurred the two researchers discussed the reasons for their decision, referring back to the eligibility criteria throughout the process, until they reached consensus.

Data collection process

Characteristics of the intervention programs undertaken in Asian countries were retrieved from original publications independently by two of the researchers (KH and EB/KF). Comparing characteristics of the interventions (eg for exercise – type, duration, frequency) provided an opportunity to explore factors that may have contributed to the success or failure of different interventions in Asian populations.

Study quality assessment for additional RCTS (post 2012 Cochrane review)

The Cochrane Collaboration's risk of bias tool [26] was used by two researchers (KF and EB) independently to assess methodological quality of each paper (study quality of the Cochrane review papers can be accessed in the Cochrane review). Categories assessed by the tool include sequence generation, allocation concealment, blinding (staff, participants and outcome assessors), incomplete outcome data, selective outcome reporting and other sources of bias [26]. Risk of bias was assessed as "low risk", "medium risk" or "high risk" [26]. Where disagreement occurred between the two assessors a third assessor (EJB) also assessed those domains, discussed with the two independent assessors their outcomes and came to a consensus.

Analysis

Extracted data from all the Asian studies were tabulated, and overall sample demographics and characteristics of the interventions undertaken in Asian countries summarized, to assist in interpretation of the primary outcomes.

The primary analyses for this paper involved meta-analysis of falls outcomes (e.g. number of falls) by intervention type (according to the ProFaNE classification) where there were two or more randomized controlled trials conducted in Asian countries with comparable data.

For the meta-analyses, number of participants and events (falls, fallers, non-injurious falls) for each group were sourced from all original articles (including Cochrane review articles). RevMan 5.3 software was used to conduct the analyses and generate the forest plots, using a Mantel-Haenszel's fixed effect model (with odds ratios and 95% confidence intervals calculated) [27]. Visual inspection of the forest plots and the I^2 statistic were used to assess heterogeneity. When heterogeneity was deemed as high ($I^2 > 50\%$) a random effects model was applied. Where a study had two intervention groups and one control group, the intervention groups (i.e.

dichotomous data) were summed (as recommended by Cochrane) for both the outcome (e.g. number of falls) and the sample size independently [27]. Subgroup analysis was undertaken where two or more programs/services within the intervention type could be grouped, for example Tai Chi within exercise interventions. Studies with group differences in baseline characteristics were omitted. Statistical significance was considered at $p \leq 0.05$. Some studies reported more than one single intervention arm compared to a control arm. In these cases, if separate falls data were reported for each of the single interventions, then these were reported separately in the tables and meta-analyses.

Results

Of the 159 randomized controlled trials reported in the 2012 Cochrane review [3], 15 (9%) were conducted in Asian countries (Japan, $n = 6$; [28–33] Taiwan $n = 6$; [34–39] Thailand, $n = 2$; [40, 41] and China/Hong Kong, $n = 1$ [42]). The remaining 144 published randomized trials were from non-Asian countries, including the United States of America ($n = 34$, 21%), Australia ($n = 27$, 17%), the United Kingdom ($n = 27$, 17%), Canada ($n = 12$, 8%), the Netherlands ($n = 9$, 6%), New Zealand ($n = 6$, 4%), Germany ($n = 6$, 4%), and other parts of Europe ($n = 16$, 11%). Several other studies were conducted across two or more countries ($n = 4$), although none of these included countries in Asia. Table 1 reports the number of randomized trial interventions classified by ProFaNE classification type, between Asian and non-Asian populations.

An additional 11 RCTs were identified that were conducted in Asian populations, met the inclusion criteria for the community based falls prevention Cochrane review, and were published after the 2012 Cochrane review (Japan $n = 6$; [43–48] Taiwan $n = 2$; [49, 50] and China/Hong Kong, [51] Malaysia [52] and Singapore [53] $n = 1$ each). See Fig. 1 for study selection flowchart to identify papers after 2012. Seven of the additional studies published since the 2012 review were exercise interventions.

Combining the Asian studies in the 2012 Cochrane review and subsequent Asian RCTS, the only single intervention types that had more than one randomized controlled trial conducted in Asia were exercise ($n = 15$), and environment (home hazard assessment and modification, $n = 2$ [38, 48]). There were also five multiple intervention studies [34, 37, 41, 50, 53] (two with two or more single intervention arms included as separate single interventions in the single intervention analyses [34, 37]) and four multifactorial intervention studies conducted in Asia [35, 36, 39, 40].

Samples

Overall, the sample size for the studies in Asia included in the 2012 Cochrane review were relatively small (median 150, minimum 52, maximum 1043) (Table 2), whereas for the full Cochrane review [3], the median sample size across all randomized trials irrespective of where the trial was conducted was 230 (minimum 10, maximum 9940). The more recent Asian randomised controlled trials (published since the Cochrane review) have been substantially larger than those pre-2012 (median sample size 196, minimum 68, maximum 710). The samples were mixed in terms of the age inclusion criteria, with two studies including people aged over 50 years [28, 52], five had samples aged ≥60 years [37, 39, 41, 49, 51], 14 had samples aged ≥65 years [29, 32–36, 38, 42, 44, 45, 47, 48, 50, 53], three samples were aged ≥70 years [30, 40, 46], one sample was aged ≥73 years [31], and one sample was aged ≥75 years [43]. The average age of participants across the included studies in Asia was 75.1 years, and samples were on average 71.9% female (five studies had female only samples [29–31, 46, 51]).

Intervention types

Table 1 reports details of individual intervention types conducted in Asia relative to the rest of the world, while Table 2 summarizes the main characteristics of the trials conducted in Asia, grouped by intervention type, including details of whether or not the intervention was effective in reducing the rate of falls or number of people falling.

Single interventions

Exercise interventions

Fifteen trials from Asia reported results for an exercise intervention as a single intervention, with only six reporting one or more positive fall related outcomes [31, 42–44, 46, 49]. The exercise trials generally had small sample sizes (60% $n < 70$, median sample size $n = 105$). Three of the effective exercise programs used Tai Chi (20% of all exercise trials) either as the sole intervention [42, 49] or combined with other balance and strength exercises [31]. In contrast, only five of the 51 (10%) exercise trials from the 2012 Cochrane review conducted outside of Asia involved Tai Chi. The other effective exercise approaches in Asian studies used an obstacle course [43], a multi-target stepping program [44], and a group balance and strength training program [46]. Only one of the ineffective exercise trials used Tai Chi, [34] other ineffective programs used home programs, [29, 38, 52] group programs (balance [47] or multimodal [45]), a group exercise program combined with trail walking, [33] combined group and home based program, [28] and a square stepping program [32]. Effective exercise programs ranged from 24 to 52 weeks duration, and had from 1 to 8 times per week of recommended exercise (included both supervised and home-based sessions).

In a separate review, Sherrington and colleagues reviewed 54 randomized controlled trials evaluating exercise interventions to reduce falls, and identified several key criteria that appeared to differentiate effective from ineffective falls prevention trials [54]. Two of the key

Table 1 Study types based on the ProFaNE intervention classification, for randomized trials conducted in Asia, and elsewhere in the world

	Number of randomized controlled trials (RCTs) in the 2012 Cochrane review				Additional RCTs conducted in Asia post 2012 Cochrane review	RCTs conducted in Asia (in the Cochrane review, and published subsequently) with significant improved falls outcome[#]
INTERVENTION TYPE	Asia	Rest of world	Total	Cochrane review conclusion (one or both falls outcomes) – all countries (see Footer for notes re ✓ / x)	Asia	Asia
Single intervention						
Exercise (includes studies that had other types of interventions in separate arms of the study, relative to a control group, so long as results were reported separately for comparison of the exercise and control interventions)	8 [28, 29, 31–34, 38, 42]	51	59	• Multi component group exercise ✓ • Multi component home exercise ✓ • Tai Chi ✓	7 [43–47, 49, 52]	• Multi-component group exercise ✓ • Tai Chi ✓ • Multi-target stepping program/obstacle course program ✓
Medication (drug target)	1 [30]	15	16	• Vitamin D supplementation (for those with baseline low vitamin D) ≈ • Psychotropic withdrawal ✓ • Prescribing modification ✓	0	• Vitamin K2, vitamin D2 and calcium supplementation (combined) ✓
Surgery (including cataract surgery, cardiac pacing)	0	5	5	• Cataract surgery (first eye) ✓ • Cardiac pacing (for those with carotid sinus hyper-sensitivity) ✓	0	x
Fluid or nutrition therapy	0	3	3	x	0	x
Psychological interventions (including cognitive behavioral therapy)	1 [37]	1	2	x	0	x
Environment / assistive technology	1[38]	12	13	• Home safety assessment and modification ✓ (more effective for those with high falls risk, and when done by occupational therapists)	1 [48]	• Home safety assessment and modification ✓
Knowledge interventions	1 [34]	4	5	x	0	x
Other (vibration intervention without exercise)	0	0	0	–	1 [51]	• Low magnitude/high frequency vibration (standing still on machine) ✓
Multiple interventions	3 [34, 37, 41]	15	18	Data not pooled because of heterogeneity of interventions	2 [50, 53]	• Education + geriatric clinic assessment ✓
Multiple interventions including an exercise intervention	2 [34, 37]	14	16	Separate meta-analysis for exercise related multiple interventions not conducted in the Cochrane review	2 [50, 53]	x
Multifactorial interventions	4 [35, 36, 39, 40]	36	40	RaR 0.76 (0.67–0.86) for rate of falls outcome	0	• Improved pre-operative, post-operative and post discharge management (multifactorial, multidisciplinary) for patients after hip fracture

Note – for multiple intervention studies where individual outcome data were reported for a single intervention as well as a multiple intervention arm, these are reported in both intervention types in the Table
✓ = significant reduction in fall rate or falls risk, or both; x = no significant effect on falls outcomes; ≈ = borderline significance
[#]One or more randomised trials conducted in Asia with significantly reduced falls outcome (falls, falls risk, fall injuries) – refer to Table 2 for details of individual studies

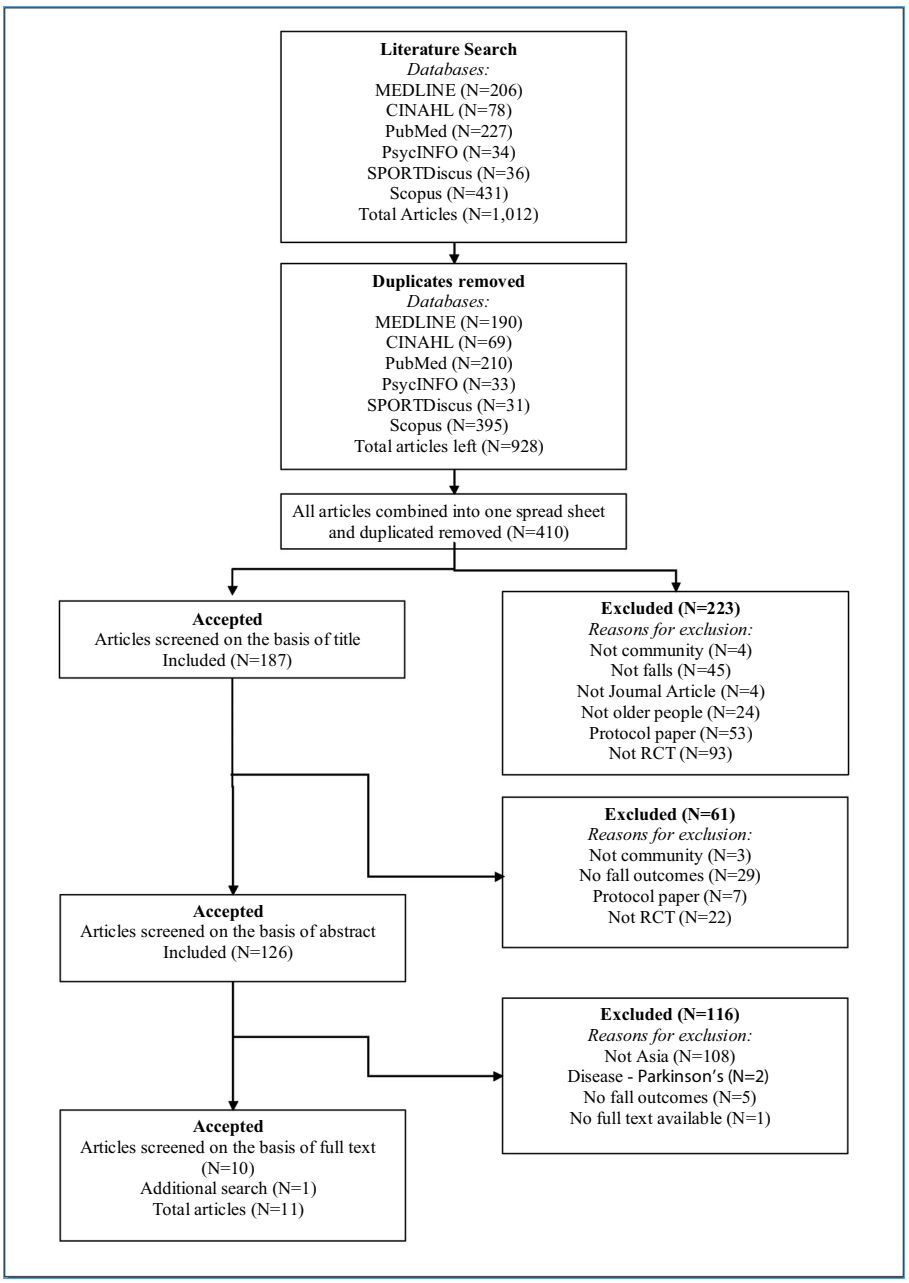


Fig. 1 Study selection flowchart for randomised trials published after the 2012 Cochrane

criteria related to: (1) the exercise intervention having a moderate to high challenge to balance; and (2) a minimum of 50 h overall exercise dosage. Of the six effective exercise trials in Asia, four (67%) met both of these criteria [31, 42, 46, 49]. The other two effective exercise studies [43, 44] met the "challenge to balance" criteria, but involved substantially less than the 50 h duration. Three of the ineffective exercise studies met both of Sherrington's criteria, but were likely to be underpowered (sample sizes of 57 [29], 68 [52] and 93 [47]). Most of the remaining ineffective exercise interventions did not meet the dosage criteria, although the majority met the balance challenge criteria. Although adherence to the exercise programs was reported in ten of the studies, the method of reporting was variable, making comparisons difficult. The study by Huang and colleagues reported high attrition in the Tai Chi group (52%) over the five month intervention [34]. All except

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria[a]
SINGLE INTERVENTIONS							
Exercise							
Suzuki et al., 2004 (Japan) Ø[31]	*N* = 52; Women only. Mean age = 78; Participating in gerontology longitudinal study. RCT.	Group exercise program strength, balance, walking, Tai Chi, supplemented by home exercise program.	Pamphlet and advice on falls prevention.	Group – 10 × 1 hour session × 6 months; Home – 30 min 3 x/week.	75.3% average attendance at exercise classes. Follow-up at 8 and 20 months for falls data. 20 month data reported in the Cochrane review.	✓ • F – RaR = 0.35 [0.14, 0.88] • FR – RR = 0.25 [0.08, 0.78] Intervention group significant improvements on tandem walking, Functional Reach and knee extension power.	Bal ✓ Dose ✓
Woo et al., 2007 (Hong Kong, China) [42]	*N* = 180 (90 men and 90 women); Mean age 69 years, range 65–74. RCT with block randomization by gender.	(1) 24 form Yang style Tai Chi (2) Resistance exercises with thera-band.	Usual care.	(1) Tai Chi 24 form, 3 x/week × 12 months (2) Resistance exercise 3 x/week 12 months (arm lifting, hip abduction, heel raise, hip flexion, hip extension, squatting ankle dorsiflexion).	High compliance – Tai Chi 81%, resistance training 76%. 12 month follow-up.	(1) ✓ (2) x • (1) FR – RR = 0.48 [0.29, 0.80]## • (2) FR - 0.77 [0.52, 1.14] No significant between group differences for strength, balance or flexibility measures. Sub-group analysis in women identified significantly reduced rate of loss of total hip bone mineral density in both exercise groups relative to control group.	(1) Bal ✓ Dose ✓ (2) Bal? (not clear starting position or hand support) Dose ✓
Lin et al., 2007 (Taiwan, rural/ agricultural area) [38]	*N* = 150, recent fallers. 51% female. Mean age = 76.8 years. RCT.	(1) Home exercise with physio (individualized flexibility, strength & balance exercises) (2) Home safety assessment and modification by public health worker (see below for results)	Education and social visit every 2 weeks with public health worker + falls prevention brochure.	(1) 40–60 min, 3 x/week × 4 months, physio visit every 2 weeks (2) 30–40 min visit every 2 weeks to perform safety assessment and make recommendations.	Assessment at 2 and 4 months for quality of life, depression and physical performance measures. Adherence to exercise program not reported.	x (exercise) • F – RaR = 0.67 [0.35, 1.28] Exercise group had significant improvements on Functional Reach, Tinetti POMA, and fear of falling than the Education group.	(1) Home exercise program Bal ✓ Dose x
Shigematsu et al., 2008 (Japan) [32]	*N* = 68. 63% female. Mean age 69 years, range 65–74. RCT.	Square stepping exercise, included forward, backward, lateral, and oblique stepping patterns on a thin felt mat, added challenge after familiarity by walking on toes; and increased complexity of step pattern.	Supervised outdoor walking program - 40 min, 1 x/week × 12 weeks, emphasis on increasing steps.	70 min, 2 x/week × 12 weeks.	Adherence: Square stepping exercise – 91% of sessions, outdoor walking – 84% of sessions. Falls data followed up for 8 months.	x • F – RaR = 0.70 [0.23, 2.13] • FR – RR = 0.64 [0.21, 1.95] Significant improvement in the Square stepping exercise group on leg extension power, forward/backward tandem walking, stepping with both feet, simple/choice reaction time, and perceived health status.	Bal ✓ Dose x
Iwamoto et al., 2009 (Japan) Ø [28]	N = 68. Attending orthopaedic clinics. 90% female. Age > 50 (mean age = 76.4 years). RCT.	Calisthenics, balance, power and walking exercises (home based, but 3 x/week supervision in clinic).	Usual care.	Daily exercise, with supervision in clinic 3 x/week × 30 min. Duration of exercise 5 months.	Exercise adherence reported as 100%.	x • FR – RR = 0.11 [0.01, 1.52]# Study reports significant reduction in falls, but effect not significant in Cochrane review. Significant improvement in flexibility, balance, mobility and gait measures in the exercise group relative to the control group.	Bal ✓ Dose x

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria*
Kamide et al., 2009 (Japan) Ø [29]	*N* = 57. 100% female. Attending employment agency (for light work or volunteer activity). Age > 65 (mean age = 71 years). RCT.	1 × 1 h education session (focus on osteoporosis fracture prevention and exercise) and 1 h training for home exercise; then home based exercise program (flexibility, balance, strength and impact exercises).	Usual care. Therapist contact by phone or mail each 3 months.	3 days/week × 6 months. No home visits by therapist re exercise program, but contact by phone or mail monthly to support motivation.	Exercise adherence - 82% of exercise participants completed the study. Of these, 91% performed exercises at least twice weekly. Follow-up over 12 months.	x • FR – RR = 0.38 [0.02, 7.91] Significant improvement in Timed Up and Go in exercise group relative to the control group.	Bal ✓ Dose ✓
Yamada et al., 2010 (Japan) Ø [33]	*N* = 60. % female not stated. Age > 65 (mean age approx 80 years). RCT.	Exercise class + multi-component trail walking program. Variability in how program was implemented to add challenge and motivation.	Exercise class + simple indoor walking program.	Exercise class 1 x/week × 16 weeks (60 min, included aerobic, strength, balance and flexibility exercises). Trail walking involved walking to set flags in order, changing direction, focus on speed. 30 min/session. Indoor walking program was 30 min/session.	Adherence – median for both group 100%	x • F – RaR = 0.45 [0.14, 1.49] • FR – RR = 0.45 [0.18, 1.13] Study reported reduced falls at 6 months, not sustained at 12 months. Cochrane review utilized 12 months falls data. Significant improvement in Timed Up and Go, walking task, and dual task gait tasks for Trail walking group relative to indoor walking group.	Trail walking program Bal ✓ Dose x NB – both groups received multimodal exercise class.
Huang et al., 2010 (Taiwan) Ø [34]	*N* = 261 randomized, *N* = 163 follow-up. 48% female after loss to follow-up. Age > 65 (mean age = 71 years). Cluster RCT (by village). 5 month intervention and 12 month post intervention follow-up.	(1) Education (falls risk factors, and reducing falls risk) (2) Tai Chi (3) Combined education + Tai Chi as above.	Not described.	(1) 5 × 1 h group sessions across 5 months. (2) 40 min sessions, 3 x/week × 5 months. (3) Combined education and Tai Chi program.	18 month follow-up for falls data. High drop-out rates over 5 month intervention period – education (52%), Tai Chi (52%), Education + Tai Chi (34%), control (6%).	(2) x • FR – RR = 0.51 [0.02, 12.49] Cochrane review used raw data at 5 months only, as 18 month raw data not provided. Cochrane review reports all interventions as non-significant (although combined education and Tai Chi reported as effective in reducing falls at 5 months, and all three interventions as effective in reducing falls at 18 months in the paper). Secondary measures only compared pre –post (within group).	(2) Tai Chi Bal ✓ Dose x
Yamada et al. 2012 (Japan) Ø [43]	*N* = 157. 87.8% female. Age ≥ 75 years (mean age 85.5 years).	All participants (intervention and control groups) received 45 min of group training sessions 1 x weekly (strength, balance, cardiovascular, flexibility exercises). In addition, the intervention group undertook a complex obstacle course negotiation program each session (finding marker to walk to, direction changes, avoiding obstacles).	Same main exercise class as intervention group, but undertook an additional simple obstacle course negotiation program (6 trials / session of 15 m walkway with obstacles interspersed along walkway).	24 weeks, once weekly sessions. Two trials of finding 15 markers/session in addition to common exercise program once weekly, 45 min duration).	Median adherence in both groups – 96%. No significant difference between groups on balance and mobility measures after intervention (except for a complex obstacle negotiation task, with the Intervention group achieving significantly greater improvement than the Control group).	✓ Fallers – 2 in intervention group (2.8%), and 19 (26.0%) in the control (simple obstacle course group). IRR for falls in the control group relative to intervention group was 9.37 (2.26–38.77). IRR for fall-related fractures in the control group relative to intervention group was 7.89 (1.01–61.49).	Bal ✓ Dose x (comparing difference in exercise time between two groups)
Yamada et al., 2013 (Japan) Ø [44]	*N* = 264. 57.3% female. Age ≥ 65 years (mean age 76.7 years).	All participants (Intervention and Control groups) received 30 min of group training session (2 x weekly × 30 min, aerobic, strength, balance and flexibility	Same main exercise class as intervention group, but undertook an additional	Twice weekly for 24 weeks. Total time spent walking on the mat during the Multi-task	Intervention group achieved significant improvement relative to Control group in	✓ Fallers: 13 Intervention participants (11.6%) and 39 (33.0%) in the	Bal ✓ Dose x

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria*
		exercises). In addition, the Intervention group undertook a Multi-task Stepping activity each session, that involved varied stepping pattern along a walkway, at comfortable speed.		Stepping Intervention was 1 to 2 min/task, repeated 4 times/session (total additional time of Multi-task Stepping/ session was 5 to 7 min).	walk time and Timed Up and Go.	Control group fell during the 12-month follow-up period. IRR for falls in the Intervention group relative to the Control group was 0.35 (0.19–0.66). Fall-related fractures: 3 participants in the Intervention group had fall related fractures compared to 13 participants in the Control group RR for fall-related fractures in the Intervention group relative to the Control group was 0.22 (0.06–0.80).	
Ohtake et al. 2013 (Japan) [45]	*N* = 196. 83.5% female. Age > 65 years (mean age = 83.6 years).	Both intervention and control group received a health education program on falls prevention. Intervention group also undertook a group based exercise program (strength, balance and flexibility).	Health education program on falls prevention (same program also delivered to intervention group)	8 week exercise program once weekly (20–30 min), together with 1–2 x weekly home exercise program	8.9% of the exercise group dropped out after baseline assessment. 97% participation rate in the group exercise sessions, participants also did on average 3.8 days/ week of home exercise.	x Small number of fallers in each group – Intervention group n = 7 (7.6%); and in the control group *n* = 9 (12.1%) ($p = 0.323$). Significant improvement in Functional Reach and falls efficacy in the intervention group relative to control group.	Bal ✓ Dose x
Kim et al., 2014 (Japan) Ø [46]	*N* = 105. 100% female. Age > 70 years (mean age = 77.8 years). Participants had one or more falls in the past year.	Group based strength and balance exercise program.	3 month health education sessions (60 mins each month).	3 month group program twice weekly × 60 min, then 4–12 months 1 x monthly group exercise program supplemented with home exercise ≥3 times weekly.		✓ At 12 months: Falls – I 19.6%, C 40.4%, (OR 2.78, 1.17–6.96); Repeated falls - I 20%, C 33.3%, (OR 1.85, 0.33–7.38); Injurious falls – I 80%, C 62%, (OR 0.82, 0.22–3.05). The exercise group significantly improved in one leg standing time, knee extension strength and ankle dorsiflexion strength over the 12 months, compared to the education and excluded group who showed no significant improvements.	Bal ✓ Dose ✓
Hirase et al., 2015 (Japan) Ø [47]	*N* = 93 older adults using community day centers. 69.9% female. Age > 65 years (mean age 82.1 years)	Group based programs: (1) Foam rubber balance exercises (2) Stable surface exercises. Both interventions were supplemented with a daily home exercise program (2–3 exercises).	Continued activities at the day centres, but did not perform balance or strengthening exercises.	4 months program, Once weekly 60 min exercise class supplemented with home exercise program (for both intervention groups).	7.5% of participants withdrew from the study. High adherence to the exercise programs: 95.5%, 93.3% of all possible classes in the foam rubber and stable surface groups.	x Mean number (SD) of falls for the foam rubber, stable surface, and control groups was 0.24 (0.51), 0.59 (1.94), and 0.90 (1.45) ($p = 0.07$). Significant improvement for both exercise groups compared to the control group for balance and sit to stand measures. Significant improvement for foam exercise group compared to firm surface exercise group for some balance tests at 2, 3 and 4 months.	Bal ✓ Dose ✓
	N = 68. 57.4% female.		Maintain usual activities	16 week program, 20–30 min/day, ≥4 times/week,	91% I group completed 16 week program	x	Bal ✓ Dose ✓

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria[2]
Ashari et al., 2016 (Malaysia) [52]	Age > 50 (mean age = 63.7 years) Inclusion criteria of impaired turning performance (Neurocom Force Platform)	Individualised home based exercise program, based on Otago Exercise Program. 6–8 balance and strengthening exercises (including 2 turning exercises).		and walking program ≥3 times / week.		Fallers in 16 week program – I 5.8%, C 8.8% (underpowered, no significance testing). I significantly improved relative to C on turning measures, Timed Up and Go (single and dual task) and static stance sway.	
Hwang et al., 2016 (Taiwan) Ø [49]	*N* = 456. 67% female. Age ≥ 60 years (mean age = 72.4 years). Participants had one or more Emergency Dept presentation due to a fall 6 or more months prior to study.	Compared two 24 week home-based exercise interventions: (1) Yang style Tai Chi (2) Lower Extremity Training (LET - balance, strength and flexibility exercises, individualised) Both groups had one supervised session/week and were encouraged to self practice daily at home. After 6 months (when supervised exercise ceased) participants asked to continue self practice daily.	No control group	(1) Tai Chi - 60 min supervised session (10-min warm-up followed by a review of previous movements, introduction of new movements, and 5 min of relaxation) (2) Lower Extremity Training - 60 min physio supervised session (10-min warm-up, 45 min of exercise, and a 5-min cool-down).	Results reported for 6 month intervention period, and subsequent 12 months. Adherence: 78% of Tai Chi participants and 72% of the LET group participated in ≥20 of the 24 (83%) sessions. During the 6-month intervention, 50% of the Tai Chi participants and 67% of the LET group independently practiced the exercise program ≥7 times per week.	✓ At 6 months-falls: F: IR (Tai Chi/LET) 0.30 (0.15–0.60) FR: RR (Tai Chi/LET) 0.76 (0.66–0.87). At 6 months - Injurious falls: F: IR (Tai Chi/LET) 0.33 (0.16–0.68) FR: RR (Tai Chi/LET) 0.86 (0.77–0.96). At 18 months: falls and injurious falls remained significantly reduced for the Tai Chi group compared to the LET group. For the Tai Chi group, handgrip strength, Tinetti balance and gait, depression, and cognition scores improved significantly during the 6-month intervention. For the Lower Extremity Training group, handgrip strength, Tinetti balance and gait, fear of falling, depression, and cognition scores improved significantly during the 6-month intervention.	Bal ✓ Dose ✓
Medication							
Sato et al., 2005 (Japan) [30]	*N* = 200. 100% female. Age > 70 years (mean age = 78 years). Ambulatory women recruited from an out-patient department with probable Alzheimer's disease. RCT.	45 mg menatetrenone (vitamin K2), 1000 IU ergocalciferol (vitamin D2), and 600 mg calcium.	Usual care.	Daily medication for 2 years.	Significant reduction in fractures.	✓ • FR – RR = 0.13 [0.04, 0.43]	
Psychological intervention							
Huang et al., 2011 (Taiwan) Ø [37]	*N* = 186. 59% female Age ≥ 60 years. RCT.	(1) Cognitive behavior therapy group, intervention based on previous program, [69] but added newly	Usual care.	(1) 60–90 min weekly × 8 weeks	Outcomes assessed at 2 and 5 months.	x • F – RaR = 1.00 [0.37, 2.72] • FR – RR = 1.00 [0.40, 2.51]	

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria[2]
		developed fear of falling management model developed by first author. (2) Cognitive behavior therapy group + Tai Chi		(2) CBT as above + Tai Chi 60 min 5 x/ week × 8 weeks.	NOTE: Tai Chi group total exercise dosage 40 h.		
Environment/assistive technology intervention							
Lin et al., 2007 (Taiwan, rural/agricultural area) [38]	See above.	(1) Home exercise with physio (individualized flexibility, strength & balance exercises) (see above for outcomes) (2) Home Safety Assessment and Modification by public health worker.	See above.		For the Home Assessment and Modification intervention, 14 inexpensive modifications (of a list of 28 options) were implemented within the first week of the intervention. Other recommended modifications were recommended to the family by the assessor (a public health worker). No data provided on adherence to home modifications provided, nor for uptake of the additional recommended home modifications.	(2) ✓(Home modification)[##] • F – RaR = 0.46 [0.22, 0.95] No significant differences for the Home Assessment and Modification group on the WHOQOL-BREF domains relative to the Education group.	
Kamei et al., 2015 (Japan) Ø [48]	*N* = 130. 84.6% female. Age > 65 years (mean age 75.8 years).	Both groups undertook the same 4 x weekly falls prevention multifactorial Program (120 min each) covering physical and mental assessment interview; (ii) blood pressure check; (iii) education regarding fall risk factors, food and nutrition, foot self-care; and (iv) exercise sessions focussed on strength, coordination and balance. Intervention group also received a home hazard checklist, a training program on home hazard awareness and modification	4 weekly multifactorial program as described for intervention group.	4 weeks intervention, 120 min/session.	16.4% of the intervention group did not attend sessions regularly and withdrew from the study. Intervention group significantly improved falls prevention awareness and home modifications.	x 10.9% reduction in all falls in intervention group compared to control group. Time to first fall: HR = 0.591 (0.305–1.147), $p = 0.116$. Indoor falls – reduced by 11.7% in Intervention group relative to control group HR = 0.397 (0.151–1.045), $p = 0.052$.	
Knowledge intervention							
Huang et al., 2010 (Taiwan) Ø [34]	See above.	Education intervention included separate sessions on medications, nutrition, safe home environment, and footwear. It included a component of each session for revision.				x • FR – RR = 1.62 [0.11, 24.16] The Education group achieved improved score post relative to pre intervention on indoor environment score, fear of falling, and Timed Up and Go. Significance of differences in change between groups with the intervention relative to control was not provided.	
Other							

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria[a]
Leung et al., 2014 (China) Ø [51] NOTE: Some vibration studies incorporate exercise during vibration – this intervention involved participants standing with knees straight – no exercise	*N* = 710. 100% female (post-menopausal). Aged >60 years (mean age 72.9 years). Cluster RCT. Participants recruited through community centres for older people.	Low magnitude high frequency vibration – standing upright without knee bending on a purpose built vibration platform that provided vertical synchronous vibration at 35 Hz, 0.3 g.	Habitual lifestyle, participated in normal interest group activities run by the community centres.	18 months, 5 x/week × 20 min standing on vibration platform.	29.7% of vibration group were lost to follow-up at 18 months (most of these declined to continue participation). No serious adverse events, though nine vibration participants and seven control participants complained of leg pain; and five vibration and one control participant complained of dizziness.	✓ Fall or fracture incidence: I – 18.6% C – 28.7%. Adjusted Incident Rate Ratio for falls or fractures: 0.54, 95%CI 0.37–0.78, $p = 0.001$. Significant improvement for vibration group on secondary measures including leg muscle strength and balance.	
Multiple interventions							
Assantachai et al., 2002 (Thailand) Ø [41]	*N* = 1043. 64% female. Age > 60 (mean age = 68). Cluster RCT (by community).	Received information leaflet describing risk factors for falls and strategies to reduce risk. Risk factors covered included nutritional advice (including calcium intake), activities of daily living, hypertension, special sense function and high risk medications. Also offered free access to geriatric clinic for any health problem patients wanted reviewed.	Usual care	No information provided regarding the proportion of the intervention group who took up the offer of free access to the geriatric clinic, what type of interventions were provided for those who accessed it, and their adherence.		✓ • FR – RR = 0.77 [0.63, 0.94]	
Huang et al., 2010 (Taiwan) Ø [34]	See above (Study had three intervention groups and control – Education only, Tai Chi only and combined Tai Chi and Education).	Combined program incorporated 5 x education sessions over 5 months and a Tai Chi (13 forms) exercise program – 40 min/session, 3 x/week for 5 months.			High drop-out rates over 5 month intervention period - Education + Tai Chi group (34%). NOTE: Tai Chi group total exercise dosage 40 h.	x • FR – RR = 1.68 [0.16, 17.67] Cochrane review used raw data at 5 months only, as 18 month raw data not provided. Cochrane review reports all interventions as non-significant (although combined education and Tai Chi group reported as reducing falls at 5 and 18 months in the paper).	
Huang et al., 2011 (Taiwan) Ø [37]	See above (Cognitive Behavioral Therapy + Tai Chi) (also compared to Cognitive Behavior Therapy alone)	Cognitive Behavioral Therapy program based on previous program, [69] but added newly developed fear of falling management model developed by first author. This arm of the intervention combined the cognitive behavior therapy with a Tai Chi exercise program.	Usual care.	Cognitive Behavior Therapy (CBT) and Tai Chi combined program incorporated 60–90 min weekly × 8 weeks for the CBT and 60 min 5 x / week × 8 weeks for the Tai Chi component.	Outcomes assessed at 2 and 5 months. NOTE: Tai Chi group total exercise dosage 40 h.	x • F – RaR = 0.38 [0.10, 1.47] • FR – RR = 0.40 [0.11, 1.45] The combined Cognitive Behavior Therapy + Tai Chi group achieved significantly improved falls efficacy, improved mobility, higher social support satisfaction, and quality of life than the control or cognitive behavior therapy alone groups.	
Lee et al., 2013 (Taiwan) Ø [50]	*N* = 616. 55% female. Age > 65 years (mean age 75.7 years).	All intervention participants received: 1. Strength, balance, cardiovascular and flexibility group exercise	Health education brochures, medication reviews and medical referrals without direct exercise interventions	3 month multifactorial intervention	Attrition rate for 3 month intervention period: I 10.9%; C 13.5%.	x For 12 month followup period:	

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria*
	Participants had high falls risk, with any of: (1) recurrent falls in the previous year; (2) medical history associated with high falls risk (ie, stroke, Parkinson's disease, head injury, fractures due to falls); and (3) fell only once in the previous year, and had gait or balance problems (poor TUG score)	program, supplemented with home exercise program. 2. Health education sessions and brochures on falls prevention. 3. Home hazards evaluation/ modification. 4. Medication review 5. Ophthalmology/ other specialty consults.	or structured health education sessions.			F: I 128 (0.41 falls/ person year), C 132 (0.44 falls/person year) ($p = 0.692$) FR: HR 0.90 (0.66–1.23). At end of 3 month intervention: Significant improvement in intervention group relative to control group for Physiological Profile Assessment, reaction time, balance, mobility and depression (although control group improved significantly more than intervention group on knee strength and proprioception).	
Ng et al., 2015 (Singapore) [53]	$N = 246$. 61% female. Age > 65 years (mean age 70). Factorial design RCT, with nutrition, exercise and cognitive training groups, and combined intervention group. Recruited frail and pre-frail older people based Cardiovascular Health Study (frailty phenotype) criteria.	Nutrition group: nutritional supplementation with commercial formula, iron and folate, Vit B6 and B12, calcium and vitamin D supplements daily. Exercise: Strength and balance group program (weeks 1–12) then home program weeks 13–24. Cognitive training: Cognitive enhancing activities including verbal recall, mazes, problem solving etc. Combination group: all of the above interventions	Standard care + placebo supplement liquid + placebo capsule and tablets (identical appearance to intervention nutrition supplements), and instructions not to replace their meals with the supplements.	6 month intervention period. Nutrition: supplements daily. Exercise: Weeks 1–12, twice weekly group sessions; weeks 13–24 home exercise 2 h weekly. Cognitive training: Weeks 1–12 - 2 h/week; weeks 13–24 – fortnightly 2 h booster sessions.	Low dropout rate (8% for nutritional supplement; 10% cognitive training; 4% for exercise; 6% for combination, and 8% for control).	x Small numbers of fallers/ group, analysis only provided across all groups (12 months, $p = 0.67$). Fallers/group (12 months): Nutrition: 4 fallers (8.6%) Exercise: 3 fallers (6.3%) Cognitive training: 2 fallers (4.1%) Combination: 2 fallers (4.1%) Control: 5 fallers (10.4%)	
Multifactorial interventions							
Jitapunkul et al., 1998 (Thailand) [40]	$N = 160$. 65% female. Age ≥ 70 years (mean age = 76 years). Participants recruited from RCT.	Home visit by non-professional with a structured health questionnaire. Referral to a nurse or geriatrician if function declined or ≥1 fall in 3 months, with subsequent nurse or geriatrician home visit to assess, educate, prescribe, or make other referrals.	Usual care. Assessment at end of 3 year period.	Home visit at study commencement, then three monthly visits × 3 years.	Intervention group had significantly less rate of functional decline (Chula ADL Index and Barthel ADL Index) over the study period.	x • FR – RR = 0.52 [0.14, 1.94]	
Huang and Acton, 2004 (Taiwan) Ø [35]	$N = 120$. 46% female. Age ≥ 65 years (mean age = 72 years). RCT.	Falls prevention information brochure + individualized falls prevention information (medication and home safety focus) – brochure and verbal.	Falls prevention information brochure.	Three home visits by nurse in 4 months (a) for initial assessment), (b) to work through individualized risk factors (medication and home safety), and (c) re-assessment at 4 months.		x • FR – RR = 0.12 [0.01, 1.76] Improved knowledge of medications and reduced home hazards in intervention group. Only 2 months follow-up after intervention provided.	
Huang et al., 2005 (Taiwan) Ø [36]	$N = 141$. 69% female. Age > 65 years (mean age = 77 years).	Enhanced discharge planning by experienced gerontological nurse, including visits on wards, home visit, and phone contacts post discharge. Included discharge	Usual discharge planning (no brochures, no written discharge summaries, no home visit, no telephone contact).	Visits on wards at least every 2 days, home visit within 3–7 days of discharge, and once weekly phone calls post discharge.	Positive outcomes for the intervention group included significantly reduced hospital length of stay.	x • FR – RR = 0.67 [0.22, 2.01]	

Table 2 Comparison of main characteristics of interventions shown to be effective and not effective in reducing falls (F, rate of falls) or falls risk (FR, proportion of fallers) in studies conducted in Asia *(Continued)*

	Sample and design	Details of intervention	Control activity	Duration and frequency	Other comments	Effective (✓) or Ineffective (x) falls outcomes	Meets Sherrington's criteria*
	Hip fracture patients recruited and randomized at hospital discharge. RCT. Patients interviewed at discharge, 2 weeks post-discharge and 3 months post-discharge.	and falls prevention brochure. Involved patient, family and other health care professionals. Included nurse follow-up with physicians.					
Shyu et al., 2010 (Taiwan) Ø [39]	*N* = 162. 69% female. Age ≥ 60 years (mean age = 78 years). Patients admitted to hospital for hip arthroplasty or internal fixation. RCT. Excluded patients with cognitive impairment, marked functional impairment pre-operatively, or those who were terminally ill.	Three components to intervention: (1) geriatrician review; (2) rehabilitation; and (3) discharge planning service.	Usual care, described as limited interdisciplinary involvement, usually no home visit, and no in-home physiotherapy.	(1) Geriatrician/geriatric nurse review and recommendations pre-operatively, and geriatric nurse review and recommendations post-operatively. (2) Focus on early post-operative rehabilitation, and in-home rehabilitation. (3) Discharge planning was coordinated by a geriatric nurse, and included a pre-discharge home visit and recommendations, and follow-up phone calls. On average, there was 1 x geriatrician visit, 5.4 geriatric nurse visits, 3.1 physical therapist visits, and 1 rehabilitation physician visit during hospitalization; and 9.9 geriatric nurse and 3.0 physical therapist home visits after return home.	Intervention group had significant improvement relative to control group on Activities of Daily Living, walking ability, reduced depression and better SF36 scores (two year follow-up).	✓ • FR – RR = 0.56 [0.34, 0.93]	

NB: for multiple intervention studies where results have been reported against a control group for individual interventions, these have been included in the single intervention component of the table as well
RaR = Rate Ratio; RR = Risk Ratio
✓=yes; x= no
* Criteria based on Sherrington's review and meta-analysis [54] (for exercise studies only: (1) moderate to high challenge to balance; and (2) at least 50 h total dosage
#Reported as non-significant falls outcome in Cochrane review, although paper reported significant reduction in incidence of falls
##Reported as non significant falls outcome in the published paper, however reported as significant reduction in falls outcome/s in the Cochrane review
Ø Indicates that study was conducted with a primary focus on prevention of falls (identified in aim, hypothesis, or as primary outcome / used for power calculation)
POMA = Problem Oriented Mobility Assessment; RCT = Randomized Controlled Trial; Bal = Balance; IU=International Units; CBT = Cognitive Behavioral Therapy; ADL = Activities of Daily Living

two of the studies reported significant improvements in the secondary balance, strength, or mobility related measures in the exercise intervention relative to the control group [34, 42].

Medication interventions

Medication interventions can include approaches to review or reduce medications overall, or specifically target high risk medications such as psychotropic medications; or provide supplementation to improve fall related outcomes [3]. Only one Asian study investigated a medication related intervention, by evaluating the effect of two years supplementation with vitamin K2, vitamin D2 and calcium, compared to usual care, in older women with probable Alzheimer's disease [30]. There was no effect on falls outcomes, but a significant reduction in fractures in the medication intervention group.

Psychological interventions

One study investigated the effectiveness of Cognitive Behavioral Therapy (CBT) alone, and in combination with Tai Chi, relative to a control group receiving usual care [37]. Neither CBT alone or combined with Tai Chi reduced falls outcomes. CBT alone did not result in significant improvements in any secondary measures, including falls efficacy, relative to the control group (see multiple interventions for outcomes for the combined Tai Chi and CBT intervention group).

Environment/assistive technology interventions

Two studies investigated home assessment and modification interventions. In a Japanese study, Kamei et al. evaluated the effect of a home hazard checklist and a training program on home hazard awareness for older people, superimposed on a multifactorial assessment and intervention program received by both the intervention and control group (both groups received a falls risk factor education program, exercise, blood pressure review, and physical and cognitive assessments) [48]. Although the intervention group significantly increased their falls prevention awareness and home modifications implemented, there was no significant reduction in falls. Another randomized trial in Taiwan compared a home assessment and modification intervention against an education intervention, and an exercise intervention [38]. The home modification program was conducted by a public health worker, and involved a standard assessment, provision of 14 standard, inexpensive modifications (eg removal of loose mats, marking of step edges, rectification of poor lighting), and recommendations regarding another 14 modifications if required. Although the home safety assessment and modification intervention was reported by Lin et al. [38] as not achieving a significant reduction in rate of falls, the reduction was significant in the Cochrane review analysis [3]. There was no information provided about the range of additional recommendations made, nor the level of adherence with the implemented or recommended home modifications.

Knowledge interventions

One randomized trial from Taiwan investigated the effect of knowledge based interventions provided to older people on reducing risk of falls [34]. The education intervention involved five one hour group education sessions over five months, targeting separate important risk factors at each session [34]. While the intervention achieved improved knowledge about falls risk, there was no reduction in falls outcomes.

Other single interventions

There were a number of additional interventions that have been shown to be effective in reducing falls, falls risk or falls injuries in the 2012 falls prevention in the community setting Cochrane review, which have not been investigated in any randomized controlled trials in Asian countries [3]. These include (a) medication prescription review; and high-risk medication withdrawal (eg psychotropic medications); (b) cataract surgery (first eye); (c) changing from bifocal or multifocal glasses to distance glasses for outdoors mobility; (d) cardiac pacing surgery (for carotid artery hypersensitivity); and (e) footwear (anti-slip shoe device for icy conditions). Vitamin D supplementation, which was considered effective in reducing falls in at-risk populations by the Cochrane review, has also not been evaluated for its effect on falls in Asian populations as a single vitamin supplement (Sato and colleagues evaluated supplementation of vitamin K2, vitamin D2 and calcium as a single intervention) [30]. Some of these interventions may be inappropriate in many parts of Asia (eg anti-slip shoe device for icy conditions). However, the other intervention types are likely to have direct or perhaps modified applicability for older people in Asia.

One recent trial in Asia has been classified as "Other" under the single intervention studies – an 18 month investigation of low-magnitude high-frequency vibration program in post-menopausal women in China/Hong Kong ($n = 710$) [51]. Vibration interventions often utilize exercises while performing vibration, [55] and so may be classified under exercise interventions, however the study by Leung and colleagues had participants standing with straight knees on the vibrating platform for the duration of the vibration procedure (20 min/session). The group receiving the vibration therapy had significantly reduced Hazard Ratio for the

combined outcome of "falls or fractures" ([0.56, 0.40–0.78, $p = 0.001$).

Multiple interventions

Five randomized controlled trials evaluated the effect of a multiple intervention approach to reducing falls [34, 37, 41, 50, 53]. Only one of the trials was effective, providing education (brochure targeting a number of important falls risk factors) and free access to a geriatric clinic as required for the intervention group [41]. Although effective, no details were provided about the uptake, type of interventions, or adherence to recommended interventions for the intervention group. The ineffective interventions included a combined Tai Chi exercise program and education program (each of these two components were also evaluated in isolation); [34] a combined Tai Chi and Cognitive Behavioral Therapy program; [37] a strength, balance, and fitness exercise program combined with health education, home assessment and modification, medication review, and ophthalmology or other specialty consultations; [50] and a factorial design study with participants receiving one or more of nutrition, exercise, and cognitive training interventions [53]. Tai Chi was shown to be an effective exercise intervention to reduce falls when used as a single intervention in three Asian studies, [31, 42, 49] and in non-Asian countries [3]. In both of the ineffective multiple intervention studies incorporating Tai Chi, the Tai Chi component did not incorporate the 50 h of exercise recommended to improve likelihood of achieving a significant reduction in falls [54].

Multifactorial interventions

Four of the randomized controlled trials conducted in Asia utilized a multifactorial falls prevention intervention [35, 36, 39, 40]. Two of these targeted a high-risk population (patients with hip fracture returning home after surgery), and utilized improved discharge planning and post discharge follow-up, including home visit/s [36, 39]. The study by Shyu and colleagues also incorporated a strong interdisciplinary care model pre and post-surgery [39]. The trial conducted by Jitapunkul and colleagues utilized a regular (three monthly) home visit health screening process by a non-health professional, with subsequent referral to a nurse or geriatrician if there was recent functional decline or ≥1 fall in the preceding three months. Similarly, the study by Huang and Acton utilized a targeted falls risk brochure, followed up by targeted risk factor management support by a visiting nurse (focussing on medication and home safety) [35]. Only the trial by Shyu and colleagues was effective in reducing falls in the intervention group [39].

Quality of the studies

Risk of bias was reported for all studies published in the 2012 Cochrane review. The Cochrane Collaboration's tool for assessing the risk of bias was used to evaluate quality of the RCTs published since Gillespie et al.'s Cochrane review. Most of the recent studies had low to medium risk of bias (Table 3). Two studies had low risk of bias across all domains, [49, 52] with all the other studies having at least one section that was unclear (authors did not provide enough evidence). The only high risk of bias was for sequence generation and blinding in the study by Ohtake et al. [45] because participants were categorized by the day of the week and they did not blind the participants, personnel or outcome assessors. The risk of bias for the studies was viewed as low to medium due to sections of data in a number of the categories not being available, and are known to be essential for conducting high quality RCTs (blinding, allocation concealment).

Meta-analyses

For intervention types with two or more studies from the Asian region with comparable data available, a meta-analysis was conducted. Only the exercise ($n = 15$) intervention of the single interventions, and the multifactorial interventions ($n = 4$) were included in the meta-analyses (Fig. 2). The Ashari et al. study [52] was not included in the exercise meta-analysis due to significant differences between the groups at baseline. Woo and colleagues [42] had two intervention groups and therefore dichotomous (outcome and sample) data for the intervention groups only were combined, as described in the methods. Results from the meta-analysis indicate that exercise achieved significant reduction in number of fallers (OR: 0.43 [0.34–0.53]), number of falls (OR: 0.35 [0.21–0.57]) and number of fallers injured (OR: 0.50 [0.35–0.71]). Heterogeneity was at appropriate levels for the exercise intervention meta-analyses.

A separate subgroup meta-analysis was conducted for the Tai Chi exercise interventions. Results indicated Tai Chi achieved significant reduction in the number of falls (OR: 0.24 [0.13–0.47]) and number of fallers (OR: 0.46 [0.30–0.70]), although it must be noted that there was high heterogeneity for the number of fallers ($I^2 = 67\%$) and therefore these results should be used with caution. Separate subgroup analysis was not able to be undertaken for the two home exercise program studies due to the Ashari and colleagues study [52] having differences at baseline discussed previously. No other sub-group analyses were possible.

The meta-analysis of the multifactorial interventions did not reach significance for number of fallers (OR: 0.57 [0.23–1.44]). Although there were two home assessment and medication trials, they reported different falls

Table 3 Quality of the studies

Study	Sequence generation	Allocation concealment	Blinding	Incomplete outcome data	Selective outcome reporting	Free of other bias
Ashari et al. 2016	+	+	+	+	+	+
Hirase et al. 2015	?	+	?	+	+	?
Hwang et al. 2016	+	+	+	+	+	+
Kamai et al. 2015	?	?	?	+	+	?
Kim et al. 2014	+	?	+	+	+	?
Lee et al. 2013	+	+	?	+	+	?
Leung et al. 2014	+	+	+	+	+	?
Ng et al. 2015	+	+	+	+	+	?
Ohtake et al. 2013	–	?	–	?	+	?
Yamada et al. 2012	+	+	?	?	+	?
Yamada et al. 2013	?	?	+	?	+	?

Note. Bias was scored as low risk (+), or high risk (–) or unclear (?). [26]

outcome data, so meta-analysis was not able to be performed. Similarly, for the falls risk outcome for the multifactorial intervention classification, there were not two studies reporting this outcome to allow meta-analysis. Similar to the Cochrane review, meta-analysis of the multiple interventions was not possible because of the diversity of the intervention types combined as multiple interventions.

Discussion

Falls among older people remain a major public health problem world-wide. Substantial inroads are being made, with growing research evidence of single, multiple and multifactorial interventions being effective in reducing falls. However, the results of this focussed review of the subgroup of falls prevention randomized trials for the community setting conducted in countries in Asia,

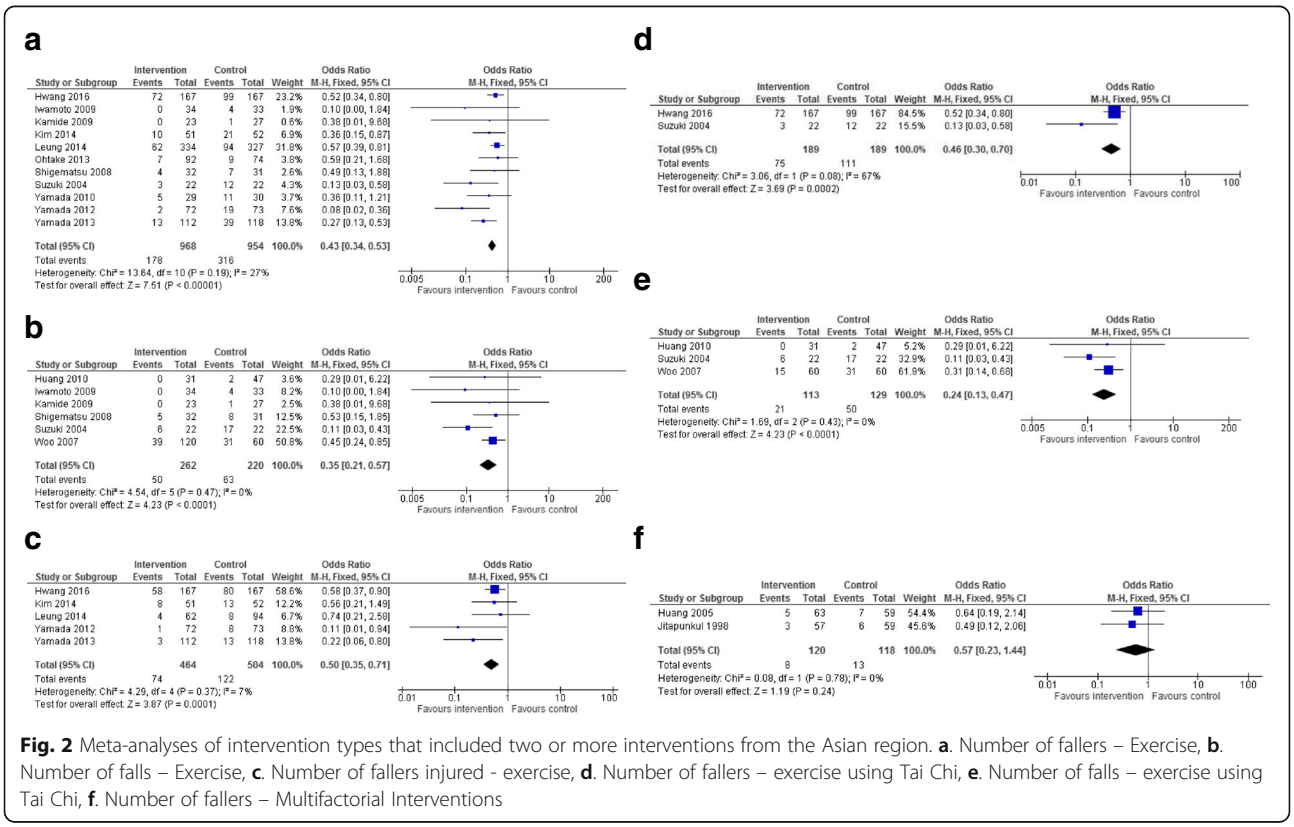


Fig. 2 Meta-analyses of intervention types that included two or more interventions from the Asian region. **a**. Number of fallers – Exercise, **b**. Number of falls – Exercise, **c**. Number of fallers injured - exercise, **d**. Number of fallers – exercise using Tai Chi, **e**. Number of falls – exercise using Tai Chi, **f**. Number of fallers – Multifactorial Interventions

where 60% of the world's population live, indicates substantial gaps. Not only has there been limited falls prevention research conducted in Asian countries, but where studies have been conducted, sample sizes are generally small, and only 11 of the 30 interventions evaluated in the 26 studies in Asia achieved a significant reduction in one or more falls outcomes (37%). The meta-analysis results where more than one study was able to be pooled for an intervention type (exercise and multifactorial interventions) indicated that only exercise was effective in reducing falls in Asian populations.

Exercise is by far the most researched intervention type in Asian countries (as it is world-wide), with over half of the studies reported having an exercise component. Not surprisingly, Tai Chi was commonly investigated, and was effective in reducing falls outcomes in three studies in Asian countries, and in the sub-group meta-analysis. Although Tai Chi has been shown to be acceptable [56] and to reduce falls-related outcomes in non-Asian countries, [3] it does seem to be a preferred exercise approach for investigation and implementation in Asian countries. Adherence to the successful Tai Chi programs conducted in Asia was high (75–81%), though similarly high adherence to Tai Chi has been reported in studies conducted in other countries (68–80%) [57, 58]. Of note though, nine of the 15 studies investigating exercise in Asian countries were not effective, and the majority of these did not meet the criteria recommended by Sherrington [54] to be likely to be effective in reducing falls (having a moderate challenge to balance, and minimum of 50 h exercise dosage). Three studies that did meet these criteria were not effective, but were substantially underpowered to identify a reduction in falls related outcomes [29, 47, 52]. Other exercise approaches that were effective in single studies were an obstacle course, [43] a multi-target stepping program, [44] and a group balance and strength training program [46].

In an updated meta-analysis by Sherrington published around the time of this paper's publication, more stringent dosage criteria (greater than three hours per week) were identified as achieving the greatest benefit in reducing falls (23% reduction) [59]. Higher exercise dosages, and sustained exercise (lifelong behaviour change) are clearly more desirable. However, low levels of sustained participation have been reported in falls prevention exercise programs, [60] suggesting that strategies such as starting off with lower dosages and gradually building up, and embedding behaviour change elements into exercise programs to support sustained and increased participation may be required to achieve this high dosage. Future exercise studies should aim to adopt a method that meets the updated Sherrington criteria while being culturally relevant, as well as being adequately powered for falls outcomes.

These results have substantial implications for falls prevention research and practice in Asia. It is often assumed that interventions shown to be effective in randomized trials or meta-analyses in one country will be generalizable elsewhere. However, researchers in the falls prevention area, [23, 61] and in other areas of health (eg hepatocellular carcinoma) [62] have called for local research in Asian and or developing countries to address key gaps and differences. In this context, recent research has also highlighted the diverse range of factors that may complicate or reduce likely effectiveness of directly translating falls prevention approaches found to be effective in non-Asian countries into Asian countries [18, 19, 23]. Some of these differing factors include: (1) role of family (including filial piety); (2) indoor (including floor surfaces) and outdoor environments; (3) regularly worn footwear in many parts of Asia differ from what is considered the ideal footwear for falls prevention; (4) lifestyle factors such as incidental and formal exercise approaches that are routine and acceptable, diet, and sunlight exposure; (5) health services and systems, and patient expectations of specific health practitioners; (6) differing understanding of prevention and active engagement in prevention and intervention approaches; and (7) differences in concern about falls influencing behaviours [23, 61, 63, 64].

Given the factors outlined above, direct translation of interventions from non-Asian studies to Asian countries may warrant careful consideration. There are a number of implications for researchers, practitioners, policy and planning personnel, and research funders who may be involved in future falls prevention research in Asia. Firstly, in the area where the meta-analysis of Asian studies indicated effective interventions (Tai Chi and other exercise approaches), and where there is at least one effective randomized trial (home modifications, multiple and multifactorial interventions), that these interventions could be considered for broader implementation into practice. Two other interventions were shown to be effective in special populations: vitamin K2, vitamin D2 and calcium supplementation for women with probable Alzheimer's disease; and a multifactorial intervention for post hip fracture surgery patients. Even in these areas where there is some evidence of effectiveness, there is a need for research with larger samples (as the majority in this review were small samples, which limits the rigor and confidence in study findings). Furthermore, in areas like exercise, the different exercise intervention types (other than Tai Chi) have been grouped together in this review (because of the small number of studies), so there remains scope for further research exploring other exercise modalities, particularly those that may be most acceptable to Asian populations.

Secondly, for other areas of practice where there is research evidence of effective interventions in non-Asian

populations but not in Asian populations, a number of options are available. For those interventions where research has been conducted in Asian countries but was not shown to be effective (eg knowledge/education interventions, and psychological interventions) there is a clear need to review these unsuccessful methodologies in the context of local factors that may influence their uptake and effectiveness. For example, critical elements to a successful falls prevention education program include that participants understand that falls are preventable, and that changes in behaviour, even at later ages, can still improve risk of future falls. However, research indicates that in some Asian cultures (eg China), fatalistic beliefs about falls is a major barrier that would need to be overcome [22]. Researchers and practitioners need to undertake research to improve understanding of these beliefs, and strategies that may influence these beliefs. Recent research in Australia has shown a World Café approach to be valuable in informing understanding of factors that would influence uptake and sustained engagement in falls prevention among older people, [65] and these factors have been introduced into a peer education program that was effective in increasing intention to engage in falls prevention activities [66]. Innovative, culturally relevant approaches to understanding these factors in Asian countries, and approaches to achieve sustained behaviour change (which may include these or other culturally relevant approaches) are foundational to achieving improved falls related outcomes. In using results from these type of local studies, practitioners can implement these interventions in a blended manner that retains as much of the original intervention approach as possible, but with local tailoring. For other types of falls prevention interventions that have not been investigated in Asian samples (eg medication review/reduction) there would be merit in establishing local factors as described above, prior to researchers in Asia seeking funding for local research to evaluate a culturally tailored intervention's effectiveness in reducing falls or falls injuries. In the context of medication review, inclusion of traditional and herbal medicines (which are widely used in some Asian countries), as well as interactions with pharmacy medicines would be important considerations. There may be value for this research to be collaborative with researchers who have implemented effective randomized controlled trials in non-Asian countries.

Finally, there may be other novel, locally relevant intervention types that have not been researched, that may warrant funding being sought for these interventions to be evaluated in an Asian context.

For the purposes of this research we have grouped countries under the broad umbrella classification of "Asia", however it is important to recognize that there is considerable diversity between some of the countries in Asia, and even within some countries (for example, in Malaysia where there are three significant ethnic populations), or between urban and rural populations, where socioeconomic and other differences in some Asian countries can be stark. Specific understanding of local, cultural and societal factors that may influence acceptability of interventions need to be considered when implementing falls prevention interventions in Asian countries.

Although we have identified only a relatively small number of randomized controlled trials investigating falls prevention approaches in Asian countries, there appears to be a steady growth in the number, size and quality of the studies published more recently (since the 2012 Cochrane review) [3]. There are also two protocol papers published for studies that are underway in Asia, including a multifactorial intervention for older people with recent history of falls or injuries from falls in Malaysia ($n = 300$), [67] and an evaluation of the use of exercise using Nintendo® Wii in Singapore ($n = 80$), [68] that will add to the small but growing volume of falls prevention research in this region. Of note, a number of the more recent studies are substantially larger than the median sample size of 150 for studies from Asian countries reported in the Cochrane review, which strengthens confidence in the study findings. However, the majority focussed specifically on exercise interventions only, and all were from more well developed countries in Asia.

There were several limitations to this review. A moderate limitation given the focus of this review on studies conducted in Asian countries is that only studies published in English have been included, therefore some Asian studies published in languages other than English may not have been identified. Another limitation is the varying falls outcomes published, which did not allow for additional meta-analyses to be undertaken. It would be beneficial for future falls prevention studies to include standardized outcomes to make direct comparisons and meta-analyses possible.

Conclusions

In summary, this focussed review of community setting falls prevention randomized controlled trials conducted in Asia found limited evidence of a small number of effective intervention types, relative to the strong evidence across a range of intervention types from non-Asian countries. Exercise approaches have had the strongest level of investigation, and several exercise approaches, in particular Tai Chi, have been shown to be effective. There is a need for substantial investment in large, adequately powered randomized controlled trials evaluating falls prevention interventions across Asia, in particular that incorporate tailoring of intervention approaches to the local Asian context, in order to reduce the projected escalating impact of falls in this rapidly aging part of the world.

Additional file

Additional file 1: PRISMA checklist. (DOC 64 kb)

Abbreviations

CBT: Cognitive Behavioral Therapy; OR: Odds Ratio; ProFaNE: Prevention of Falls Network Europe; RCT: randomised controlled trial

Acknowledgements

The research team acknowledge the contribution of Eileen Boyle (EJB) for assistance with reviewing some of the post 2012 falls preventions studies in Asia.

Funding

There was no funding associated with this publication.

Availability of data and materials

The main study data is the data extraction materials and quality ratings of included papers, most of which are included in the manuscript tables. Any other supporting data relating to this review is available from the authors.

Authors' contributions

KH, PS, and EB were involved in conception and design of the study; KH, EB, and KF were involved in acquisition and analysis of data; KH, PS, SL, WT, AA, AH, KF, EB were involved in interpretation of data. KH, EB and KF drafted the manuscript, and KH, PS, SL, WT, AA, AH, KF, EB were involved in revising the manuscript critically for important intellectual content. All authors (KH, PS, SL, WT, AA, AH, KF, EB) have provided final approval of the version of the manuscript submitted for publication, and all authors (KH, PS, SL, WT, AA, AH, KF, EB) agree to be accountable for all aspects of the work.

Ethics approval and consent to participate

Not applicable - This publication is a systematic review of already published papers. No ethics approvals or consent to participate were required for this work.

Consent for publication

Not applicable - No details, images, or videos relating to an individual person have been included in this manuscript, therefore no consent to publish is required.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]School of Physiotherapy and Exercise Science, Curtin University, GPO Box U1987, Perth, WA 6845, Australia. [2]Physical Therapy Department, Faculty of Allied Health Sciences, Thammasat University, Bangkok, Pathumthani, Thailand. [3]Department of Physical Therapy, National Cheng Kung University, Tainan, Taiwan. [4]Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong. [5]Malaysian Research Institute on Ageing, Universiti Putra Malaysia, Serdang, Selangor Darul Ehsan, Malaysia.

Received: 14 August 2017 Accepted: 12 December 2017
Published online: 05 January 2018

References

1. Global Burden of Disease Study 2013 Collaborators: Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the global burden of disease study 2013. *Lancet*.2015, 386(9995): 743–800.
2. World Health Organisation: WHO Global report on falls prevention in older age. In.: World Health Organisation, ISBN 978 92 4 156353 6; 2007.
3. Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, Lamb SE. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev. 2012;9:CD007146.
4. Pointer S: Trends in hospitalised injury, Australia: 1999–00 to 2012–13. In. Canberra: Australian Institute of Health and Welfare (AIHW) 2015.
5. Hill AD, Pinto R, Nathens AB, Fowler RA. Age-related trends in severe injury hospitalization in Canada. J Trauma Acute Care Surg. 2014;77(4):608–13.
6. Orces CH, Alamgir H. Trends in fall-related injuries among older adults treated in emergency departments in the USA. Inj Prev. 2014;20(6):421–3.
7. Ezeh AC, Bongaarts J, Mberu B. Global population trends and policy options. Lancet. 2012;380(9837):142–8.
8. Population Reference Bureau: World population data sheet 2014. In. Washington DC, USA - website: http://www.prb.org/publications/datasheets/2014/2014-world-population-data-sheet/data-sheet.aspx; 2016.
9. Tey NP, Siraj SB, Kamaruzzaman SB, Chin AV, Tan MP, Sinnappan GS, Muller AM. Aging in multi-ethnic Malaysia. Gerontologist. 2015;
10. Chongsuvivatwong V, Phua KH, Yap MT, Pocock NS, Hashim JH, Chhem R, Wilopo SA, Lopez AD. Health and health-care systems in southeast Asia: diversity and transitions. Lancet. 2011;377(9763):429–37.
11. United Nations: World population prospects: the 2015 revision, Key Findings and Advance Tables.. In.: United Nations, Department of Economic and Social Affairs, Population Division 2015.
12. Dolinis J, Harrison J, Andrews G. Factors associated with falling in older Adelaide residents. Aust N Z J Public Health. 1997;21:462–8.
13. Campbell J, Borrie M, Spears G. Risk factors for falls in a community-based prospective study of people 70 years and older. J Gerontol. 1989;44:M112–7.
14. Koski K, Luukinen H, Laippala P, Kivela SL. Risk factors for major injurious falls among the home-dwelling elderly by functional abilities. A prospective population-based study. Gerontology. 1998;44(4):232–8.
15. Kwan MM, Close JC, Wong AK, Lord SR. Falls incidence, risk factors, and consequences in Chinese older people: a systematic review. J Am Geriatr Soc. 2011;59(3):536–43.
16. Chu LW, Chi I, Chiu AY. Incidence and predictors of falls in the chinese elderly. Ann Acad Med Singap. 2005;34(1):60–72.
17. Shi J, Zhou BY, Tao YK, Yu PL, Zhang CF, Qin ZH, Sun ZQ. Incidence and associated factors for single and recurrent falls among the elderly in an urban community of Beijing. Biomed Environ Sci. 2014;27(12):939–49.
18. Han BH, Ferris R, Blaum C. Exploring ethnic and racial differences in falls among older adults. J Community Health. 2014;39(6):1241–7.
19. Jagnoor J, Keay L, Ivers R. A slip and a trip? Falls in older people in Asia. Injury. 2013;44(6):701–2.
20. Karuncharernpanit S, Hendricks J, Toye C. Perceptions of exercise for older people living with dementia in Bangkok, Thailand: an exploratory qualitative study. Int J Older People Nursing. 2016;11(3):166–75.
21. Jagnoor J, Keay L, Jaswal N, Kaur M, Ivers R. A qualitative study on the perceptions of preventing falls as a health priority among older people in northern India. Inj Prev. 2014;20(1):29–34.
22. Horton K, Dickinson A. The role of culture and diversity in the prevention of falls among older Chinese people. Can J Aging. 2011;30(1):57–66.
23. Romli MH, Tan MP, Mackenzie L, Lovarini M, Suttanon P, Clemson L. Falls amongst older people in Southeast Asia: a scoping review. Public Health. 2017;145:96–112.
24. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.
25. Lamb SE, Becker C, Gillespie LD, Smith JL, Finnegan S, Potter R, Pfeiffer K. Reporting of complex interventions in clinical trials: development of a taxonomy to classify and describe fall-prevention interventions. Trials. 2011;12:125.
26. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
27. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions5. 1.0; 2011 In; 2011. Available from: http://training.cochrane.org/handbook.
28. Iwamoto J, Suzuki H, Tanaka K, Kumakubo T, Hirabayashi H, Miyazaki Y, Sato Y, Takeda T, Matsumoto H. Preventative effect of exercise against falls in the elderly: a randomized controlled trial. Osteoporos Int. 2009;20(7):1233–40.
29. Kamide N, Shiba Y, Shibata H. Effects on balance, falls, and bone mineral density of a home-based exercise program without home visits in community-dwelling elderly women: a randomized controlled trial. J Physiol Anthropol. 2009;28(3):115–22.
30. Sato Y, Kanoko T, Satoh K, Iwamoto J. Menatetrenone and vitamin D2 with calcium supplements prevent nonvertebral fracture in elderly women with Alzheimer's disease. Bone. 2005;36(1):61–8.

31. Suzuki T, Kim H, Yoshida H, Ishizaki T. Randomized controlled trial of exercise intervention for the prevention of falls in community-dwelling elderly Japanese women. J Bone Miner Metab. 2004;22(6):602–11.
32. Shigematsu R, Okura T, Nakagaichi M, Tanaka K, Sakai T, Kitazumi S, Rantanen T. Square-stepping exercise and fall risk factors in older adults: a single-blind, randomized controlled trial. J Gerontol A Biol Sci Med Sci. 2008;63(1):76–82.
33. Yamada M, Tanaka B, Nagai K, Aoyama T, Ichihashi N. Trail-walking exercise and fall risk factors in community-dwelling older adults: preliminary results of a randomized controlled trial. J Am Geriatr Soc. 2010;58(10):1946–51.
34. Huang HC, Liu CY, Huang YT, Kernohan WG. Community-based interventions to reduce falls among older adults in Taiwan - long time follow-up randomised controlled study. J Clin Nurs. 2010;19(7–8):959–68.
35. Huang TT, Acton GJ. Effectiveness of home visit falls prevention strategy for Taiwanese community-dwelling elders: randomized trial. Public Health Nurs. 2004;21(3):247–56.
36. Huang TT, Liang SH. A randomized clinical trial of the effectiveness of a discharge planning intervention in hospitalized elders with hip fracture due to falling. J Clin Nurs. 2005;14(10):1193–201.
37. Huang TT, Yang LH, Liu CY. Reducing the fear of falling among community-dwelling elderly adults through cognitive-behavioural strategies and intense tai chi exercise: a randomized controlled trial. J Adv Nurs. 2011;67(5):961–71.
38. Lin MR, Wolf SL, Hwang HF, Gong SY, Chen CY. A randomized, controlled trial of fall prevention programs and quality of life in older fallers. J Am Geriatr Soc. 2007;55(4):499–506.
39. Shyu YI, Liang J, CC W, JY S, Cheng HS, Chou SW, Chen MC, Yang CT, Tseng MY. Two-year effects of interdisciplinary intervention for hip fracture in older Taiwanese. J Am Geriatr Soc. 2010;58(6):1081–9.
40. Jitapunkul S. A randomised controlled trial of regular surveillance in Thai elderly using a simple questionnaire administered by non-professional personnel. J Med Assoc Thail. 1998;81(5):352–6.
41. Assantachai P, Chatthanawaree W, Thamlikitkul V, Praditsuwan R, Pisalsarakij D. Strategy to prevent falls in the Thai elderly: a controlled study integrated health research program for the Thai elderly. J Med Assoc Thail. 2002;85(2):215–22.
42. Woo J, Hong A, Lau E, Lynn H. A randomised controlled trial of tai chi and resistance exercise on bone health, muscle strength and balance in community-living elderly people. Age Ageing. 2007;36(3):262–8.
43. Yamada M, Aoyama T Fau - Arai H, Arai H Fau - Nagai K, Nagai K Fau - Tanaka B, Tanaka B Fau - Uemura K, Uemura K Fau - Mori S, Mori S Fau - Ichihashi N, Ichihashi N: Complex obstacle negotiation exercise can prevent falls in community-dwelling elderly Japanese aged 75 years and older. Geriatrics and Gerotology International 2012, 12(3):461–467.
44. Yamada M, Higuchi T, Nishiguchi S, Yoshimura K, Kajiwara Y, Aoyama T. Multitarget stepping program in combination with a standardized multicomponent exercise program can prevent falls in community-dwelling older adults: a randomized, controlled trial. J Am Geriatr Soc. 2013;61(10):1669–75.
45. Ohtake M, Morikagi Y, Suzuki I, Kanoya Y, Sato C. Effects of exercise on the prevention of conditions leading to the need for long-term care. Aging Clin Exp Res. 2013;25(1):49–57.
46. Kim H, Yoshida H, Suzuki T. Falls and fractures in participants and excluded non-participants of a fall prevention exercise program for elderly women with a history of falls: 1-year follow-up study. Geriatr Gerontol Int. 2014; 14(2):285–92.
47. Hirase T, Inokuchi S, Matsusaka N, Okita M. Effects of a balance training program using a foam rubber pad in community-based older adults: a randomized controlled trial. J Geriatr Phys Ther. 2015;38(2):62–70.
48. Kamei T, Kajii F, Yamamoto Y, Irie Y, Kozakai R, Sugimoto T, Chigira A, Niino N. Effectiveness of a home hazard modification program for reducing falls in urban community-dwelling older adults: a randomized controlled trial. Jpn J Nurs Sci. 2015;12(3):184–97.
49. Hwang HF, Chen SJ, Lee-Hsieh J, Chien DK, Chen CY, Lin MR. Effects of home-based tai chi and Lower extremity training and self-practice on falls and functional outcomes in older fallers from the emergency department-a randomized controlled trial. J Am Geriatr Soc. 2016;64(3):518–25.
50. Lee HC, Chang KC, Tsauo JY, Hung JW, Huang YC, Lin SI. Effects of a multifactorial fall prevention program on fall incidence and physical function in community-dwelling older adults with risk of falls. Arch Phys Med Rehabil. 2013;94(4):606–15. 615 e601
51. Leung KS, Li CY, Tse YK, Choy TK, Leung PC, Hung VW, Chan SY, Leung AH, Cheung WH. Effects of 18-month low-magnitude high-frequency vibration on fall rate and fracture risks in 710 community elderly–a cluster-randomized controlled trial. Osteoporos Int. 2014;25(6):1785–95.
52. Ashari A, Hamid TA, Hussain MR, Hill KD. Effectiveness of individualized home-based exercise on turning and balance performance among adults older than 50 yrs: a randomized controlled trial. Am J Phys Med Rehabil. 2016;95(5):355–65.
53. Ng TP, Feng L, Nyunt MS, Feng L, Niti M, Tan BY, Chan G, Khoo SA, Chan SM, Yap P, et al. Nutritional, Physical, Cognitive, and Combination Interventions and Frailty Reversal Among Older Adults: A Randomized Controlled Trial. Am J Med. 2015;128(11):1225–36. e1221
54. Sherrington C, Tiedemann A, Fairhall N, Close J, Lord S. Exercise to prevent falls in older adults: an updated meta-analysis and best practice recommendations. N S W Public Health Bulletin. 2011;22(3–4):78–83.
55. Sitja-Rabert M, Martinez-Zapata MJ, Fort-Vanmeerhaeghe A, Rey-Abella F, Romero-Rodriguez D, Bonfill X. Whole body vibration for older persons: an open randomized, multicentre, parallel, clinical trial. BMC Geriatr. 2011;11:89.
56. Hill K, Choi W, Smith R, Condron J. Tai chi in Australia: an acceptable and effective approach to improve balance and mobility in older people? A pilot study. Australas J Ageing. 2005;24:9–13.
57. Li F, Harmer P, Fisher KJ, McAuley E, Chaumeton N, Eckstrom E, Wilson NL. Tai chi and fall reductions in older adults: a randomized controlled trial. J Gerontol A Biol Sci Med Sci. 2005;60(2):187–94.
58. Hill K, Choi W, Smith R, Condron J. Tai chi in Australia: acceptable and effective approach to improve balance and mobility in older people? Australas J Ageing. 2005;24(1):9–13.
59. Sherrington C, Michaleff ZA, Fairhall N, Paul SS, Tiedemann A, Whitney J, Cumming RG, Herbert RD, Close JC, Lord SR. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. [Epub ahead of print]. Br J Sports Med. 2016;
60. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. Age Ageing. 2012;41(1):16–23.
61. Kalula SZ, Scott V, Dowd A, Brodrick K. Falls and fall prevention programmes in developing countries: environmental scan for the adaptation of the Canadian falls prevention curriculum for developing countries. J Saf Res. 2011;42(6):461–72.
62. Kong NH, Chow PK. Conducting randomised controlled trials across countries with disparate levels of socio-economic development: the experience of the Asia-Pacific hepatocellular carcinoma trials group. Contemp Clin Trials. 2013;36(2):682–6.
63. Dsouza S, Rajashekar B, Dsouza H, Kumar K. Falls in Indian older adults: a barrier to active ageing. Asian J Gerontol Geriatr. 2014;9(1):1–8.
64. Kwan MM, Tsang WW, Lin SI, Greenaway M, Close JC, Lord SR. Increased concern is protective for falls in Chinese older people: the chopstix fall risk study. J Gerontol A Biol Sci Med Sci. 2013;68(8):946–53.
65. Bulsara C, Khong L, Hill K, Hill A. Investigating community perspectives on falls prevention information seeking and delivery: older person perceptions regarding preferences for falls prevention education using a world cafe approach. J Community Psychol. 2016;44(7):937–44.
66. Khong LAM, Berlach RG, Hill KD, Hill AM. Can peer education improve beliefs, knowledge, motivation and intention to engage in falls prevention amongst community-dwelling older adults? European journal of ageing. 2017;14(3):243–55.
67. Tan PJ, Khoo EM, Chinna K, Hill KD, Poi PJ, Tan MP. An individually-tailored multifactorial intervention program for older fallers in a middle-income developing country: Malaysian falls assessment and intervention trial (MyFAIT). BMC Geriatr. 2014;14:78.
68. Kwok BC, Mamun K, Chandran M, Wong CH. Evaluation of the Frails' fall efficacy by comparing treatments (EFFECT) on reducing fall and fear of fall in moderately frail older adults: study protocol for a randomised control trial. Trials. 2011;12:155.
69. Tennstedt S, Howland J, Lachman M, Peterson E, Kasten L, Jette A. A randomized, controlled trial of a group intervention to reduce fear of falling and associated activity restriction in older adults. J Gerontol B Psychol Sci Soc Sci. 1998;53(6):P384–92.

Marley et al. BMC Musculoskeletal Disorders (2017) 18:482
DOI 10.1186/s12891-017-1836-2

BMC Musculoskeletal Disorders

RESEARCH ARTICLE Open Access

The effectiveness of interventions aimed at increasing physical activity in adults with persistent musculoskeletal pain: a systematic review and meta-analysis

Joanne Marley[1,4], Mark A. Tully[2,3*], Alison Porter-Armstrong[1], Brendan Bunting[1], John O'Hanlon[4], Lou Atkins[5], Sarah Howes[1] and Suzanne M. McDonough[1,3,6]

Abstract

Background: Individuals with persistent musculoskeletal pain (PMP) have an increased risk of developing co-morbid health conditions and for early-mortality compared to those without pain. Despite irrefutable evidence supporting the role of physical activity in reducing these risks; there has been limited synthesis of the evidence, potentially impacting the optimisation of these forms of interventions. This review examines the effectiveness of interventions in improving levels of physical activity and the components of these interventions.

Methods: Randomised and quasi-randomised controlled trials were included in this review. The following databases were searched from inception to March 2016: CENTRAL in the Cochrane Library, Cochrane Database of Systematic Reviews (CDSR), MEDLINE, Embase, CINAHL, PsycINFO and AMED. Two reviewers independently screened citations, assessed eligibility, extracted data, assessed risk of bias and coded intervention content using the behaviour change taxonomy (BCTTv1) of 93 hierarchically clustered techniques. GRADE was used to rate the quality of the evidence.

Results: The full text of 276 articles were assessed for eligibility, twenty studies involving 3441 participants were included in the review. Across the studies the mean number of BCTs coded was eight (range 0–16); with 'goal setting' and 'instruction on how to perform the behaviour' most frequently coded. For measures of subjective physical activity: interventions were ineffective in the short term, based on very low quality evidence; had a small effect in the medium term based on low quality evidence (SMD 0.25, 95% CI 0.01 to 0.48) and had a small effect in the longer term (SMD 0. 21 95% CI 0.08 to 0.33) based on moderate quality evidence. For measures of objective physical activity: interventions were ineffective - based on very low to low quality evidence.

Conclusions: There is some evidence supporting the effectiveness of interventions in improving subjectively measured physical activity however, the evidence is mostly based on low quality studies and the effects are small. Given the quality of the evidence, further research is likely/very likely to have an important impact on our confidence in effect estimates and is likely to change the estimates. Future studies should provide details on intervention components and incorporate objective measures of physical activity.

Keywords: Physical activity, Low back pain, Osteoarthritis, Musculoskeletal pain, Chronic pain, Persistent pain, Behaviour change techniques, Systematic review

* Correspondence: m.tully@qub.ac.uk; http://go.qub.ac.uk
[2]Centre for Public Health, Queens University Belfast, Royal Victoria Hospital, Grosvenor Road, Belfast BT12 6BA, UK
[3]UKCRC Centre of Excellence for Public Health (Northern Ireland), Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queens University Belfast Room 02020, Institute of Clinical Science B, Royal Victoria Hospital, Grosvenor Road, Belfast, BT 12 6BJ, UK
Full list of author information is available at the end of the article

Background

Epidemiological studies suggest one in five people across Europe suffer from persistent pain [1, 2]. Most persistent pain arises from musculoskeletal disorders, such as low back pain and osteoarthritis; both of which are considered leading causes of disability, worldwide [3]. It can be expected that with aging populations, the health, economic, and social problems associated with these conditions are likely to rise [1, 2, 4]. In addition to causing considerable disability, persistent musculoskeletal pain (PMP) also increases an individual's risk of developing other health conditions including; depression, obesity, heart disease [5–7], cancer [8] and indeed early mortality [7–9]. Despite this, efforts to address these broader health implications of PMP are somewhat lacking.

Description of the intervention

Clinical guidelines widely endorse exercise and/or physical activity (PA) in the management of PMP [10–17]. This is largely due to the positive impact these interventions can have on reducing pain and disability. However, improving levels of PA can lead to broader health benefits: with even small changes in PA levels leading to substantial health gains [18, 19].

PA can be defined as any movement produced by skeletal muscles resulting in energy expenditure, it occurs across several domains including: social and domestic activities, commuting, recreational and leisure activities [20]. PA may or may not include exercise: exercise is a subset of PA tending to be planned, structured or repetitive [20] with a specific purpose such as improving strength, it has been recommended that the terms PA and exercise are not confused [21].

How the intervention might work

Improving levels of PA requires behaviour change. Behaviour change interventions are coordinated sets of activities designed to change specified patterns of behaviour [22]. Behaviour change techniques (BCTs) are the components of interventions that effect change [23]. Taxonomies of BCTs have been used to describe intervention content in a number of PA behaviour change interventions [24–28]. Across these interventions and in line with NICE recommendations for individual level behaviour change [29], some consistent techniques appear to be associated with effective interventions e.g. self-monitoring behaviour, providing feedback, and goal setting.

Why it is important to do this review

PA and exercise interventions are often recommended in the management of PMP as they can have a positive effect on pain and disability levels. However, the extent to which these interventions actually result in changes to behaviour and consequently increased levels of physical activity is less clear. Although individual studies have demonstrated it is possible to increase PA levels in those with back pain [30] or osteoarthritis [31, 32], the results of systematic reviews are conflicting and limited. In adults with osteoarthritis a systematic review concluded that self-management programmes achieve small improvements in subjectively measured PA in the short-term [32]: whereas, a review of PA interventions in adults with PMP reported no improvements in objectively measured PA [33]. Furthermore, the BCTs used within these forms of interventions and the relationship if any, to outcomes has not yet been systematically explored.

Objectives

This systematic review investigated the effectiveness of any form of intervention with a clear aim of increasing PA in adults with PMP. Possible associations between BCTs or intervention characteristics and intervention effects were also investigated.

The objectives of this review are to:

1. Determine the effectiveness of interventions in increasing PA levels in adults with PMP.
2. Identify BCTs used within interventions.
3. Determine if particular BCTs or other intervention characteristics (intensity, recruitment route, type of PA, etc.) are associated with greater effect sizes.

Methods

The full protocol for this review has been published [34].

Population

Randomised and quasi-randomised controlled trials in adults (≥18) with PMP (pain lasting ≥3 months), in the axial skeleton or large peripheral joints were included. We excluded studies focusing on fibromyalgia, inflammatory and/or autoimmune disorders and perioperative patients, which may require a different management strategy.

Types of interventions

All interventions that had a clear aim of increasing PA in adults with PMP were eligible for inclusion. We excluded site specific rehabilitative exercise interventions unless it was clear the intervention also addressed habitual PA. We included trials with a comparative control group and trials with multiple intervention arms. We did not include population or community-wide interventions.

Types of outcome measures

The primary outcome of interest was PA measured by self-reported or objective measures; questionnaires, recall diaries, pedometers or actigraphy. Measurements of adherence or attendance at classes alone, were not

sufficient. The secondary outcome of interest was adverse incidents.

Search methods for identification of studies

Search strategies were developed for each electronic database and were based on the initial Medical Literature Analysis and Retrieval System Online (MEDLINE) strategy (Additional file 1). We searched the Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library, the Cochrane Database of Systematic Reviews (CDSR) in the Cochrane Library, Ovid MEDLINE(R) Daily Update, Ovid MEDLINE(R), Ovid MEDLINE (R) - includes new records, not yet fully indexed, Ovid Embase, EBSCO Cumulative Index to Nursing and Allied Health Literature (CINAHL), Ovid PsycINFO, AMED (Allied and Complementary Medicine). All databases were searched from inception to March 2016.

Reference lists of systematic reviews and articles retrieved from the search were scanned for additional references.

Data collection and analysis

Selection of studies

Results from the searches were imported into EndNote (X7) bibliographic software (Thomson Reuters, Philadelphia, PA, USA) and duplicates removed. Titles and abstracts obtained from the search were independently screened by two authors (JM 100%, MAT 70% and SMcD 30%). Articles not meeting the inclusion criteria and outside the scope of the review were removed. Full text reports of the remaining publications were retrieved. Two review authors (JM, SMcD) used a standardised form tested prior to use, to select trials eligible for inclusion. Non-English papers were assessed and, where necessary, translated in part or in full.

Data extraction and management

Data was extracted independently by two reviewers (JM, SMcD) using a customised form tested prior to use. Relevant data was extracted for methodological issues, intervention characteristics, study design, study characteristics and adverse events. Intervention content was coded according to the BCTTv1 [35]. Two coders (JM, SH) independently coded BCTs, inter-rater reliability was assessed using the prevalence-adjusted bias-adjusted Kappa (PABAK) statistic [36]. PABAK adjusted for the high frequency of agreement on absent BCTs. Values of 0.60–0.79 indicated 'substantial' reliability and 0.80 and above 'outstanding' reliability [37].

Assessment of risk of bias in included studies

Two reviewers (JM, SMcD) independently assessed studies for risk of bias (ROB), using the Cochrane risk of bias tool [38]. An additional domain was added to determine if studies were adequately powered. For cluster randomised controlled trials, five additional domains were assessed, as recommended by Cochrane (16.3.2) [38].

Quality of the evidence

The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach was used to interpret and evaluate the quality of the evidence [39, 40]. The methods and recommendations described in the Cochrane handbook [38] and by the GRADE working group [33] were used to assess the quality of a body of evidence using five domains: risk of bias, inconsistency, indirectness of evidence, imprecision of effect estimates and potential publication bias. Data for each outcome was entered into GRADEpro to create 'Summary of Findings' table and footnotes were used to justify all decisions on the downgrading of the quality of the evidence.

The definitions described by the GRADE working group were used to grade the quality of evidence as follows:

- High – Further research is very unlikely to change our confidence in the estimate of effect.
- Moderate – Further research is likely to have an important impace on our confidence in the estimate of effect and may change the estimate.
- Low – Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- Very low – Any estimate of effect is very uncertain.

Measures of treatment effect

Continuous outcomes were analysed using post intervention measures, we reported effect sizes using the standardised mean difference (SMD) as outcomes were reported across different scales. For comparisons of the results we categorised studies into effect sizes according to Cohen's classification; SMD; 0.2 < 0.3 as small, 0.3–0.8 as moderate, >0.8 as large [41]. *P*-values of <0.05 and confidence intervals that excluded null values were considered statistically significant.

Unit of analysis issues

Where studies involved multiple intervention groups we followed recommendations suggested by the Cochrane collaboration (16.5.4) [38] by combining similar intervention groups to perform a single pairwise comparison.

Where studies reported PA domains separately or reported more than one PA outcome, data were extracted for each, however, for the effect size analysis, measures of overall PA were given preference, if these were not available leisure time PA was given preference'.

To facilitate exploration of results not suitable for quantitative synthesis we grouped studies by effect size

using an aggregate of subjective and objective measures (objective measures given preference to subjective where available) at the post intervention time point.

Dealing with missing data

Attempts were made to contact original investigators to request missing data.

The frequency and duration of the intervention was used to calculate an estimated overall intervention contact time 'intensity'. The calculation was based on the full intervention being delivered as planned. If the duration of a session was not reported or the data was unobtainable from authors, we allocated 20 min for telephone follow up and 45 min for face to face interventions.

Assessment of heterogeneity

Diversity across the studies was qualitatively assessed in terms of the intervention, participant demographics, outcome measures and follow-up. Data was assessed for statistical heterogeneity using RevMan version 5.3 using the I^2 statistic, values of I^2 ranging from 30% to 60% were considered to represent moderate heterogeneity and 50% to 90% substantial heterogeneity [38].

Data synthesis

Separate meta-analyses were completed for subjective and objective outcome data at three time points; short term (not longer than 12 weeks' post-randomisation), medium term (not longer than 6 months' post randomisation) and long term (greater than 6 months post randomisation). Outcomes were analysed using the SMD, with the inverse variance method to calculate the overall effect and standard error, a random effects model was applied to incorporate heterogeneity.

Subgroup analysis and investigation of heterogeneity

We performed the following pre-specified subgroup analysis:

- Clinical subgroups: classified as 'persistent low back pain' and 'osteoarthritis'
- Frequency and duration of intervention (intensity) classified as 'higher' or 'lower' relative to the median number of contact hours across the studies

The following subgroups were planned but not conducted as the data generated was deemed insufficient.

- BCTs
- Recruitment routes

Descriptive statistics were therefore used to explore possible associations between these factors and other intervention characteristics and intervention effects.

Sensitivity analysis

A sensitivity analysis was performed to check if excluding studies with a higher ROB affected results. The threshold for sensitivity analysis was set for studies meeting at least 50% of the criteria of the ROB assessment, excluding blinding of participants and providers.

Results

Results of the search

The electronic searches returned 18,953 records, (Fig. 1) after de-duplication in the referencing software, 11,323 title and abstracts were screened against the inclusion criteria. In total 276 records were identified as potentially relevant, and the full text reports were retrieved. Twenty-six studies were initially agreed for inclusion; six studies were subsequently found to contain unusable

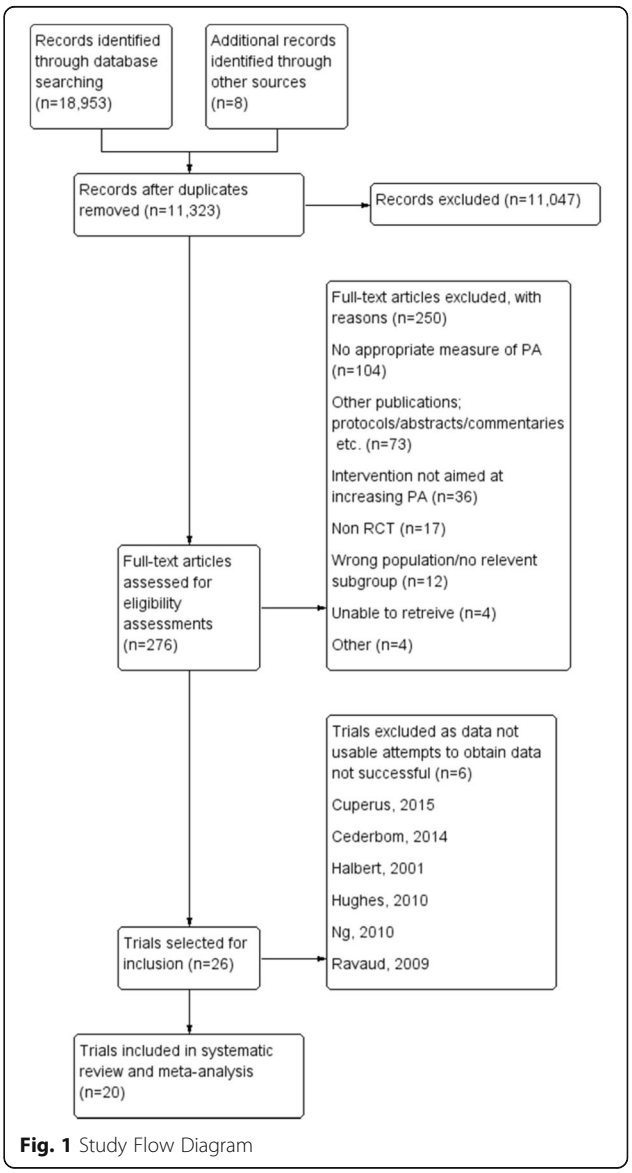


Fig. 1 Study Flow Diagram

outcome data, requests to obtain the data were not successful (Fig. 1). Twenty studies had sufficient data to be included in a meta-analysis [30–32, 42–58]. Nine authors were contacted regarding studies that were deemed to have potentially usable data; six replied, four authors provided the information needed to include their study [43, 44, 52, 58].

Eight non-English language studies were translated but none were eligible for inclusion.

Excluded studies

A total of ($n = 250$) studies were excluded from the review. Exclusions were most often due to no or unacceptable measures of PA and studies having no clear aim of increasing PA (Fig. 1).

Characteristics of included studies

Ten studies were described as randomised controlled trials (RCTs), three were cluster RCTs [43, 44, 55], five feasibility or pilot RCTs [30, 32, 48, 51, 56] and one was a controlled clinical trial [42]. Sullivan et al. [57] reported a one year follow-up of patients who had participated in an RCT [59]. The maximum number of groups within studies was three, [44, 46, 47, 52].

Participants in included studies

The studies involved 3441 suitable participants (4875 in total) (Table 1), over half were female (approx. 59.2%). Thirteen studies focused on osteoarthritis (1874 participants; $n = 7$ knee, $n = 5$ hip and/or knee, $n = 1$ generalised) and seven on persistent low back pain ($n = 1567$ participants). The mean age of participants with osteoarthritis ranged from 61 to 73.8 years, and for persistent low back pain from 40.4 years to 51.9 years.

Interventions

Table 2 summarises modes of delivery, intervention content, provider and intensity for each intervention. Most studies incorporated more than one mode of delivery but have been described according to what was considered the 'primary' delivery mode. Most interventions were provided by healthcare professionals (12/20), other providers included exercise and fitness professionals and a counsellor. Intervention contact times ranged from <1 h for a educational pamphlet [32] to approximately 200 h of contact time [46] occurring over a twelve month intervention. The median number of contact hours was 8.3 h. Walking was the most common form of PA, followed by multicomponent programmes utilising a mixture of aerobic, strengthening and/or general flexibility exercises. All of the interventions incorporated some form of educative component relating to the role of PA in managing PMP.

A total of 160 BCTs (mean per study 8, range 0–16) were coded across the 20 studies (Table 2). The most frequently coded techniques were 'goal setting (behaviour)' and 'instruction on how to perform the behaviour' (65%) followed by 'behavioural practice/rehearsal' and 'self-monitoring of the behaviour' (55%). A mean PABAK score of (0.9) indicated outstanding agreement on identification of BCTs.

Control groups

The content of control groups varied (Table 1); seven studies referred to control groups as 'treatment as usual' or some form of 'standard care' [30, 43, 49, 52, 55–57]. Two studies [45, 58] used waiting list control groups. A clinical guideline posted to GP's was used as a control in the study by Becker et al. [44]. Pamphlets were used as a control in the study by Brosseau et al. [46] and a copy of the 'Arthritis Help book' was given to controls in the study by Hughes et al. [50]. Two studies used self-management programmes in their intervention, but provided it as a stand-alone intervention for controls; [31, 47]. Two studies directly compared two forms of back rehabilitation programmes of varying intensity and content [42, 54]. In the study by Williams et al. [32] the control booklet content differed to the intervention booklet. Krein et al. [53] provided controls with an uploading pedometer and reminder emails to upload data but not access to the web-based intervention, available to the intervention group. In two studies [48, 51] in addition to exercise classes, intervention groups received additional intervention components.

Outcome measures

Across the 20 studies 13 scales or tools for measuring PA were identified (Table 1) twelve studies reported subjective PA; five objective PA and three reported both. Self-reported measures of PA included estimates of total PA and estimates of frequency, intensity and time in different domains of activity. Only two tools were used in more than one study; the International Physical Activity Questionnaire, [32, 51, 52], and the Freiburg Questionnaire of PA, [44, 54]. Objective measures of PA included steps per day or total PA and/or time in different intensities of PA, measured by accelerometers and/or pedometers.

Follow-up (post randomisation) (Table 1)

The longest follow up was 18 months [46] six months after a twelve month intervention. Eleven studies reported outcomes at 12 months [42–45, 48, 50, 52–54, 56, 57] however, the latter two studies involved interventions that lasted the 12 months. Four studies reported outcomes at 6 months [30, 31, 51, 58] and one at 3 months [32]. One study had only post-intervention

Table 1 Characteristics of included studies ($n = 20$)

Author/Year	Study Design	No of Participants	Gender	Age Range	Condition	Intervention	Control Condition	Recruitment Route	PA Outcome	Longest follow-up
Alaranta, 1994 [42]	Controlled Clinical Trial	293	F160 M133	40.4 (4.8) Control 40.5 (4.6) Intervention	PLBP	Home training programme + Inpatient rehabilitation with education	Inpatient rehabilitation 40–50% less strenuous	Finnish Social Security Insurance Institution	Subjective - Leisure time PA (strenuousness)	12 months
Allen, 2016 [43];	Cluster RCT	300 (patients) 30 (providers)	F28 M272	61.6 (9.2)	OA hip/knee	Patients - Physical activity and weight management counselling Healthcare providers received treatment recommendations	Usual care	Medical records veterans affairs	CHAMPS	12 months (12 month intervention)
Becker, 2008 [44];	Cluster RCT	1378 (chronic pain subgroup 332)	F801 M577 (entire group no figures for subgroup)	49.1 (13.3) guideline group 47.4 (13.5) guideline + MC 50.2 (14.3) Control	LBP (mixed)	Practitioner education – guideline implementation Practitioner education – guideline implementation + MI	Guideline delivered via post	Primary Care GP's	Freiburg Questionnaire	12 months
Bossen, 2013 [45];	RCT	199	F129 M70	64 (6.6) All 61 (5.9) Intervention 63 (5.4) Control	OA hip/knee	Web based intervention to increase PA using behavioural graded activity	Waiting list	Volunteers from newspapers and websites	PASE and Subgroup ACTi graph	12 months
Brosseau, 2012; [46];	RCT	222	153F 69M	63.9 (103) Walking 63.9 (8.2) Walking + Booklet 62.3 (6.8) Control	OA Knee	Walking group Walking and behavioural education	Self-directed received educational pamphlet	Unclear	7 day Par (recall)	18 months
Farr, 2010;[47];	RCT	293	F218 M75	55.5 (7.3) Resistance training 55.8 (6.1) Self-Management 54.2 (7.3) Combined	OA Knee	Resistance training + self-management	Self-management	General community mass mailings, media ads and local physicians	ACTi graph 7 days	9 months
Focht, 2014;[48];	RCT (pilot)	80	F67 M13	63.5 (6.86)	OA Knee	Group mediated cognitive behavioural exercise intervention	Traditional centre based exercise	Direct referral State Medical Centre Rheumatologists, ads Arthritis Foundation groups	Accelerometer (PA Lifecorder plus) 7 days	12 months
Hiyama, 2012;[49];	RCT	40	32F 19M	71.9 (5.2) Walking 73.8 (5.7) Control	OA Knee	Instructed to increase number of steps, physical therapy + programme of walking	Physical therapy + advice re walking	Unclear - community dwelling females	Pedometer (steps per day)	4 weeks
Hughes, 2006;[50]	RCT (block randomisation)	215	363F 56M	71.1 (59 -91 yrs)	OA hip/knee	Education, exercise and fitness walking	Arthritis self-help book and information on exercise programmes in community	Senior centres, newsletters, local media, presentations to senior groups	Total minutes exercised	12 months

Table 1 Characteristics of included studies ($n = 20$) *(Continued)*

Author/Year	Study Design	No of Participants	Gender	Age Range	Condition	Intervention	Control Condition	Recruitment Route	PA Outcome	Longest follow-up
Hunter, 2012;[51];	RCT (feasibility)	51	167F M79	43.2 (13.5) Exercise 42.4 (11.3) Exercise Auricular Acupuncture	PLBP	Exercise and acupuncture	Exercise	Primary Care GP's, Physiotherapy waiting list and University population	IPAQ (ActivPal - steps per day)	6 months
Hurley, 2015:[52]:	RCT	246	40F	45.4 (11.4)	PLBP	Walking programme Exercise class	Usual physiotherapy	Physiotherapy departments	IPAQ	12 months
Krien, 2013;[53];	RCT	229	29F 200M	51.2 (12.5) Walking 51.9 (12.8) Enhanced Usual Care	PLBP	Walking group	Enhanced usual care	Individuals referred for back class and medical record system	Pedometer (steps per day)	12 months
McDonough, 2013; [30];	RCT (feasibility)	56	31F 25M	51 (42 – 60 yrs) Exercise 48 (43 – 55 yrs) Exercise Walking Programme	PLBP	Education and advice and walking group	Usual care	Physiotherapy waiting lists primary care	MGROC PA (ActivPal - steps per day)	6 months
Meng, 2011;[54];	RCT	360	231F 129M	50.2 (7.6) Intervention 49.5 (7.7) Control	PLBP	Biopsychosocial back school programme (inpatient)	Traditional back school (setting unclear)	Orthopaedic hospital - patients had applied for inpatient rehabilitation	Freiburger Questionnaire	12 months
Pisters, 2010 [55]	RCT Cluster (analysis of secondary outcomes)	200	F154 M46	64.8 (7.9)	OA Hip or Knee	Behavioural graded activity and operant conditioning and exercise therapy	Usual physiotherapy (per clinical guidelines)	Physiotherapists and press releases in local newspapers	PA SQUASH - Converted using METs total hrs. Per week in health enhancing PA	65 weeks (14.9 months)
Schlenk, 2011;[56];	RCT (feasibility)	26	F25 M1	63.2 (9.8)	OA Knee (overweight)	Counselling, exercise, fitness walking programme	Usual care	Rheumatology practices, arthritis disease network registry, self-referral	Diary - Minutes walked per week and other aerobic PA minutes	12 months
Sullivan, 1998;[57];	RCT (follow-up)	102 (52 in this follow-up)	F85 M17 (f44 m8)	70.38 (9.11) Intervention 68.48 (11.32) Control	OA Knee	Supervised fitness walking and supportive education	Standard medical care, weekly interviews about function and daily activity	Community clinics, private clinics - rheumatology	Recall - Average distance walked per week	12 months
Talbot, 2003;[31];	RCT	34	F26 M8	69.59 (6.74) Pedometer 70.76 (4.71) Education	OA Knee	Arthritis self-management programme + walking programme	Arthritis self-management programme	Senior Centres and ads in local papers	Pedometer (steps per day) + Accelerometer	6 months
Trudeau, 2015;[58];	RCT	228 (Subgroup 94)	F72 M156	49.9 (11.6)	Arthritis (all – subgroup data OA spine, large	Web-based painAction programme, informative articles, self-check assessments etc.	Waiting list control	Flyers in surgeries, Pain association members, google adwords,	Aerobic exercise minutes (all)	6 months

Table 1 Characteristics of included studies ($n = 20$) *(Continued)*

Author/Year	Study Design	No of Participants	Gender	Age Range	Condition	Intervention	Control Condition	Recruitment Route	PA Outcome	Longest follow-up
					peripheral joints via author)			ClinicalTrials.gov. PainEDU.org health professionals		
Williams, 2011 [32]	RCT (feasibility)	119	F76 M43	68.2 (8.1) Intervention 68.6 (8.5) control	OA Hip or Knee	'New' Advice booklet – emphasis on addressing exercise related beliefs	Arthritis UK booklet	GP Practices	IPAQ	3 months

PA Physical Activity, *MI* Motivational Interviewing, *IPAQ* International physical activity questionnaire, *MGROC* Modified global rating of change (physical activity), *SQUASH* Short questionnaire to assess health enhancing physical activity, *PASE* Physical activity scale for elderly, *7 day PAR* 7-day physical activity recall, *CHAMPS* Community healthy activities model programme for seniors, *OA* Osteoarthritis, *LBP* Low back pain, *PLBP* Persistent low back pain, *RCT* Randomised controlled trial

Table 2 Interventions, quality assessment, BCTs - studies grouped post intervention using aggregated outcome measures

Author, Year	Hiyama, 2012;	Hughes, 2006;	Alaranta, 1994;	Focht, 2014;	Pisters, 2010;	Farr, 2010;	Allen, 2016;	Meng, 2011;	Becker, 2008;	Sullivan, 1998;
Effect Size SMD 95% CI	1.96 [1.19, 2.73]	0.87 [0.58, 1.15]	0.77 [0.53, 1.01]	0.56 [0.07, 1.06]	0.51 [0.21, 0.80]	0.29 [−0.03, 0.61]	0.28 [0.04, 0.53]	0.25 [0.02, 0.48]	0.17 [−0.07, 0.41]	0.12 [−0.50, 0.74]
ROB assessment	Lower	Higher	Higher	Lower	Lower	Higher	Lower	Lower	Higher	Higher
Mode of delivery										
Automated Web-based										
Inpatient Programme			x					x		
Centre-based	x	x		x	x	x				x
Home-based		+	+	x	x					
Community-based										
Other							x		x	
Session structure										
Individual	?	+	+		x	x	x	x	x	+
Group based		x	x	x		x		x		x
Type of PA										
Multicomponent Exercise Programme		x	x	x		x		x		
Walking	x	x		x						x
User Selected					x		x		x	
Other/Unclear							x		x	
Provider										
Physiotherapist	x	x			x					x
Nurse									x	
Doctor									x	
Fitness Professional				x		x				
Multidisciplinary			x					x		
Other						? SM	?			
Estimated Intervention Contact Time (hrs)	3	36	111	36	11.5	134	6	50	1.5	24
No. of BCT's coded	3	12	3	16	9	5	16	1	0	6

+ to a lesser extent, ? unclear from study description,/not explicit, *SM* self-management, *WP* walking programme, *EC* exercise class., *SMD* standardised mean difference, *CI* confidence intervals, *ROB* risk of bias (meeting at least 50% of domains assessed, excluding blinding participants and providers)

Table 2 Interventions, quality assessment, BCTs - studies grouped post intervention using aggregated outcome measures *(Continued)*

Author, Year	Williams, 2011;	Brosseau, 2012;	Trudeau, 2015;	Hunter, 2012;	Bossen, 2013;	Schlenk, 2011;	McDonough, 2013;	Krien, 2013;	Hurley, 2015;	Talbot, 2003;
Effect Size SMD 95% CI	0.11 [−0.31, 0.53]	0.10 [−0.27, 0.48]	0.07 [−0.35, 0.49]	0.06 [−0.60, 0.72]	0.02 [−0.50, 0.54]	−0.00 [−0.77, 0.77]	−0.00 [−0.74, 0.73]	−0.03 [−0.35, 0.30]	−0.29 [−0.59, 0.01]	−0.32 [−1.0, 0.35]
ROB assessment	Lower	Lower	Lower	Lower	Lower	Lower	Lower	Lower	Lower	Higher
Mode of delivery										
Automated Web-based			x		x			x		
Inpatient Programme										
Centre-based				x		x			x	
Home-based							x		x	x
Community-based		x								
Other	x									
Session structure										
Individual		x				x	x		x (WP)	x
Group based		x		x					x (EC)	x
Type of PA										
Multicomponent Exercise Programme				x		x			x (EC)	
Walking		x				x	x	x	x (WP)	x
User Selected	x				x					
Other/Unclear	x		x							
Provider										
Physiotherapist				x		x	x	+	x	
Nurse						x				x
Doctor										
Fitness Professional		x								
Multidisciplinary										
Other	x		x		x			x		
Estimated Intervention Contact Time (hrs)	0.5	200.5	4.3	8	1.166	7.5	3.5	8.6	8	12.15
No. of BCT's coded	2	14	5	7	12	10	11	8	15	5

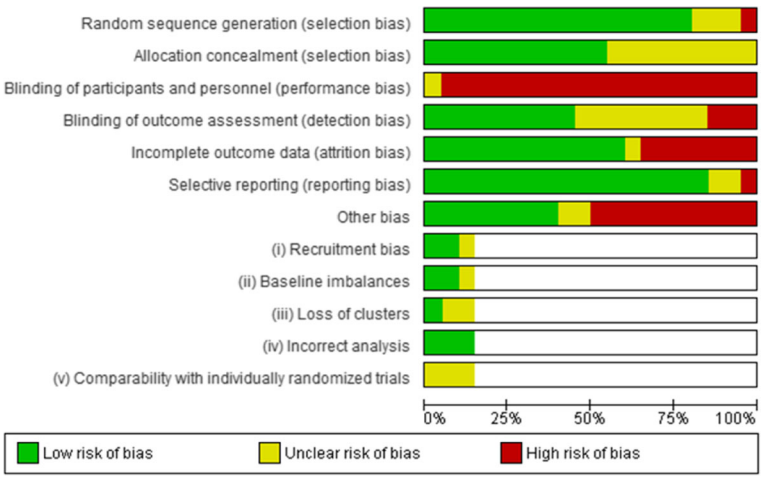


Fig. 2 Risk of bias summary of all studies assessed using Cochrane risk of bias tool

outcomes at four weeks [49] and one study reported outcomes at nine months [47]. Pisters et al. [55] reported outcomes at 65 weeks, the intervention duration was described as 12 weeks however booster sessions were provided to participants up until week 55.

Risk of bias in included studies (Figs. 2 and 3)

The ROB in the included studies is summarised in Figs. 2 and 3. Blinding, inadequately powered studies and attrition bias were considered the greatest ROB in the included studies. Due to the difficultly in blinding participants and providers in PA interventions, the risk of performance bias was considered high in all but one study which involved posting an intervention or control pamphlet to participants [32], the review authors felt there was insufficient information in the report to support a judgement of high or low ROB for this study. The majority of studies included in the review were not sufficiently powered, only nine reported conducting a power calculation for their primary outcome [32, 43–45, 48, 52–55]. Only two studies [45, 55] conducted power calculations for PA outcomes. Attrition bias was considered high in just over one third of the included studies (35%).

Risk of bias in cluster randomised controlled trials

Three studies utilised cluster RCTs [43, 44, 55], summarised in (Figs. 2 and 3). Two studies [43, 55] were judged to be of unclear ROB in relation to loss of clusters, this was due to the loss of clusters not being reported or discussed in the analysis or results. ROB on comparability with individually randomised trials was unclear in all three studies, this was largely due to a lack of reporting of comparability or the influence of clustering on intervention effects.

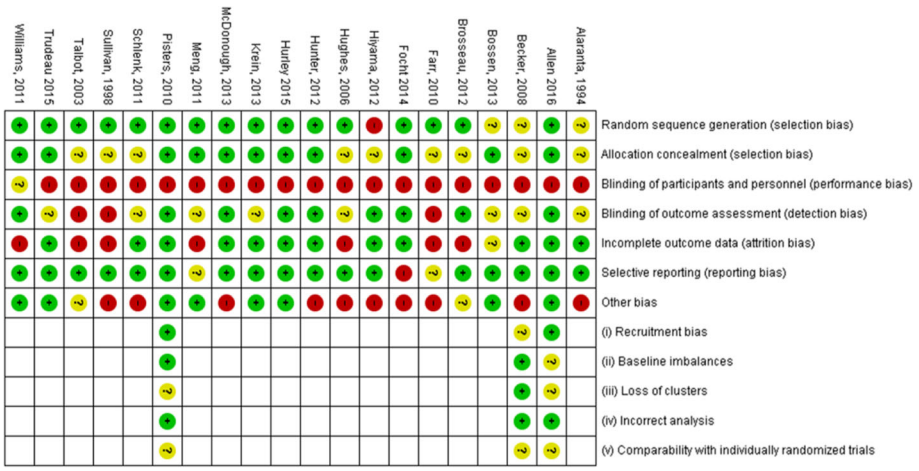


Fig. 3 Risk of bias in individual studies

Effects of interventions: Meta-analysis

Meta-Analysis 1: Effects of Intervention versus control on subjectively measured PA.

Fifteen studies reported continuous measures of subjective self-reported PA [30, 32, 42–46, 50–52, 54–58].

Short term: no longer than 12 weeks post randomisation.

Nine studies (1096 participants) reported short term subjective PA outcomes (Fig. 4) [30, 32, 42, 45, 50–52, 57, 58]. Based on very low quality evidence the pooled effects of the interventions showed no demonstrable effect (SMD 0.24, 95% CI -0.07, 0.55). The quality of the evidence was downgraded from high to very low quality due to substantial statistical heterogeneity ($I^2 = 83\%$), wide confidence intervals around the effect estimate and ROB (Table 3).

Medium term: greater than 3 months, not more than 6 months post randomisation.

Nine studies (1309 participants) reported medium term measures (Fig. 4) [30, 44, 50–52, 54–56, 58]. Based on low quality evidence the pooled effects of the studies at the medium term was significant with a small effect size (SMD 0.25, 95% CI 0.01, 0.48). The quality of the evidence was downgraded from high due to the substantial heterogeneity in the observed effects ($I^2 = 72\%$) and weighting of studies at high ROB included in the analysis (Table 3).

Long term: greater than 6 months post randomisation.

Eleven studies (1872 participants) reported long term follow-up measures (Fig. 4) [42–46, 50, 52, 54–57]. Based on moderate quality evidence the pooled effects were small and statistically significant (SMD 0.21, 95% CI 0.08, 0.33) heterogeneity was moderate in the observed effects ($I^2 = 40\%$). The quality of the evidence was downgraded from high to moderate due to the

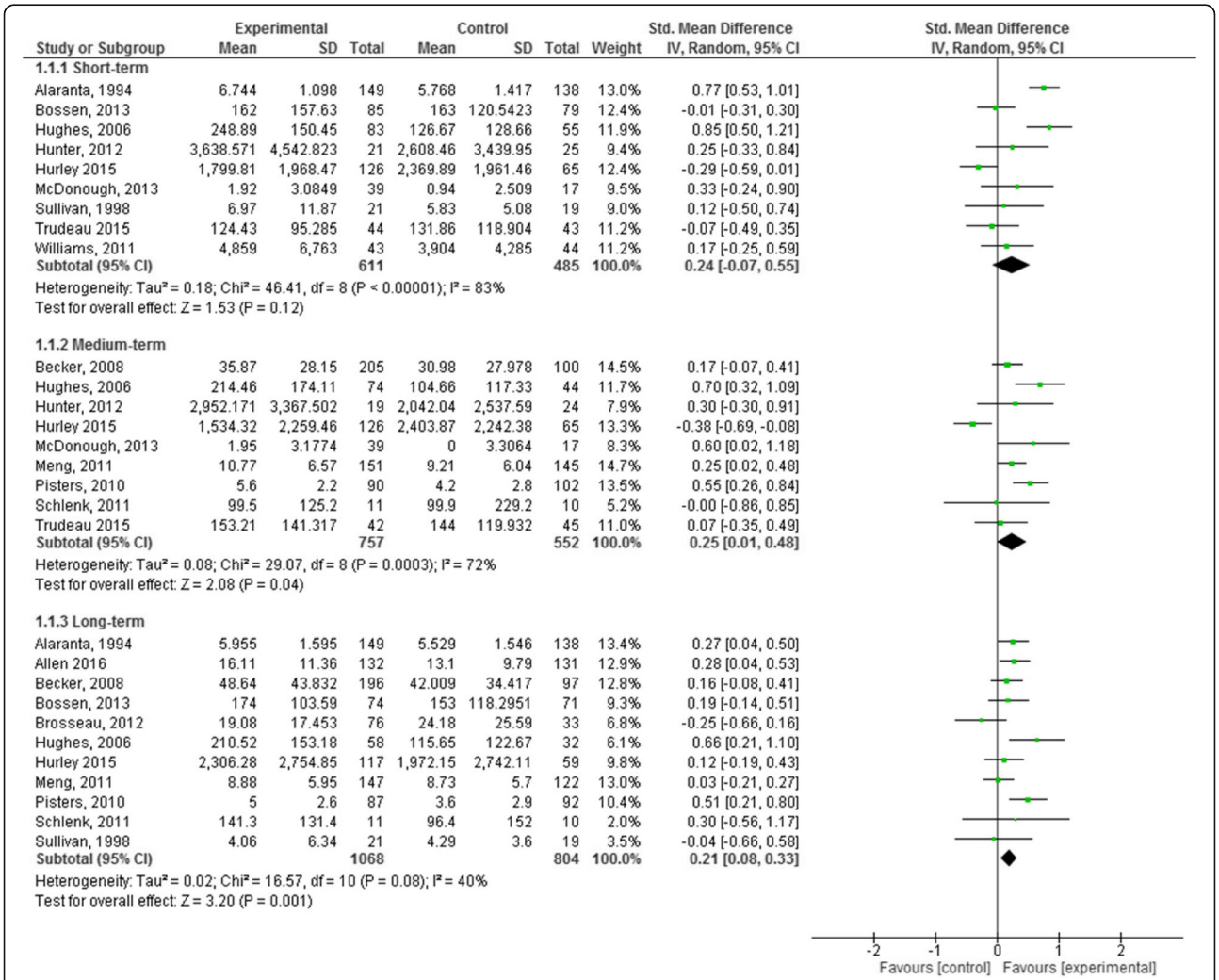


Fig. 4 Forest plot of comparison: 1 Effects of intervention versus control on subjectively measured physical activity: short-term, medium-term and long-term

Table 3 Summary of quality of evidence using the GRADE approach

Quality assessment							№ of patients		Effect	Quality
№ of studies	Study design	Risk of bias (a)	Inconsistency (b)	Indirectness (c)	Imprecision (d)	Other considerations (e)	Interventions	control	Absolute(95% CI)	
Short-term Subjective Physical Activity										
9	randomised trials	serious	serious	not serious	serious	none	611	485	SMD 0.24 SD higher (−0.07 lower to 0.55 higher)	⊕○○○VERY LOW
Medium-Term Subjective Physical Activity (follow up: range 12 weeks to 6 months)										
9	randomised trials	serious	serious	not serious	not serious	none	757	552	SMD 0.25 SD higher (0.01 higher to 0.48 higher)	⊕⊕○○LOW
Long-Term Subjective Physical Activity (follow up: >6 months)										
11	randomised trials	serious	not serious	not serious	not serious	none	1068	804	SMD 0.21 SD higher (0.08 higher to 0.33 higher)	⊕⊕⊕○MODERATE
Short-Term Objective Physical Activity										
7	randomised trials	serious	serious	not serious	serious	none	255	186	SMD 0.31 SD higher (−0.11 lower to 0.74 higher)	⊕○○○VERY LOW
Medium-Term Objective Physical Activity (follow up: range 12 weeks to 6 months)										
4	randomised trials	not serious	not serious	not serious	very serious	none	135	110	SMD −0.02 SD lower (−0.40 lower to 0.36 higher)	⊕⊕○○LOW
Long-Term Objective Physical Activity (follow up: range 6+ months)										
4	randomised trials	serious	not serious	not serious	serious	none	251	184	SMD 0.22 SD higher (−0.02 lower to 0.46 higher)	⊕⊕○○LOW

CI Confidence interval, *SMD* Standardised mean difference

a. Risk of Bias – Using weighting shown in RevMan analysis a serious downgrade is applied where 25% or more of the results are derived from studies judged to be at high risk of bias (see methods for details), a very serious downgrade is applied where 50% of weighting is derived from studies at high risk of bias

b. Inconsistency – a serious downgrade was applied if there is substantial statistical heterogeneity indicated by an (I^2) of 50 to 90%. A very serious downgrade is applied if there was substantial heterogeneity and there was inconsistency arising from the populations, interventions or outcomes

c. Indirectness – a serious downgrade is applied if there was indirectness in one of population, intervention, comparator or outcome. A very serious downgrade was applied if there was indirectness in more than one area

d. Imprecision –a serious downgrade is applied when the total population size is less than 400 (provided there is more than one study). Or, if the 95% CI includes 0 (no effect) or the upper and lower confidence interval cross an effect size (SMD) of 0.5 in either direction. A very serious downgrade is applied where there is a small population and imprecision of the effect estimate

e. Where there was sufficient papers (10) a funnel plot was prepared and inspected, a serious downgrade was applied if this suggested a publication bias

weighting applied to studies judged as high ROB in the analysis (Table 3).

Meta-analysis 2: Effects of intervention versus control on objectively measured PA

Eight studies reported objective measures of PA [30, 31, 45, 47–49, 51, 53].

Short term: no longer than 12 weeks post randomisation.

Seven studies (441 participants) reported short term measures (Fig. 5, Table 3) [30, 31, 45, 47–49, 51]. Based on very low quality evidence, the pooled effect was positive but not significant (SMD 0.31, 95% CI -0.11, 0.74) with substantial heterogeneity ($I^2 = 76\%$). The quality of the evidence was downgraded from high to very low due to wide confidence intervals in the effect estimates and the weighting applied to studies judged as high ROB in the analysis (Table 3).

Medium term: greater than 3 months, not more than 6 months' post randomisation.

Four studies (245 participants) reported medium term measures (Fig. 5) [30, 31, 51, 53]. Based on low quality evidence, the pooled effect was negative (SMD -0.02, 95% CI -0.40, 0.36) with moderate heterogeneity in the observed effects ($I^2 = 41\%$). The quality of the evidence was downgraded due to the small number of participants included in the analysis and wide confidence intervals that included no effect.

Long term: greater than 6 months post randomisation.

Four studies (435 participants) reported long term follow-up measures (Fig. 5) [45, 47, 48, 53]. Based on low quality evidence, the pooled effect was positive but not significant (SMD 0.22, 95% CI -0.02, 0.46) with low heterogeneity in the observed effects ($I^2 = 29\%$). The quality of the evidence was downgraded from high to low due to imprecision of the effect estimates as evidenced by the confidence intervals included no effect and the weighting applied in the analysis to studies at high ROB.

Sensitivity analysis

We examined the pooled effects for the two types of outcomes (subjective and objective) at each time point by an assessment of the ROB. When limited to studies with a lower ROB, effect sizes were not significant at any timepoint.

Subgroup Analyses: To increase statistical power for the planned subgroup analysis we used subjective measures of PA ($n = 16$ studies).

Subgroup analysis 1: Clinical conditions osteoarthritis and low back pain:

Effects were demonstrated for the osteoarthritis subgroup only, effects sizes were moderate in the medium-term (SMD 0.41, 95% CI 0.10, 0.72) and small in the longer term (SMD 0.29, 95%CI 0.08, 0.49).

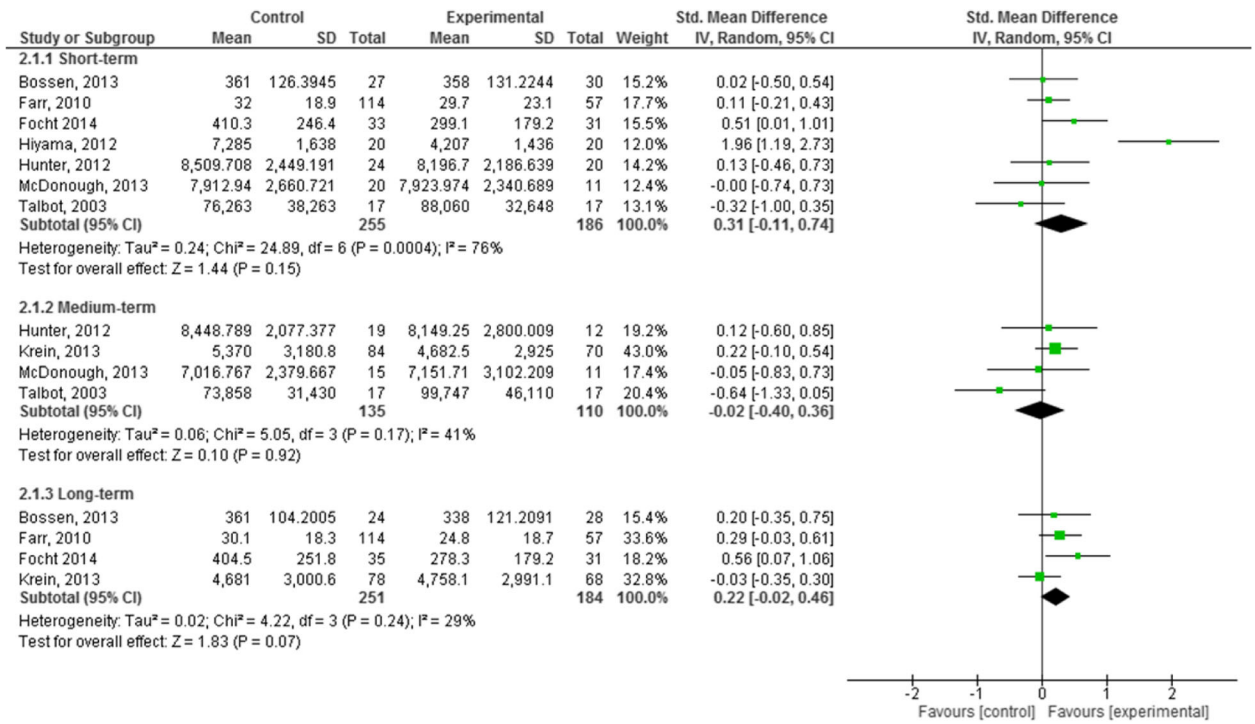


Fig. 5 Forest plot of comparison: 2 Effects of intervention versus control on objectively measured physical activity: short-term, medium term and long-term

Subgroup analysis 2: Intervention Intensity:

Only interventions that were of higher intensity, relative to the median calculated contact hours of the interventions (8.3 h) reached important effect sizes (seven studies). Higher intensity interventions resulted in moderate effect sizes for short term (SMD 0.66 95% CI 0.41, 0.91) and medium term (SMD 0.47 95% CI 0.20, 0.74) outcomes, and small effect sizes for longer term outcomes (SMD 0.25 95% CI 0.02, 0.48).

Influence of BCTS and recruitment route

It was not possible to conduct the quantitative subgroup analysis of BCTs and recruitment routes as the data generated from the review was not sufficient to permit valid comparisons. Descriptive statistics were used to describe possible associations between these factors and other intervention characteristics. To facilitate this exploration, all studies were grouped by effect size, post intervention (Fig. 6).

Behaviour change techniques

Seven studies demonstrated statistically significant small to large effect sizes on post intervention PA (Table 2). Across these studies, 60 BCTs were coded with a mean of 8.57 per study, range (1–16). In total 28 unique BCTs were identified, the most commonly coded were 'goal setting behaviour', and 'instruction on how to perform the behaviour' featuring in 71.4% of studies. 'Self-monitoring behaviour', 'social support (unspecified)', and 'framing/reframing' were also coded frequently and were present in over half of the included studies (57%).

Thirteen studies demonstrated no effect, or negligible effects (<0.2) post intervention (Table 2). Across these studies 100 BCTs were coded with a mean of 7.7 per study, range (0–15) with 31 unique BCTs present. The most commonly coded BCTs were; 'goal setting behaviour', 'information on health consequences' 'instruction on how to perform the behaviour' and 'behavioural practice/rehearsal' which featured in 61.5% of the studies.

Recruitment route and other intervention characteristics: (Tables 1and 2)

No notable differences were observed with regards to the influence of recruitment route, type of PA, mode of delivery and post-intervention effect sizes.

In seven studies demonstrating positive effects, five (71.4%) were delivered by healthcare professionals (2 multidisciplinary and 3 by physiotherapists). In comparison, studies with no effect (<0.2) were less frequently delivered by healthcare professionals (53.8%).

Secondary outcomes

Adverse incidents

Only six studies made explicit statements regarding adverse incidents; two studies, although not explicitly stated, documented adverse incidents. Allen et al. [43] reported four adverse incidents unrelated to the intervention; one study [51] reported no adverse incidents related to the exercise components. Relatively minor musculoskeletal complaints were reported in three studies [30, 52, 53]. Allergic reactions to pedometer clips [30] and minor cardiovascular events [53] were also reported. One author [52] noted that half of the participants in a walking group who developed increases in musculoskeletal complaints withdrew from the study. A fall resulting in a hip fracture sustained during a session was reported in one study [57] and three withdrawals due to increasing back pain were reported [42].

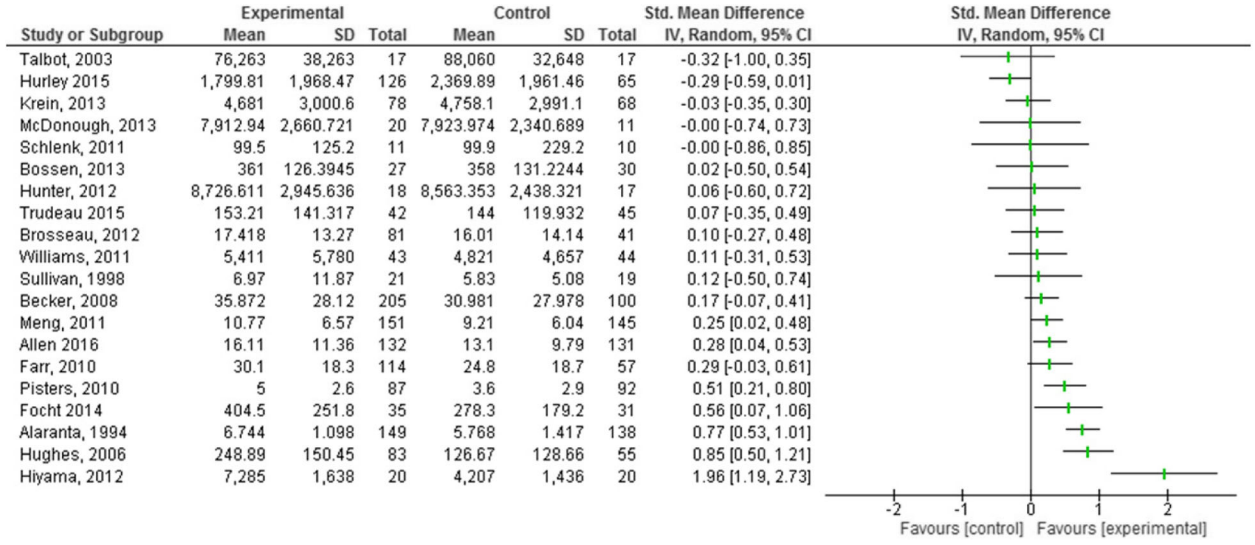


Fig. 6 Forest plot: Studies grouped by effect size (aggregated subjective and objective measures) post intervention

Discussion

Summary of findings

This is the first systematic review and meta-analysis examining the effectiveness of interventions in improving subjective and/or objective levels of PA in adults with PMP and possible associations between BCTs and other intervention characteristics on effect sizes.

In builds on the findings of two similar reviews; Williamson et al. [60] who assessed the effectiveness of behavioural PA interventions in participants with lower-limb osteoarthritis, and Oliveira et al. [33] who assessed the effectiveness of interventions in increasing objectively measured PA in chronic musculoskeletal pain. In contrast to the latter study this review makes a clear distinction between therapeutic exercise programmes and interventions specifically aimed at increasing PA levels or 'habitual PA behaviours'.

With respect to subjective PA, interventions were ineffective in the short term (up to 12 weeks, very low quality evidence); or had a small effect medium term (3–6 months: SMD 0.25, 95%CI 0.01 to 0.48, low quality evidence) and long term (SMD 0.21 95% CI 0.08 to 0.33, moderate evidence). Given the quality of the evidence further research is likely or very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Analysis of the evidence for objective outcomes showed that interventions were not effective at any time point. These observations were based on very low to low quality evidence therefore the estimate of effect is very uncertain and further research is very likely to change the estimate.

Subgroup analyses indicated that interventions were more effective in improving PA levels in adults with osteoarthritis compared to those with persistent low back pain. Intervention effects were also consistently higher in interventions with a greater number of contact hours (> 8.3 h). These subgroup analyses should be interpreted with caution; as differences may not relate to their classifications. However, subgrouping participants by condition was clinically plausible and intervention intensity has previously been associated with effectiveness.

Comparison of subjective outcomes with published literature

Two reviews examining long term outcomes of PA interventions: a Cochrane review of face-to-face interventions to promote PA [61] and a systematic review of PA interventions for adults aged 55–70 years [62]: both reported significant, but very small effects (SMD 0.19) at 12 months. Similarly, this review found small effects for outcomes measured beyond six months (SMD 0.21 95% CI 0.08, 0.33). These findings may indicate that individuals with pain respond to PA interventions in a similar manner to non-pain populations.

In a subgroup analysis Williamson et al. [60] found intervention effects were greatest between 6 and 12 months (SMD 0.53, 95% CI 0.41 to 0.65) and that the effectiveness of interventions declined over time, reporting no significant benefit compared to controls in outcomes beyond 12 months. Similarly, in our osteoarthritis sub-group we found a moderate effect size for medium term outcomes (>3 months ≤6 months) (SMD 0.41, 95% CI 0.10, 0.72) that diminished over time (>6 months) (SMD 0.29, 95% CI 0.08, 0.49). These findings may suggest that individuals with osteoarthritis make changes to their PA levels gradually. However, without ongoing support or increased efforts directed towards maintenance of PA, individuals with osteoarthritis may struggle to sustain increased levels of PA.

Comparison of objective outcomes with published literature

In line with our own findings of no detectable effect on objectively measured PA, Oliveira et al. [33] also found no effect on short, intermediate or long term objective outcomes. Williamson et al. [60] were unable to conduct a meta-analysis using objective measures due to a lack of studies reporting objective measures. In contrast to our findings, the review of interventions aimed at increasing PA in adults aged 55 to 70 years, found larger effects for objective measures (steps per day) (SMD 1.08; 95%CI 0.16, 1.99) at 12 months [62]. A possible explanation for this difference could be that the participants included in this review by Hobbs et al. [62] were essentially 'healthy populations' in contrast, our review and that of Williamson et al. [60] and Oliveira et al. [33] all involved participants with PMP.

Intervention characteristics

We found interventions with a higher number of contact hours resulted in greater effect sizes. Similarly in a post hoc meta-regression, Williamson et al. [60] also found, that a higher number of contact hours had a significant influence on intervention effectiveness. In contrast Hobbs et al. [62] found less intensive interventions were more effective than higher intensity interventions. A plausible explanation for these contrasting findings, is that those with PMP may need additional interventional support, in order to successfully change their PA behaviours in comparison to healthy populations.

In this review the influence of BCTs on PA outcomes is unclear but the findings are consistent with those of previous reviews. Bishop et al. [63] published a review and meta-analysis exploring the effects of contextual and BCT content of control and target interventions in 42 trials included in a Cochrane review of interventions to improve adherence to exercise for chronic musculoskeletal pain [64]. In keeping with the findings from our review, among the most frequently coded BCT's were

'instruction on how to perform the behaviour' and 'behavioural practice and rehearsal'. A finding also reported by Keogh et al. [65] who reviewed BCTs utilised in chronic low back pain self-management programmes. We found 'self-monitoring of the behaviour' was amongst the most frequently coded techniques in interventions with greater effect sizes, a finding not replicated in the either the Bishop et al. [63] or the Keogh et al. [65] reviews, but consistent to findings of PA reviews in healthy populations [24], older adults [66], and in obese adults [28]. As our review was more narrowly focused on habitual PA as opposed to adherence to exercise or self-management, this finding (although tentative) lends some support to the evidence that this technique may be particularly useful in PA interventions.

Interventions included in this review were generally multifaceted often involving several modes of delivery with varying degrees of complexity. It was difficult to draw firm conclusions regarding which characteristics of interventions are associated with more effective interventions.

Few studies provided explicit statements regarding adverse incidents; where they were reported they were largely limited to minor musculoskeletal complaints. Although risk of adverse incidents in PA interventions is generally regarded as low; it is plausible that exacerbating pain may have a deleterious effect on participation, particularly in those with PMP.

Completeness and quality of the evidence

The quality of the evidence within this review ranged from moderate to very low across the different timepoints and outcomes. Effect sizes at best are small and limited to subjective measures. Key limiting factors leading to downgrading the quality of the evidence were, ROB, statistical heterogeneity in the observed effects and imprecision as evidenced by wide confidence intervals. With respect to ROB many studies were designed to identify changes in pain and function/disability as their primary outcomes and were thus underpowered to detect changes in physical activity levels; as such the results of this review should be interpreted with caution. Furthermore, a number of studies failed to provide adequate detail regarding blinding of outcome assessors and allocation concealment. In cluster randomised controlled trials it was often unclear if authors had considered the effect of trial design and the influence clustering may have had on results and whether this was considered when comparing effects with other trials.

Whilst the use of validated measures of PA, was in itself a strength, a more standardised approach to reporting PA data would have permitted a more robust statistical analysis, strengthening the evidence. Self-report measures are known to be prone to recall bias: it has been suggested that as both the intervention and control groups complete the measure any misclassification should be non-differential [67]. However, it could be argued, that using self-report measures in interventions where participants and providers are also unlikely to be blinded the potential of recall bias is increased. Only three studies included subjective and objective measures; this approach might be considered ideal given the relative strengths and limitations of each.

Descriptions of intervention content varied greatly impacting on the number of BCTs that could be reliably reported as occurring within an intervention. In this review we only coded BCTs clearly delivered to the participants and directed towards the target behaviour. As reported by others, [24, 63] this approach, although more rigorous, may result in less BCTs being coded than were actually delivered.

The variation noted across the control conditions could have influenced effect-estimates with smaller between group effects associated with comparisons against more active control treatments [68]. However, we did not detect this when reviewing individual effect size comparisons.

Six studies initially assessed as suitable for inclusion did not report means, standard deviations or sample sizes and requests to obtain this data from study authors were unsuccessful; this data could have added to the quality of the evidence in this review.

Study participants were largely recruited from primary or secondary care (General Practitioners, physiotherapy clinics): it is very possible that the effects seen in those recruited via these settings, differ to those accessing for example, specialist pain services.

Potential biases in the review process

Studies were primarily excluded from the review because a suitable measure of PA was not reported. This may reflect a selective reporting bias; however, it is suggested this is more likely to reflect the changing emphasis of healthcare interventions, particularly the drive towards self-management and a public health approach to managing long term conditions. Although databases were searched from inception only two studies included in the review were published prior to 2003 [42, 57].

Conclusions

Implications for practice

Based on the findings of this review it is not possible to conclude which characteristics of interventions are more effective. However, based on observational analysis and in line with findings of previous reviews, integration of behavioural techniques such as; 'self-monitoring of the behaviour', 'instruction on how to perform the behaviour' and 'goal setting (behaviour)' may be indicated. Higher intensity interventions - in terms of the estimated

contact time with the intervention, may be more effective than less intensive interventions.

The emphasis of PA and exercise interventions in PMP has largely been directed at reducing pain and disability. However, these interventions may have little impact on the overall level of PA an individual engages. Targeted behaviour change interventions are likely to be required to address the risk of morbidity and mortality in this population.

Implications for research

Persistent pain, like many other non-communicable diseases is influenced by several determinants of health such as; socioeconomic status, education, employment and mental health [69]. There is a need for future studies to adopt methods to encourage and secure participation from individuals representing the broad spectrum of persistent pain patients. In particular, those accessing specialist pain services were under represented in this review. Individuals accessing specialist pain services are often deemed to be on the more severe end of the pain spectrum and typically report much higher levels of disability and poorer health related quality of life scores [2]. We agree with previous suggestions [70] that health inequalities may actually be increased because of differences in responses to recruitment. A clear finding from this review is the need to standardise the measurement of PA in PMP populations.

To improve the quality of evidence, future studies should be sufficiently powered, collect longer term follow up data and report on cost-effectiveness. Study authors should report methods for blinding outcome assessors clearly. Providing access to supplementary data such may improve the quality of coding and reporting of intervention content. Future reviews should consider incorporating meta-regression or moderator analysis to explore if specific components or characteristics of interventions are associated with more effective interventions.

Differences between published protocol and review

The review authors reappraised the decision to include unpublished studies and included only those that had been published.

Study authors were amended: SH was added to review team and coded intervention content. LA was added to the review team and provided expert input on aspects relating to coding of BCTs.

The review team agreed to limit the extraction of secondary outcomes to adverse incidents relating to the intervention. There were two main reasons; firstly, to maintain the focus and specificity of the review. Secondly a number of systematic reviews have recently been published describing many of the secondary measures; pain, disability and function, it was felt that extracting these outcomes would be of little additional value to readers of the review.

ROB: The validity of the PA outcome measure is not added as an additional domain within the ROB. This data was included in the data extraction forms and is discussed in relation to outcome measures. An additional domain of sample size calculation for the primary outcome (not specifically for PA) was added to the ROB table and a priori agreements were made during piloting of the ROB table with regards to agreed cut-offs for attrition bias.

The GRADE approach was adopted post-protocol to rate the quality of evidence generated within the review process.

Additional files

Additional file 1: Search Strategy for Medline. (DOCX 13 kb)

Abbreviations

BCT: Behaviour change technique; GRADE: Grading of recommendations, assessment, development and evaluations; PA: Physical Activity; PABAK: Prevalence-adjusted bias-adjusted Kappa; PMP: Persistent musculoskeletal pain; RCT: Randomised controlled trials; ROB: Risk of bias; SMD: Standardised mean difference

Acknowledgements

The authors would like to express their gratitude to Miss Mary Rose Holman, subject librarian at Ulster University, for her help in developing the search strategies and Dr. Ian Bradbury, Frontier Science, Scotland for comments and advice on statistical analysis and assessing risk of bias in cluster randomized trials.

Funding

This research was funded by the Public Health Agency, HSC R&D Division through a Doctoral Fellowship awarded to JM.
SMcD and MAT are co-funded by the UKCRC Centre of Excellence for Public Health (Northern Ireland), a UKCRC Public Health Research Centre of Excellence. Funding from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, Research and Development Office for the Northern Ireland Health and Social Services, and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

Availability of data and materials

The datasets used and/or analysed during this study are available from the corresponding author on reasonable request.

Authors' contributions

JM: conception and design of study, developed initial search strategy, collected background data, completed title and abstract screening, ROB assessments and statistical analysis, prepared first draft of manuscript. SMcD: conception and design of study, refinement of search strategy, reviewing drafts, inputting on methodology, title and abstract screening, ROB assessments, and intellectual content. MAT: refining search strategy, title and abstract screening, BCT coding advice, critical revisions, reviewing methodology and intellectual content. BB: provided input on methodology and statistical analysis. APA: refining search strategy and inputting on methodology. SH: completed dual coding of BCTs. LA: completed critical reviews and inputted on aspects of the review pertaining to BCTs. JO'H: inputted on the design of the study, provided content expertise from a clinical perspective. All authors critically reviewed the manuscript and approved the final version submitted for publication.

Ethics approval and consent to participate

Not Applicable.

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests, two authors of this review are authors of studies included in the review (SMcD and MAT).

Publisher's Note

Author details
[1]Centre for Health and Rehabilitation Technologies, Institute of Nursing and Health, School of Health Sciences, Ulster University, Shore Road, Newtownabbey, Co Antrim BT37 0QB, UK. [2]Centre for Public Health, Queens University Belfast, Royal Victoria Hospital, Grosvenor Road, Belfast BT12 6BA, UK. [3]UKCRC Centre of Excellence for Public Health (Northern Ireland), Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queens University Belfast Room 02020, Institute of Clinical Science B, Royal Victoria Hospital, Grosvenor Road, Belfast, BT 12 6BJ, UK. [4]Belfast Health and Social Care Trust, Chronic Pain Service, Belfast City Hospital, 51 Lisburn Road, Belfast BT9 7AB, UK. [5]Centre for Behaviour Change, University College London, 1-9 Torrington Place, London, UK. [6]Honorary Research Professor, School of Physiotherapy, University of Otago, Dunedin, New Zealand.

Received: 1 June 2017 Accepted: 13 November 2017
Published online: 22 November 2017

References

1. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain. 2006;10(4):287–333.
2. EFIC. Pain proposal improving the current and future management of chronic pain. A European Consensus Report, 2010.
3. IHME. The global burden of disease: generating evidence, guiding policy. European Union and European free trade association regional edition. Seattle: IHME, 2013: Institute For Health Metrics And Evaluation University of Washington; 2013.
4. IHME. The Global Burden Of Disease: Generating Evidence, Guiding Policy. European Union And European Free Trade Association Regional Edition. Seattle, WA: IHME, 2013: Institute For Health Metrics And Evaluation University of Washington.
5. Andersson HI. Increased mortality among individuals with chronic widespread pain relates to lifestyle factors: a prospective population-based study. Disabil Rehabil. 2009;31(24):1980–7.
6. Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. Ann Rheum Dis. 2004;63(4):408–14.
7. Nuesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Juni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. BMJ. 2011;342:d1165.
8. McBeth J, Symmons DP, Silman AJ, Allison T, Webb R, Brammah T, Macfarlane GJ. Musculoskeletal pain is associated with a long-term increased risk of cancer and cardiovascular-related mortality. Rheumatology. 2009;48(1):74–7.
9. Macfarlane GJ, McBeth J, Silman AJ. Widespread body pain and mortality: prospective population based study. BMJ. 2001;323(7314):662–5.
10. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G, Pain CBWGoGfCLB. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J. 2006;15(Suppl 2):S192–300.
11. BPS. Core standards for pain management services in the UK. CSPMS. UK: British Pain Society; 2015.
12. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, Doherty M, Geenen R, Hammond A, Kjeken I, Lohmander LS, Lund H, Mallen CD, Nava T, Oliver S, Pavelka K, Pitsillidou I, da Silva JA, de la Torre J, Zanoli G, Vliet Vlieland TP. European league against R. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis. 2013;72(7):1125–35.
13. NICE. Low back pain: early management of persistent non-specific low back pain. CG88. London: National Institute for Health and Care Excellence; 2009.
14. NICE. Osteoarthritis care and management in adults. CG177. London: National Institute for Health and Care Excellence; 2014.
15. NICE. Non-specific low back pain and sciatica: management. CG [in development]. London: National Institute for Health and Care Exellence; 2016.
16. SIGN. Management of Chronic Pain. Scottish Intercollegiate Guidelines Network: Scotland; 2013.
17. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M, Hochberg M, Hunter DJ, Kwoh K, Lohmander LS, Tugwell P. OARSI recommendations for the management of hip and knee osteoarthritis: part III: changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthr Cartil. 2010;18(4):476–99.
18. Warburton DE, Bredin SS. Reflections on physical activity and health: what should we recommend? Can J Cardiol. 2016;32(4):495–504.
19. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ. 2006;174(6):801–9.
20. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 1985;100(2):126–31.
21. WHO. Physical activity factsheet. Online: World Health Organisation, 2016.
22. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci. 2011;6:42.
23. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol. 2008;27(3):379–87.
24. Bird EL, Baker G, Mutrie N, Ogilvie D, Sahlqvist S, Powell J. Behavior change techniques used to promote walking and cycling: a systematic review. Health Psychol. 2013;32(8):829–38.
25. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychol. 2011;30(1):99–109.
26. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, Schwarz P, Group IS. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health. 2011;11:119.
27. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol. 2009;28(6):690–701.
28. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act. 2013;10:29.
29. NICE. Behaviour change: individual approaches. PH49. London: National Institute for Health and Care Excellence; 2014.
30. McDonough SM, Tully MA, Boyd A, O'Connor SR, Kerr DP, O'Neill SM, Delitto A, Bradbury I, Tudor-Locke C, Baxter GD, Hurley DA. Pedometer-driven walking for chronic low back pain: a feasibility randomized controlled trial. Clin J Pain. 2013;29(11):972–81.
31. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. J Am Geriatr Soc. 2003;51(3):387–92.
32. Williams NH, Amoakwa E, Belcher J, Edwards RT, Hassani H, Hendry M, Burton K, Lewis R, Hood K, Jones J, Bennett P, Linck P, Neal RD, Wilkinson C. Activity increase despite arthritis (AIDA): phase II randomised controlled trial of an active management booklet for hip and knee osteoarthritis in primary care. Br J Gen Pract. 2011;61(589):e452–8.
33. Oliveira CB, Franco MR, Maher CG, Lin CC, Morelhao PK, Araujo AC, Negrao Filho RF, Pinto RZ. Physical activity interventions for increasing objectively measured physical activity levels in chronic musculoskeletal pain: systematic review. Arthritis Care Res. 2016;
34. Marley J, Tully MA, Porter-Armstrong A, Bunting B, O'Hanlon J, McDonough SM. A systematic review of interventions aimed at increasing physical activity in adults with chronic musculoskeletal pain–protocol. Syst Rev. 2014;3:106.
35. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, Eccles MP, Cane J, Wood CE. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med. 2013;46(1):81–95.
36. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46(5):423–9.

37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.
38. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated march 2011]: the Cochrane collaboration, 2011. .
39. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, Treweek S, Mustafa RA, Rada G, Rosenbaum S, Morelli A, Guyatt GH, Oxman AD. GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: introduction. BMJ. 2016;353. doi:10.1136/bmj.i2016.
40. Oxman AD. Grading quality of evidence and strength of recommendations: GRADE working group. BMJ. 2004;328(7454):1490.
41. Cohen J. Statistical power analysis or the behavioural sciences: Lawrence Earlbaum associates, 1988.
42. Alaranta H, Rytokoski U, Rissanen A, Talo S, Ronnemaa T, Puukka P, Karppi S, Videman T, Kallio V, Slatis P. Intensive physical and psychosocial training program for patients with chronic low back pain: a controlled clinical trial. Spine. 1994;19(12):1339–49.
43. Allen KD, Yancy WS Jr, Bosworth HB, Coffman CJ, Jeffreys AS, Datta SK, McDuffie J, Strauss JL, Oddone EZ. A combined patient and provider intervention for Management of Osteoarthritis in veterans: a randomized clinical trial. Ann Intern Med. 2016;164(2):73–83.
44. Becker A, Leonhardt C, Kochen M, Keller S, Wegscheider S, Baum E, Donner-Banzhoff N, Pfingsten M, Hildebrandt J, Heinz-Dieter BW, Chenot J. Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial. Spine. 2008;33(5):473–80.
45. Bossen D, Veenhof C, Van Beek KE, Spreeuwenberg PM, Dekker J, De Bakker DH. Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. J Med Internet Res. 2013;15(11):e257.
46. Brosseau L, Wells GA, Kenny GP, Reid R, Maetzel A, Tugwell P, Huijbregts M, McCullough C, Angelis G, Chen L. The implementation of a community-based aerobic walking program for mild to moderate knee osteoarthritis: a knowledge translation randomized controlled trial: part II: clinical outcomes. BMC Public Health. 2012;12:1073.
47. Farr JN, Going SB, McKnight PE, Kasle S, Cussler EC, Cornett M. Progressive resistance training improves overall physical activity levels in patients with early osteoarthritis of the knee: a randomized controlled trial. Phys Ther. 2010;90(3):356–66.
48. Focht BC, Garver MJ, Devor ST, Dials J, Lucas AR, Emery CF, Hackshaw KV, Rejeski WJ. Group-mediated physical activity promotion and mobility in sedentary patients with knee osteoarthritis: results from the IMPACT-pilot trial. J Rheumatol. 2014;41(10):2068–77.
49. Hiyama Y, Yamada M, Kitagawa A, Tei N, Okada S. A four-week walking exercise programme in patients with knee osteoarthritis improves the ability of dual-task performance: a randomized controlled trial. Clin Rehabil. 2012;26(5):403–12.
50. Hughes SL, Seymour RB, Campbell RT, Huber G, Pollak N, Sharma L, Desai P. Long-term impact of fit and strong! On older adults with osteoarthritis. Gerontologist. 2006;46(6):801–14.
51. Hunter R, McDonough S, Bradbury I, Liddle S, Walsh D, Dhamija S, Glasgow P, Gormley G, McCann S, Park J. Exercise and auricular acupuncture for chronic low-back pain: a feasibility randomized-controlled trial. Clin J Pain. 2012;28(3):259–67.
52. Hurley DA, Tully MA, Lonsdale C, Boreham CA, van Mechelen W, Daly L, Tynan A, McDonough SM. Supervised walking in comparison with fitness training for chronic back pain in physiotherapy: results of the SWIFT single-blinded randomized controlled trial (ISRCTN17592092). Pain. 2015;156(1):131–47.
53. Krein SL, Kadri R, Hughes M, Kerr EA, Piette JD, Holleman R, Kim HM, Richardson CR. Pedometer-based internet-mediated intervention for adults with chronic low back pain: randomized controlled trial. J Med Internet Res. 2013;15(8):e181.
54. Meng K, Seekatz B, Roband H, Worringen U, Vogel H, Faller H. Intermediate and long-term effects of a standardized back school for inpatient orthopedic rehabilitation on illness knowledge and self-management behaviors: a randomized controlled trial. Clin J Pain. 2011;27(3):248–57.
55. Pisters M, Veenhof C, de BD SF, Dekker J. Behavioural graded activity results in better exercise adherence and more physical activity than usual care in people with osteoarthritis: a cluster-randomised trial. Aus J Physiother. 2010;56(1):41–7.
56. Schlenk E, Lias J, Sereika S, Dunbar-Jacob J, Kwoh C. Improving physical activity and function in overweight and obese older adults with osteoarthritis of the knee: a feasibility study. Rehabilitation Nursing. 2011;36(1):32–42.
57. Sullivan T, Allegrante JP, Peterson MG, Kovar PA, MacKenzie CR. One-year followup of patients with osteoarthritis of the knee who participated in a program of supervised fitness walking and supportive patient education. Arthritis Care Res. 1998;11(4):228–33.
58. Trudeau KJ, Pujol LA, DasMahapatra P, Wall R, Black RA, Zacharoff K. A randomized controlled trial of an online self-management program for adults with arthritis pain. J Behav Med. 2015;38(3):483–96.
59. Kovar PA, Allegrante JP, MacKenzie CR, Peterson MG, Gutin B, Charlson ME. Supervised fitness walking in patients with osteoarthritis of the knee. A randomized, controlled trial. Ann Intern Med. 1992;116(7):529–34.
60. Williamson W, Kluzek S, Roberts N, Richards J, Arden N, Leeson P, Newton J, Foster C. Behavioural physical activity interventions in participants with lower-limb osteoarthritis: a systematic review with meta-analysis. BMJ Open. 2015;5(8). doi:10.1136/bmjopen-2015-007642.
61. Richards J, Hillsdon M, Thorogood M, Foster C. Face-to-face interventions for promoting physical activity. Cochrane Database Syst Rev. 2013;9: CD010392.
62. Hobbs N, Godfrey A, Lara J, Errington L, Meyer TD, Rochester L, White M, Mathers JC, Sniehotta FF. Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. BMC Med. 2013;11:75.
63. Bishop FL, Fenge-Davies AL, Kirby S, Geraghty AW. Context effects and behaviour change techniques in randomised trials: a systematic review using the example of trials to increase adherence to physical activity in musculoskeletal pain. Psychol Health. 2015;30(1):104–21.
64. Jordan JL, Holden MA, Mason EE, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. Cochrane Database Syst Rev. 2010;(1):Cd005956.
65. Keogh A, Tully MA, Matthews J, Hurley DA. A review of behaviour change theories and techniques used in group based self-management programmes for chronic low back pain and arthritis. Man Ther. 2015. doi:10.1016/j.math.2015.03.014.
66. French DP, Olander EK, Chisholm A, Mc SJ. Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. Ann Behav Med. 2014;48(2):225–34.
67. Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev. 2005;1:Cd003180.
68. Karlsson P, Bergmark A. Compared with what? An analysis of control-group types in Cochrane and Campbell reviews of psychosocial treatment efficacy with substance use disorders. Addiction. 2015;110(3):420–8.
69. Goldberg DS, McGee SJ. Pain as a global public health priority. BMC Public Health. 2011;11:770.
70. Foster CE, Brennan G, Matthews A, McAdam C, Fitzsimons C, Mutrie N. Recruiting participants to walking intervention studies: a systematic review. Int J Behav Nutr Phys Act. 2011;8:137.

Casey et al. Trials (2018) 19:194
https://doi.org/10.1186/s13063-018-2543-5

Trials

STUDY PROTOCOL Open Access

Exercise combined with Acceptance and Commitment Therapy (ExACT) compared to a supervised exercise programme for adults with chronic pain: study protocol for a randomised controlled trial

Máire-Bríd Casey[1*], Keith Smart[2], Ricardo Segurado[1], Conor Hearty[3], Hari Gopal[3], Damien Lowry[4], Dearbhail Flanagan[5], Lance McCracken[6] and Catherine Doody[1]

Abstract

Background: Acceptance and Commitment Therapy (ACT) is a form of cognitive behavioural therapy, which may be beneficial for people with chronic pain. The approach aims to enhance daily functioning through increased psychological flexibility. Whilst the therapeutic model behind ACT appears well suited to chronic pain, there is a need for further research to test its effectiveness in clinical practice, particularly with regards to combining ACT with physical exercise.

Methods/design: This prospective, two-armed, parallel-group, single-centre randomised controlled trial (RCT) will assess the effectiveness of a combined Exercise and ACT programme, in comparison to supervised exercise for chronic pain. One hundred and sixty patients, aged 18 years and over, who have been diagnosed with a chronic pain condition by a physician will be recruited to the trial. Participants will be individually randomised to one of two 8-week, group interventions. The combined group will take part in weekly psychology sessions based on the ACT approach, in addition to supervised exercise classes led by a physiotherapist. The control group will attend weekly supervised exercise classes but will not take part in an ACT programme. The primary outcome will be pain interference at 12-week follow-up, measured using the Brief Pain Inventory-Interference Scale. Secondary outcomes will include self-reported pain severity, self-perception of change, patient satisfaction, quality of life, depression, anxiety and healthcare utilisation. Treatment process measures will include self-efficacy, pain catastrophising, fear avoidance, pain acceptance and committed action. Physical activity will be measured using Fitbit Zip™ activity trackers. Both groups will be followed up post intervention and again after 12 weeks. Estimates of treatment effects at follow-up will be based on an intention-to-treat framework, implemented using a linear mixed-effects model. Individual and focus group qualitative interviews will be undertaken with a purposeful sample of participants to explore patient experiences of both treatments.

Discussion: To our knowledge, this will be the first RCT to examine whether combining exercise with ACT produces greater benefit for patients with chronic pain, compared to a standalone supervised exercise programme.

Trial registration: www.ClinicalTrials.gov, ID: NCT03050528. Registered on 13 February 2017.

Keywords: Chronic pain, Exercise, Acceptance and commitment therapy, Multidisciplinary pain programme, Physical therapy, Physiotherapy, Psychological therapy

* Correspondence: maire-brid.casey@ucdconnect.ie
[1]School of Public Health, Physiotherapy and Sports Science, Health Sciences Building, University College Dublin, Belfield, Dublin 4, Ireland
Full list of author information is available at the end of the article

Background

Chronic pain is a major health problem, reported to affect 19% of adult Europeans [1] and up to 35.5% of Irish adults [2]. The economic burden is significant, with a recent survey in Ireland estimating the total cost of treating chronic pain at €5.34 billion per year [3]. This survey of 1204 Irish people also reported that health-related quality of life was significantly lower in people with chronic pain compared to people without pain, and that depression was significantly higher [2].

Chronic pain has been defined as an unpleasant sensory or emotional experience, associated with actual or potential tissue damage, which persists for over 3 months' duration [4]. The multidimensional nature of chronic pain presents significant challenges for patients and healthcare professionals. There is a plethora of treatment options available but the effects of these interventions on pain and disability are modest and improvements are typically short term [5, 6]. Traditional biomedical interventions, such as surgery and spinal injections, have not been shown to be superior to conservative treatments for chronic pain and they carry greater risks [7, 8]. Exercise interventions and psychological therapies, such as cognitive behavioural therapy (CBT), are examples of conservative treatments that are known to be effective for patients with chronic pain [5, 6, 9–12]. These interventions can be provided individually or they can be effectively combined in the form of a multidisciplinary biopsychosocial rehabilitation programme [13].

Physical activity is an important outcome to target with chronic pain interventions, as in addition to the physical limitations imposed by pain, there are strong associations with cardio-metabolic and respiratory conditions [14]. Exercise, including aerobic, strengthening and aquatic exercise has been shown to reduce pain and improve physical function and quality of life [15–17], but the quality of the evidence is low and further studies with larger samples are required [12]. No particular type of exercise has been shown to be superior to another [5, 18] and research suggests that group-based physiotherapy interventions incorporating exercise are just as effective for pain and disability as individual treatment [19]. Patient adherence to treatment should be promoted by providing individualised exercises within supervised programmes, and supplementing with home exercises [20].

There is a large evidence base related to psychological treatments for chronic pain, with CBT being the dominant intervention. A Cochrane review concluded that CBT has small to moderate effects on pain, disability, mood and catastrophising [6]. The authors noted that whilst there have been improvements in the methodological quality of studies in recent years, there has been no change in the overall effects of the interventions and they recommend against further randomised controlled trials (RCTs) examining the efficacy of CBT.

Acceptance and Commitment Therapy (ACT) is a psychological therapy that encourages participants to change their relationship with their thoughts and physical sensations through mechanisms of acceptance, mindfulness and value-based action [21]. Systematic reviews of RCTs featuring ACT for adults with chronic pain have reported that ACT is effective for enhancing general function and decreasing distress, compared to inactive treatment comparisons [22, 23]. One RCT included in these reviews compared ACT with CBT and found no significant differences in improvement between the two treatments; however, greater levels of satisfaction were reported by the ACT participants [24]. Another RCT reported equivalent reductions in pain and disability with ACT, when compared with applied relaxation for chronic pain [25]. Whilst there is growing evidence to support the effectiveness of ACT, it has been acknowledged that there are currently only a small number of high-quality studies and further RCTs have been recommended, in particular with active treatment comparisons [9, 22].

There are currently no RCTs that have examined the effectiveness of exercise combined with ACT for chronic pain. Furthermore, in the RCTs published to date, ACT as a standalone therapy has not been shown to be effective in enhancing physical activity [26, 27]. When CBT and ACT were compared for chronic pain, no significant improvement in physical activity was observed for either psychological approach [27] and the authors suggest that tailored interventions, with greater emphasis on exercise, may complement psychological treatment for chronic pain. To our knowledge, this will be the first RCT to assess the effectiveness of a combined Exercise and ACT intervention for chronic pain.

Research objectives

The primary objective is to determine whether a combined Exercise and ACT group-based intervention is effective for reducing pain interference at 12-week follow-up, in patients with chronic pain, compared to a physiotherapy-led supervised exercise programme.

Secondary objectives

1. To investigate whether exercise combined with ACT has a positive impact on study participants compared to a supervised exercise programme, with regard to the self-reported secondary outcomes: pain severity, self-perception of change, patient satisfaction, quality of life, depression, anxiety and healthcare costs, and treatment process outcomes: self-efficacy, pain catastrophising, fear avoidance, pain acceptance and

committed action following treatment and at 12-week follow-up
2. To examine whether exercise combined with ACT has a significant effect on objective physical activity measures (step count, distance travelled and active minutes) post treatment, compared to a supervised exercise programme
3. To explore the experiences of a purposeful sample of participants of both interventions with embedded qualitative interviews

Methods/design

Study design

The ExACT trial is a two-armed, single-centre, parallel-group, randomised controlled superiority trial.

Setting

Participants will be recruited from a consultant-led pain clinic and musculoskeletal out-patient clinics within a secondary care setting of a large academic teaching hospital in Dublin, Ireland. Treatments will take place in the pain clinic and in the physiotherapy department of the hospital.

Participants

A total of 160 participants will be randomised to the combined Exercise and ACT or supervised exercise groups over a 20-month period. Adults (aged 18 years and older) with any type of chronic pain condition, other than cancer pain, (diagnosed by a physician), which is persisting for over 12 weeks' duration and who report a score of ≥ 2 on the Brief Pain Inventory-Interference Scale (BPIIS) are eligible for inclusion in the study. Participants must also be able to provide informed consent and communicate effectively in the English language. Figure 1 shows the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) Diagram for the trial.

Exclusion criteria are as follows: need for further diagnostic evaluation (determined by a physician), presence of major medical or psychiatric disorder that would impede ability to participate with treatment), presence of active cancer or cancer-related pain, unstable inflammatory condition (e.g. rheumatoid arthritis, gout), presence of substance misuse, surgical or interventional procedure (e.g. spinal cord stimulator, rhizotomy, epidural or intra-articular injection) within the last 3 months, concurrent participation (or participation within the last 3 months) in a supervised exercise programme or a course of psychological or physiotherapy treatment, previous participation in any multidisciplinary pain management programme or presence of any contraindication to participation in a gym or pool-based exercise programme (e.g. shortness of breath at rest, unstable diabetes or epilepsy, recent myocardial infarction, stroke, pulmonary embolism, asthma attack, weight > 125 kg (19.5 stone) or waist circumference > 50 in. (restriction due to hydrotherapy evacuation equipment).

Participant identification, recruitment and consent

Adults, who attend hospital out-patient clinics for treatment of chronic pain will be screened for study eligibility by a physician. The number of patients who undergo screening will be recorded in order to quantify the number of patients who are deemed eligible or ineligible for the study and how many patients decline to participate. The reasons for ineligibility will be recorded. Those who meet the eligibility criteria will be informed about the study by their physician and written information in the form of a Patient Information Leaflet will be provided. Patients who express interest in participating in the study will be contacted by telephone by the lead researcher (MBC). Any questions will be clarified on the telephone and patients who remain interested will be invited to attend an individual face-to-face appointment with the lead researcher in the hospital pain clinic. Baseline outcome measures will be sent in advance by post and the patients will be asked to bring the completed questionnaires with them on the day of their appointment.

Informed consent will be obtained in writing by the lead researcher, prior to participation in the study. Patients will be informed that they are under no obligation to participate and they may withdraw their consent at any time without need for explanation. Where possible the reasons for withdrawal from the trial will be recorded. Patients who do not wish to take part in the study will continue to have treatment as usual. A sample size target of 160 participants over a 20-month period has been set. Recruitment will be monitored throughout the trial and if expected rates of recruitment are not being achieved, additional patients may be recruited via the physiotherapy waiting list or paper triage of referral letters and patient databases, performed by physicians, with subsequent eligibility screening by the lead researcher.

Interventions

The study interventions are described below and are written with reference to the TIDieR guidelines for better reporting of interventions [28].

Combined Exercise and Acceptance and Commitment Therapy (ExACT)

ExACT is a multidisciplinary pain programme combining exercise and psychological therapy. It is a face-to-face, group-based treatment, with up to ten individuals per group. Participants will attend a total of eight sessions, once per week, with each session lasting 3.5 h. Each day will begin with a 2-h psychology session held in the hospital pain clinic. The sessions will follow the psychological

	STUDY PERIOD						
	ENROLMENT	**BASELINE APPOINTMENT**	**ALLOCATION**	**POST ALLOCATION**			
TIMEPOINT		*No more than 4 weeks prior to commencing intervention*		*1 week pre-baseline*	*Baseline*	*Post intervention follow up*	*12-week follow-up*
ENROLMENT:							
Identification	X						
Eligibility screening	X						
Completion of baseline measures		X					
Provision of activity tracker		X					
Informed consent		X					
Randomisation			X				
Allocation			X				
INTERVENTIONS							
ExACT (combined exercise and acceptance and commitment therapy)					←	→	
Standalone supervised exercise					←	→	
DATA COLLECTION							
Demographics		X					
Primary outcome measure		X				X	X
Secondary outcome measures		X				X	X
Treatment process measures		X				X	X
Objective physical activity data				←	—	→	
Treatment attendance					←	→	
Adverse events					←	→	

Fig. 1 Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) Diagram

approach Acceptance and Commitment Therapy (ACT) and are designed to promote psychological flexibility through methods that encourage openness, awareness and engagement. The overall aim is to reduce pain-avoidant behaviours, in the service of living a rich and meaningful life [21]. Group discussions, experiential exercises and mindfulness practice (see Table 1) will be led by a psychologist who has been trained in ACT and is experienced in treating chronic pain. The content of the sessions will be adapted from an ACT treatment manual used in a recently published study [29] and available to members of the Association for Contextual Behavioral Science (ACBS) from the website http://www.contextualscience.org/better_living_with_illness. Written supplementary material will be provided each week and participants will be encouraged to spend time reviewing the material at home.

Following a break for lunch, participants will attend a 1.5-h supervised exercise class in the physiotherapy department of the hospital. The classes will be delivered by

Table 1 Summary of the content of the Acceptance and Commitment Therapy (ACT) component of the combined intervention

Session	Content
1	Introductions and basic foundations of treatment, present the goal of ACT – shifting focus from pursuit of symptom reduction to improving function
2	Review of previous treatment history – creative hopelessness exercise including primary and secondary suffering Introduce openness as a skill area – acceptance as an alternative to avoidance
3	Recap of acceptance and continued focus on enhancing openness. Introduce process of defusion. Passenger on the bus experiential exercise
4	Focus on engagement: values awareness and assessment Experiential values exercise
5	Further values clarification work Committing to action that improves and enriches one's life
6	Focus on awareness – contact with the present moment, perspective taking and self-awareness as distinct from fusion with thought content and perception of self
7	Treatment review Walking mindfulness exercise
8	Wrap up and conclusions Relapses and set-backs: preparation not prevention

a physiotherapist and will feature two components: (1) education/advice and (2) exercise (see Table 2). The education/advice sessions will be interactive and will take the form of group discussions of approximately 30-min duration, covering topics such as goal-setting, understanding pain, physical activity and pacing. The physiotherapist will answer questions and facilitate discussion related to relevant issues that the participants bring to the group. The exercise sessions will take place in either a pool or a gym setting (four sessions of each). The aquatic sessions will include a warm up, gentle aerobic exercise, buoyancy-assisted and resisted movements, and informal ball games. The gym programme will feature a combination of gentle aerobic exercise, stretches and strengthening exercises. The specific exercises will be chosen by the physiotherapist and examples will include, but will not be limited to: cycling on a static bike and treadmill walking, pulleys, sit to stands, step-ups, wall

Table 2 Summary of the content of the supervised exercise component of both interventions

Session	Education/advice (30 min)	Exercise (1 h)
1	Introduction to exercise Pool orientation Induction to gym programme Demonstration of gym exercises	Gym exercise: Gentle warm up – walking and stretches Brief gym circuit Cool down
2	Group discussion on goal setting Provision of individual home exercise programme HEP created by physiotherapist based on patient's individual goals	Hydrotherapy Warm up Gentle aerobic and buoyancy assisted and resisted exercises
3	Understanding pain Group will be shown the YouTube video 'understanding pain in 5 minutes' followed by a group question/answer session	Hydrotherapy Warm up Aerobic, strengthening exercises and informal pool games
4	Group discussion about physical activity, introduction to pacing and principles of graded exposure Time to answer any questions from participants	Hydrotherapy Continuation of above and gentle progression of exercises
5	Group discussion on pacing and graded exposure including potential challenges that may be arising regarding putting principles into practice	Hydrotherapy Continuation of above and gentle progression of exercises
6	Continued group discussion on progress and problem solving	Gym session – participants are free to perform exercises from their individualised exercise programme or other exercises of their own choosing under the guidance of the physiotherapist
7	Group discussion on progress and problem solving. Introduce topic of maintaining behaviour change	Gym session Continuation of above and progression of exercise under guidance of physiotherapist
8	Wrap-up session including preparation for maintaining an active lifestyle and managing setbacks	Gym session Continuation of above and progression of exercise under guidance of physiotherapist

squats, wall press-ups, seated flexion, trunk rotation in standing, knee rolling/trunk rotation on plinth, bridging, knees to chest. The exercise programmes will be individualised, based on each participant's personal goals and a written exercise programme, compiled by the physiotherapist will be provided. The programmes will be progressed and modified for each individual as deemed appropriate by the physiotherapist. Participants will be encouraged to carry out their exercises at home or in their local pool. Throughout the aquatic and gym exercise sessions, there will be an emphasis on reducing threat and fear of movement. The physiotherapist will encourage improved physical activities in a manner that gradually increases physical function and enhances enjoyment of physical activity. The physiotherapist leading the supervised exercise programmes will have over 7 years of experience treating patients with chronic pain, including the delivery of group exercise programmes to patients with similar conditions to the trial participants. The physiotherapist will not have had formal training in ACT, ensuring that only the participants of the combined group will be exposed to this form of psychological therapy.

Standalone supervised exercise

The standalone, supervised exercise programme will also consist of eight face-to-face, 1.5-h sessions delivered by a physiotherapist, on a weekly basis, to groups of up to ten participants. The intervention will mirror the supervised exercise component of the combined treatment as described above.

Treatment adherence and other interventions

Attendance at both interventions will be recorded by the treating clinicians. Participants will be encouraged to inform the administrative staff in the pain clinic by telephone or email if they are unable to attend and, where possible, the reasons for absence will be recorded. Attendance rates will be reported with the trial results.

All study participants will continue to attend routine medical appointments with their general practitioner or hospital consultants for the duration of the trial. These appointments will be recorded and reported. Other than the trial interventions, participants will be asked to refrain from additional treatment provided by allied health practitioners, such as psychologists, counsellors, physiotherapists or complementary therapists, during the 8-week treatment period. Any medication changes and the administration of any additional interventions during the course of the trial will be recorded and reported, and reasons for same will be documented. Patients will not be denied any treatments that a physician deems necessary to administer urgently.

Treatment fidelity

Assessment of treatment fidelity is an important component in ensuring transparency in clinical research and increasing confidence that the intervention is delivered as described [30]. Treatment fidelity of the ACT intervention in this trial will be assessed by a health psychologist (NL), who is highly experienced in the delivery of group psychological interventions for chronic pain using ACT. All eight ACT sessions, from one treatment group will be audio-recorded and sent to NL for review. An ACT treatment fidelity tool that has been modified for chronic pain will be used to evaluate the intervention [31]. The psychologist delivering the ACT intervention will complete written notes, detailing the content covered within each session and a brief note outlining any relevant observations. These will be reviewed alongside the audio-recordings.

The treating physiotherapist will also complete checklists after each exercise session and will record any additional relevant details. Treatment fidelity of the physiotherapy components of the trial will be assessed by another member of the project team (KS), a practising clinical specialist physiotherapist, who will review the checklists and notes at monthly intervals.

Randomisation

Randomisation will take place after baseline measures have been assessed as recommended by the Consolidated Standards of Reporting Trials (CONSORT) guidelines [32]. Randomisation will be coordinated by the trial supervisor (CD), who will have no involvement in eligibility screening, enrolment or treatment processes. On receipt of signed consent forms and baseline measures, participants will be given a unique code and randomised using an online randomisation database [33]. The computer-generated randomisation schedule will apply a permuted block design to ensure that the groups are balanced periodically. The block size will be concealed until after the primary endpoint has been analysed. The randomisation list will remain with the trial supervisor for the full duration of the study. The list will be stored in an encrypted file on a password-protected computer in the trial supervisor's office to ensure concealment of allocation. Allocation of participants will be communicated to administrative staff in the pain clinic by the trial supervisor via email. The administrative staff will store this allocation list in an encrypted file, on a password-protected computer in the pain clinic administrative office. Participants will be informed of their group allocation in writing by the administrative staff, who will send notification in sealed, opaque envelopes.

Ethics

Ethical approval for the study has been granted by the Mater Misericordiae University Hospital Institutional Review Board (Ref No. 1/378/1864) and ethical exemption has been accepted by the UCD Human Research Ethics Committee – Sciences (Ref No. LS-E-17-03-Casey-Doody). The trial will be performed in accordance with the Declaration of Helsinki [34]. No significant adverse events are anticipated during this trial, but will be monitored and any adverse events that occur will be recorded and reported.

This protocol has been written in accordance with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) Statement [35] (See Additional file 1: ExACT Trial SPIRIT 2013 Checklist). Any significant modifications to the protocol will require a formal protocol amendment, agreed on by the project team and approved by the MMUH Institutional Review Board. Minor administrative changes to the protocol will be documented in a memorandum.

Data collection methods

Data will be collected via self-report questionnaires at baseline (t_0), post intervention (t_1) and at 12-week follow-up (t_2). Physical activity patterns will be measured objectively using Fitbit Zip™ (available from: http://www.fitbit.com) activity trackers and recorded from 1 week prior to commencing treatment through to completion of the intervention.

Baseline

A booklet of self-report questionnaires will be sent to participants via post, in advance of a scheduled face-to-face appointment with the lead researcher. This appointment will take place prior to randomisation and no more than 4 weeks prior to starting the intervention. Participants will be requested to complete the questionnaires at home and return them in person on the day of the appointment, when any questions related to the questionnaires can be clarified. The questionnaires will be checked for missing data by the lead researcher and completed by the participant where possible. The questionnaires will later be given to the trial supervisor who will de-identify and code them. No information about treatment allocation will be included on the questionnaires and they will be returned to the lead researcher who will enter the data in a secure, web-enabled information management system. One week of baseline physical activity data (daily step count, active minutes and distance travelled) will be collected prior to starting the interventions. The Fitbit Zip™ activity tracker will be provided to participants at the baseline appointment, with a request to begin wearing the device at least 1 week prior to commencement of the intervention and to continue wearing it for the full duration of the 8-week intervention. An individual Fitbit account will be created for each participant using a unique email address, purposefully created for the study. Data for each participant will be retrieved remotely by the lead researcher via the Fitbit website.

Post intervention follow-up

The treating physiotherapist will provide follow-up questionnaires to all trial participants on the last day of the interventions. The participants will be provided with a private space to complete the questionnaires at the end of the last day and the physiotherapist will not be present with the participants whilst they fill in the questionnaires. The questionnaires will be placed in an opaque envelope and sealed before sending to the trial supervisor by registered post. The trial supervisor will check the questionnaires for missing data and will follow-up by post or with telephone calls to participants where possible. The questionnaires will be de-identified as outlined, before returning to the lead researcher who will enter the data into the electronic database.

Twelve-week follow-up

The 12-week follow-up time point has been selected as the primary endpoint for establishing effectiveness of the trial intervention. Participants will be contacted via telephone by the administrative staff in the pain clinic and questionnaires will be administered via post with an enclosed stamped addressed envelope for return to the trial supervisor (CD). The trial supervisor will follow up on any missing data again by telephone and a follow-up letter will be sent to participants who fail to return the questionnaires within 2 weeks. A final reminder telephone call will be made by the administrative staff in the pain clinic to participants who have not returned questionnaires after a further week.

Blinding

Blinding of patients or the treating health professionals will not be possible due to the nature of the interventions. However, a position of clinical equipoise will be maintained, with patients advised verbally and in the Patient Information Leaflet, that they are being offered one of two treatments that are believed to be helpful for chronic pain but it is not known if one treatment is superior to the other. The lead researcher (MBC) will be blinded to group allocation when entering and analysing the data and the statistician (RS) analysing data will also be blinded. Methods to ensure maintenance of this blinding will include de-identification and coding of questionnaires by the trial supervisor (CD), who will be the only person to have access to the locked codes used for treatment allocation. The trial supervisor will also be

responsible for randomisation and follow-up of missing data post intervention and at 12-week follow-up. Day-to-day communication and management of scheduling after randomisation will be coordinated by the administrative staff in the hospital pain clinic. Un-blinding of trial participants will occur only after creation of a final locked analysis dataset when the last patient has provided data at 12-week follow-up.

Outcomes

The outcome measures included in this trial are based on the IMMPACT recommendations for outcome measures for use in chronic pain clinical trials [36]. A recent systematic review [22] recommended formally defining outcome measures as primary, secondary and treatment process measures in future RCTs featuring ACT.

Demographic data

Data collected will include age, gender, education level, relationship status and work status. Details regarding pain history will be collected including diagnosis (if applicable), and duration of pain.

Primary outcome Pain interference at 12-week follow-up has been chosen as the primary outcome based on the IMMPACT recommendations and also a systematic review of ACT for chronic pain, which suggests using a measure of physical or social functioning, rather than pain or emotional functioning as a primary outcome [22]. The 12-week follow-up time point has been specified as the primary outcome as this has been suggested to present a low risk of bias in chronic pain trials [37]. Pain interference will be assessed with the Brief Pain Inventory-Interference Scale (BPIIS). This is a seven-item self-report questionnaire that measures the extent to which pain interferes with functions such as general activity, walking ability, normal work, mood, relations with people, enjoyment of life and sleep. The Brief Pain Inventory (BPI) has been shown to be a valid tool for assessing pain interference, with acceptable internal consistency [38]. Excellent test-retest reliability has been reported in a chronic pain cohort [39] and a reduction of 1 point on the interference scale has been recommended as a clinically meaningful change [40].

Secondary outcome measures We hypothesise that participation in the interventions will influence many health dimensions and the following secondary outcomes will be assessed:

Pain intensity: Brief Pain Inventory (BPI) – Pain severity subscale

Pain intensity will be measured with the pain severity subscale of the BPI. Reductions of pain intensity of between 10 and 20% have been reported to represent a clinically meaningful change [40].

Patient Global Impression of Change (PGIC) scale

The PGIC scale will measure participants' perceived level of improvement or lack thereof, due to the intervention. The PGIC has strong clinical relevance to the individual with good face and test-retest reliability [41]. The percentages of participants endorsing each of the responses will be reported as per the IMMPACT recommendations [40].

Patient satisfaction with treatment

This will be measured using a single question (question 7) from the Client Satisfaction Questionnaire-8 (CSQ-8), which is designed to measure client satisfaction with services [42]. The question will ask 'In an overall, general sense, how satisfied are you with the service you have received?' and four potential responses will be provided (very satisfied, mostly satisfied, indifferent or mildly dissatisfied and quite dissatisfied). The percentages of participants endorsing each of the responses will be reported.

Health-related quality of life: EuroQoL (EQ-5D-5 L)

The EQ-5D-5 L assesses quality of life in five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is scored out of a possible five levels of severity (no problems, slight problems, moderate problems, severe problems and extreme problems). The digits applied to each dimension are then combined in a five-digit number that describes the respondent's health state. The EQ-5D-5 L has been shown to have good construct validity and responsiveness in a chronic pain cohort [43]. The EQ-5D-5 L will also be used to generate QALYs (Quality-adjusted Life Years), which will be required for a cost-consequence analysis.

Mood: Patient Health Questionnaire-9 (PHQ-9) and General Anxiety Disorder-7 (GAD-7)

Symptoms of depression will be assessed using the PHQ-9 [44], which is a nine-item questionnaire generating scores ranging from 0 to 27. A score of ≥ 10 is indicative of probable depressive disorder. The GAD-7 [45] assesses symptoms of anxiety experienced during the last two weeks. Both questionnaires are validated and commonly used to identify and measure symptoms of depression and anxiety in patients with chronic illness.

Health economics

A cost-consequence analysis will be performed and data related to costs and QALYs will be reported alongside outcomes. Patient healthcare resource utilisation data

will be collected at baseline and at 12-week follow-up, using a self-report questionnaire that will record concomitant care (general practitioner and other healthcare professional contacts, emergency department visits and number of days of hospital in-patient stays related to pain management), investigations and pain interventions during the preceding 3-month period. The costs of providing the interventions will be calculated in terms of direct contact time with healthcare professionals and QALYs will be generated using the EQ-5D-5 L.

Medication

Current medications will be recorded at each time point with the BPI, which features a specific question related to medication usage. Any changes to medications will be reported.

Adverse events

The occurrence of any adverse events will be monitored by the treating clinicians throughout the 8-week intervention period. Any adverse events that occur will be recorded by the lead researcher and reported with the study results.

Treatment process measures

Self-efficacy: Pain Self Efficacy Questionnaire (PSEQ)

Self-efficacy refers to a person's confidence in their ability to perform activities despite pain [46] and has been identified as an important mediator in the relationship between pain and disability [47]. The PSEQ features ten items that produce a total score between 0 to 60, with higher scores indicating greater self- efficacy. Analyses have shown the PSEQ to have strong psychometric properties [48].

Pain catastrophising: Pain Catastrophising Scale (PCS)

Catastrophisation is defined as an elevated negative cognitive response to painful stimuli [49]. Change in pain catastrophising has been shown to mediate reductions in pain and disability [50]. The PCS consists of 13 items that refer to thoughts and feelings related to pain. Respondents are asked to rate the degree to which they experience each item on a 5-point scale 0 (not and all) to 4 (all the time) and items are summed to give a potential total score of 52. There are also three subscales within the PCS; rumination, magnification and helplessness. The PCS has been shown to have adequate to excellent internal consistency and validity [49, 51].

Fear avoidance: Tampa Scale of Kinesiophobia (TSK)

Fear avoidance beliefs have been shown to be associated with higher levels of disability [52] and worse prognosis in patients with low back pain [53]. The TSK is a 17-item questionnaire, which assesses fear of movement and re-injury. It has been reported to be a reliable and valid measure of fear of movement in individuals with chronic pain [54].

Chronic Pain Acceptance Questionnaire – 8 (CPAQ-8) and Committed Action Questionnaire – 8 (CAQ-8)

Pain acceptance and committed action are components of the ACT model and may be potential mediators of change, related to ACT. The CPAQ-8 is a shortened version of the original 20-item CPAQ, with two subscales; activity engagement and pain willingness. Each item is scored from 0 (never true) to 6 (always true). The CPAQ has been shown to be valid and reliable, with good internal consistency and sensitivity to change [55] and the shortened version has demonstrated a sound factor structure and similar psychometric properties to the CPAQ [56, 57]. The CAQ-8 is a shortened version of the original 18-item Committed Action Questionnaire, which assesses an individual's persistence and flexibility in acting in the direction of valued goals [58]. The items are rated from 0 (never true) to 6 (always true). The CAQ-8 has shown comparable reliability and validity to the original version [59].

Physical activity outcomes

Physical activity will be measured using a Fitbit Zip™ wearable activity tracker. The trackers provide an objective indicator of physical activity behaviour and avoid common sources of error in subjective measurement (e.g. self-report measurement). The Fitbit Zip™ has an internal memory that can store data for up to 30 days and data can be transferred wirelessly to the Fitbit website via a smartphone application or by computer using the dongle provided. The ease of download of information from the Fitbit website will enable the data to be captured remotely.

Data collected will include average weekly step count, distance travelled and active minutes. Participants will be provided with the activity tracker at the baseline appointment and instructed in how to use it. The activity trackers will be worn by participants for 1 week prior to starting the interventions and for the full duration of the interventions (8 weeks in total). The main time points for analysis will be the baseline week and the last week of the intervention.

The Fitbit Zip™ has been found to be a valid measure of free-living physical activity in healthy adults [60]. A recent study comparing the reliability and validity of ten consumer activity trackers reported excellent test-retest reliability of the Fitbit Zip™, which was also reported to be the most valid of the ten trackers [61]. To our knowledge this will be the first RCT to collect physical activity data related to the effect of a combined ACT and exercise intervention on physical activity patterns measured objectively in a chronic pain cohort.

Sample size

Sample size was estimated using a target power of 80%, at a type I error rate of 0.05 and was calculated relative to the primary outcome measure; Pain severity subscale of the BPI. The statistical test assumed was an independent samples t test for group differences in the change from baseline to subsequent assessment, assuming that the randomisation ensures no systematic baseline or other covariate group differences. The minimal clinically significant difference for the interference scale of the BPI is 1 unit (standard deviation of improvement of 2 units) [40].

Calculation produced a suggested sample size of 64 per group. Allowing for potential attrition rate of 20% our final sample size is 80 participants per group.

Statistical analysis

Outcome analyses will be conducted by a professional academic statistician (RS) who will be blinded to treatment group allocation. A statistical analysis plan (SAP) will be drafted outlining the precise model to be applied and finalised before the last patient assessment is completed. No interim analyses will be conducted. Descriptive statistics will be calculated for all outcome measures at each time point, including for continuous variables: means, standard deviations, or medians with ranges of scores; and for categorical variables: frequencies and percentages.

Analyses of effectiveness of primary and secondary outcomes

Descriptive and inferential statistics will be obtained using the appropriate statistical methods that seek to address our identified objectives. The primary analysis will compare the effect of the interventions on the primary outcome; pain interference at 12 weeks post completion of the intervention. All outcome analyses will be conducted according to an intention-to-treat principle, i.e. all randomised participants will be included in the main analysis and will be analysed as randomised, regardless of protocol adherence. Secondary analysis will include the analysis of the primary outcome post intervention and analysis of the secondary and treatment process outcomes detailed previously post intervention and at 12-week follow-up. Linear mixed models on the outcome measures over time will be fitted to evaluate the effectiveness of both interventions, which intrinsically adjusts for pre-treatment scores. Statistical significance will be assessed from a p value < 0.05 from the group by time interaction term. For all tests, two-sided p values will be used, which will be reported to four decimal places with p values < 0.001 reported as $p < 0.001$. The Bonferroni method will be used to appropriately adjust the overall level of significance for multiple secondary outcomes as applicable. In the case of a significant result, planned contrasts of the group effects at post treatment and at 12-week follow-up will be used to investigate the direction and pattern of effects, and outlined in advance in the SAP. As a key component of the reporting of the analyses of outcomes, the mean changes (irrespective of statistical significance) and correlations of the measures between the assessment time periods will be obtained. An up-to-date version of SPSS will be used to conduct the analyses.

Missing data

Careful attention will be paid to ensure that all participants are fully assessed at all time points. Baseline data will be checked for missing data by the lead researcher at the baseline assessment and participants will be encouraged to complete any missing answers. It is hoped that this process will help minimise missing data at follow-up as the senior researcher will be able to clarify any ambiguous questions face-to-face at baseline. However, the trial supervisor will follow up on any missing data by telephone or post at the subsequent time points. For the purpose of secondary analysis and as a sensitivity analysis for the primary outcome, multiple imputation will be considered for any measure with over 5% missing data, using a chained equations method robust to non-normally distributed data. This will be fully reported in line with the updated CONSORT recommendations [32]. All primary and secondary outcomes (excluding physical activity data), treatment process measures and pre-selected baseline covariates (age, gender, work and educational status, pain-intensity, anxiety, depression, self-efficacy and catastrophising) will be included in the imputation model. These baseline covariates have been selected based on studies that have examined predictors of outcome of multidisciplinary treatment in chronic pain [62, 63] .

Sensitivity analysis

The following sensitivity analyses will be undertaken and reported:

1. *A per-protocol analysis*: the per-protocol analysis will exclude participants found to be ineligible after randomisation and those who attend less than 50 % of the intervention. Both intention-to-treat and per--protocol analysis sets will be reported and superiority will be determined only if demonstrated with the primary intention to treat analysis
2. *Multiple imputation of missing data*: the results from a complete case analysis will be compared to those from imputed data to assess whether they change the interpretation of findings

Analysis of physical activity data

The following data will be collected for each trial participant at baseline and on completion of the treatment: average daily step count, distance travelled and active minutes. Only those participants who have worn the device for at least 4 out of 7 days during the baseline and final week will be included in the analyses. The number of participants reaching the global recommendations for physical activity for health [64] will also be recorded. Descriptive statistics will be obtained for the Fitbit variables at baseline and by treatment arm. Linear mixed models will be used to analyse the change in measures between groups.

Methodology for the embedded qualitative study

Participants

Embedded qualitative interviews within this RCT will assist with interpretation of the findings of the study. Focus groups and individual, semi-structured, face-to-face interviews will be conducted with a purposeful sample of participants from both study arms, after the 12-week follow-up time point. Focus groups will be conducted initially and it is anticipated that ongoing analysis of focus group data may stimulate further research questions, which may be more appropriately investigated through individual interviews. People of different genders and ages and with different levels of pain intensity, pain interference, pain acceptance, fear avoidance, depression and anxiety will be invited to attend the focus groups. Depending on the trial progress, it may be useful to purposefully sample and interview a range of individual participants; for example, participants who have dropped out of the programme, participants who have not responded to the programme, participants who have responded well to the programme etc. to further explore and understand the reasons for same. These particular topics would be best investigated in a one-to-one interview. Selected participants will be sent a postal invitation to take part in the qualitative study by the pain clinic administrative staff and a copy of the trial Patient Information Leaflet will be included with the letter. The participants will be asked to respond by telephone or email, confirming whether or not they would like to take part. Those who opt to participate will be sent an appointment to attend either a focus group or individual interview. Travel expenses will be provided to participants attending from outside the local area.

Data collection

The qualitative data will be collected after the 12-week follow-up time point, and no longer than 6 months post completion of the intervention. The focus groups and individual interviews will be conducted by the lead researcher (MBC), trial supervisor (CD) or an external research assistant. The interviews will be semi-structured using a topic guide but participants will be encouraged to speak openly and freely about their experiences, positive and negative in relation to the interventions. Interviews will be scheduled for up to 90 min. The number of participants invited to the focus groups and individual interviews will be determined by ongoing data analysis and theme saturation.

Analysis

The focus group and individual interview audiotapes will be transcribed verbatim by a member of the project team, omitting any names, locations or information that could identify any individual. The de-identified transcripts will be analysed using an interpretative phenomenological approach [65]. Interpretative phenomenological analysis (IPA) is a qualitative analytic approach, commonly used in the field of health psychology [66]. The aim of IPA is to examine how people make sense of lived experiences. The approach is most often concerned with events that have significance to an individual, such as a major life experience, which would prompt a considerable amount of thinking and feeling as a person reflects on its meaning [65]. This study features an open research question, focussed on peoples' experiences and views of the featured interventions for chronic pain. This type of research question is well suited to an IPA method, which aims to both give voice to the opinions of participants and to make sense of them by offering an interpretation [67]. Standards of verification will be adhered to including member checks, peer debriefing, external audit, negative case analysis, rich description including citations from the interview transcripts identified to participant and line number [68].

Data management – data entry, coding, security, storage

In order to ensure patient confidentiality, all questionnaire data will be de-identified after collection and referred to only by a unique code assigned by the trial supervisor (CD). The trial supervisor will be the only person to hold the 'key' to re-identify the data for the full duration of the trial. The researchers who will have access to the de-identified datasets via the secure online database website are MBC, CD, KS and RS for the purpose of data entry (MBC), checking (CD and KS) and analysis (RS). All files will be encrypted and accessed via password-protected computers. Hard copies of the de-identified questionnaires will be stored in a locked cabinet in an office in the MMUH Pain Clinic.

Electronic data collected from Fitbit Zip^{TM} activity tracker will be stored on the Fitbit website and accessed only by the lead researcher (MBC) and trial supervisor (CD). Anonymised transcripts of the qualitative interviews will be accessed by the lead researcher and trial supervisor. Audio files of the interviews will be accessed

by the lead researcher (MBC) who will transcribe and de-identify the data. The audio files will be destroyed once they have been transcribed. Audio files of the ACT sessions that are recorded for the purpose of fidelity assessment will be password protected and accessed only by the psychologists (DL and NL). These files will also be destroyed on completion of the fidelity assessment.

Dissemination

The results of this trial will be published in peer-reviewed journals and will be disseminated at relevant conferences.

Discussion

This paper describes the protocol of the ExACT trial, a RCT comparing the effectiveness of exercise combined with ACT, to a supervised exercise programme in reducing pain interference in a heterogeneous adult patient population with chronic pain. We have endeavoured to address recommendations that have been made to enhance the quality of research in the field of ACT, including choice of outcome measures and comparison to an active control group [22]. The inclusion of objective measurement of physical activity, using wearable activity trackers is novel in this field of research and highly relevant, considering the prevalence of co-morbidities in chronic pain patients. The inclusion criteria for the study are broad, in recognition that chronic pain is a heterogeneous condition, and with the aim of maximising the generalisability of the findings. Through the embedded qualitative study, we aim to provide insight into patients' experiences and views of the interventions.

There are a number of limitations inherent in this trial of what have been termed 'complex interventions' [69, 70], most prominently concerning the inability to blind study participants and clinicians to treatment allocation. We have attempted to address issues of additional biases as far as has been practically possible through the provision of randomisation and concealment of allocation, strategies to minimise and manage incomplete outcome data, assessment of intervention fidelity, an adequate sample size with appropriate duration of follow-up and a priori specification of all primary and secondary outcomes as detailed in this study protocol.

To our knowledge, this will be the first RCT to assess the effectiveness of a combined Exercise and ACT intervention for chronic pain. The study results will add to current knowledge in the field of chronic pain management and will have the potential to inform the delivery of effective treatments for patients.

Trial status

This trial is currently recruiting participants. It is anticipated that recruitment will be ongoing until the end of 2018.

Additional file

Additional file 1: ExACT Trial SPIRIT 2013 Checklist. (DOC 122 kb)

Abbreviations

ACT: Acceptance and Commitment Therapy; BPI: Brief Pain Inventory; CAQ: Committed Action Questionnaire; CPAQ: Chronic Pain Acceptance Questionnaire; EQ-5D-5 L: EuroQoL 5D-5 L; ExACT: Exercise and Acceptance and Commitment Therapy; IPA: Interpretative phenomenological analysis; PCS: Pain Catastrophising Scale; PSEQ: Pain Self Efficacy Questionnaire; QALY: Quality-adjusted Life Year; RCT: Randomised controlled trial; TSK: Tampa Scale for Kinesiophobia

Acknowledgements

The authors would like to acknowledge Ms. Nóirín Nealon Lennox, Health Psychologist, University of Limerick for her contribution to the assessment of treatment fidelity of the psychological intervention featured in this trial.

Funding

This study is funded by the Health Research Board, Research Training Fellowship (Grant reference no. 2016–1685). The funding agency have no role in the study design, collection, management, analysis or interpretation of data, writing of the report or the decision to submit the report for publication.

Availability of data and materials

The final datasets generated and analysed in the current study will be available from the corresponding author on reasonable request. The datasets will be made available on a public repository following publication of papers related to the relevant data.

Authors' contributions

MBC and CD were responsible for the study conception, design and funding acquisition. CH, HG, DL and KS contributed to the conceptualisation and design of the study. RS oversaw the trial SAP, conducted the sample size calculation and will assist with the study analysis. DL and DF contributed to the design of the interventions, which they will deliver to trial participants. LMcC is an expert in the field of ACT research and RCTs and is an official collaborator on this study in an advisory capacity. With respect to this protocol, the manuscript was drafted by MBC and was critically revised by CD, KS, RS, DL and LMcC. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Ethical approval has been granted from the Mater Misericordiae University Hospital Institutional Review Board and the University Hospital Dublin Human Research Ethics Committee. Written consent will be collected from all participants prior to enrolment.

Consent for publication

Patients will be informed, prior to consenting to participate in the trial, that the results of the study may be presented at academic conferences or published in peer-reviewed journals. Participants will be assured that their confidentiality will be maintained at all times and they will not be identifiable in any publications.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]School of Public Health, Physiotherapy and Sports Science, Health Sciences Building, University College Dublin, Belfield, Dublin 4, Ireland. [2]Physiotherapy Department, St Vincents University Hospital, Elm Park, Dublin 4, Ireland. [3]Department of Pain Medicine, Mater Misericordiae University Hospital, Eccles St, Dublin 7, Ireland. [4]Psychology Department, Mater Misericordiae University Hospital, Eccles St, Dublin 7, Ireland. [5]Physiotherapy Department, Mater

Misericordiae University Hospital, Eccles St, Dublin 7, Ireland. [6]Psychology Department, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, 5th Floor Bermondsey Wing, Guy's Campus, London SE1 9RT, United Kingdom.

Received: 8 August 2017 Accepted: 13 February 2018
Published online: 22 March 2018

References

1. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain. 2006;10(4):287–333.
2. Raftery MN, Sarma K, Murphy AW, De la Harpe D, Normand C, McGuire BE. Chronic pain in the Republic of Ireland—Community prevalence, psychosocial profile and predictors of pain-related disability: results from the Prevalence, Impact and Cost of Chronic Pain (PRIME) study, Part 1. Pain. 2011;152(5):1096–103.
3. Raftery MN, Ryan P, Normand C, Murphy AW, de la Harpe D, McGuire BE. The economic cost of chronic noncancer pain in Ireland: results from the PRIME study, Part 2. J Pain. 2012;13(2):139–45.
4. Merskey H. Pain terms: a list with definitions and notes on usage recommended by the IASP subcommittee on taxonomy. Pain. 1979;6: 249–52.
5. van Middelkoop M, Rubinstein SM, Verhagen AP, Ostelo RW, Koes BW, van Tulder MW. Exercise therapy for chronic nonspecific low-back pain. Best Pract Res Clin Rheumatol. 2010;24(2):193–204.
6. Williams AC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev. 2012;11:Cd007407.
7. Mannion AF, Brox JI, Fairbank JC. Consensus at last! Long-term results of all randomized controlled trials show that fusion is no better than non-operative care in improving pain and disability in chronic low back pain. Spine J. 2016; 16(5):588–90.
8. Harris IA, Buchbinder R. Time to reconsider steroid injections in the spine? Med J Aust. 2013;199(11):754.
9. Scottish Intercollegiate Guidelines Network (SIGN). Management of chronic pain. Edinburgh: SIGN publication no. 136; 2013.
10. Carville SF, Arendt-Nielsen S, Bliddal H, Blotman F, Branco JC, Buskila D, et al. EULAR evidence-based recommendations for the management of fibromyalgia syndrome. Ann Rheum Dis. 2008;67(4):536–41.
11. Bertozzi L, Gardenghi I, Turoni F, Villafane JH, Capra F, Guccione AA, et al. Effect of therapeutic exercise on pain and disability in the management of chronic nonspecific neck pain: systematic review and meta-analysis of randomized trials. Phys Ther. 2013;93(8):1026–36.
12. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database Syst Rev. 2017;1:CD011279.
13. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets R, Ostelo R, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. BMJ. 2015;350:h444.
14. Simões D, Araújo FA, Severo M, Monjardino T, Cruz I, Carmona L, et al. Patterns and consequences of multimorbidity in the general population: there is no chronic disease management without rheumatic disease management: multimorbidity patterns and adverse outcomes. Arthritis Care Res. 2016;69(1):12–20.
15. Barker AL, Talevski J, Morello RT, Brand CA, Rahmann AE, Urquhart DM. Effectiveness of aquatic exercise for musculoskeletal conditions: a meta-analysis. Arch Phys Med Rehabil. 2014;95(9):1776–86.
16. Hauser W, Klose P, Langhorst J, Moradi B, Steinbach M, Schiltenwolf M, et al. Efficacy of different types of aerobic exercise in fibromyalgia syndrome: a systematic review and meta-analysis of randomised controlled trials. Arthritis Res Ther. 2010;12(3):R79.
17. Bennell KL, Hinman RS. A review of the clinical evidence for exercise in osteoarthritis of the hip and knee. J Sci Med Sport. 2011;14(1):4–9.
18. Smith BE, Littlewood C, May S. An update of stabilisation exercises for low back pain: a systematic review with meta-analysis. BMC Musculoskelet Disord. 2014;15:416.
19. O'Keeffe M, Hayes A, McCreesh K, Purtill H, O'Sullivan K. Are group-based and individual physiotherapy exercise programmes equally effective for musculoskeletal conditions? A systematic review and meta-analysis. Br J Sports Med. 2017;51(2):126–32.
20. Jordan JL, Holden MA, Mason EE, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. Cochrane Database Syst Rev. 2010;1:CD005956.
21. Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. Behav Ther. 2004; 35(4):639–65.
22. Hann KEJ, McCracken LM. A systematic review of randomized controlled trials of Acceptance and Commitment Therapy for adults with chronic pain: outcome domains, design quality, and efficacy. J Contextual Behav Sci. 2014;3(4):217–27.
23. Hughes LS, Clark J, Colclough JA, Dale E, McMillan D. Acceptance and Commitment Therapy (ACT) for chronic pain: a systematic review and meta-analyses. Clin J Pain. 2017;33(6):552–68.
24. Wetherell JL, Afari N, Rutledge T, Sorrell JT, Stoddard JA, Petkus AJ, et al. A randomized, controlled trial of acceptance and commitment therapy and cognitive-behavioral therapy for chronic pain. Pain. 2011;152(9):2098–107.
25. Kemani MK, Olsson GL, Lekander M, Hesser H, Andersson E, Wicksell RK. Efficacy and cost-effectiveness of Acceptance and Commitment Therapy and applied relaxation for longstanding pain: a randomized controlled trial. Clin J Pain. 2015;31(11):1004–16.
26. Kangasniemi AM, Lappalainen R, Kankaanpaa A, Tolvanen A, Tammelin T. Towards a physically more active lifestyle based on one's own values: the results of a randomized controlled trial among physically inactive adults. BMC Public Health. 2015;15(1):260.
27. VanBuskirk K, Roesch S, Afari N, Wetherell JL. Physical activity of patients with chronic pain receiving Acceptance and Commitment Therapy or Cognitive Behavioural Therapy: 1. Behav Chang. 2014;31(2):131.
28. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687.
29. Brassington L, Ferreira NB, Yates S, Fearn J, Lanza P, Kemp K, et al. Better living with illness: a transdiagnostic acceptance and commitment therapy group intervention for chronic physical illness. J Contextual Behav Sci. 2016; 5(4):208–14.
30. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. J Public Health Dent. 2011;71:S52–63.
31. Plumb JC, Vilardaga R. Assessing treatment integrity in acceptance and commitment therapy: strategies and suggestions. Int J Behav Consultation Ther. 2010;6(3):263–95.
32. Schulz KF, Altman DG, Moher D, Group C, the CG. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010;8(1):18.
33. 2001–2017. Available from: http://www.sealedenvelope.com. Last accessed 2 Mar 2018.
34. WMA. WMA Declaration of Helsinki – Ethical principles for medical research involving human subjects. 1964. Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/. Last accessed 2 Mar 2018.
35. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.
36. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005;113(1):9–19.
37. Moore RA, Eccleston C, Derry S, Wiffen P, Bell RF, Straube S, et al. 'Evidence' in chronic pain – establishing best practice in the reporting of systematic reviews. Pain. 2010;150(3):386–9.
38. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for chronic nonmalignant pain. J Pain. 2004;5(2):133–7.
39. Mendoza T, Mayne T, Rublee D, Cleeland C. Reliability and validity of a modified Brief Pain Inventory short form in patients with osteoarthritis. Eur J Pain. 2006;10(4):353–61.
40. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT Recommendations. J Pain. 2008;9(2):105–21.
41. Kamper SJ, Maher CG, Mackay G. Global Rating of Change Scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009;17(3):163–70.
42. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann. 1979;2(3):197–207.

43. Obradovic M, Lal A, Liedgens H. Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. Health Qual Life Outcomes. 2013;11(1):110.
44. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
45. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092–7.
46. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev. 1977;84(2):191–215.
47. Costa Lda C, Maher CG, McAuley JH, Hancock MJ, Smeets RJ. Self-efficacy is more important than fear of movement in mediating the relationship between pain and disability in chronic low back pain. Eur J Pain. 2011;15(2):213–9.
48. Nicholas MK. The Pain Self-efficacy Questionnaire: taking pain into account. Eur J Pain. 2007;11(2):153–63.
49. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess. 1995;7(4):524–32.
50. Smeets RJEM, Vlaeyen JWS, Kester ADM, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. J Pain. 2006;7(4):261–71.
51. Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The Pain Catastrophizing Scale: further psychometric evaluation with adult samples. J Behav Med. 2000;23(4):351–65.
52. Waddell G, Newton M Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain. 1993;52(2):157–68.
53. Wertli MM, Rasmussen-Barr E, Weiser S, Bachmann LM, Brunner F. The role of fear avoidance beliefs as a prognostic factor for outcome in patients with nonspecific low back pain: a systematic review. Spine J. 2014;14(5):816-36.e4.
54. Roelofs J, Goubert L, Peters ML, Vlaeyen JWS, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. Eur J Pain. 2004;8(5):495–502.
55. Wicksell RK, Olsson GL, Melin L. The Chronic Pain Acceptance Questionnaire (CPAQ)—further validation including a confirmatory factor analysis and a comparison with the Tampa Scale of Kinesiophobia. Eur J Pain. 2009;13(7):760–8.
56. Rovner GS, Arestedt K, Gerdle B, Borsbo B, McCracken LM. Psychometric properties of the 8-Item Chronic Pain Acceptance Questionnaire (Cpaq-8) in a Swedish chronic pain cohort. J Rehabil Med. 2014;46(1):73–80.
57. Baranoff JA, Hanrahan SJ, Burke ALJ, Connor JP. Changes in acceptance in a low-intensity, group-based Acceptance and Commitment Therapy (ACT) chronic pain intervention. Int J Behav Med. 2016;23(1):30–8.
58. McCracken LM. Committed action: an application of the psychological flexibility model to activity patterns in chronic pain. J Pain. 2013;14(8):828–35.
59. McCracken LM, Chilcot J, Norton S. Further development in the assessment of psychological flexibility: a shortened Committed Action Questionnaire (CAQ-8): committed action in chronic pain. Eur J Pain. 2015;19(5):677–85.
60. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fitbit Zip™ physical activity monitor as a measure of free-living physical activity. BMC Res Notes. 2014;7(1):952.
61. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabi. 2015;7(24).
62. de Rooij A, van der Leeden M, Roorda LD, Steultjens MP, Dekker J. Predictors of outcome of multidisciplinary treatment in chronic widespread pain: an observational study. BMC Musculoskelet Disord. 2013;14:133.
63. Boonstra AM, Reneman MF, Waaksma BR, Schiphorst Preuper HR, Stewart RE. Predictors of multidisciplinary treatment outcome in patients with chronic musculoskeletal pain. Disabil Rehabil. 2015;37(14):1242–50.
64. World Health Organisation. Global recommendations on physical activity for health. 2010.
65. Smith JA, Larkin M, Flowers P. Interpretative phenomenological analysis: theory method and research. 1st ed. Sage; 2009. p. 79-107.
66. Brocki JWA. A critical evaluation of interpretative phenomenological analysis (IPA) in health psychology. Psychol Health. 2006;21:87–108.
67. Thompson A, Larkin M. Interpretative Phenomenological Analysis. In: Thompson A, Harper D, editors. Qualitative research methods in mental health and psychotherapy: a guide for students and practitioners. Oxford: Wiley; 2012. p. 99–116.
68. Creswell J. Qualitative inquiry and research design. Thousand Oaks: Sage; 1998.
69. Campbell-Yeo M, Ranger M, Johnston C, Fergusson D. Controlling bias in complex nursing intervention studies: a checklist. Can J Nurs Res. 2009;41(4):32–50.
70. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008;337:a1655.

Cavalcante et al. Sports Medicine - Open (2017) 3:42
DOI 10.1186/s40798-017-0111-2

Sports Medicine - Open

SYSTEMATIC REVIEW Open Access

Aerobic but not Resistance Exercise Can Induce Inflammatory Pathways via Toll-Like 2 and 4: a Systematic Review

Paula Andréa Malveira Cavalcante[1,3,4,7*], Marcos Fernandes Gregnani[2,3,4], Jessica Salles Henrique[5,6], Fábio Henrique Ornellas[1,3,4] and Ronaldo Carvalho Araújo[1,2,3,4]

Abstract

Background: Only a few studies have addressed the relationship between toll-like receptors 2 and 4 (TLR2 and TLR4) and the production of local and systemic cytokines in response to physical exercise, and they have produced conflicting results. We aimed to determine whether acute and chronic exercise outcomes are associated with changes in TLR2 and TLR4 expression and signaling and if so, the mechanisms that connect them.

Methods: PubMed database were consulted. This systematic review selected 39 articles, 26 involving humans and 13 based on rodents.

Results: In acute resistance exercise studies, 75% reported a decrease in TLR4 or TLR2 expression and 25% did not find differences. For chronic resistance exercise studies, 67% reported a reduction of expression and 33% did not find differences. Studies of both types reported reductions in pro-inflammatory cytokines. In acute aerobic exercise studies, 40% revealed a decline in the expression of the receptors, 7% reported no significant difference, 40% showed an increase, and 13% did not evaluate their expression. Fifty-eight percent of studies of chronic aerobic exercise revealed a reduction in expression, 17% did not find a difference, and 25% reported increases; they also suggested that the expression of the receptors might be correlated with that of inflammatory cytokines. In studies on combined exercise, 50% reported a decline in receptors expression and 50% did not find a difference.

Conclusions: The majority of the articles (54%) link different types of exercise to a decline in TLR4 and TLR2 expression. However, aerobic exercise may induce inflammations through its influence on these receptor pathways. Higher levels of inflammation were seen in acute sessions (40%) than regular sessions (25%).

Keywords: TLR2, TLR4, Toll-like, Exercise, Training, Aerobic, Resistance, Inflammation

Key Points

- It is known that regular exercise acts as an anti-inflammatory agent by down-regulating TLR4 in immune cells. Paradoxically, acute, extended, or intense exercise can be harmful to the immune system.
- The molecular mechanisms by which various types of physical exercise modulate the TLR2 and TLR4 pathways are still not fully understood.
- Physical exercise reduced the expression of TLR2 and TLR4. However, aerobic exercise is potentially inflammatory when compared with resistance exercise.

Background

The connections between lifestyle factors and health have been the subject of intense research, partly motivated by alarming changes in the health landscape of industrialized societies. One clear trend is that moderate exercise benefits health in many ways, while extremes of too little or excessive exercise have been linked to chronic diseases. Many of these have an immune component—individuals with very sedentary lifestyles often

* Correspondence: paulaacavalcante@gmail.com
[1]Medicine (Nephrology) Program, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil
[3]Laboratory of Exercise Genetics and Metabolism, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil
Full list of author information is available at the end of the article

fall prey to low-grade chronic inflammations [1–4]. Over the long term, this condition can lead to type 2 diabetes, cardiovascular diseases, particular types of cancer, chronic respiratory diseases, and other serious health problems. Physicians have called this constellation a worldwide epidemic [5]. The immune system can also be disrupted by excessive exercise. While progress has been made, there remain many gaps in our understanding of the mechanisms that connect the types and amounts of a person's activity to immune responses and disease.

The prevalence of inflammations suggests a logical point of departure for such studies. Inflammation involves complex interactions at the molecular and cellular levels that can arise in any vascular tissue as a result of traumatic, infectious, post-ischemic, toxic, or autoimmune injuries [6]. Toll-like receptors play a role in many of these conditions; they are known to make significant contributions to obesity [7, 8], type 2 diabetes [9], non-alcoholic steatosis [10], cardiovascular disease [11, 12], cerebral ischemia [13, 14], Alzheimer's disease [15], rheumatoid arthritis [16], and other diseases. This review examined recent work that suggests they also help modulate the effects of different levels of physical activity on states of health and disease.

TLRs are type I transmembrane proteins involved in both innate and adaptive immune system responses [17, 18]. These receptors mediate the recognition of pathogen-associated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs)—specific molecules released by damaged or necrotic cells [18, 19]. The immune activities of TLRs are generally modulated through signaling via the NF-kB pathway. Responses begin with the stimulation of the receptor by an external signal. This alters the cytoplasmic regions of TLRs, which contain Toll/interleukin-1 (IL-1) receptor (TIR) domains. Stimulation causes these domains to recruit adaptor proteins in a process that ultimately activates the nuclear transcription factor NF-kB [17]. This releases NF-kB for transport to the cell nucleus, where it triggers the transcription of cytokines including IL-1β, IL-6, and IL-8 interleukins; TNF-α [20–22]; and other elements [23] that play key roles in the immune system responses. Alongside cytokines, NF-kB induces the expression of growth factors and other molecules involved in stress response, cell proliferation, and cell cycle progression [24–26].

TLRs are expressed in the immune cells including macrophages, dendritic cells (DCs), B cells, and specific types of T cells. They are also present in non-immune cells such as fibroblasts and epithelial cells [27] and in the tissues of the ovary, prostate, placenta, testicles, lungs, liver, and skeletal muscle [28].

The toll-like receptors TLR2 and TLR4 have received particular attention due to their ability to identify molecular patterns exhibited by several invasive pathogens [18]. They also seem to play an important role in the anti-inflammatory effects observed in physically active individuals [29]. Regular exercise has been determined to have anti-inflammatory effects [2, 29–34] by downregulating TLR4 in the immune cells. A bit paradoxically, at the other end of the activity spectrum, acute, extended, or intense exercise can have a negative impact on the immune system [35–42]. But the molecular mechanisms by which exercise modulates the TLR2 and TLR4 pathways are still not fully understood.

One plausible link comes from the demonstration that TLR2 and TLR4 are activated by the extracellular non-esterified fatty acids (NEFAs). Concentrations of extracellular NEFAs undergo transient increases during aerobic exercise (AE). If levels are chronically elevated, however, TLRs may induce the production of pro-inflammatory cytokines in macrophages, adipocytes, liver, and skeletal muscle cells. This suggests that the receptors may participate in the development of insulin resistance [43]. Yet, they also have protective effects against insulin resistance, which may be explained by the down-regulation of TLR expression that occurs during physical exercise [43].

Here, this review investigated the existing literature on the inflammatory and anti-inflammatory effects of different types of physical exercise with a focus on systematically collecting connections to TLR2 and TLR4 modulation and signaling. To accomplish this, the results were divided into single sessions of acute exercise and chronic exercise, based on periodicity. Additionally, this review identified key biomarkers and analyzed the combined TLR2 and TLR4 responses to markers involved in the process of inflammation process, including anti- and pro-inflammatory cytokines, adaptor proteins, and the transcription factor NF-kB.

Inflammatory Effects of Physical Exercise

Analyzing the modulation of inflammation patterns permits insights into specific underlying physiological mechanisms. As a controllable model of stress, physical exercise is a good tool to analyze inflammatory responses [44].

Physical exercise permits the control of variables related to activity such as volume, intensity, frequency, and duration. These factors have led to its adoption as a good strategy to study alterations that occur due to inflammations caused by stress and their implications for health [45–47]. Local and systemic cytokine production in response to physical exercise resembles the cytokine response to infections, trauma, and sepsis [44, 45, 48]. There is evidence that very strenuous physical exercise can cause substantial tissue damage and initiate an inflammatory reaction and excessive immunosuppression, in a way

that highly resembles features observed in clinical sepsis [49]. However, trauma, infection, and septic complications can produce an uncontrollable inflammatory response with long-term detrimental or fatal consequences. In physical exercise, although the inflammatory cascade has obvious similarities, the response appears to be limited [44].

Usually, the process of inflammation has an overall positive effect on the organism. Short-term, acute inflammation allows the body to survive progressive tissue destruction by promoting healing [50, 51]. On the other hand, if destruction and repair are not properly coordinated, inflammation may lead to persistent tissue damage. The mechanisms by which acute inflammation starts and develops are well understood, but little is known about the causes of chronic inflammation and its association with molecular and cellular pathways [51].

A comparison can also be made between chronic inflammation and strenuous physical exercise in which pro-inflammatory pathways seem to be activated [38, 41, 52]. In response to heavy exercise, inflammation stimulates tissue monocyte production, and platelet hyperactivity promotes fibrinogen biosynthesis and induces the formation of the microparticle and the accumulation of erythrocytes to trigger a prothrombotic state. In fact, vigorous aerobic exercise may be atherogenic and atherothrombotic due to the overproduction of mitochondrial-free radicals in the skeletal and myocardial muscle. On the other hand, both moderate AE and low-load resistance exercise (RE) may reduce inflammation and improve fibrinolysis. [52].

An elegant study [53] found associations between all causes of mortality and doses of jogging. Light and moderate joggers had a lower mortality than sedentary non-joggers, while there was no significant statistical difference between mortality in strenuous joggers and the sedentary group. In this analysis, high running loads in sports such as marathons, ultramarathons, triathlons, and long high-intensity bike rides can cause negative effects such as acute inflammations; in the long term, these activities may lead to chronic inflammation, irregular fibrosis formation, alterations in the size of the cardiac chambers, and atrial fibrillation [54]. Moreover, long-distance runners may have increased levels of atherosclerosis and coronary disease due to constant training throughout the year [54]. In atherosclerosis, the endothelial permeability is increased by the oxidative damage that promotes the entry of lipoproteins in the subendothelial space, resulting in inflammation [55]. When the lipoproteins are oxidative, they interact with TLR4 in particular and promote cardiovascular disease [56].

According to the American College of Sports Medicine (ACSM) and the American Heart Association [57], the minimum recommendation for physical exercise for adults and seniors aiming to avoid chronic disease is 30 min of moderate aerobic activity per day, five times a week; 20 min per day of intense activity, three times a week; or a combination of moderate and vigorous activity. These guidelines also suggest that high loads of AE may be necessary for some groups to prevent a transition to an estimation that they are overweight or a diagnosis of obesity. However, they also recommend limiting vigorous physical training to 60 min a day, for a weekly total of no more than 5 h, including 1 to 2 days without high-intensity exercise per week [58, 59]. Strenuous AE has been shown to induce an excess of reactive oxygen species (ROS) [60]; can modulate TLR4 signal transduction at many levels [61]; stimulate pro-inflammatory transcription factors such as NF-kB, AP-1, and Nrf2 [62, 63]; and promote inflammation [64].

NADPH oxidase 4 (NOX4), involved in redox signaling in vascular cells, has direct interactions with TLR4 in both for the generation of endogenous and exogenous ROS-mediated by LPS and the activation of NF-kB [65]. In addition, high levels of ROS in the muscles can provoke a hyperactivation of the innate immune system in cells such as macrophages and neutrophils [66], and it leads to the production of several peroxides and aldehydes that are potentially toxic to the cells [67], also affecting T cell polarization and contributing to pro-inflammatory cytokine secretion [68]. It is already known that ROS production and neutrophil counts change in athletes involved in activities such as running, jumping, throwing, combined events (triathlon, heptathlon, and decathlon), swimming, cycling, and soccer, but only high-intensity exercise induces oxidative damage in lymphocytes [69]. In contrast, moderate-intensity AE stimulates the combat of excessive ROS by maintaining redox balance in the muscle [70]. A study [71] of soccer players showed a significant correlation between leukocyte ROS production and creatine kinase (CK) values, considered a qualitative marker for microtrauma skeletal muscle.

In fact, the physiological effects of strenuous AE, for example, participation in triathlons, include a large increase in CK, C-reactive protein (CRP), cortisol, and aldosterone and a decrease in testosterone levels [72]. Moreover, after strenuous exercise, increased levels of LPS may trigger an increase in the production of pro-inflammatory cytokines [73–76]. Long periods of AE [72] or short acute sessions of strenuous physical exercise [41] can disturb homeostasis and enhance inflammation. Consistent with this, Rodrigues-Miguelez et al. [39] found an increase in TLR4 and pro-inflammatory cytokines such as TNF-α and IL-1β in acute AE sessions; however, the effects were reversed with regular training in reasonable doses.

TNF-α represents a group of peptides that are released into the bloodstream in response to the endotoxin

stimulation during infectious processes. TNF-α has a catabolic effect [77] and plays a role in the loss of muscle mass that usually appears in chronic diseases such as rheumatoid arthritis and cancer [78]. TNF-α genesis in low-grade systemic inflammation is thought to occur mainly in the adipose tissue [79–81]. Furthermore, systemic inflammation and high concentrations of pro-inflammatory cytokines act on the hypothalamic-pituitary-adrenal axis and can increase serum concentrations of cortisol [82, 83]. Physical exercise and nutrition modulate the cortisol response. Variables such as intensity, lactate accumulation, total volume, and resting period determine the level of cortisol released to stimulate glycogenolysis and gluconeogenesis [84, 85]. Moderate- to high-intensity exercise can cause increases in circulating levels of cortisol. On the other hand, low-intensity exercise (40% VO2max) reduces circulating levels of cortisol [84]. In the study by Lira et al. [76], TLR-4 and NF-kBp65 were increased in animals from both groups (overtraining and resting after overtraining). Additionally, a decrease in the performance and an increase in the production of corticosterone and endotoxin were observed in overtraining groups compared to both control and trained groups, indicating that chronically high levels of plasma cortisol can increase inflammation in the epididymal adipose tissue.

Thereby, an excess of physical (blood cortisol levels) and oxidative stress (intracellular ROS accumulation) can generate temporary immune dysfunction [86]. In contrast, physical exercise at moderate intensities regulates the immune system and reduces oxidative stress [87]. Figure 1 presents a simplified comparison of some mechanisms that can be activated by strenuous physical exercise and by regular exercise performed at moderate intensity.

Anti-inflammatory Effects of Physical Exercise

It is well known that regular physical exercise has anti-inflammatory effects [8, 29–31, 88–93]. Therefore, regular physical exercise, as well as a physically active lifestyle, may be useful as a treatment for a range of chronic diseases and conditions characterized by low-grade systemic inflammation [3, 94].

However, the link between physical exercise and TLRs is still a matter of debate. Although the pro-inflammatory effects of TLR2 and TLR4 signaling

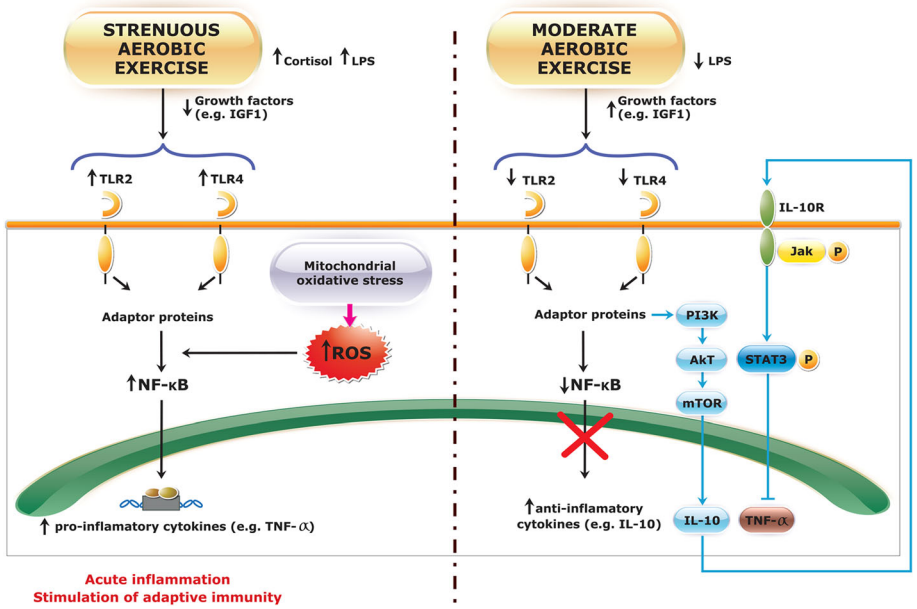


Fig. 1 Signaling involving TLR2 and TLR4 in strenuous and moderate aerobic exercise. Excess physical exercise increases LPS levels and contributes to TLR2, TLR4, and NF-kB upregulation. As a consequence, there is an increase in circulating pro-inflammatory cytokines. Stimuli of exercise stress transmit nerve impulses to the brain, raising the levels of counter-regulatory hormones such as cortisol. Accordingly, high mitochondrial oxidative stress induced by strenuous aerobic exercise causes excessive intracellular ROS formation that also upregulates NF-kB expression, intensifying the acute inflammation state. Under these excessive stress conditions, adaptive immunity can be triggered by the increase in costimulatory molecules in antigen-presenting cells, thus activating T cells. In contrast, the regular physical exercise of moderate intensity reduces LPS, TLR2, TLR4, and NF-kB expression. Under these conditions, NF-kB does not translocate to the cell nucleus. Instead, the anti-inflammatory pathway PI3K/AKT/mTOR is activated, promoting the production of anti-inflammatory cytokines such as IL-10 that inactivate TNF-α. Physical exercise at a moderate intensity also has a compensatory effect against the exacerbated production of reactive oxygen and nitrogen species responsible for the oxidative damage. Elevated production of IGF-1 is observed after exercise. IGF-1 provides an anti-inflammatory effect on the skeletal muscle cells, reducing the expression of the pro-inflammatory cytokines through a decrease of TLR4 expression

have been well studied, anti-inflammatory responses due to the activation of these receptors are still not fully understood [95]. For this reason, this article will briefly address a number of molecules that act directly during the processes of adaptation to physical exercise—including hormones, myokines, and chemical molecules such as ROS.

The skeletal muscle can function as an endocrine organ due to its production of growth hormones and cytokines known as myokines, which are induced by an exercise stimulus [96, 97]. One of the best-known exercise-induced adaptations [98, 99] is an increase in circulating levels of insulin-like growth factor 1 (IGF-1). Elevated levels of circulating IGF-1 have been observed after exercise, probably in response to hepatic secretion stimulated by growth hormone (GH) [85].

The first evidence that IGF-1 is a potent modulator of TLR4 (protein expression) in the skeletal muscles was provided by Lee [31]. The author demonstrated that IGF-1 stimulation had anti-inflammatory effects on the skeletal muscle and suppressed TLR4 signaling. Treatment with IGF-1 attenuated the amounts of endogenous IL-6 and TNF-α, indicating that IGF-1 had an anti-inflammatory effect on the skeletal muscle cells by reducing the expression of pro-inflammatory cytokines under baseline conditions through a down-regulation of the expression of TLR4. This led to a hypothesis that cells with low levels of TLR4 are less responsive to ligands that stimulate endogenous inflammation, such as the heat shock protein, and thus contribute to a lower basal response of pro-inflammatory cytokines [31]. In addition to the anti-inflammatory effects of IGF-1, regular AE promotes the remodeling of mitochondrial networks with significant improvements in both the quality and quantity of the mitochondria [100]. This results in positive changes in the respiratory capacity and oxygen extraction of trained subjects [100, 101].

Likewise, there is an increase in angiogenesis, the formation of new capillaries from pre-existing ones. High levels of VEGF—resulting from endurance training—offer favorable conditions for an increase in the density of the muscle capillaries [100]. Furthermore, a moderate level of AE reduces pro-atherogenic cytokines such as TNF-α and IFN-γ and simultaneously increases atheroprotective cytokines such as IL-4, IL-10, and TGF-β [102].

The anti-inflammatory effects of regular exercise might be mediated by a reduction of visceral fat mass followed by a decline in the release of adipocytokines, as well by the anti-inflammatory environment induced by exercise [103]. This environment consists of three variables: cortisol and adrenaline release from suprarenal glands, an increase in the production and release of IL-6 and other myokines from skeletal muscle, and a decrease in amounts of TLR (cell surface protein and mRNA expression) - in monocytes and macrophages, and as a consequence, the inhibition of the release of pro-inflammatory cytokines [103].

In fact, there is evidence that exercise is responsible for reducing the expression of these receptors at both mRNA expression and protein levels [2, 29, 30, 32, 93]. In diet-induced obesity rats (DIO), both acute aerobic exercise (AAE) and chronic aerobic exercise (CAE) led to a significant suppression of the TLR4 signaling pathway in liver, muscle, and adipose tissue, reduced LPS in serum, and improved insulin signaling [9]. However, the anti-inflammatory responses induced by TLR4 activation have not been characterized as clearly. In contrast to TLR4 pro-inflammatory signaling at the cell surface, TLR4 signaling from endosomal compartments induces the secretion of the anti-inflammatory cytokine IL-10 [95].

During physical exercise, a transient increase in IL-6 in circulation appears to be responsible for a further increase in the levels of circulating anti-inflammatory cytokines such as IL-10 and IL-1ra [104–106]; this also stimulates the release of cortisol from the adrenal glands [106]. Increases in IL-6 levels during exercise are transient and return to resting levels usually within 1 h after exercise [107]. This phenomenon may occur because IL-6 production is modulated by the glycogen content in muscles [108], which function as an energy sensor [97].

The anti-inflammatory effects of TLR2 and TLR4 during exercise are mediated by the PI3K/AKT/mTOR pathway after an activation of adaptor proteins, leading to the production of IL-10 (Fig. 1) [95], an anti-inflammatory cytokine produced by Th1 cells, monocytes, and macrophages that is present in higher concentrations after physical exercise and acts as a potent inhibitor of pro-inflammatory cytokines [109, 106].

IL-10/IL-10R signaling is mediated by the activation of the JAK/STAT pathway through the phosphorylation of the Tyk2/JAK1 tyrosine, which results in the activation of STAT3 [110]. This mechanism is independent of the toll-like pathway. An analysis of the IL-10/TNF-α ratio is often used as an indicator of inflammatory conditions [32, 111]. This is evidence that IL-10 acts as a natural antagonist of TNF-α and is able to inhibit NF-κβ signaling [110, 112], as shown in Fig. 1.

Methods

This review consulted the PubMed database in a search involving seven keywords: “exercise,” “training,” “physical activity,” “TLR,” “TLR2,” “TLR4,” and “toll-like,” To cross-reference the words, 12 groups were created to link terms associated with exercise (“exercise,” “training,” “physical activity”) to toll-like terms (“TLR,” “TLR2,” “TLR4,” and “toll-like”), building groups formed from two individual keywords linked by the Boolean operator

Table 1 Eligibility codes

Eligibility codes	Description
I	Included articles
D	Duplicate articles
E1	Non-English articles
E2	Articles that did not provide enough information
E3	Literature review articles
E4	Articles that did not cover Toll-like receptors
E5	Articles studying TLRs other than TLR2 and TLR4
E6	Articles without exercise protocols
E7	Articles that used animal models other than humans, rats, and mice
E8	Articles that involved diet, supplementation, or drugs

"AND." This produced groups organized as follows: group 1: "exercise" and "TLR"; group 2: "exercise" and "TLR2"; group 3: "exercise" and "TLR4"; group 4: "exercise" and "toll-like"; group 5: "training" and "TLR"; group 6: "training" and "TLR2"; group 7: "training" and "TLR4"; group 8: "training" and "toll-like"; group 9: "physical activity" and "TLR"; group 10: "physical activity" and "TLR2"; group 11: "physical activity" and "TLR4"; and group 12: "physical activity" and "toll-like."

Only studies carried out directly in animal models (human, rat, and mouse) were included. For scientific substantiation, 119 scientific articles were also consulted in addition to the 39 studies which met the criteria of eligibility for this review.

Criteria which excluded articles from this review, described in Table 1, fell into categories as follows: non-English articles; literature reviews; articles that did not cover Toll-like receptors (TLRs); articles studying TLRs other than TLR2 and TLR4; articles without exercise protocols; experimental articles that did not use humans, mice, or rats; and finally, articles that involved diet, supplementation, or drugs. To do so, codes to link the eligibility criteria of all of the items found in the search were created.

Initially, 1385 articles were found. After an update, the search ended up with 1548 articles from the PubMed database. The updated search was carried out in October 2015. The search group distribution can be seen in Table 2. Figure 2 shows a flowchart of the article selection process, as well as how the articles were linked to the search theme. The total number of articles found and the distribution of the excluded articles are also carefully detailed.

Results and Discussion

To investigate the roles of TLR2 and TLR4 behavior in the inflammatory and anti-inflammatory effects of exercise, the results were distributed according to the type of exercise (resistance, aerobic, and combined) and frequency of training (acute or chronic), taking the exclusion criteria into account.

Considering the total of 39 studies that met the eligibility requirements for this review, 28 articles were based on the samples from a disease-free setting and 11 samples related to a disease. Three articles studied the effects of exercise and TLR2 and TLR4 on obesity [8, 113, 114], one on pre-diabetes [115], one on low back pain [116], two on cerebral ischemia [13, 14], one on pulmonary inflammation [117], one on Alzheimer's disease [15], one on chronic fatigue syndrome [36], and one on multiple sclerosis and fibromyalgia [118].

As shown in Table 3, 21 of the 39 eligible articles (54%) showed a reduction in TLR4 and/or TLR2 at the levels of both cell surface protein and mRNA expression,

Table 2 Distribution of the number of articles per studied groups

Groups	Keywords	Number of articles	Number of articles (after an update)
1	"Exercise" and "TLR"	46	54
2	"Exercise" and "TLR2"	19	23
3	"Exercise" and "TLR4"	64	71
4	"Exercise" and "Toll-like"	111	123
5	"Training" and "TLR"	158	181
6	"Training" and "TLR2"	89	97
7	"Training" and "TLR4"	154	178
8	"Training" and "Toll-like"	372	410
9	"Physical activity" and "TLR"	66	72
10	"Physical activity" and "TLR2"	43	46
11	"Physical activity" and "TLR4"	91	104
12	"Physical activity" and "Toll-like"	172	189
	Total	1.385	1.548

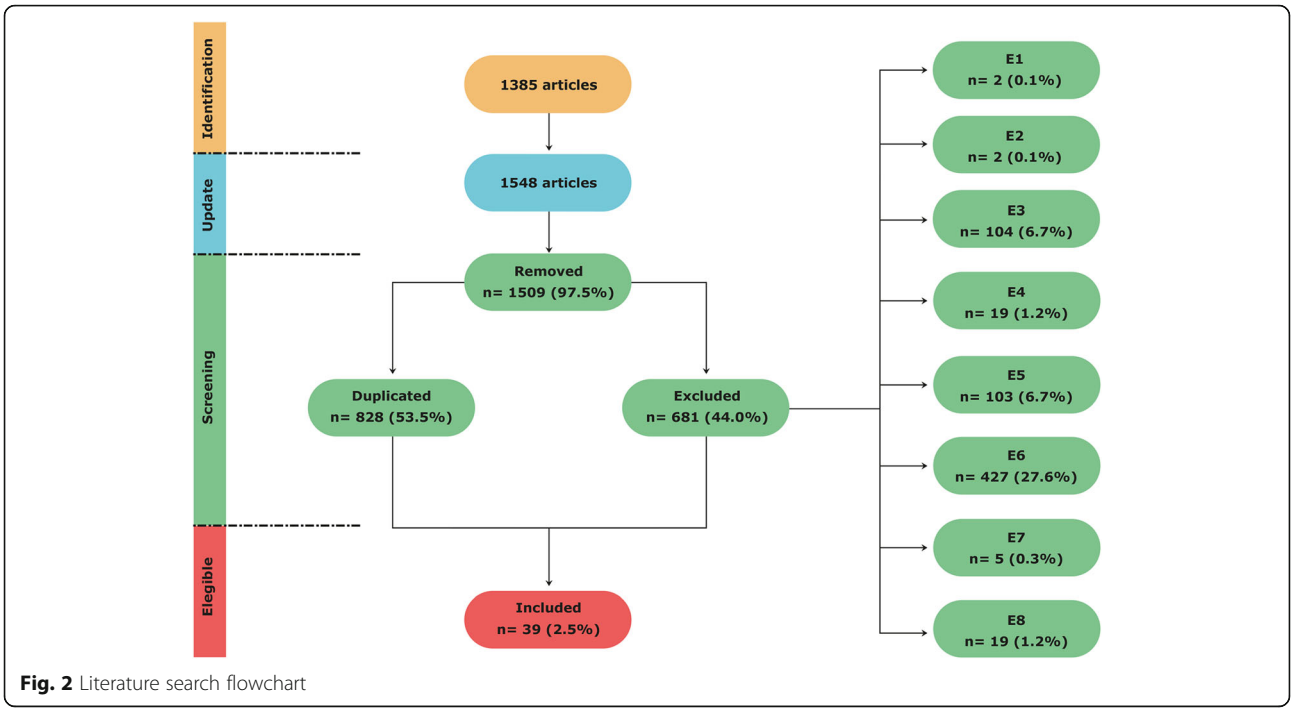


Fig. 2 Literature search flowchart

7 (18%) did not show statistically significant differences, 2 articles (5%) did not test TLR4 and/or TLR2 expression but were included in this review for the evaluation of downstream targets of the receptor pathways, and 9 articles (23%) reported an increase in TLR2 and/or TLR4 (gene expression or protein levels) after AE sessions.

The results were also analyzed by subgroups and divided according to the type and frequency of training (Table 3 and Fig. 3). For chronic resistance exercise (CRE), four articles (67%) reported a reduction of TLR4 and/or TLR2 expression and two (33%) did not show any significant change. For acute resistance exercise (ARE), three articles (75%) revealed a decrease in the expression of these receptors and one study (25%) failed to find a significant difference. For CAE, seven articles (58%) reported a reduction in TLR4 and/or TLR2 expression, two studies (17%) did not find a significant difference, and three articles (25%) found an increase in the expression of TLR4 and/or TLR2. For AAE, six experiments (40%) showed a decrease, one (7%) did not show any difference, six (40%) reported an increase, and two articles (13%) tested neither TLR2 nor TLR4 expression. Regarding combined exercise (CE), one study (50%) reported a reduction in the expression of the receptors and one study (50%) revealed no significant difference.

Resistance Exercise and Inflammation

Six articles that studied TLR4 and/or TLR2 behavior with CRE were identified (Table 4). Two studies found a reduction of TLR4 and TLR2 in terms of protein expression [92, 119], two revealed a decrease in mRNA expression [116, 120], and two did not find a statistically significant difference [8, 121]. Three articles [29, 88, 122] showed reductions in the protein and gene expression of TLR4 after an ARE session, and one article [123] did not show a significant difference in TLR2 (protein levels), as shown in Table 5. This systematic review

Table 3 Results of TLR2 and TLR4 expression of all eligible articles divided by type and frequency of exercise

	Physical exercise	↓TLR2 and 4		↔TLR2 and 4		↑TLR2 and 4		No results		Total
		n	%	*n*	%	*n*	%	*n*	%	
Total		21	54	7	18	9	23	2	5	39
Subgroups by exercise types	Chronic resistance	4	67	2	33					6
	Acute resistance	3	75	1	25					4
	Chronic aerobic	7	58	2	17	3	25			12
	Acute aerobic	6	40	1	7	6	40	2	13	15
	Combined	1	50	1	50					2

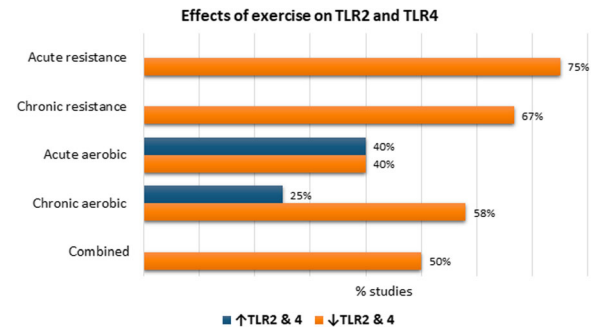


Fig. 3 Effects of exercise on TLR2 and TLR4. For chronic resistance exercise, 67% of studies reported a reduction of TLR4 and/or TLR2 expression. For acute resistance exercise, 75% of studies revealed a decrease in the expression of these receptors. For chronic aerobic exercise, 58% of studies reported a reduction in TLR4 and/or TLR2 expression and 25% found an increase in the expression of TLR4 and/or TLR2. For acute aerobic exercise, 40% of studies showed a decrease and 40% reported an increase. Regarding combined exercise, 50% of studies reported a reduction in the expression of the receptors

showed that resistance exercise (RE), whether acute or chronic, could act as a regulator of inflammation. In this subset of the literature, we observed no increases in the expression of TLR4 and/or TLR2 or pro-inflammatory cytokines after exercise.

Some studies [30, 124, 125] corroborate the results of this review and suggest that CRE may have anti-inflammatory effects. In contrast, ARE may stimulate changes in metabolic demand and promote inflammatory responses, whose occurrences is fundamentally determined by the exercise protocol [126, 127]. In this analysis, ARE transiently increases circulating levels of CK and pro-inflammatory cytokines, e.g., TNF [126] and IL1β [127]. Some studies that were not eligible for this review [128, 129] have shown that ARE induced microdamage in the skeletal muscle, along with an increase in inflammation markers such as IL-6, IL-8, monocyte chemotactic protein-1 (MCP-1), CK, and CRP when performed at high levels of stress.

The ten eligible studies of CRE and ARE [8, 29, 88, 92, 116, 119–123], tested different frequencies, intensities, and durations of exercise, none of these methods, however, produced changes in levels of TLR2 and/or TLR4. In these studies, intensities ranged from 60 to 80% of 1 RM with a gradual increase [119], or 6–14 RM [121]. In one study [92], the training volume followed a criterion of progression. Another study [120] used 80, 90, and 95% of maximal volitional strength capacity (MVSC), with low training volume as the criterion.

Regarding the inflammation markers that were subjected to the analysis here, neither acute nor chronic RE increased levels of pro-inflammatory cytokines such as TNF-α or IL-6. Eight studies tested TNF-α, and the majority [8, 88, 92, 116, 120, 122] found a significant decline of this cytokine. Two studies [29, 119] found no difference in this marker. Four studies analyzed levels of IL-6 after RE. Two studies [8, 116] found a drop in levels, but no significant difference appeared in the studies by Zanchi et al. [120] and McFarlin et al. [29].

The results showed that the RE protocols for both chronic and acute training adopted by the authors did not generate a pro-inflammatory response. Instead, three studies analyzed by this review [92, 119, 120] established an inverse relationship between the TLR2 and TLR4

Table 4 Modulation of TLR2 and TLR4 after chronic resistance exercise

Authors	Sample	Disease	Frequency, intensity, and duration	Post-exercise results		
				TLR	Cytokine	Other
Zanchi et al. 2010 [120]	Wistar rats	No disease	2 days/week, 80–95% MVSC, 12 weeks	↓TLR4	↓TNF-α ↔IL-6 ↔IL-10 ↑IL-10/TNF-α ratio ↔IL-15	↔Hsp70
Cheng et al. 2015 [116]	Adults	Low back pain	3 days/week, no information, 4 weeks	↓TLR4	↓TNF-α ↓IL-6 ↓IFN-γ ↓IL-1β ↓IL-8	↓NF-kBp65, ↓p53, ↑SIRT1, ↑PGC-1α, ↑PPAR-γ, ↑FoxO1, ↑FoxO3, ↑IKB, ↑SOD
Rodriguez-Miguelez et al. 2014 [119]	Elderly	No disease	2 days/week, 60–80% 1 RM, 8 weeks	↓TLR2, ↓TLR4	↔TNF-α ↑IL-10	↓MyD88, ↓p65, ↓phospho-p38/p38, ↓IKKi/IKKε, ↓TRIF, ↓phospho-IRF3/IRF3, ↓phospho-IRF7/IRF7, ↓Hsp60, ↑Hsp70, ↑phospho-ERK1/2, ↓CRP.
Rodriguez-Miguelez et al. 2015 [92]	Elderly	No disease	2 days/week, 20–35 Hz, 8 week	↓TLR2, ↓TLR4	↓TNF-α ↑IL-10	↓MyD88, ↓p65, ↓TRIF, ↓Hsp60, ↑Hsp70, ↓CRP
Prestes et al. 2015 [121]	Elderly	No disease	2 days/week, 6–14 RM, 16 weeks	↔TLR4	↔IL-1β ↔IL-10 ↔IL-1ra ↔ IL-15	↔BDNF, ↔irisin ↑functional capacity, ↑neuromuscular function, ↔body composition
Phillips et al. 2012 [8]	Elderly	Obesity	3 days/week, 8–12 RM, 12 weeks	↔TLR4	↓TNF-α ↓IL-6 ↑IL10	↓CRP, ↓leptin, ↑LPS-IL10, ↑LPS-TNF, ↔body composition

Table 5 Modulation of TLR2 and TLR4 after acute resistance exercise

Authors	Sample	Disease	Intensity and duration	Post-exercise results		
				TLR	Cytokine	Other
Millard et al. 2013 [123]	Adults	No disease	120–150 beats/min, 68.8 s (up and down 150 steps)	↔TLR2	↔IFN-γ	↑CD3−/CD56+NK, ↓NK CD56bright Short exercise did not affect NK cytotoxicity.
Fernandez-Gonzalo et al. 2012 [88]	Adults	No disease	40–50 MVIC, 18 acute eccentric bouts	↓TLR4	↓TNF-α	↓CD14, ↓MyD88, ↓TRIF, ↓TRAF6, ↓p65, ↓phospho-IκB, ↓phospho-ERK1/2, ↓CRP, 2 h after the 2nd acute session.
Fernandez-Gonzalo et al. 2014 [122]	Adults	No disease	40–50 MVIC, 18 acute eccentric bouts	↓TLR4	↓TNF-α	↓CD14, ↓MyD88, ↓TRIF, ↓TRAF6, ↓p65, ↓phospho-IκB, ↓phospho-ERK1/2, ↓CRP, 2 h after the 2nd acute session.
McFarlin et al. 2004 [29]	Elderly	No disease	80% 1 RM, 1 bout/3 sets/10 repetitions	↓TLR4	↔TNF-α ↔IL-6 ↔IL-1β	↔CD14

receptors and IL-10. In the five studies that investigated IL-10 with RE, four [8, 92, 119, 120] found an increase in this marker and one study found no significant difference [121]. It is known that IL-10 levels are higher after chronic exercise, and this anti-inflammatory cytokine acts as a natural TNF-α antagonist [106, 109].

Aerobic Exercise and Inflammation

A total of 12 articles verified that TLR4 and TLR2 undergo changes in response to CAE (Table 6). Four studies verified a significant decrease in TLR4 and/or TLR2 [13, 15, 76, 115] in terms of protein levels, two studies [117, 130] showed reductions in mRNA expression, and one indicated decreases at both the gene and protein level [14]. Two studies [74, 114] revealed an increase in TLR4 and/or TLR2 (gene and protein), one study reported increased mRNA expression [131], and two studies [113, 132] did not find any significant difference in TLR4 expression.

In 15 studies, a relationship between AAE and TLR2 and/or TLR4 was identified (Table 7). Three studies [133–135] found a significant reduction of TLR4 and/or TLR2 (protein levels), and two revealed a decrease in mRNA expression [136, 137]. Four studies [35, 39, 40, 42] found an increase in the protein levels of these receptors, and two studies [36, 37] increased mRNA expression. One study did not find a significant difference [138], and one study reported a significant decline in TLR4 (mRNA expression) in multiple sclerosis but found no difference in cases of fibromyalgia [118]. Two studies [139, 140] did not analyze TLR2 or TLR4 expression.

As demonstrated by the results from the analysis of TLR2 and TLR4 behavior, this review showed that in 23% of all of the articles that were analyzed, AE was associated with increases in inflammation. These results differ from previous studies that tested the expression of these receptors in RE. Ten months of CAE was more effective than strength and flexibility exercises in reducing inflammatory markers such as CRP, IL-6, and IL-18 in the elderly [141].

Most studies found that CAE reduced the levels of TLR2 and/or TLR4 [13–15, 76, 115, 117, 130]. However, the major immunological benefits came with exercise performed at a moderate intensity [13–15, 76, 117, 130, 132]. On the other hand, Zheng et al. [131] observed an increase in TLR2 (gene expression) and inflammatory cytokines such as TNF-α and IL-6 in the regular moderate intensity exercise group (badminton), with or without stimulation from microbial antigens. However, cytokine levels were suppressed after non-microbial antigen stimulation. The authors attributed this result to possible improvements in the body's resistance to invasion by pathogens in response to regular exercise, indicating that an increase of these receptors does not necessarily indicate a negative impact on health, though further research is still needed to address this possibility.

The chronic low-grade inflammatory profile (CLIP) is a common feature of the normal aging process, and it is also involved in the pathogenesis of several age-related diseases [142]. CLIP has already been recognized as a factor that plays a causative role in the development of sarcopenia. TNF-α and IL-6 are the most commonly reported inflammatory parameters in these studies [143]. Additionally, human aging is associated with metabolic endotoxemia and high levels of signaling of the RST4-NFkB-MAPK pathway in the muscle. These factors may play a role in the types of insulin resistance mediated by aging and muscle loss [74]. In this analysis, Ghosh et al. [74] observed an increase in TLR4 (mRNA and protein levels) in older people but not in younger participants. The study examined people engaged in a progressive

Table 6 Modulation of TLR2 and TLR4 after chronic aerobic exercise

Authors	Sample	Disease	Frequency, intensity, and duration	Post-exercise results		
				TLR	Cytokine	Other
Ma et al. 2013 [13]	Wistar rats	Cerebral ischemia	5 days/week, 12 m/min, 3 days–2 weeks	↓TLR4, ↓TLR2		↓NFkB e ↓MyD88
Lira et al. 2010 [76]	Wistar rats	No disease	5 days/week, 15–25 m/min, 11 weeks	↓ TLR4 (TR group), ↑TLR4 (R group)	TR group: ↔ TNF-α ↔IL-6 ↔IL-10	TR group (trained) ↓NFkBp65. OT group (overtrained) and R (resting overtrained): ↓performance decline, ↓testosterone, ↑corticosterone, ↑endotox. ↑IL-6, ↑IL-10, ↑NFkBp65
Fashi et al. 2015 [117]	Wistar rats	Pulmonary infection	5 days/week, mean speed of the group workload, 4 weeks	↓TLR4	↓TNF-α	↓NF-kB (exe group+PM10)
Jun et al. 2014 [130]	Sprague-Dawley rats	Ovariectomized rats	5 days/week, 18–26 m/min, 16 weeks	↓TLR4	↓TNF-α ↔IL-6	↓MCP-1 in adipose tissue (moderate trained group).
Holland et al. 2015 [132]	Sprague-Dawley rats	No disease	1/day, 30 m/min, 10 days	↔TLR4	↓TNF-α ↔IL-6 ↔IFNy ↔ IL10	Moderate training: ↔NFkB, ↔CCL2, ↔IL10, ↔NFkBp65
Zwagerman et al. 2010 [14]	Sprague-Dawley rats	Stroke	5 days/week, 30 m/min, 3 weeks	↓TLR4		↓Cerebral infarction volume
Choi et al. 2014 [15]	Sprague-Dawley rats	Alzheimer's disease	5 days/week, 2–8 m/min, 6 weeks	↓TLR4	↓TNF-α ↓IL-1α	↓NF-kB, in the STZ-exe group. ↑Cognitive function
Zheng et al. 2015 [131]	Adults (members of a university badminton club)	No disease	3 days/week, no information, 26–32 days	↑TLR2, ↔TLR4, with or without microbial antigen stimulation	↑TNF-α ↑IL-6 with or without microbial antigen stimulation	
Robinson et al. 2015 [115]	Adults	Pre-diabetes	1/day, 65–90% peak heart rate, 2 weeks	↓TLR4, ↓TLR2	↔TNF-α ↔IL-6 ↔IL-1β ↔ IL10	↓Fasting glucose in group MICT (moderate-intensity continuous training).
Nickel et al. 2011 [114]	Adults (amateur marathon runners)	Obesity	Training documented with respect to intensity, duration, and kilometers run per week by a written individual protocol, 10 weeks	↑TLR2 in LNE group (lean-non-elite). ↑TLR4 (all groups).	↔TNF-α	↑oxLDL in LE (lean-elite); ↓oxLDL in ONE (obese-non-elite).
Nickel et al. 2012 [113]	Adults (amateur marathon runners)	Obesity	Training documented with respect to intensity, duration, and kilometers run per week by a written individual protocol, 10 weeks	↔TLR2, ↔TLR4	↑TNF-α (24 h post-marathon) ↑IL-6 and ↑IL-10 (immediately after the run)	↑BDCA-1, ↓BDCA2, ↓TLR7, ↑PCR, ↔oxLDL
Ghosh et al. 2015 [74]	Adults and elderly	No disease	3–4 days/week, 65–80% VO2max, 16 weeks	↑TLR4 (aged individuals)		In elderly: ↑NF-kBp65, ↑NF-kBp50, ↑pJNK, ↑endotoxin, ↔pERK, ↔p-p38, ↑insulin resistance.

regime of the intensity and volume of training, ranging from 65 to 80% of VO2max, and an increase in the duration and number of sessions. Their results provide evidence that higher LPS flow in the elderly can play a critical role in age-related sarcopenia and insulin resistance.

Studies that did not fit our criteria [54, 58, 144, 59, 145] suggested that CAE performed under conditions of

Table 7 Modulation of TLR2 and TLR4 after acute aerobic exercise

Authors	Sample	Disease	Intensity and duration	Post-exercise results		
				TLR	Cytokine	Other
Rosa et al. 2011 [40]	Wistar rats	No disease	70% VO2max, 50 min	↑TLR4		↑MyD-88, ↑TRAF6, ↑NF-kBp65
Rodriguez-Miguelez et al. 2015 [39]	Wistar rats	No disease	16 m/min, 90 min/18 bouts/5 min/bout	↑TLR4	↑TNF-α ↑IL α-1β	↑HIF-1α, ↑VEGF, ↑eNOS, ↑MPO.
Liao et al. 2010 [136]	Sprague-Dawley rats	No disease	25 m/min, 1–2 h	↓TLR4	↑TNF-α	↑TNF-α, ↑NFkB, ↑p65, ↑ROS, ↑endotoxina
Zbinden-Foncea et al. 2012 [42]	Mice	No disease	70% of FCmax, two bouts of 60 min	↑TLR2, ↑TLR4		↑NEFA, ↑p38MAPK, ↑JNK.
Tanaka et al. 2010 [138]	Mice	No disease	9 m/min to exhaustion, 1 acute bout	↔TLR4	↓TNF-α	
Ortega et al. 2009 [140]	Adults	No disease	70% VO2 max,1 h			Hsp72-induced stimulation of neutrophil chemotaxis disappeared when TLR2 was blocked.
Lancaster et al. 2005 [133]	Adults	No disease	65% VO2max, 1.5 h	↓TLR4, ↓TLR2	↓IL-6	
Booth et al. 2010 [35]	Adults	No disease	60 km distance in the cycle the fastest possible time. Heart rate (bpm) and power output (watts) were monitored	↑TLR2, ↑TLR4		↓HLA.DR
Simpson et al. 2009 [135]	Adults	No disease	75% VO2max, 45 min	↓TLR4, ↓TLR2		↓HLA.DR
Neubauer et al. 2013 [37]	Adults	No disease	Borg 6–20, 10 km	↑TLR4	↑IL-6 ↔IL-1β ↑IL-10 ↑IL-1ra	↑IRAK3, ↑creatin kinase 3 h after, ↑plasma myoglobin 3 h after, ↑neutrophil 3 h after
Oliveira and Gleeson 2010 [134]	Adults	No disease	75% VO2peak, 1.5 h	↓TLR4		TLR4 returned to basal levels within 4 h after exercise, ↔TLR2.
Radom-Aizik et al. 2014 [137]	Adults	No disease	82% VO2peak, 2-min bouts	↓TLR4	↓TNF-α	↓CD36 e ↑EREG genes and ↑CXCR4
Light et al. 2009 [36]	Adults	Chronic fatigue syndrome	70% age-predicted maximal heart rate, 5–9 min	↑TLR4	↑IL6 ↑IL1β ↑IL-10 ↑IL13 ↑IL8 ↑IL12	↑Pain ↑mental fatigue. ↑α2-A, ↑RNAm of β-2 receptor in leucocytes, ↑COMT RNAm
White et al. 2012 [118]	Adults	Multiple sclerosis (ME) and fibromyalgia (SDC)	70% of age-predicted maximal heart rate, 20 min	ME: ↓TLR4 SDC: ↔TLR4	ME: after 8 h ↔IL-6 ↑IL-10 SDC: after 48 h ↔IL-6 ↑IL-10	↑Fatigue ↑pain, ↑adrenergic receptors.
Li and Geib 2013 [139]	Adults and elderly	No disease	1 h Tai Chi		↑IL-13	↓CD14+CD16+

high stress leads to inflammation in participants of all ages. They observed that long-distance runners might have increased levels of atherosclerosis and coronary heart diseases due to a training regime that went uninterrupted over many years [54]. Additionally, endotoxemia was found in 68% of athletes after a long-distance triathlon, and LPS levels were associated with higher levels of CRP [75]. A recent study showed that 24 h of continuous ultramarathon activity resulted in a higher level of LPS and increased levels of circulating

pro-inflammatory cytokines [146]. In fact, prolonged intense physical exercise leads to elevated concentrations of counter-regulatory hormones in plasma such as cortisol and catecholamines related to low immunity [147]. In addition, high levels of muscle oxidative stress lead to an excessive production of ROS and inflammation [60]. In contrast, regular moderate physical exercise can compensate for oxidative stress [148].

Short acute sessions of physical exercise may disturb homeostasis and increase inflammation [41], as verified by some of the articles reviewed here [35, 37, 39, 40, 42]. With the exception of the study by Light et al. [36], which tested an AAE protocol at moderate intensity and in samples obtained from individuals with disease, studies based on different strenuous exercise protocols consistently led to increases in TLR4, TLR2, and pro-inflammatory cytokines [35, 37, 39, 40, 42]. Rodrigues-Migueles et al. [39] found an increase in TLR4 (protein) and pro-inflammatory cytokines in AAE sessions. However, all of these effects were extinguished by CAE through a weekly exercise protocol of increasing intensity and duration.

In studies which reported increases in TLR2, TLR4, and pro-inflammatory cytokines after acute sessions, IL-10 was tested in only three experiments, all of which revealed a significant increase in the expression of this cytokine [36, 37, 118]. This was probably caused by a transient increase in IL-6 which then led to a subsequent increase in levels of IL-10 [104, 106]. However, other studies [133–135] indicated that AAE had beneficial effects, as observed through a decline in terms of protein levels of TLR2 and/or TLR4 and at the mRNA expression [118, 137]. Radom-Aizik et al. [137] verified that AAE not only prevents the normal effects of aging in terms of atherosclerosis but also reduces its symptoms in a manner that promotes cardiovascular health despite the global stress response that is generally evoked by this activity.

One exception is a study by Liao et al. [136], which showed a reduction in TLR4 (gene expression), but also showed an increase in inflammatory responses as exhibited by high levels of TNF-α, NF-kB, and LPS. The reason for the down-regulation of TLR4 is not clear, but the authors believe that this may be related to high levels of ROS. Here, from our review of the literature, we suggest that increases in circulating LPS and an excessive generation of ROS are the main actors in the acute inflammatory process generated by excessive AE. However, more studies are needed to complete the mechanistic picture that links these effects and other aspects of inflammatory responses in AE.

Combined Exercise and Inflammation

Only two studies [93, 149] relating TLR2 and/or TLR4 to CE (combining aerobic and resistance exercises in single sessions) were found. One study [93] demonstrated a significant decline in TLR4, and the other [149] did not find a difference in TLR4 (Table 8).

The Timmerman et al. [149] study analyzed the response of 12 weeks of exercise training on the part of aged, physically inactive subjects who performed AE for 20 min and RE for 30 min. No significant differences in TLR4 (protein expression) were found in the trained group compared to the controls, but a decline in TNF-α was observed. Stewart et al. [93] compared CE effects in adult and aged participants and showed a significant decline in TLR4 as well as IL-6 in the physically inactive groups compared to controls; however, levels of TLR2 were not significantly changed.

Another experiment [150] verified a decline in CRP in both trained and active control groups and concluded that AE and RE may be applied in the same session as a potential therapeutic intervention for adults and aged individuals to avoid some chronic diseases. Therefore, this review suggests that AE and RE in combination protect against the negative effects of AE.

Exercise, Disease, and Inflammation

The majority of the studies eligible for this review show that both AE [13–15, 113–115, 117] and RE [8, 116] can act as excellent auxiliary treatments for chronic disease. However, we found no article that tested ARE in samples from patients with diseases.

One of the important features of obesity-induced inflammation is a phenotypic change in the populations of macrophages and T cells present in the adipose tissue. This is reflected in levels of the production of anti- and pro-inflammatory cytokines [151]. It has been suggested that free saturated fatty acids can induce inflammation

Table 8 Modulation of TLR2 and TLR4 after combined exercise (aerobic and resistance)

Authors	Sample	Disease	Frequency, intensity, and duration	Post-exercise results		
				TLR	Cytokine	Other
Stewart et al. 2005 [93]	Adults and elderly	No disease	3 days/week, 70–80% 1 RM and 50–70% of heart rate reserve, 12 weeks	↓TLR4, ↔TLR2	↔TNF-α ↔IL-1β ↓IL-6	
Timmerman et al. 2008 [149]	Elderly	No disease	3 days/week, 70–80% 1 RM and 70–80% of heart rate reserve, 12 weeks	↔TLR4	↓TNF-α	

through the activation of macrophages, TLR2, and TLR4 in the adipose tissue, culminating in the activation of NF-kB and an increased expression of pro-inflammatory cytokines such as TNF-α or IL-6 [7, 9, 151].

The study by Phillips et al. [8] in post-menopausal obese women showed that CRE did not decrease TLR4 in terms of mRNA expression but reduced inflammatory markers such as TNF-α and IL-6. In another study related to obesity, 10 days of either moderate (MICT) or high intensity (HIIT) CAE in inactive overweight women promoted improvements in glucose control and cardiorespiratory capacity and a decrease in TLR2 and TLR4 (protein content) [115].

Most studies in this review that tested the levels of TLR2 and/or TLR4 receptors in a disease context used moderate load protocols, with the exception of the study by Nickel et al. [114], which studied marathon runners and found an increase in the mRNA expression and protein levels of these receptors. In this study, TLR2 was significantly increased in lean-non-elite athletes when compared to the obese-non-elite and lean-elite groups, and TLR4 increased in all groups in response to exercise. However, levels of the systemic cytokines TNF-α and IL-6 remained stable. Interestingly, oxidized low-density lipoprotein (oxLDL) levels in obese athletes were reduced and associated with higher adiponectin levels, in contrast to increased levels of oxLDL found in the group of lean-elite athletes [114]. This can be understood from the fact that TLR4 plays a crucial role in cellular responses to oxLDL exposure and the activation of NF-κB [152, 153]. Wang et al. [152] showed that the activation of the TLR4/NF-κB signaling pathway was a potential mechanism for oxLDL-induced apoptosis in cardiomyocytes.

Higher levels of this low-density lipoprotein (LDL) are usually associated with an increased risk for atherosclerosis [114], and marathon runners may, in fact, have increased levels of atherosclerosis [54]. LDL, when modified by enzymes such as phospholipases, gives rise to oxidized low-density lipoprotein (oxLDL), which contributes to the formation and progression of atherosclerotic plaques [152, 154]. oxLDL is known to be immunogenic and activates endothelial cells, monocytes, macrophages, and T cells [155]. Furthermore, oxLDL is toxic at higher concentrations and thus could be a cause of cell death in lesions [156]. The plasma level of oxLDL was shown to be a predictor of mortality in patients with chronic congestive heart failure [157] and induced severe cell damage in ventricular myocytes [158].

This review also found articles that generally analyzed TLR2 and/or TLR4 expression in relation to other diseases. The study by Zwagerman et al. [14], for example, found that in addition to reduced levels of TLR4 (gene and protein), CAE reduced the frequency of cerebral infarction. Another study [36] analyzed chronic fatigue syndrome in acute AE sessions at moderate intensity for 25 min. In addition to an increase in the mRNA expression of TLR4 and pro-inflammatory cytokines, symptoms such as pain and physical and mental fatigue became worse after exercise, suggesting a dysregulation of the immune and sympathetic nervous systems.

Conclusions

This is the first systematic review of the literature that addresses the roles of TLR2 and TLR4 receptors in various types of exercise. Our main finding is evidence for an accentuation in the inflammatory processes orchestrated by these receptors in both AAE and CAE. The results also suggest that the expression of the receptors is correlated with that of anti- and pro-inflammatory cytokines. Taken together, these data open new perspectives for studies aimed at a better understanding of the response of inflammatory processes to physical exercise.

An analysis of the pathways involving TLR2 and TLR4 reveal something about the way specific types of physical exercise are related to differences in the types of inflammatory responses they stimulate. The results indicate that AE is potentially inflammatory; a smaller number of studies revealed that acute exercise has anti-inflammatory effects, compared to studies of chronic exercise.

Our analysis showed that in RE, TLR2 and TLR4 expression and signaling adopt an anti-inflammatory pattern. Studies that met our criteria for inclusion indicated that acute or chronic sessions reduced TLRs as well as inflammatory cytokines, particularly TNF-α, and promoted increases in IL-10, which can be considered a beneficial adaptation for both healthy people and those affected by certain diseases.

The same results were obtained when differences in the populations and intensities of exercise were taken into account. This indicates that RE can be broadly used to prevent or minimize the potentially deleterious effects of TLR expression and that the intensity can be manipulated to achieve other goals, such as increasing body strength, without a loss of benefits vis-à-vis the overall inflammatory profile.

For AE, the intensity of exercise is a crucial factor—better responses were achieved under moderate intensities. But overall, whether the effects of AE will be positive or negative depends on a person's other physiological characteristics, so they must be taken into account.

Generally, CE seems to be a good choice in most situations due to its positive effects on TLR expression and signaling. In other words, the possible negative "side effects" of AE can be overcome through the positive impact of RE. This combination of training strategies appears to improve a person's general inflammatory profile while maintaining the cardiovascular and metabolic

benefits of AE. In most cases, this leads to better adaptations. But because the number of studies addressing the effects of TLR2 and TLR4 in CE is very small, further research is needed for both amateurs and elite athletes.

Abbreviations

AAE: Acute aerobic exercise; AE: Aerobic exercises; ARE: Acute resistance exercise; Arg1: Arginase-1; CAE: Chronic aerobic exercise; CE: Combined exercise; CK: Creatine kinase; CRE: Chronic resistance exercise; CRP: C-reactive protein; DAMPs: Damage-associated molecular patterns; IGF-1: Insulin-like growth factor 1; LPS: Lipopolysaccharides; MAPK: Mitogen-activated protein kinase; PAMPS: Pathogen-associated molecular patterns; RE: Resistance exercise; ROS: Reactive oxygen species; TLR: Toll-like receptor; TLR2: Toll-like receptor 2; TLR4: Toll-like receptor 4; TNF-α: Tumor necrosis factor alpha; VEGF: Vascular endothelial growth factor

Acknowledgments

We thank Russ Hodge for assisting in writing the manuscript.

Availability of Data and Materials

Not applicable.

Authors' Contributions

PAMC: substantial contributions to the conception; design and drafting of the work; survey of the literature; preparation of tables and creation of figures; analysis; interpretation of data; critical review. MFG: contributions to the conception; review of tables; analysis; critical review. JSH: contributed to the analysis; English translation; critical review. FHO: contributions to the conception; analysis; critical review. RCA: contributions to the conception; analysis, interpretation of data; critical review. All authors read and approved the final manuscript.

Funding

The authors have received no payment or services from a third party for any aspect of the submitted work.

Authors' Information

PAMC: Ph.D. Student in Medicine (Nephrology) Program at Federal University of São Paulo, Department of Biophysics at Federal University of São Paulo, and Laboratory of Exercise Genetics and Metabolism at Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil with a collaborative period in Berlin, Germany at Max-Delbrück-Centrum für Molekulare Medizin; Master of Physical Education at São Judas Tadeu University, São Paulo, SP, Brazil. **MFG**: Ph.D. Student in Molecular Biology Program at Federal University of São Paulo, Department of Biophysics at Federal University of São Paulo, and Laboratory of Exercise Genetics and Metabolism at Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil; Master of Molecular Biology at Federal University of São Paulo, São Paulo, SP, Brazil. **JSH**: Master Degree Student in Neurology/Neuroscience Program at Federal University of São Paulo and Exercise Neurophysiology Laboratory at Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil; Graduation in Physical Education at University of Santos (UNIMES), São Paulo, SP, Brazil. **FHO**: Ph.D. Student in Medicine (Nephrology) Program at Federal University of São Paulo, Department of Biophysics at Federal University of São Paulo, and Laboratory of Exercise Genetics and Metabolism at Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil; Master of Health Sciences at Faculty of Medical Sciences of Santa Casa (FCMSCSP), São Paulo, SP, Brazil. **RCA**: Post-doctorate in Biological Sciences at Federal University of São Paulo (UNIFESP), SP, Brazil; Ph.D. in Biological Sciences (Molecular Biology) at Federal University of São Paulo (UNIFESP), SP, Brazil with a collaborative period in Berlin, Germany at Max-Delbrück-Centrum für Molekulare Medizin; Master of Biological Sciences (Physiology and Pharmacology) at Federal University of Minas Gerais (UFMG), MG, Brazil. He is currently a professor of the Biophysics Department of the Federal University of São Paulo (EPM). He has an experience in the field of Biophysics, working mainly on the following subjects: physical exercise, obesity, kinins, transgenic animals, kallikrein, inflammation, and blood pressure.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

Paula Andréa Malveira Cavalcante, Marcos Fernandes Gregnani, Jessica Salles Henrique, Fábio Henrique Ornellas, and Ronaldo Carvalho Araújo declare that they have no conflicts of interest related to this manuscript.

Publisher's Note

Author details

[1]Medicine (Nephrology) Program, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil. [2]Molecular Biology Program, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil. [3]Laboratory of Exercise Genetics and Metabolism, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil. [4]Department of Biophysics, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil. [5]Neurology/Neuroscience Program, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil. [6]Exercise Neurophysiology Laboratory, Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil. [7]Rua Pedro de Toledo, 669/9and., 04039-032, São Paulo, SP, Brazil.

Received: 4 September 2017 Accepted: 15 November 2017
Published online: 28 November 2017

References

1. Bermudez EA, Rifai N, Buring J, Manson JE, Ridker PM. Interrelationships among circulating interleukin-6, C-reactive protein, and traditional cardiovascular risk factors in women. Arterioscler Thromb Vasc Biol. 2002; 22(10):1668–73.
2. Gleeson M, McFarlin B, Flynn M. Exercise and Toll-like receptors. Exerc Immunol Rev. 2006;12:34–53.
3. McFarlin BK, Flynn MG, Campbell WW, Craig BA, Robinson JP, Stewart LK, et al. Physical activity status, but not age, influences inflammatory biomarkers and Toll-like receptor 4. J Gerontol Ser A Biol Med Sci. 2006; 61(4):388–93.
4. Scheidt-Nave C, Bismar H, Leidig-Bruckner G, Woitge H, Seibel MJ, Ziegler R, et al. Serum interleukin 6 is a major predictor of bone loss in women specific to the first decade past menopause. J Clin Endocrinol Metab. 2001; 86(5):2032–42. https://doi.org/10.1210/jcem.86.5.7445.
5. Daar AS, Singer PA, Persad DL, Pramming SK, Matthews DR, Beaglehole R, et al. Grand challenges in chronic non-communicable diseases. Nature. 2007;450(7169):494–6. https://doi.org/10.1038/450494a.
6. Nathan C. Points of control in inflammation. Nature. 2002;420(6917):846–52. https://doi.org/10.1038/nature01320.
7. Fessler MB, Rudel LL, Brown JM. Toll-like receptor signaling links dietary fatty acids to the metabolic syndrome. Curr Opin Lipidol. 2009;20(5):379–85. https://doi.org/10.1097/MOL.0b013e32832fa5c4.
8. Phillips MD, Patrizi RM, Cheek DJ, Wooten JS, Barbee JJ, Mitchell JB. Resistance training reduces subclinical inflammation in obese, postmenopausal women. Med Sci Sports Exerc. 2012;44(11):2099–110. https://doi.org/10.1249/MSS.0b013e3182644984.
9. Oliveira AG, Carvalho BM, Tobar N, Ropelle ER, Pauli JR, Bagarolli RA, et al. Physical exercise reduces circulating lipopolysaccharide and TLR4 activation and improves insulin signaling in tissues of DIO rats. Diabetes. 2011;60(3): 784–96. https://doi.org/10.2337/db09-1907.
10. Miura K, Ohnishi H. Role of gut microbiota and Toll-like receptors in nonalcoholic fatty liver disease. World J Gastroenterol. 2014;20(23):7381–91. https://doi.org/10.3748/wjg.v20.i23.7381.
11. Dybdahl B, Wahba A, Lien E, Flo TH, Waage A, Qureshi N, et al. Inflammatory response after open heart surgery: release of heat-shock protein 70 and signaling through toll-like receptor-4. Circulation. 2002;105(6):685–90.
12. Laberge MA, Moore KJ, Freeman MW. Atherosclerosis and innate immune signaling. Ann Med. 2005;37(2):130–40. https://doi.org/10.1080/07853890510007304.
13. Ma Y, He M, Qiang L. Exercise therapy downregulates the overexpression of TLR4, TLR2, MyD88 and NF-κB after cerebral ischemia in rats. Int J Mol Sci. 2013;14(2):3718–33. https://doi.org/10.3390/ijms14023718.
14. Zwagerman N, Plumlee C, Guthikonda M, Ding Y. Toll-like receptor-4 and cytokine cascade in stroke after exercise. Neurol Res. 2010;32(2):123–6. https://doi.org/10.1179/016164109x12464612122812.

15. Choi DH, Kwon IS, Koo JH, Jang YC, Kang EB, Byun JE, et al. The effect of treadmill exercise on inflammatory responses in rat model of streptozotocin-induced experimental dementia of Alzheimer's type. Journal of exercise nutrition & biochemistry. 2014;18(2):225–33. https://doi.org/10.5717/jenb.2014.18.2.225.
16. Pierer M, Rethage J, Seibl R, Lauener R, Brentano F, Wagner U, et al. Chemokine secretion of rheumatoid arthritis synovial fibroblasts stimulated by Toll-like receptor 2 ligands. Journal of immunology (Baltimore, Md: 1950). 2004;172(2):1256–65.
17. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol. 2010;11(5):373–84. https://doi.org/10.1038/ni.1863.
18. Mukherjee S, Karmakar S, Babu SP. TLR2 and TLR4 mediated host immune responses in major infectious diseases: a review. Braz J Infect Dis. 2016;20(2):193–204. https://doi.org/10.1016/j.bjid.2015.10.011.
19. Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity. 2011;34(5):637–50. https://doi.org/10.1016/j.immuni.2011.05.006.
20. Banchereau J, Steinman RM. Dendritic cells and the control of immunity. Nature. 1998;392(6673):245–52. https://doi.org/10.1038/32588.
21. Means TK, Golenbock DT, Fenton MJ. The biology of Toll-like receptors. Cytokine Growth Factor Rev. 2000;11(3):219–32.
22. Medzhitov R. Toll-like receptors and innate immunity. Nat Rev Immunol. 2001;1(2):135–45. https://doi.org/10.1038/35100529.
23. Cristofaro P, Opal SM. Role of Toll-like receptors in infection and immunity: clinical implications. Drugs. 2006;66(1):15–29.
24. Ghosh S, May MJ, Kopp EB, NF-kappa B. Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol. 1998;16:225–60. https://doi.org/10.1146/annurev.immunol.16.1.225.
25. Pahl HL. Activators and target genes of Rel/NF-kappaB transcription factors. Oncogene. 1999;18(49):6853–66. https://doi.org/10.1038/sj.onc.1203239.
26. Siebenlist U, Franzoso G, Brown K. Structure, regulation and function of NF-kappa B. Annu Rev Cell Biol. 1994;10:405–55. https://doi.org/10.1146/annurev.cb.10.110194.002201.
27. Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell. 2006;124(4):783–801. https://doi.org/10.1016/j.cell.2006.02.015.
28. Chang ZL. Important aspects of Toll-like receptors, ligands and their signaling pathways. Inflamm Res. 2010;59(10):791–808. https://doi.org/10.1007/s00011-010-0208-2.
29. McFarlin BK, Flynn MG, Campbell WW, Stewart LK, Timmerman KL. TLR4 is lower in resistance-trained older women and related to inflammatory cytokines. Med Sci Sports Exerc. 2004;36(11):1876–83.
30. Flynn MG, McFarlin BK, Phillips MD, Stewart LK, Timmerman KL. Toll-like receptor 4 and CD14 mRNA expression are lower in resistive exercise-trained elderly women. J Appl Physiol. 2003;95(5):1833–42. https://doi.org/10.1152/japplphysiol.00359.2003.
31. Lee WJ. IGF-I exerts an anti-inflammatory effect on skeletal muscle cells through down-regulation of TLR4 signaling. Immune network. 2011;11(4):223–6. https://doi.org/10.4110/in.2011.11.4.223.
32. Petersen AM, Pedersen BK. The anti-inflammatory effect of exercise. J Appl Physiol. 2005;98(4):1154–62. https://doi.org/10.1152/japplphysiol.00164.2004.
33. Taaffe DR, Harris TB, Ferrucci L, Rowe J, Seeman TE. Cross-sectional and prospective relationships of interleukin-6 and C-reactive protein with physical performance in elderly persons: MacArthur studies of successful aging. J Gerontol A Biol Sci Med Sci. 2000;55(12):M709–15.
34. Walsh NP, Gleeson M, Pyne DB, Nieman DC, Dhabhar FS, Shephard RJ, et al. Position statement. Part two: maintaining immune health. Exerc Immunol Rev. 2011;17:64–103.
35. Booth S, Florida-James GD, McFarlin BK, Spielmann G, O'Connor DP, Simpson RJ. The impact of acute strenuous exercise on TLR2, TLR4 and HLA.DR expression on human blood monocytes induced by autologous serum. Eur J Appl Physiol. 2010;110(6):1259–68. https://doi.org/10.1007/s00421-010-1616-2.
36. Light AR, White AT, Hughen RW, Light KC. Moderate exercise increases expression for sensory, adrenergic, and immune genes in chronic fatigue syndrome patients but not in normal subjects. J Pain. 2009;10(10):1099–112. https://doi.org/10.1016/j.jpain.2009.06.003.
37. Neubauer O, Sabapathy S, Lazarus R, Jowett JBM, Desbrow B, Peake JM, et al. Transcriptome analysis of neutrophils after endurance exercise reveals novel signaling mechanisms in the immune response to physiological stress. J Appl Physiol. 2013;114(12):1677–88. https://doi.org/10.1152/japplphysiol.00143.2013.
38. Nieman DC. Exercise, infection, and immunity. Int J Sports Med. 1994;15(Suppl 3):S131–41. https://doi.org/10.1055/s-2007-1021128.
39. Rodriguez-Miguelez P, Lima-Cabello E, Martinez-Florez S, Almar M, Cuevas MJ, Gonzalez-Gallego J. Hypoxia-inducible factor-1 modulates the expression of vascular endothelial growth factor and endothelial nitric oxide synthase induced by eccentric exercise. J Appl Physiol. 2015;118(8):1075–83. https://doi.org/10.1152/japplphysiol.00780.2014.
40. Rosa JC, Lira FS, Eguchi R, Pimentel GD, Venancio DP, Cunha CA, et al. Exhaustive exercise increases inflammatory response via Toll like receptor-4 and NF-kappaBp65 pathway in rat adipose tissue. J Cell Physiol. 2011;226(6):1604–7. https://doi.org/10.1002/jcp.22490.
41. Simpson RJ, Bosch JA. Special issue on exercise immunology: current perspectives on aging, health and extreme performance. Brain Behav Immun. 2014;39:1–7. https://doi.org/10.1016/j.bbi.2014.03.006.
42. Zbinden-Foncea H, Raymackers JM, Deldicque L, Renard P, Francaux M. TLR2 and TLR4 activate p38 MAPK and JNK during endurance exercise in skeletal muscle. Med Sci Sports Exerc. 2012;44(8):1463–72. https://doi.org/10.1249/MSS.0b013e31824e0d5d.
43. Francaux M. Toll-like receptor signalling induced by endurance exercise. Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme. 2009;34(3):454–8. https://doi.org/10.1139/h09-036.
44. Shek PN, Shephard RJ. Physical exercise as a human model of limited inflammatory response. Can J Physiol Pharmacol. 1998;76(5):589–97.
45. Moldoveanu AI, Shephard RJ, Shek PN. The cytokine response to physical activity and training. Sports Med. 2001;31(2):115–44. https://doi.org/10.2165/00007256-200131020-00004.
46. Nakajima T, Kurano M, Hasegawa T, Takano H, Iida H, Yasuda T, et al. Pentraxin3 and high-sensitive C-reactive protein are independent inflammatory markers released during high-intensity exercise. Eur J Appl Physiol. 2010;110(5):905–13. https://doi.org/10.1007/s00421-010-1572-x.
47. Okutsu M, Suzuki K, Ishijima T, Peake J, Higuchi M. The effects of acute exercise-induced cortisol on CCR2 expression on human monocytes. Brain Behav Immun. 2008;22(7):1066–71. https://doi.org/10.1016/j.bbi.2008.03.006.
48. Pedersen BK, Ostrowski K, Rohde T, Bruunsgaard H. The cytokine response to strenuous exercise. Can J Physiol Pharmacol. 1998;76(5):505–11.
49. Shephard RJ, Shek PN. Immune responses to inflammation and trauma: a physical training model. Can J Physiol Pharmacol. 1998;76(5):469–72.
50. Jeon J-W, Ha U-H, Paek S-H. In vitro inflammation inhibition model based on semi-continuous Toll-like receptor biosensing. PLoS One. 2014;9(8):e105212. https://doi.org/10.1371/journal.pone.0105212.
51. Weiss U. Inflammation. Nature. 2008;454(7203):427.
52. Chen YW, Apostolakis S, Lip GY. Exercise-induced changes in inflammatory processes: implications for thrombogenesis in cardiovascular disease. Ann Med. 2014;46(7):439–55. https://doi.org/10.3109/07853890.2014.927713.
53. Schnohr P, O'Keefe JH, Marott JL, Lange P, Jensen GB. Dose of jogging and long-term mortality: the Copenhagen City Heart Study. J Am Coll Cardiol. 2015;65(5):411–9. https://doi.org/10.1016/j.jacc.2014.11.023.
54. Lavie CJ, Lee DC, Sui X, Arena R, O'Keefe JH, Church TS, et al. Effects of running on chronic diseases and cardiovascular and all-cause mortality. Mayo Clin Proc. 2015;90(11):1541–52. https://doi.org/10.1016/j.mayocp.2015.08.001.
55. Sima AV, Stancu CS, Simionescu M. Vascular endothelium in atherosclerosis. Cell Tissue Res. 2009;335(1):191–203. https://doi.org/10.1007/s00441-008-0678-5.
56. Bjorkbacka H. Multiple roles of Toll-like receptor signaling in atherosclerosis. Curr Opin Lipidol. 2006;17(5):527–33. https://doi.org/10.1097/01.mol.0000245258.25387.ec.
57. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc. 2007;39(8):1423–34. https://doi.org/10.1249/mss.0b013e3180616b27.
58. Lavie CJ, O'Keefe JH, Sallis RE. Exercise and the heart—the harm of too little and too much. Current sports medicine reports. 2015;14(2):104–9. https://doi.org/10.1249/jsr.0000000000000134.
59. O'Keefe JH, Patil HR, Lavie CJ, Magalski A, Vogel RA, McCullough PA. Potential adverse cardiovascular effects from excessive endurance exercise. Mayo Clin Proc. 2012;87(6):587–95. https://doi.org/10.1016/j.mayocp.2012.04.005.
60. Lee S, Kim M, Lim W, Kim T, Kang C. Strenuous exercise induces mitochondrial damage in skeletal muscle of old mice. Biochem Biophys Res Commun. 2015;461(2):354–60. https://doi.org/10.1016/j.bbrc.2015.04.038.
61. Kong X, Thimmulappa R, Kombairaju P, Biswal S. NADPH oxidase-dependent reactive oxygen species mediate amplified TLR4 signaling and sepsis-

induced mortality in Nrf2-deficient mice. Journal of immunology (Baltimore, Md: 1950). 2010;185(1):569–77. https://doi.org/10.4049/jimmunol.0902315.
62. Imhoff BR, Hansen JM. Extracellular redox status regulates Nrf2 activation through mitochondrial reactive oxygen species. The Biochemical journal. 2009;424(3):491–500. https://doi.org/10.1042/bj20091286.
63. Sen CK, Packer L. Antioxidant and redox regulation of gene transcription. FASEB J. 1996;10(7):709–20.
64. Moriuchi H, Moriuchi M, Fauci AS. Nuclear factor-kappa B potently up-regulates the promoter activity of RANTES, a chemokine that blocks HIV infection. Journal of immunology (Baltimore, Md: 1950). 1997;158(7):3483–91.
65. Park HS, Jung HY, Park EY, Kim J, Lee WJ, Bae YS. Cutting edge: direct interaction of TLR4 with NAD(P)H oxidase 4 isozyme is essential for lipopolysaccharide-induced production of reactive oxygen species and activation of NF-kappa B. Journal of immunology (Baltimore, Md: 1950). 2004;173(6):3589–93.
66. Lorne E, Zmijewski JW, Zhao X, Liu G, Tsuruta Y, Park Y-J, et al. Role of extracellular superoxide in neutrophil activation: interactions between xanthine oxidase and TLR4 induce proinflammatory cytokine production. Am J Physiol Cell Physiol. 2008;294(4):C985–C93. https://doi.org/10.1152/ajpcell.00454.2007.
67. Alessio HM, Goldfarb AH, Cutler RG. MDA content increases in fast- and slow-twitch skeletal muscle with intensity of exercise in a rat. Am J Phys. 1988;255(6 Pt 1):C874–7.
68. Jackson SH, Devadas S, Kwon J, Pinto LA, Williams MS. T cells express a phagocyte-type NADPH oxidase that is activated after T cell receptor stimulation. Nat Immunol. 2004;5(8):818–27. https://doi.org/10.1038/ni1096.
69. Mougios V. Reference intervals for serum creatine kinase in athletes. Br J Sports Med. 2007;41(10):674–8. https://doi.org/10.1136/bjsm.2006.034041.
70. Mangner N, Linke A, Oberbach A, Kullnick Y, Gielen S, Sandri M, et al. Exercise training prevents TNF-α induced loss of force in the diaphragm of mice. PLoS One. 2013;8(1):e52274. https://doi.org/10.1371/journal.pone.0052274.
71. Becatti M, Mannucci A, Barygina V, Mascherini G, Emmi G, Silvestri E, et al. Redox status alterations during the competitive season in elite soccer players: focus on peripheral leukocyte-derived ROS. Intern Emerg Med. 2017; https://doi.org/10.1007/s11739-017-1653-5.
72. Baur DA, Bach CW, Hyder WJ, Ormsbee MJ. Fluid retention, muscle damage, and altered body composition at the Ultraman Triathlon. Eur J Appl Physiol. 2016;116(3):447–58. https://doi.org/10.1007/s00421-015-3291-9.
73. Bosenberg AT, Brock-Utne JG, Gaffin SL, Wells MT, Blake GT. Strenuous exercise causes systemic endotoxemia. J Appl Physiol. 1988;65(1):106–8.
74. Ghosh S, Lertwattanarak R, Garduno Jde J, Galeana JJ, Li J, Zamarripa F, et al. Elevated muscle TLR4 expression and metabolic endotoxemia in human aging. J Gerontol A Biol Sci Med Sci. 2015;70(2):232–46. https://doi.org/10.1093/gerona/glu067.
75. Jeukendrup AE, Vet-Joop K, Sturk A, Stegen JH, Senden J, Saris WH, et al. Relationship between gastro-intestinal complaints and endotoxaemia, cytokine release and the acute-phase reaction during and after a long-distance triathlon in highly trained men. Clinical science (London, England: 1979). 2000;98(1):47–55.
76. Lira FS, Rosa JC, Pimentel GD, Tarini VA, Arida RM, Faloppa F, et al. Inflammation and adipose tissue: effects of progressive load training in rats. Lipids Health Dis. 2010;9:109. https://doi.org/10.1186/1476-511x-9-109.
77. Beutler B, Mahoney J, Le Trang N, Pekala P, Cerami A. Purification of cachectin, a lipoprotein lipase-suppressing hormone secreted by endotoxin-induced RAW 264.7 cells. J Exp Med. 1985;161(5):984–95.
78. Espat NJ, Copeland EM, Moldawer LL. Tumor necrosis factor and cachexia: a current perspective. Surg Oncol. 1994;3(5):255–62.
79. Coppack SW. Pro-inflammatory cytokines and adipose tissue. Proc Nutr Soc. 2001;60(3):349–56.
80. Hotamisligil GS. Inflammation and metabolic disorders. Nature. 2006; 444(7121):860–7. https://doi.org/10.1038/nature05485.
81. Hotamisligil GS, Shargill NS, Spiegelman BM. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. Science (New York, NY). 1993;259(5091):87–91.
82. Lakier SL. Overtraining, excessive exercise, and altered immunity: is this a T helper-1 versus T helper-2 lymphocyte response? Sports medicine (Auckland, NZ). 2003;33(5):347–64.
83. Smith LL. Tissue trauma: the underlying cause of overtraining syndrome? Journal of strength and conditioning research / National Strength & Conditioning Association. 2004;18(1):185–93.
84. Hill EE, Zack E, Battaglini C, Viru M, Viru A, Hackney AC. Exercise and circulating cortisol levels: the intensity threshold effect. J Endocrinol Investig. 2008;31(7):587–91. https://doi.org/10.1007/bf03345606.
85. Kraemer WJ, Ratamess NA. Hormonal responses and adaptations to resistance exercise and training. Sports medicine (Auckland, NZ). 2005; 35(4):339–61.
86. Jin CH, Paik IY, Kwak YS, Jee YS, Kim JY. Exhaustive submaximal endurance and resistance exercises induce temporary immunosuppression via physical and oxidative stress. Journal of exercise rehabilitation. 2015;11(4):198–203. 10.12965/jer.150221.
87. Radak Z, Chung HY, Koltai E, Taylor AW, Goto S. Exercise, oxidative stress and hormesis. Ageing Res Rev. 2008;7(1):34–42. https://doi.org/10.1016/j.arr.2007.04.004.
88. Fernandez-Gonzalo R, De Paz JA, Rodriguez-Miguelez P, Cuevas MJ, Gonzalez-Gallego J. Effects of eccentric exercise on Toll-like receptor 4 signaling pathway in peripheral blood mononuclear cells. J Appl Physiol. 2012;112(12):2011–8. https://doi.org/10.1152/japplphysiol.01499.2011.
89. Ford ES. Does exercise reduce inflammation? Physical activity and C-reactive protein among U.S. adults. Epidemiology (Cambridge, Mass). 2002;13(5):561–8. https://doi.org/10.1097/01.ede.0000023965.92535.c0.
90. Giannopoulou I, Fernhall B, Carhart R, Weinstock RS, Baynard T, Figueroa A, et al. Effects of diet and/or exercise on the adipocytokine and inflammatory cytokine levels of postmenopausal women with type 2 diabetes. Metab Clin Exp. 2005;54(7):866–75. https://doi.org/10.1016/j.metabol.2005.01.033.
91. Oberbach A, Tonjes A, Kloting N, Fasshauer M, Kratzsch J, Busse MW, et al. Effect of a 4 week physical training program on plasma concentrations of inflammatory markers in patients with abnormal glucose tolerance. European journal of endocrinology / European Federation of Endocrine Societies. 2006;154(4):577–85. https://doi.org/10.1530/eje.1.02127.
92. Rodriguez-Miguelez P, Fernandez-Gonzalo R, Collado PS, Almar M, Martinez-Florez S, de Paz JA, et al. Whole-body vibration improves the anti-inflammatory status in elderly subjects through Toll-like receptor 2 and 4 signaling pathways. Mech Ageing Dev. 2015;150:12–9. https://doi.org/10.1016/j.mad.2015.08.002.
93. Stewart LK, Flynn MG, Campbell WW, Craig BA, Robinson JP, McFarlin BK, et al. Influence of exercise training and age on CD14+ cell-surface expression of toll-like receptor 2 and 4. Brain Behav Immun. 2005;19(5):389–97. https://doi.org/10.1016/j.bbi.2005.04.003.
94. Pedersen BK, Saltin B. Evidence for prescribing exercise as therapy in chronic disease. Scand J Med Sci Sports. 2006;16(Suppl 1):3–63. https://doi.org/10.1111/j.1600-0838.2006.00520.x.
95. Li J, Lee DSW, Madrenas J. Evolving bacterial envelopes and plasticity of TLR2-dependent responses: basic research and translational opportunities. Front Immunol. 2013;4 https://doi.org/10.3389/fimmu.2013.00347.
96. Febbraio MA, Pedersen BK. Contraction-induced myokine production and release: is skeletal muscle an endocrine organ? Exerc Sport Sci Rev. 2005; 33(3):114–9.
97. Pedersen BK, Steensberg A, Fischer C, Keller C, Keller P, Plomgaard P, et al. Searching for the exercise factor: is IL-6 a candidate? J Muscle Res Cell Motil. 2003;24(2–3):113–9.
98. de Rezende Gomes M, Santana de Oliveira Pires I, Alves de Castro I, Tirapegui J. Effect of moderate physical exercise on plasma and tissue levels of insulin-like growth factor-1 in adult rats. Nutr Res. 2004;24(7):555–64. https://doi.org/10.1016/j.nutres.2004.04.003.
99. Eliakim A, Nemet D. Exercise training, physical fitness and the growth hormone-insulin-like growth factor-1 axis and cytokine balance. Medicine and sport science. 2010;55:128–40. https://doi.org/10.1159/000321977.
100. Booth FW, Ruegsegger GN, Toedebusch RG, Yan Z. Endurance exercise and the regulation of skeletal muscle metabolism. Prog Mol Biol Transl Sci. 2015; 135:129–51. https://doi.org/10.1016/bs.pmbts.2015.07.016.
101. Egan B, Zierath JR. Exercise metabolism and the molecular regulation of skeletal muscle adaptation. Cell Metab. 2013;17(2):162–84. https://doi.org/10.1016/j.cmet.2012.12.012.
102. Smith J, Dykes R, Douglas JE, Krishnaswamy G, Berk S. Long-term exercise and atherogenic activity of blood mononuclear cells in persons at risk of developing ischemic heart disease. JAMA. 1999;281(18):1722–7. https://doi.org/10.1001/jama.281.18.1722.
103. Gleeson M, Bishop NC, Stensel DJ, Lindley MR, Mastana SS, Nimmo MA. The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease. Nat Rev Immunol. 2011;11(9):607–15. https://doi.org/10.1038/nri3041.

104. Ostrowski K, Rohde T, Asp S, Schjerling P, Pedersen BK. Pro- and anti-inflammatory cytokine balance in strenuous exercise in humans. J Physiol. 1999;515(Pt 1):287–91.
105. Ostrowski K, Schjerling P, Pedersen BK. Physical activity and plasma interleukin-6 in humans—effect of intensity of exercise. Eur J Appl Physiol. 2000;83(6):512–5. https://doi.org/10.1007/s004210000312.
106. Steensberg A, Fischer CP, Keller C, Moller K, Pedersen BK. IL-6 enhances plasma IL-1ra, IL-10, and cortisol in humans. Am J Physiol Endocrinol Metab. 2003;285(2):E433–7. https://doi.org/10.1152/ajpendo.00074.2003.
107. Meckel Y, Eliakim A, Seraev M, Zaldivar F, Cooper DM, Sagiv M, et al. The effect of a brief sprint interval exercise on growth factors and inflammatory mediators. Journal of strength and conditioning research/National Strength and Conditioning Association. 2009;23(1):225–30. https://doi.org/10.1519/JSC.0b013e3181876a9a.
108. Pedersen BK, Fischer CP. Beneficial health effects of exercise—the role of IL-6 as a myokine. Trends Pharmacol Sci. 2007;28(4):152–6. https://doi.org/10.1016/j.tips.2007.02.002.
109. Petersen AM, Pedersen BK. The role of IL-6 in mediating the anti-inflammatory effects of exercise. J Physiol Pharmacol. 2006;57(Suppl 10):43–51.
110. Hubo M, Trinschek B, Kryczanowsky F, Tuettenberg A, Steinbrink K, Jonuleit H. Costimulatory molecules on immunogenic versus tolerogenic human dendritic cells. Front Immunol. 2013;4 https://doi.org/10.3389/fimmu.2013.00082.
111. Lira FS, Rosa JC, Zanchi NE, Yamashita AS, Lopes RD, Lopes AC, et al. Regulation of inflammation in the adipose tissue in cancer cachexia: effect of exercise. Cell Biochem Funct. 2009;27(2):71–5. https://doi.org/10.1002/cbf.1540.
112. Schottelius AJ, Mayo MW, Sartor RB, Baldwin AS Jr. Interleukin-10 signaling blocks inhibitor of kappaB kinase activity and nuclear factor kappaB DNA binding. J Biol Chem. 1999;274(45):31868–74.
113. Nickel T, Emslander I, Sisic Z, David R, Schmaderer C, Marx N, et al. Modulation of dendritic cells and Toll-like receptors by marathon running. Eur J Appl Physiol. 2012;112(5):1699–708. https://doi.org/10.1007/s00421-011-2140-8.
114. Nickel T, Hanssen H, Emslander I, Drexel V, Hertel G, Schmidt-Trucksass A, et al. Immunomodulatory effects of aerobic training in obesity. Mediat Inflamm. 2011;2011:308965. https://doi.org/10.1155/2011/308965.
115. Robinson E, Durrer C, Simtchouk S, Jung ME, Bourne JE, Voth E, et al. Short-term high-intensity interval and moderate-intensity continuous training reduce leukocyte TLR4 in inactive adults at elevated risk of type 2 diabetes. J Appl Physiol. 2015;119(5):508–16. https://doi.org/10.1152/japplphysiol.00334.2015.
116. Cheng YY, Kao CL, Ma HI, Hung CH, Wang CT, Liu DH, et al. SIRT1-related inhibition of pro-inflammatory responses and oxidative stress are involved in the mechanism of nonspecific low back pain relief after exercise through modulation of Toll-like receptor 4. J Biochem. 2015;158(4):299–308. https://doi.org/10.1093/jb/mvv041.
117. Fashi M, Agha Alinejad H, Asilian Mahabadi H. The effect of aerobic exercise in ambient particulate matter on lung tissue inflammation and lung cancer. Iranian journal of cancer prevention. 2015;8(3):e2333. 10.17795/ijcp2333.
118. White AT, Light AR, Hughen RW, Vanhaitsma TA, Light KC. Differences in metabolite-detecting, adrenergic, and immune gene expression after moderate exercise in patients with chronic fatigue syndrome, patients with multiple sclerosis, and healthy controls. Psychosom Med. 2012;74(1):46–54. https://doi.org/10.1097/PSY.0b013e31824152ed.
119. Rodriguez-Miguelez P, Fernandez-Gonzalo R, Almar M, Mejias Y, Rivas A, de Paz JA, et al. Role of Toll-like receptor 2 and 4 signaling pathways on the inflammatory response to resistance training in elderly subjects. Age (Dordr). 2014;36(6):9734. https://doi.org/10.1007/s11357-014-9734-0.
120. Zanchi NE, Lira FS, de Siqueira Filho MA, Rosa JC, de Oliveira Carvalho CR, Seelaender M, et al. Chronic low frequency/low volume resistance training reduces pro-inflammatory cytokine protein levels and TLR4 mRNA in rat skeletal muscle. Eur J Appl Physiol. 2010;109(6):1095–102. https://doi.org/10.1007/s00421-010-1456-0.
121. Prestes J, da Cunha Nascimento D, Tibana RA, Teixeira TG, Vieira DC, Tajra V, et al. Understanding the individual responsiveness to resistance training periodization. Age (Dordr). 2015;37(3):9793. https://doi.org/10.1007/s11357-015-9793-x.
122. Fernandez-Gonzalo R, De Paz JA, Rodriguez-Miguelez P, Cuevas MJ, Gonzalez-Gallego J. TLR4-mediated blunting of inflammatory responses to eccentric exercise in young women. 2014;2014:479395. doi:https://doi.org/10.1155/2014/479395.
123. Millard A-L, Valli PV, Stussi G, Mueller NJ, Yung GP, Seebach JD. Brief exercise increases peripheral blood NK cell counts without immediate functional changes, but impairs their responses to ex vivo stimulation. Front Immunol. 2013; https://doi.org/10.3389/fimmu.2013.00125.
124. Greiwe JS, Cheng B, Rubin DC, Yarasheski KE, Semenkovich CF. Resistance exercise decreases skeletal muscle tumor necrosis factor alpha in frail elderly humans. FASEB J. 2001;15(2):475–82. https://doi.org/10.1096/fj.00-0274com.
125. Phillips MD, Flynn MG, McFarlin BK, Stewart LK, Timmerman KL. Resistance training at eight-repetition maximum reduces the inflammatory milieu in elderly women. Med Sci Sports Exerc. 2010;42(2):314–25. https://doi.org/10.1249/MSS.0b013e3181b11ab7.
126. Fatouros I, Chatzinikolaou A, Paltoglou G, Petridou A, Avloniti A, Jamurtas A, et al. Acute resistance exercise results in catecholaminergic rather than hypothalamic-pituitary-adrenal axis stimulation during exercise in young men. Stress (Amsterdam, Netherlands). 2010;13(6):461–8. https://doi.org/10.3109/10253891003743432.
127. Przybyla B, Gurley C, Harvey JF, Bearden E, Kortebein P, Evans WJ, et al. Aging alters macrophage properties in human skeletal muscle both at rest and in response to acute resistance exercise. Exp Gerontol. 2006;41(3):320–7. https://doi.org/10.1016/j.exger.2005.12.007.
128. Barnes JN, Trombold JR, Dhindsa M, Lin HF, Tanaka H. Arterial stiffening following eccentric exercise-induced muscle damage. J Appl Physiol. 2010;109(4):1102–8. https://doi.org/10.1152/japplphysiol.00548.2010.
129. Stupka N, Tarnopolsky MA, Yardley NJ, Phillips SM. Cellular adaptation to repeated eccentric exercise-induced muscle damage. J Appl Physiol. 2001;91(4):1669–78.
130. Jun JK, Lee WL, Park HG, Lee SK, Jeong SH, Lee YR. Moderate intensity exercise inhibits macrophage infiltration and attenuates adipocyte inflammation in ovariectomized rats. Journal of exercise nutrition and biochemistry. 2014;18(1):119–27. https://doi.org/10.5717/jenb.2014.18.1.119.
131. Zheng Q, Cui G, Chen J, Gao H, Wei Y, Uede T, et al. Regular exercise enhances the immune response against microbial antigens through up-regulation of Toll-like receptor signaling pathways. Cell Physiol Biochem. 2015;37(2):735–46. https://doi.org/10.1159/000430391.
132. Holland AM, Hyatt HW, Smuder AJ, Sollanek KJ, Morton AB, Roberts MD, et al. Influence of endurance exercise training on antioxidant enzymes, tight junction proteins, and inflammatory markers in the rat ileum. BMC research notes. 2015;8:514. https://doi.org/10.1186/s13104-015-1500-6.
133. Lancaster GI, Khan Q, Drysdale P, Wallace F, Jeukendrup AE, Drayson MT, et al. The physiological regulation of Toll-like receptor expression and function in humans. J Physiol. 2005;563(Pt 3):945–55. https://doi.org/10.1113/jphysiol.2004.081224.
134. Oliveira M, Gleeson M. The influence of prolonged cycling on monocyte Toll-like receptor 2 and 4 expression in healthy men. Eur J Appl Physiol. 2010;109(2):251–7. https://doi.org/10.1007/s00421-009-1350-9.
135. Simpson RJ, McFarlin BK, McSporran C, Spielmann G, Hartaigh B, Guy K. Toll-like receptor expression on classic and pro-inflammatory blood monocytes after acute exercise in humans. Brain Behav Immun. 2009;23(2):232–9. https://doi.org/10.1016/j.bbi.2008.09.013.
136. Liao P, Zhou J, Ji LL, Zhang Y. Eccentric contraction induces inflammatory responses in rat skeletal muscle: role of tumor necrosis factor-alpha. American Journal of Physiology Regulatory, Integrative and Comparative Physiology. 2010;298(3):R599–607. https://doi.org/10.1152/ajpregu.00480.2009.
137. Radom-Aizik S, Zaldivar FP Jr, Haddad F, Cooper DM. Impact of brief exercise on circulating monocyte gene and microRNA expression: implications for atherosclerotic vascular disease. Brain Behav Immun. 2014;39:121–9. https://doi.org/10.1016/j.bbi.2014.01.003.
138. Tanaka Y, Kawanishi N, Shiva D, Tsutsumi N, Uchida M, Kitamura H, et al. Exhaustive exercise reduces tumor necrosis factor-alpha production in response to lipopolysaccharide in mice. Neuroimmunomodulation. 2010;17(4):279–86. https://doi.org/10.1159/000290044.
139. Li H, Geib RW. Exploring the use of five color flow cytometry to examine the effect of acute tai chi practice on pro inflammatory monocyte subtypes—biomed 2013. Biomed Sci Instrum. 2013;49:209–15.
140. Ortega E, Hinchado MD, Martin-Cordero L, Asea A. The effect of stress-inducible extracellular Hsp72 on human neutrophil chemotaxis: a role during acute intense exercise. Stress (Amsterdam, Netherlands). 2009;12(3):240–9. https://doi.org/10.1080/10253890802309853.
141. Kohut ML, McCann DA, Russell DW, Konopka DN, Cunnick JE, Franke WD, et al. Aerobic exercise, but not flexibility/resistance exercise, reduces serum IL-18, CRP, and IL-6 independent of beta-blockers, BMI, and psychosocial factors in older adults. Brain Behav Immun. 2006;20(3):201–9. https://doi.org/10.1016/j.bbi.2005.12.002.

142. Tilstra JS, Clauson CL, Niedernhofer LJ, Robbins PD. NF-kappaB in Aging and Disease. Aging and disease. 2011;2(6):449–65.
143. Beyer I, Mets T, Bautmans I. Chronic low-grade inflammation and age-related sarcopenia. Current Opinion in Clinical Nutrition and Metabolic Care. 2012;15(1):12–22. https://doi.org/10.1097/MCO.0b013e32834dd297.
144. McCullough PA, Lavie CJ. Coronary artery plaque and cardiotoxicity as a result of extreme endurance exercise. Mo Med. 2014;111:9Y94.
145. O'Keefe JH, Lavie CJ, Guazzi M. Part 1: potential dangers of extreme endurance exercise: how much is too much? Part 2: screening of school-age athletes. Prog Cardiovasc Dis. 2015;57(4):396–405. https://doi.org/10.1016/j.pcad.2014.11.004.
146. Gill SK, Hankey J, Wright A, Marczak S, Hemming K, Allerton DM, et al. The impact of a 24-h ultra-marathon on circulatory endotoxin and cytokine profile. Int J Sports Med. 2015;36(8):688–95. https://doi.org/10.1055/s-0034-1398535.
147. Pedersen BK, Bruunsgaard H, Jensen M, Krzywkowski K, Ostrowski K. Exercise and immune function: effect of ageing and nutrition. Proc Nutr Soc. 1999;58(3):733–42.
148. Radak Z, Taylor AW, Ohno H, Goto S. Adaptation to exercise-induced oxidative stress: from muscle to brain. Exerc Immunol Rev. 2001;7:90–107.
149. Timmerman KL, Flynn MG, Coen PM, Markofski MM, Pence BD. Exercise training-induced lowering of inflammatory (CD14+CD16+) monocytes: a role in the anti-inflammatory influence of exercise? J Leukoc Biol. 2008;84(5):1271–8. https://doi.org/10.1189/jlb.0408244.
150. Stewart LK, Flynn MG, Campbell WW, Craig BA, Robinson JP, Timmerman KL, et al. The influence of exercise training on inflammatory cytokines and C-reactive protein. Med Sci Sports Exerc. 2007;39(10):1714–9. https://doi.org/10.1249/mss.0b013e31811ece1c.
151. Johnson AR, Milner JJ, Makowski L. The inflammation highway: metabolism accelerates inflammatory traffic in obesity. Immunol Rev. 2012;249(1):218–38. https://doi.org/10.1111/j.1600-065X.2012.01151.x.
152. Wang X, Sun Y, Yang H, Lu Y, Li L. Oxidized low-density lipoprotein induces apoptosis in cultured neonatal rat cardiomyocytes by modulating the TLR4/NF-κB pathway. Sci Rep. 2016;6:27866. https://doi.org/10.1038/srep27866.
153. Yang K, Zhang XJ, Cao LJ, Liu XH, Liu ZH, Wang XQ, et al. Toll-like receptor 4 mediates inflammatory cytokine secretion in smooth muscle cells induced by oxidized low-density lipoprotein. PLoS One. 2014;9(4):e95935. https://doi.org/10.1371/journal.pone.0095935.
154. Frostegård J. Immunity, atherosclerosis and cardiovascular disease. BMC Med. 2013;11:117. https://doi.org/10.1186/1741-7015-11-117.
155. Frostegård J, Zhang Y, Sun J, Yan K, Liu A. Oxidized low-density lipoprotein (OxLDL)-treated dendritic cells promote activation of T cells in human atherosclerotic plaque and blood, which is repressed by Statins: microRNA let-7c is integral to the effect. J Am Heart Assoc. 2016;5(9) https://doi.org/10.1161/jaha.116.003976.
156. Lizard G, Monier S, Cordelet C, Gesquière L, Deckert V, Gueldry S, et al. Characterization and Comparison of the Mode of Cell Death, Apoptosis Versus Necrosis, Induced by 7β-hydroxycholesterol and 7-ketocholesterol in the cells of the vascular wall. Arterioscler Thromb Vasc Biol. 1999;19(5):1190–200. https://doi.org/10.1161/01.atv.19.5.1190.
157. Tsutsui T, Tsutamoto T, Wada A, Maeda K, Mabuchi N, Hayashi M, et al. Plasma oxidized low-density lipoprotein as a prognostic predictor in patients with chronic congestive heart failure. J Am Coll Cardiol. 2002;39(6):957–62. https://doi.org/10.1016/s0735-1097(02)01721-7.
158. Zorn-Pauly K, Schaffer P, Pelzmann B, Bernhart E, Wei G, Lang P, et al. Oxidized LDL induces ventricular myocyte damage and abnormal electrical activity—role of lipid hydroperoxides. Cardiovasc Res. 2005;66(1):74–83. https://doi.org/10.1016/j.cardiores.2004.12.009.

Pain Physician 2012; 15:ES205-ES213 • ISSN 2150-1149

Narrative Review

Dysfunctional Endogenous Analgesia During Exercise in Patients with Chronic Pain: To Exercise or Not to Exercise?

Jo Nijs, PhD[1,2,3], Eva Kosek, MD, PhD[4], Jessica Van Oosterwijck, PhD, [1,3], and Mira Meeus, PhD[1,2]

From: [1]Chronic Pain and Chronic Fatigue Research Group (CHROPIVER), Department of Human Physiology, Faculty of Physical Education & Physiotherapy, Vrije Universiteit Brussel, Brussels, Belgium. [2]Chronic Pain and Chronic Fatigue Research Group (CHROPIVER), Division of Musculoskeletal Physiotherapy, Department of Health Care Sciences, Artesis University College, Antwerp, Belgium. [3]Department of Physical Medicine and Physiotherapy, University Hospital, Brussels, Belgium. [4]Osher Center for Integrative Medicine, Stockholm Brain Institute, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden.

Address correspondence: to Jo Nijs
Vrije Universiteit Brussel
Building L-Mfys
Pleinlaan 2, BE-1050
Brussels, Belgium
E-mail: Jo.Nijs@vub.ac.be

Disclaimer: There was no external funding in the preparation of this manuscript.
Conflict of interest: None.

Manuscript received: 12/29/2011
Accepted for publication: 01/30/2012

Free full manuscript: www.painphysicianjournal.com

Background: Exercise is an effective treatment for various chronic pain disorders, including fibromyalgia, chronic neck pain, osteoarthritis, rheumatoid arthritis, and chronic low back pain. Although the clinical benefits of exercise therapy in these populations are well established (i.e. evidence based), it is currently unclear whether exercise has positive effects on the processes involved in chronic pain (e.g. central pain modulation).

Objectives: Reviewing the available evidence addressing the effects of exercise on central pain modulation in patients with chronic pain.

Methods: Narrative review.

Results: Exercise activates endogenous analgesia in healthy individuals. The increased pain threshold following exercise is due to the release of endogenous opioids and activation of (supra)spinal nociceptive inhibitory mechanisms orchestrated by the brain. Exercise triggers the release of β-endorphins from the pituitary (peripherally) and the hypothalamus (centrally), which in turn enables analgesic effects by activating μ-opioid receptors peripherally and centrally, respectively. The hypothalamus, through its projections on the periaqueductal grey, has the capacity to activate descending nociceptive inhibitory mechanisms.

However, several groups have shown dysfunctioning of endogenous analgesia in response to exercise in patients with chronic pain. Muscle contractions activate generalized endogenous analgesia in healthy, pain-free humans and patients with either osteoarthritis or rheumatoid arthritis, but result in increased generalised pain sensitivity in fibromyalgia patients. In patients having local muscular pain (e.g. shoulder myalgia), exercising non-painful muscles activates generalized endogenous analgesia. However, exercising painful muscles does not change pain sensitivity either in the exercising muscle or at distant locations.

Limitations: The reviewed studies examined acute effects of exercise rather than long-term effects of exercise therapy.

Conclusions: A dysfunctional response of patients with chronic pain and aberrations in central pain modulation to exercise has been shown, indicating that exercise therapy should be individually tailored with emphasis on prevention of symptom flares. The paper discusses the translation of these findings to rehabilitation practice together with future research avenues.

Key words: Whiplash, fibromyalgia, chronic pain, low back pain, exercise, rehabilitation, chronic fatigue syndrome, osteoarthritis, rheumatoid arthritis, sensitization, shoulder

Pain Physician 2012; 15:ES205-ES213

Chronic pain remains a challenging issue for clinicians and researchers. Over the past decades, scientific understanding of such unexplained chronic pain disorders has increased substantially. It has now become clear that the majority of cases of chronic pain can be explained by alterations in central nervous system processing of incoming messages (1). More specifically, the responsiveness of central neurons

to input from unimodal and polymodal receptors is augmented, resulting in a pathophysiological state corresponding to central sensitization, characterized by generalized or widespread hypersensitivity to a variety of stimuli (i.e. mechanical, thermal, and chemical) (2).

The term central sensitization can be used to encompass altered sensory processing in the brain (3), long-term potentiation of brain synapses (4), impaired functioning of top-down anti-nociceptive mechanisms (5), and (over)activation of top-down pain facilatory pathways which augment nociceptive transmission (3,6). Importantly, a different "pain signature" arises in the brain of those with chronic pain. This altered pain neuromatrix comprises of a) increased activity in brain areas known to be involved in acute pain sensations like the insula, anterior cingulate cortex, and the prefrontal cortex, but not in the primary or secondary somatosensory cortex (7); and b) brain activity in regions generally not involved in acute pain sensations like various brain stem nuclei, dorsolateral frontal cortex, and parietal associated cortex (7). Clinically central sensitization is characterized by non-segmental spreading of pain, "central" symptoms like concentration difficulties and fatigue, stress-intolerance and hypersensitivity to various stimuli like bright light, touch, and odors (8).

Exercise is frequently encountered as a central component of the treatment of patients with chronic pain. Exercise is an effective treatment for various chronic musculoskeletal pain disorders, including chronic low back pain (9), chronic whiplash associated disorders (10,11), osteoarthritis (12), and fibromyalgia (13,14). Although the clinical benefits of exercise therapy in these populations are well established (i.e. evidence based), it is currently unclear whether exercise therapy has positive effects on the processes involved in central sensitization. Is exercise capable of "treating" central sensitization in patients with chronic pain?

There is a strong theoretical rationale suggesting that exercise therapy can indeed "treat" central sensitization (or desensitize the central nervous system). In healthy individuals aerobic exercise of sufficient intensity (+/- 200 W or 70 % VO_{2MAX}) activates pain inhibition for up to 30 minutes post-exercise (15). Resistance exercise triggers endogenous analgesia as well, but it lasts for no more than a couple of minutes post-exercise (15). The exercise induced endogenous analgesia is presumed to be due to the release of endogenous opioids and growth factors (16,17) and activation of (supra)spinal nociceptive inhibitory mechanisms orchestrated by the brain (18,19).

Based on this theoretical rationale and on the evidence supporting the clinical benefits in various chronic musculoskeletal pain disorders, it is tempting to speculate that exercise can indeed desensitize the central nervous system. However, this hypothesis is not (yet) supported by scientific evidence. A recent systematic literature review showed that no conclusions can be made about the effect of exercise therapy on pain on pain-modulatory substances (e.g. serotonin, norepinephrine, opioids) or on its effects on altering brain activity of areas involved in pain processing in patients with musculoskeletal pain (20). Moreover, a dysfunctional response of some patients with chronic musculoskeletal pain to exercise has been shown. Several populations of chronic pain patients are unable to activate central descending nociceptive inhibition (endogenous analgesia or EA) during exercise (21-23), a dysfunction partly explaining symptom flares following exercise (22).

In what follows in this paper explains our current understanding of the biology of EA following exercise in humans. Next, it provides an overview of the studies addressing dysfunctional EA during local muscle and general aerobic exercise in patients with chronic pain. From this overview it will become clear that some chronic pain disorders (e.g. fibromyalgia) are characterized by a dysfunctional EA in response to both aerobic and local muscle exercises, while other chronic pain populations (e.g. chronic low back pain) show a normal activation of EA in response to exercise. The relevance of these findings to rehabilitation practice together with future research avenues will be discussed as well.

The Biology of Exercise-Induced Endogenous Analgesia

Several partly overlapping mechanisms are suggested to play a role in exercise-induced EA, including release of endogenous opioids and growth factors (16,17), and activation of (supra)spinal nociceptive inhibitory mechanisms orchestrated by the brain (18,19). These mechanisms might be related to cardiovascular changes (i.e. increase in heart rate and blood pressure) during exercise, a notion supported by the finding that patients with hypertension show reduced pain sensitivity (24). The interaction can be explained by similar brain stem nuclei, neurotransmitters (e.g. monoamines) and peptides (e.g. opioids) (25). The exercise-induced blood pressure increase activates arterial baroreceptors, resulting in increased supraspinal inhibition (24,25) and stimulation of brain centers involved in pain modulation (26).

In addition, Hoffman and Thoren (27) have reported that once blood pressure is displaced out of the basal range, either by physiological stimuli or pathophysiological states, the endogenous opioid system becomes activated. Exercise triggers the release of β-endorphins from the pituitary (peripherally) and the hypothalamus (centrally), which in turn enables analgesic effects by activating μ-opioid receptors peripherally and centrally, respectively (28). The hypothalamus, through its projections on the periaqueductal grey, has the capacity to activate descending nociceptive inhibitory mechanisms.

Nevertheless, it appears from animal research that multiple analgesia systems exist (opioid and non-opioid), and that properties of the exercise stressor are important in determining which system is activated during exercise. It has been shown that by manipulating the parameters of the exercise stressor, it is possible to elicit either naloxone-reversible or naloxone-insensitive EA following exercise (17,29).

The role of growth hormone and growth factors in exercise-induced EA remains unclear. Some authors hypothesize that growth hormone, via insulin-like growth factor and nerve growth factor, sensitizes rather than being involved in EA (30,31). Only one study evaluated the role of growth hormone in exercise-induced hypoalgesia, but suppressing growth hormone production during exercise did not alter exercise-induced hypoalgesia (32). β-endorphins and growth hormone are released after a certain exercise span, one hour and 10 to 40 minutes respectively, when lactate accumulation is leading to muscle acidosis (33-35).

Catecholamines also exert direct analgesic effects. Main descending inhibitory action to the spinal dorsal horn are noradrenergic. In the dorsal horn, norepinephrine, through action on alpha-2A-adrenoceptors, suppresses release of excitatory transmitters from central terminals of primary afferent nociceptors (36). In addition it may suppress postsynaptical responses of spinal pain-relay neurons (36). Besides catecholamines, the mediators of the long-term stress response, namely corticosteroids, are involved in exercise induced EA. Both opioid and non-opioid mechanisms would contribute to the development of EA induced by glucocorticoids (37). Hence, exercise can be viewed as a frequent stressor activating stress-induced analgesia.

Other possible explanations for exercise-induced EA involve an increased body awareness for somatic sensations after exercise. This awareness of more salient signals—for example, sweating and heart pounding—may divert attention away from the pain stimulus. Distraction can significantly alter pain perception (38). Furthermore, traditional gate control mechanisms, due to skin or muscle afferents competing with nociceptive afferents in the dorsal horn, may account for exercise induced EA. Finally, conditioned pain modulation, formerly referred to as diffuse noxious inhibitory controls (DNIC), may be activated subsequently to the nociceptive barrage resulting from muscle ischemia and lactate accumulation. Peripheral mechanisms are less plausible as they typically result in sensitizing agents (prostaglandins, lactate, ischemia, growth factors, etc.).

Dysfunction of Endogenous Analgesia During Muscle Contraction in Patients with Musculoskeletal Pain

Long-term, low intensity, static work is a well known risk factor for the development of work-related myalgias, and static contractions increase pain intensity in patients with myalgia (39) and fibromyalgia (39,40). Animal studies have revealed that muscle ischemia is a potent cause of sensitization of peripheral mechanonociceptors so that the increased intramuscular pressure caused by the contraction can became an effective nociceptive stimulus (41). Compared to healthy controls, patients with shoulder myalgia (42) and fibromyalgia (43) had reduced muscle blood flow during static contractions, which could lead to peripheral sensitization and explain the increased pain sensitivity reported in painful muscles in these patients (39). In accordance with this, increased sensitivity to pressure pain (i.e., increased tenderness) at the contracting muscle following static contractions was reported in fibromyalgia patients (44,45), suggesting dysfunctional EA during exercise in these patients. Indeed, healthy subjects exhibited decreased pressure pain sensitivity at the contracting muscle during and following contraction indicating that segmental or possibly plurisegmental (generalized) pain inhibitory mechanisms were activated (46). In a follow-up study, localized (at the contracting muscle) as well as generalized (at a distant resting muscle) pain inhibitory effects were seen during static contractions in healthy individuals (46). In addition, the decrease in pain sensitivity was of similar magnitude at the contracting and the distant resting muscle indicating the importance of generalized EA mechanisms (46).

To our knowledge, only a few studies have examined the effect of static contractions on pain sensitivity outside the contracting muscle. Staud et al (47) found a bilateral decrease in cutaneous (heat) and deep somatic (pressure) pain sensitivity during unilateral static

contractions sustained during 90 seconds corresponding to 30% of the individual maximal voluntary contraction force (MVC) in healthy subjects. A paradoxical increase in heat and pressure pain sensitivity was seen bilaterally in fibromyalgia patients during the unilateral contractions, providing evidence for widespread deficiency of EA or more pronounced pain facilitation in fibromyalgia patients during exercise (47).

The important question raised by Staud et al (47) regarding the importance of deficient EA versus augmented pain facilitation in pain patients during physical exercise was further addressed in another study assessing patients with shoulder myalgia and fibromyalgia, respectively, during static contractions corresponding to 20 – 25% MVC until exhaustion (maximum 5 minutes) (39). Patients and healthy controls performed static contractions with M. quadriceps femoris and M. infraspinatus. Pressure pain thresholds were assessed before and during contraction at the contracting muscle, the resting homologous contralateral muscle, and contralaterally at a distant site (M. infraspinatus during contraction of M. quadriceps and vice versa). Pressure pain thresholds increased at all sites during both contractions in healthy controls, but no increase was seen at any site during contractions in fibromyalgia patients, who even exhibited increased pain sensitivity. Myalgia patients had an increase in pressure pain thresholds at all sites during contraction of the non-painful M. quadriceps, but no increase in pressure pain thresholds was seen at any site during contraction of the painful M. infraspinatus. The authors suggested that nociceptive input from painful muscles induced central sensitization and activated descending pain facilitatory mechanisms. The facilitatory mechanisms could override the contraction-induced pain inhibition and explain the lack of generalized EA during contraction of painful muscles in myalgia patients and the increased pain sensitivity during contraction in fibromyalgia patients (39).

Interestingly, a pilot study in patients with rheumatoid arthritis indicates normal EA during static contraction in these patients (Fridén et al., submitted) and preliminary results also indicate normal function of these mechanisms in patients with osteoarthritis of the knee and hip (Kosek, Roos, Nilsdotter, manuscript in preparation). These findings are in accordance with the reported beneficial effects of exercise in these conditions (48,49).

As mentioned, many potential mechanisms have been implicated in pain regulation during muscle contractions. Conditioned pain modulation has been proposed as one possible mechanism for pain inhibition during contraction. However, the low pain ratings during contraction in healthy controls (39) make this unlikely. Furthermore, although a dysfunction of conditioned pain modulation has been shown in fibromyalgia (50), normal function of conditioned pain modulation was shown in shoulder myalgia patients (51). Exercise induced pain modulation during static contractions has also been related to arterial baroreceptor activation in humans (52). However, a normal increase in heart rate and blood pressure has been reported in fibromyalgia patients during static contractions offsetting abnormal cardiovascular response to exercise as a likely explanation for the dysfunction of EA in these patients (40,53). Finally, hormonal factors of importance for regulation of muscle blood flow and pain sensitivity could be of interest. The findings of a hypo-active sympatho-adrenal system in combination with a hyporeactive adrenal-hypothalamo-pituitary (HPA) axis in fibromyalgia patients during static contractions could contribute to the dysfunctional EA during exercise and subsequent exercise intolerance that is so characteristic for fibromyalgia patients (53).

It is concluded that muscle contractions activate generalized EA in healthy, pain-free humans and patients with either osteoarthritis and rheumatoid arthritis, but result in increased generalised pain sensitivity in fibromyalgia patients. In patients having local muscular pain (e.g. shoulder myalgia), exercising non-painful muscles activates generalized EA. However, exercising painful muscles does not change pain sensitivity either in the exercising muscle or at distant locations.

Dysfunction of Endogenous Analgesia During Aerobic Exercise in Patients with Musculoskeletal Pain

The dysfunctional EA in response to aerobic exercise was first shown in a small study of patients with chronic fatigue syndrome and healthy controls in which participants performed a graded exercise with 3 stages on a treadmill (54). Every stage of the exercise consisted of 5 minutes walking at a constant pace of 5km/h, with an increasing incline of 5°. Dysfunctional EA was demonstrated by decreased pain thresholds following exercise in patients with chronic fatigue syndrome, while pain thresholds increased in healthy controls. These findings were later replicated in 2 larger studies using various types of exercise:

1) submaximal cycle exercise with a gradual increase of 25 W every minute until 75% of the age-predicted target heart rate was achieved (22),

2) 6 short bouts of aerobic cycling interrupted by short recovery breaks (21), and
3) physiologically limited (heart rate below 80% of the anaerobic heart rate, workload below 80% of the anaerobic workload) and self-paced aerobic cycling (22).

From these studies it is concluded that neither types of aerobic exercise were able to activate EA in patients with chronic fatigue syndrome who experience chronic widespread pain. Importantly, the dysfunctional EA partly explains symptom flares following exercise in patients with chronic fatigue syndrome having chronic widespread pain (22).

A similar dysfunctional EA in response to exercise and symptom flares following exercise was shown in patients with chronic whiplash associated disorders (23), suggesting this to be a feature of central sensitization. The dysfunctional EA in patients with chronic whiplash associated disorders was observed during submaximal cycle exercise with a gradual increase of 25 W every minute until 75% of the age-predicted target heart rate was achieved, as well as during physiologically limited and self-paced aerobic cycling (23). Remarkably, in the studies outlined above the various types of aerobic exercise did activate EA in healthy sedentary controls (21-23,54) and patients with chronic low back pain (21). The latter confirms an earlier study in chronic low back pain patients (55). Thus, the mechanism of EA in patients with chronic low back pain responds normally to aerobic exercise.

Some work has been performed to unravel the mechanisms behind the dysfunctional EA during exercise in certain chronic pain disorders. Nitric oxide (NO) plays a complex role in nociceptive processing (19). Although evidence exists regarding the beneficial effects of the release of small amounts of NO during inhibition of nociceptive pathways (56), excessive amounts of NO could contribute to central sensitization. Indeed, NO is able to reduce the nociceptive inhibitory activity of the central nervous system, leading to central sensitization of dorsal horn neurones (57). A single bout of physical activity triggers release of NO (58), leading to the hypothesis that the dysfunctional EA during exercise might be due to NO release. However, NO levels were unrelated to pain processing during aerobic exercise in healthy sedentary controls, patients with chronic fatigue syndrome and chronic low back pain (21).

While endogenous opioid and adrenergic pain-inhibitory mechanisms might account for activation of EA during exercise in healthy individuals (18,19), direct evidence is lacking. Therefore, a study was undertaken to examine the contribution of endogenous opioid pain-inhibitory mechanisms during exercise in 2 chronic pain populations: rheumatoid arthritis and fibromyalgia/chronic fatigue syndrome (59). In a randomized and placebo-controlled cross-over study, we modulated endogenous opioid and serotonergic pain-inhibitory mechanisms during exercise by using selective serotonin reuptake inhibitor (SSRI; 2 mL of citalopram intravenously) during the DNIC and temporal summation model in response to exercise. SSRIs activate serotonergic descending pathways that recruit, in part, opioid peptide-containing interneurons of the dorsal horn (60). Unfortunately, significant side effects immediately after intravenous administration of citalopram resulted in early cessation of the study. Hence, currently no conclusions can be made addressing the role of serotonergic descending pathways in EA in response to exercise in chronic pain patients (59).

It is concluded that the dysfunctional EA during aerobic exercise is not characteristic for all chronic pain patients, but rather limited to those with clear evidence of central sensitization (e.g. chronic whiplash, fibromyalgia, chronic fatigue syndrome).

Analgesic Effects of Exercise Therapy in Patients with Chronic Musculoskeletal Pain: Opportunities and Challenges

Applying these Findings to the Practice of Exercise Therapy for Chronic Pain

When confronted with the cumulating evidence for a dysfunctional EA during exercise in some chronic pain disorders, one might wonder whether we should continue using exercise therapy in these patients. However, this should not be an issue. The studies summarized above address acute bouts of exercise, not findings from randomized clinical trials examining the effects of exercise as a therapeutic intervention. The studies showing dysfunctional EA during exercise in some chronic pain conditions do not contradict the clinical evidence favoring the use of exercise as an intervention for chronic pain. Exercise is an effective treatment for chronic whiplash associated disorders (10,11), fibromyalgia (13,14), chronic fatigue syndrome (61,62), osteoarthritis (48), and rheumatoid arthritis (49). Hence, its clinical use and benefits should not be questioned. To exercise or not to exercise patients with chronic pain, is no longer the question.

On the other hand, the dysfunctional EA during exercise in patients with chronic pain should not be

ignored. In fact, its clinical relevance is supported by studies showing that symptom flares following exercise are related to the dysfunctional EA during exercise (22). In addition, the dysfunctional EA during exercise might explain the low compliance with exercise interventions in chronic pain patients. Typically the early stages of exercise therapy programs are prone to dropouts.

Lack of exercise-induced analgesia implies a decreased pain threshold following exercise. This makes patients vulnerable for new nociceptive input. Exercise is typically associated with myofiber damage and substances released in response to exercise (e.g. oxidative stress, lactate), potentially providing increased nociceptive input in response to exercise (63). Hence, the dysfunctional EA during exercise increases the risk of severe symptom flares following exercise sessions. For all these reasons, we conclude that clinicians should account for the dysfunctional EA during exercise in certain chronic pain conditions.

But how? Given the dearth of studies examining the effects of exercise therapy on EA (20), this question can only be answered by applying logical (clinical) reasoning. Appropriately tailored and graded exercise therapy has been suggested as a treatment for central sensitization in patients with chronic pain (64), but evidence supporting this notion is lacking (20). Especially in the early stages of exercise therapy programs, exercise therapy should be individually tailored with emphasis on prevention of symptom flares.

This might be achieved by applying the following guidelines (Table 1): prefer aerobic exercise over eccentric or isometric muscle work, as the latter 2 are likely to increase the hyperexcitability of the central nervous system (47) and result in diminished blood flow increase in the working muscles (43). The findings from the studies explained above suggest that exercising preferably non-painful parts of the body could have pain-relieving effects in myalgia patients by reducing pain sensitivity in painful muscles, while low intensity training regimes would be expected to be favorable in fibromyalgia in order to avoid unnecessary exacerbations of pain (39).

In addition, exercise therapy for chronic pain patients should account for cognitive-emotional sensitization. Emotions, attention, expectations, depressive thoughts, and catastrophic thoughts each enhance descending facilitation (65-67), which in turn sustains the process of central sensitization. This is typically referred to as cognitive-emotional sensitization (68), which can imply increased forebrain activity that can exert powerful influences on various brainstem nuclei (69), including those identified as the origin of descending facilitatory pathways (70). Clinically cognitive-emotional sensitization is typically addressed in comprehensive pain management programs that include pain physiology education to address illness perceptions and maladaptive pain cognitions, stress management, time-contingent activity management (i.e. graded activity), and time-contingent exercise therapy (i.e. graded exercise therapy) (Table 1).

Table 1. *Practical guidelines to account for dysfunctional endogenous analgesia during exercise when applying exercise therapy in patients with chronic musculoskeletal pain.*

Keep the following guidelines in mind when applying exercise therapy in patients with chronic musculoskeletal pain and dysfunctional endogenous analgesia during exercise:

- exercise should be fun, not a burden
- Discuss the content of the exercise protocol with the patient; it should fit the needs and requests of the patient
- Use aerobic exercise as well as motor control training
- Be careful with eccentric exercise
- Include exercise of non-painful parts of the body
- Allow increased pain during and shortly following exercise but avoid continuously increasing pain intensity over time (i.e. modify exercise)
- Use a time-contingent approach with appropriate baseline
- Be conservative when setting the baseline; prefer a lower baseline to guarantee that is well within the capabilities of the patient's body
- Use multiple and long recovery breaks in between exercises
- Monitor symptom flares, especially during initiation of treatment and during grading, and adopt exercise modalities accordingly
- Minor symptom flares are natural during initial stages of exercise therapy, but should cease once an exercise routine is established
- Do not grade the exercise protocol in case of major symptom flares

Treatment of the Dysfunctional Endogenous Analgesia During Exercise by Combining Centrally Acting Drugs with Exercise Therapy?

In addition to the guidelines for designing appropriate exercise therapy programs, it seems rational to combine centrally acting drugs with exercise therapy. Unraveling the mechanisms responsible for the dysfunctional EA in response to exercise in people with chronic pain is likely to be a crucial step towards well-balanced drug + exercise treatments. In the mean time, the following suggestions seems rational given our current understanding of dysfunctional EA during exercise and chronic pain management. First, opioid use in combination with (the early stages) of graded exercise therapy might be an option in some patients with nociceptive pain. In this respect, it is important to realize that evidence indicates that opioid withdrawal is unnecessary for effective pain rehabilitation programs (71). This is important as the early pain rehabilitation programs advocated operant methods to decrease opioid consumption in the early treatment stages.

Second, activation of serotonergic and/or noradrenergic descending pathways in conjunction with graded exercise therapy might be an option. A centrally acting analgesic like Duloxetine, a selective and balanced serototin and norepinephrine reuptake inhibitor (SNRI), has proven its efficacy in a variety of chronic pain conditions characterized by central sensitization (e.g. fibromyalgia [72] and osteoarthritis [73]). It remains unclear whether these clinical effects can be reinforced by combining drug use with graded exercise therapy. Further work in this area is warranted.

Finally, the finding that peak exercise performance in healthy people improves when using acetaminophen (74) might provide a new avenue for combining analgesics with exercise therapy for patients with chronic pain and dysfunctional EA during exercise. There is evidence suggesting that acetaminophen primarily acts centrally by reinforcing descending inhibitory pathways (75), namely the serotonergic descending pain pathways.

Still, future research should examine whether these proposed combinations of drug treatment and graded exercise therapy are able to treat the dysfunctional EA in patients with chronic pain. Moreover, the combined treatment programs should not only improve EA during exercise, it should benefit the patient at the level of daily functioning and quality of life as well.

Conclusion

Exercise activates EA in healthy individuals, resulting in generalized increased pain tolerance during and immediately following exercise. This conclusion accounts for aerobic exercises like cycling, and for exercising local muscle groups. The physiological mechanisms explaining EA following exercise have not been studied in detail yet, but the available research data suggest that it is due to the release of endogenous opioids and activation of (supra)spinal nociceptive inhibitory mechanisms orchestrated by the brain. However, aerobic exercise activates pain facilitation rather than inhibition in some patients with chronic pain and central sensitization (fibromyalgia, whiplash, and chronic fatigue syndrome). Exercising local muscle groups results in increased generalised pain sensitivity in fibromyalgia patients, but recent data indicate that this might not be the case in those with osteoarthritis and rheumatoid arthritis. In shoulder myalgia, exercising non-painful muscles activates generalized EA, but exercising painful muscles does not activate EA. Further work is required to unravel the biology of the dysfunctional EA following exercise, and to establish how these findings should be applied to clinical practice.

References

1. Yunus MB. Fibromyalgia and overlapping disorders: The unifying concept of central sensitivity syndromes. *Sem Arthritis Rheumatol* 2007; 36:330-356.
2. Meyer RA, Campbell JN, Raja SN. Peripheral neural mechanisms of nociception. In: Wall PD, Melzack R (eds). *Textbook of Pain*. Third ed. Churchill Livingstone, Edinburgh, 1995, pp. 13-44.
3. Staud R, Craggs JG, Robinson ME. Perlstein WM, Price DD. Brain activity related to temporal summation of C-fiber evoked pain. *Pain* 2007; 129:130-142.
4. Zhuo M. A synaptic model for pain: Long-term potentiation in the anterior cingulated cortex. *Mol Cells* 2007; 23:259-271.
5. Meeus M, Nijs J, Van de Wauwer N, Toeback L, Truijen S. Diffuse noxious inhibitory control is delayed in chronic fatigue syndrome: An experimental study. *Pain* 2008; 139:439-448.
6. Meeus M, Nijs J. Central sensitization: A biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. *Clin Rheumatol* 2007; 26:465-473.
7. Seifert F, Maihöfner C. Central mechanisms of experimental and chronic neuropathic pain: Findings from functional imaging studies. *Cell Mol Life Sci* 2009; 66:375-390.
8. Nijs J, Van Houdenhove B, Oostendorp RAB. Recognition of central sensitization in patients with musculoskeletal pain: Application of pain neurophysiology in manual therapy practice. *Manual Ther* 2010; 15:135-141.
9. van Middelkoop M, Rubinstein SM, Ver-

hagen AP, Ostelo RW, Koes BW, van Tulder MW. Exercise therapy for chronic nonspecific low-back pain. *Best Pract Res Clin Rheumatol* 2010; 24:193-204.

10. Stewart MJ, Maher CG, Refshauge KM, Herbert RD, Bogduk N, Nicholas M. Randomized controlled trial of exercise for chronic whiplash-associated disorders. *Pain* 2007; 128:59-68.
11. Teasell RW, McClure JA, Walton D, Pretty J, Salter K, Meyer M, Sequeira K, Death B. A research synthesis of therapeutic interventions for whiplash-associated disorder (WAD): Part 4 - Noninvasive interventions for chronic WAD. *Pain Res Manag* 2010; 15:313-322.
12. Jansen MJ, Viechtbauer W, Lenssen AF, Hendriks EJ, de Bie RA. Strength training alone, exercise therapy alone, and exercise therapy with passive manual mobilisation each reduce pain and disability in people with knee osteoarthritis: A systematic review. *J Physiother* 2011; 57:11-20.
13. Mannerkorpi K, Henriksson, C. Non-pharmacological treatment of chronic widespread musculoskeletal pain. *Best Pract Res Clin Rheumatol* 2007; 21:513-534.
14. Brosseau L, Wells GA, Tugwell P, Egan M, Wilson KG, Dubouloz C-J, Casimiro L, Robinson VA, McGowan J, Busch A, Poitras S, Moldofsky H, Harth M, Finestone HM, Nielson W, Haines-Wangda A, Russell-Doreleyers M, Lambert K, Marshall AD, Veilleux L. Ottawa Panel evidence-based clinical practice guidelines for aerobic fitness exercises in the management of fibromyalgia: Part 1. *Phys Ther* 2008; 88:857-871.
15. Koltyn KF. Exercise-induced hypoalgesia and intensity of exercise. *Sports Med* 2002; 32:477-487.
16. Koltyn KF, Arbogast RW. Perception of pain after resistance exercise. *Br J Sports Med* 1998; 32:20-24.
17. Koltyn KF. Analgesia following exercise: a review. *Sports Med* 2000; 29:85-98.
18. Ray CA, Carter JR. Central modulation of exercise-induced muscle pain in humans. *J Physiol* 2007; 585:287-294.
19. Millan MJ. Descending control of pain. *Prog Neurobiol* 2002; 66:355-474.
20. Fuentes JP, Armijo-Olivo S, Magee DJ, Gross DP. Effects of exercise therapy on endogenous pain-relieving peptides in musculoskeletal pain. *Clin J Pain* 2011; 27:365-374.
21. Meeus M, Roussel N, Truijen S, Nijs J. Reduced pressure pain thresholds in response to exercise in chronic fatigue syndrome but not in chronic low back pain: An experimental study. *J Rehab Med* 2010; 42:884-890.
22. Van Oosterwijck J, Nijs J, Meeus M, Lefever I, Huybrechts L, Lambrecht L, Paul L.. Pain inhibition and post-exertional malaise in myalgic encephalomyelitis/chronic fatigue syndrome: An experimental study. *J Intern Med* 2010; 268:265-278.
23. Van Oosterwijck J, Nijs J, Meeus M, Van Loo M, Paul L. Lack of endogenous pain inhibition during exercise in patients with chronic whiplash associated disorders: An experimental study. *J Pain* 2012; 13:242-254.
24. Koltyn KF, Umeda M. Exercise, hypoalgesia and blood pressure. *Sports Med* 2006; 36:207-214.
25. Ghione S. Hypertension-associated hypalgesia. Evidence in experimental animals and humans, pathophysiological mechanisms, and potential clinical consequences. *Hypertension* 1996; 28:494-504.
26. Dworkin BR, Elbert T, Rau H, Birbaumer N, Pauli P, Droste C, Brunia CH. Central effects of baroreceptor activation in humans: Attenuation of skeletal reflexes and pain perception. *Proc Natl Acad Sci U S A.* 1994; 91:6329-6333.
27. Hoffmann P, Thoren P. Electric muscle stimulation in the hind leg of the spontaneously hypertensive rat induces a long-lasting fall in blood pressure. *Acta Physiol Scand* 1988; 133:211-219.
28. Bender T, Nagy G, Barna I, Tefner I, Kádas É, Géher P. The effect of physical activity on beta-endorphin levels. *Eur J Appl Physiol* 2007; 100:371-382.
29. Koltyn KF, Umeda M. Contralateral attenuation of pain after short-duration submaximal isometric exercise. *J Pain* 2007; 8:887-892.
30. Miura M, Sasaki M, Mizukoshi K, Shibasaki M, Izumi Y, Shimosato G, Amaya F. Peripheral sensitization caused by insulin-like growth factor 1 contributes to pain hypersensitivity after tissue injury. *Pain* 2011; 152:888-895.
31. Murase S, Terazawa E, Queme F, Ota H, Matsuda T, Hirate K, Kozaki Y, Katanosaka K, Taguchi T, Urai H, Mizumura K. Bradykinin and nerve growth factor play pivotal roles in muscular mechanical hyperalgesia after exercise (delayed-onset muscle soreness). *J Neurosci* 2010; 30:3752-3761.
32. Kemppainen P, Pertovaara A, Huopaniemi T, Johansson G. Elevation of dental pain threshold induced in man by physical exercise is not reversed by cyproheptadine-mediated suppression of growth hormone release. *Neurosci Letters* 1986; 70:388-392.
33. Kraemer WJ, Ratamess NA. Hormonal responses and adaptations to resistance exercise and training. *Sports Med* 2005; 35:339-361.
34. Schwarz L, Kindermann W. Changes in beta-endorphin levels in response to aerobic and anaerobic exercise. *Sports Med* 1992; 13:25-36.
35. Godfrey RJ, Madgwick Z, Whyte GP. The exercise-induced growth hormone response in athletes. *Sports Med* 2003; 33:599-613.
36. Pertovaara A. Noradrenergic pain modulation. *Progress in Neurobiology* 2006; 80:53-83.
37. Yarushkina NI. The role of hypothalamo-hypophyseal-adrenocortical system hormones in controlling pain sensitivity. *Neurosci Behav Physiol* 2008; 38:759-766.
38. Fillingim RB, Roth DL, Haley WE. The effects of distraction on the perception of exercise-induced symptoms. *J Psychosom Res* 1989; 33:241-248.
39. Lannersten L, Kosek E. Dysfunction of endogenous pain inhibition during exercise with painful muscles in patients with shoulder myalgia and fibromyalgia. *Pain* 2010; 151:77-86.
40. Kadetoff D, Kosek E. The effects of static muscular contraction on blood pressure, heart rate, pain ratings and pressure pain thresholds in healthy individuals and patients with fibromyalgia. *Eur J Pain* 2007; 11:39-47.
41. Mense S. The pathogenesis of muscle pain. *Current Pain & Headache Rep* 2003; 7:419-425.
42. Larsson R, Öberg Å, Larsson S-E. Changes of trapezius muscle blood flow and electromyography in chronic neck pain due to trapezius myalgia. *Pain* 1999; 79:45-50.
43. Elvín A, Siosteen A, Nilsson A, Kosek E. Decreased muscle blood flow in fibromyalgia patients during standardized muscle exercise. A contrast media enhanced colour doppler study. *Eur J Pain* 2006; 10:137-144.
44. Kosek E, Ekholm J, Hansson P. Modulation of pressure pain thresholds during and following isometric contrac-

tion in patients with fibromyalgia and in healthy controls. *Pain* 1996; 64:415-423.

45. Kosek E, Ekholm J. Modulation of pressure pain thresholds during and following isometric contraction. *Pain* 1995; 61:481-486.
46. Kosek E, Lundberg L. Segmental and plurisegmental modulation of pressure pain thresholds during static muscle contractions in healthy individuals. *Eur J Pain* 2003; 7:251-258.
47. Staud R, Robinson ME, Price DD. Isometric exercise has opposite effects on central pain mechanisms in fibromyalgia patients compared to normal controls. *Pain* 2005; 118:176-184.
48. Ageberg E, Link A, Roos EM. Feasibility of neuromuscular training in patients with severe hip or knee OA: The individualized goal-based NEMEX-TJR training program. *BMC Musculoskeletal Dis* 2010; 11:126.
49. Hurkmans E, Giesen FVd, Vlieland TV, Schoones J, Ende EVd. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database Syst Rev* 2009; Oct 7(4):CD006853.
50. Kosek E, Hansson P. Modulatory influence on somatosensory perception from vibration and heterotopic noxious conditioning stimulation (HNCS) in fibromyalgia patients and healthy subjects. *Pain* 1997; 70:41-51.
51. Leffler AS, Hansson P, Kosek E. Somatosensory perception in a remote pain-free area and function of diffuse noxious inhibitory controls (DNIC) in patients suffering from long-term trapezius myalgia. *Eur J Pain* 2002; 6:149-159.
52. Ring C, Edwards L, Kavussanu M. Effects of isometric exercise on pain are mediated by blood pressure. *Biol Psychol* 2008; 78:123-128.
53. Kadetoff D, Kosek E. Evidence of reduced sympatho-adrenal and hypothalamic–pituitary activity during static muscular work in patients with fibromyalgia. *J Rehabil Med* 2010; 42:765-772.
54. Whiteside A, Hansen S and Chaudhuri A. Exercise lowers pain threshold in chronic fatigue syndrome. *Pain* 2004; 109:497-499.
55. Hoffman MD, Shepanski MA, Mackenzie SP, Clifford PS. Experimentally induced pain perception is acutely reduced by aerobic exercise in people with chronic low back pain. *J Rehabil Res Dev* 2005; 42:183-190.
56. Luo ZD, Cizkova D. The role of nitric oxide in nociception. *Curr Rev Pain* 2000 ;4:459-466.
57. Vikman KS, Hill RH, Backstrom E, Robertson B, Kristensson K. Interferon-gamma induces characteristics of central sensitization in spinal dorsal horn neurons in vitro. *Pain* 2003; 106:241-251.
58. Jungersten L, Ambring A, Wall B, Wennmalm A. Both physical fitness and acute exercise regulate nitric oxide formation in healthy humans. *J Appl Physiol* 1997; 82:760-764.
59. Meeus M, Ickmans K, De Clerck LS, Moorkens G, Hans G, Grosemans S, Nijs J. Seretonergic descending inhibition in chronic pain: Design, preliminary findings and early cessation of a randomized controlled trial. *In Vivo* 2011; 25:1019-1025.
60. Basbaum AI, Fields HL. Endogenous pain control systems: brainstem spinal pathways and endorphin circuitry. *Annu Rev Neurosci* 1984; 7:309-338.
61. Wallman KE, Morton AR, Goodman C, Grove R, Guilfoyle AM. Randomised controlled trial of graded exercise in chronic fatigue syndrome. *Med J Austral* 2004; 180:444-448.
62. Edmonds M, McGuire H, Price J. Exercise therapy for chronic fatigue syndrome. *Cochrane Database Syst Rev* 2004 ;3:CD003200.pub2. DOI: 10.1002/14651858.CD003200.pub2.
63. Dessem D, Ambalavanar R, Evancho M, Moutanni A, Yallampalli C, Bai G. Eccentric muscle contraction and stretching evoke mechanical hyperalgesia and modulate CGRP and P2X(3) expression in a functionally relevant manner. *Pain* 2010 ;149:284-295.
64. Vierck CJ. Mechanisms underlying development of spatial distributed chronic pain (fibromyalgia). *Pain* 2006; 124:242-263.
65. Rygh LJ, Tjolsen A, Hole K, Svendsen F. Cellular memory in spinal nociceptive circuitry. *Scand J Psychol* 2002; 43:153-159.
66. Seminowicz DA, Davies KD. Cortical responses to pain in healthy individuals depends on pain catastrophizing. *Pain* 2006 ;120:297-306.
67. Ploghaus A, Becerra L, Borras C, Borsook D, Neural circuitry underlying pain modulation: Expectation, hypnosis, placebo. *Trends Cogn Sci* 2003; 7:197-200.
68. Brosschot JF. Cognitive-emotional sensitization and somatic health complaints. *Scand J Psychol* 2002; 43:113-121.
69. Lorenz J, Minoshima S, Casey KL. Keeping pain out of mind: the role of the dorsolateral prefrontal cortex in pain modulation. *Brain* 2003; 126:1079-1091.
70. Zusman M. Forebrain-mediated sensitization of central pain pathways: 'non-specific' pain and a new image for MT. *Manual Ther* 2002; 7:80-88.
71. Vowles KE, Ashworth J. Is opioid withdrawal necessary within comprehensive pain rehabilitation programs? *Pain* 2011; 152:1948-1950.
72. Arnold LM, Lu Y, Crofford LJ, Wohlreich M, Detke MJ, Iyengar S, Goldstein DJ. A double-blind, multicenter trial comparing duloxetine with placebo in the treatment of fibromyalgia patients with or without major depressive disorder. *Athritis Rheum* 2004; 50:2974-2984.
73. Chappell AS, Ossanna MJ, Liu-Seifert H, Iyengar S, Skljarevski V, Chunhong Li L, Bennett RM, Collins H. Duloxetine, a centrally acting analgesic, in the treatment of patients with osteoarthritis knee pain: A 13-week, randomized, placebo-controlled trial. *Pain* 2009; 146:253-260.
74. Mauger AR, Jones AM, Williams CA. Influence of acetaminophen on performance during time trial cycling. *J Appl Physiol* 2010; 108:98-104.
75. Pickering G; Estève V, Loriot MA, Eschalier A, Dubray C. Acetaminophen reinforces descending inhibitory pathways. *Clin Pharmacol Ther* 2008; 84:47-51.

Am J Physiol Regulatory Integrative Comp Physiol
279: R1565–R1573, 2000.

Muscle pain perception and sympathetic nerve activity to exercise during opioid modulation

DANE B. COOK,[1] PATRICK J. O'CONNOR,[2] AND CHESTER A. RAY[3]
[1]*Departments of Neuroscience and Anesthesiology, University of Medicine and Dentistry of New Jersey-New Jersey Medical School, Newark, New Jersey 07107-3000;*
[3]*Departments of Medicine and Cellular and Molecular Physiology, Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033;*
and [2]*Department of Exercise Science, University of Georgia, Athens, Georgia 30602-6554*

Received 20 March 2000; accepted in final form 28 June 2000

Cook, Dane B., Patrick J. O'Connor, and Chester A. Ray. Muscle pain perception and sympathetic nerve activity to exercise during opioid modulation. *Am J Physiol Regulatory Integrative Comp Physiol* 279: R1565–R1573, 2000.— The purpose of this experiment was to examine the effects of the endogenous opioid system on forearm muscle pain and muscle sympathetic nerve activity (MSNA) during dynamic fatiguing exercise. Twelve college-age men (24 ± 4 yr) performed graded (1-min stages; 30 contractions/min) handgrip to fatigue 1 h after the ingestion of either 60 mg codeine, 50 mg naltrexone, or placebo. Pain (0–10 scale) and exertion (0–10 and 6–20 scales) intensities were measured during the last 15 s of each minute of exercise and every 15 s during recovery. MSNA was measured continuously from the peroneal nerve in the left leg. Pain threshold occurred earlier [1.8 ± 1, 2.2 ± 1, 2.2 ± 1 J: codeine, naltrexone, and placebo, respectively] and was associated with a lower rating of perceived exertion (RPE) (2.7 ± 2, 3.6 ± 2, 3.8 ± 2: codeine, naltrexone, and placebo, respectively) in the codeine condition compared with either the naltrexone or placebo conditions. There were no main effects (i.e., drugs) or interaction (i.e., drugs × time) for either forearm muscle pain or RPE during exercise [pain: F (2, 22) = 0.69, P = 0.51]. There was no effect of drug on MSNA, heart rate, or blood pressure during baseline, exercise, or recovery. Peak exercise MSNA responses were 21 ± 1, 21 ± 2.0, and 21 ± 2.0 bursts/30 s for codeine, naltrexone, and placebo conditions, respectively. Peak mean arterial pressure responses were 135 ± 4, 131 ± 3, and 132 ± 4 mmHg for codeine, naltrexone, and placebo conditions, respectively. It is concluded that neither 60 mg codeine nor 50 mg naltrexone has an effect on forearm muscle pain, exertion, or MSNA during high- intensity handgrip to fatigue.

codeine; rating of perceived exertion; pain perception; autonomic nervous system

CERTAIN TYPES OF MODERATE-TO-HIGH intensity exercise are perceived as painful. For example, reproducible relationships between objective measures of exercise intensity and subjective judgments of leg and forearm muscle pain intensity during cycle ergometry and rhythmic handgrip have been reported (6, 14). Although it is clear that the firing rate of nociceptive afferent fibers (type III and IV) from skeletal muscle is increased in response to noxious stimuli, including exercise, the mechanisms underlying muscle pain during exercise are poorly understood (16, 17).

Endogenous opioids, such as endomorphin, enkephalins, dynorphins, and beta-endorphins have well-established analgesic actions (11, 31, 33), and opioid receptors are found on nociceptive afferent fibers as well as spinal and supraspinal sites involved in pain processing. The endogenous opioid system has been demonstrated to modulate nociceptive afferent fiber activity in both animal and human experiments (10, 26). Consequently, endogenous opioids may be involved in muscle pain during exercise. Peripheral concentrations of endogenous opioids consistently have been shown to increase during moderate and intense exercise (2, 12, 27). Nevertheless, the role of opioids in naturally occurring muscle pain experienced during exercise is unknown.

In addition to their role in pain regulation, endogenous opioids also have been implicated in the modulation of muscle sympathetic nerve activity (MSNA) responses to exercise (8, 22, 23). The opioid antagonist naloxone has been found to either increase (8) or have no effect (23) on MSNA responses to exercise. Specifically, Farrell et al. (8) reported that 1.2 mg intravenous naloxone significantly increased arterial pressure, plasma epinephrine, and muscle sympathetic nerve responses to 3 min of isometric handgrip at 25% maximal voluntary contraction (MVC). However, this finding was not reproduced by Ray and Pawelczyk (23), who did not observe an effect of naloxone on MSNA, arterial pressure, or heart rate. Moreover, preliminary data from Ray et al. (22) showed that morphine (0.075 mg/kg bolus+1 mg/h maintenance) resulted in increased resting mean arterial pressure but had no

Address for reprint requests and other correspondence: D. B. Cook, UMDNJ-New Jersey Medical School, ADMC Complex 1524A, 30 Bergen St., Univ. Heights, Newark, NJ 07107-3000 (E-mail: cookdb@umdnj.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

effect on arterial pressure, heart rate, or MSNA responses to 2 min of isometric handgrip at 30% MVC. Thus literature regarding the role of endogenous opioids in MSNA responses to exercise is both sparse and equivocal. One limitation of the studies that have employed a direct measure of sympathetic nervous system activity has been the focus on short-duration (2–3 min), low-intensity (25–30% MVC), isometric exercise that may be inadequate to stimulate the endogenous opiate system.

The primary purpose of this experiment was to examine, in a double-blind setting, the effects of codeine and naltrexone on the perception of forearm muscle pain during and after dynamic handgrip performed to fatigue. On the basis of the known influence of opioids in reducing both the activity of nociceptive afferent fibers and pain, it was hypothesized that the ingestion of codeine would result in lower pain intensity ratings compared with both the naltrexone and placebo conditions and that the ingestion of naltrexone would result in higher pain ratings compared with both the codeine and placebo conditions. The rationale for using a dynamic handgrip stimulus to fatigue was to *1*) achieve longer and more intense exercise sessions, thus increasing the likelihood of an endogenous opioid response and *2*) allow for the measurement of MSNA. Additionally, rhythmic exercise differs from the isometric protocols employed by previous investigators (8, 23) in that it does not result in continuous muscle ischemia. A second purpose was to examine the effects of codeine and naltrexone on MSNA during and after dynamic fatiguing handgrip. It was hypothesized that the ingestion of codeine would decrease MSNA responses to exercise compared with both the naltrexone and placebo conditions and that naltrexone would increase MSNA responses to exercise compared with both the codeine and placebo conditions.

METHODS

Participants. A total of 12 college-age (18–35 yr) men who were not on any medication and pain and injury free volunteered to participate in the study. A sample size of 12 provided a statistical power of 0.76 for the primary question concerning the main effect of drug on pain. This value was calculated on the basis of statistical power tables for repeated measures designs (19) and with the following assumptions: an alpha level of 0.05; a one-half SD (0.50) for the drug main effect; a correlation across exercise intensity (trials) of 0.6; and a correlation across conditions (codeine, naltrexone, and placebo) of 0.4. All participants signed a consent form approved by the Institutional Review Board at The University of Georgia. Selected characteristics (means ± SD) were as follows: age (24 ± 4 yr), height (176 ± 5 cm), and weight (75 ± 9 kg).

Procedures. The participants completed three questionnaires: a medical and a 24-h health history, and the multiple affect adjective checklist (MAACL). The 24-h and medical histories were used to inquire about the use of medications and to ensure that the participants were healthy, injury free, able to perform the maximal exercise test, and not allergic to codeine or naltrexone. None of the participants reported any forearm muscle soreness before the exercise sessions. The MAACL was employed to examine the potential role of either codeine, naltrexone, or placebo on the participant's affective state before exercise and to examine possible relationships between situational affect and muscle pain during exercise. The 132-item MAACL provides valid measures of anxiety, depression, hostility, positive affect, and sensation seeking (34). Participants were given the MAACL before each exercise session exactly 50 min after ingestion of the placebo or active drug.

MSNA. MSNA measurements were made as described by Ray et al. (24). Briefly, multiple nerve fiber recordings of MSNA were made by using a tungsten microelectrode inserted in the peroneal nerve near the head of the fibula. A reference electrode was placed subcutaneously ~2 cm from the recording electrode. Adjustments of the microelectrode were made until a site exhibiting clear spontaneously occurring sympathetic bursts was found. MSNA was expressed as burst frequency (bursts/30 s) and total activity (area/30 s). Total activity was the sum of area of all bursts in a given time period.

Arterial pressure and heart rate were measured continuously by using an Ohmeda Finapres recorder (model 2300, Englewood, NJ). The photoplethysmographic cuff was placed on the middle finger of the nonexercising arm, which was maintained at the level of the heart during testing.

For each exercise test MSNA, arterial pressure, and heart rate data were obtained before (5 min), during (~14 min), and after (5 min) exercise. In the event that the recording electrode came out of the nerve, one of the investigators quickly adjusted the electrode until the MSNA recording was reestablished. This problem occurred once in only one subject during recovery from exercise.

Codeine/naltrexone/placebo conditions. In the codeine condition the participants received 60 mg of codeine in one capsule. Studies using similar doses have demonstrated the analgesic effectiveness of 60 mg of codeine to experimental pain stimuli (18, 29), whereas others have reported little or no side effects when a single oral dose is taken (1, 20). The codeine was ingested with 8 oz of water 60 min before exercise, and consumption was witnessed by one of the investigators. The timing of drug administration was based on previous reports that maximal plasma concentrations of codeine occur ~1 h after a single oral dose of 60 mg (11, 20). The capsule was identical to the capsules used in the placebo and naltrexone conditions, a procedure designed to blind the participant as to the condition. In the naltrexone condition the participants received 50 mg of naltrexone in one capsule taken orally. Naltrexone is an opioid antagonist that exhibits no agonist effects. The naltrexone was administered with 8 oz. of water 60 min before exercise in a manner identical to the codeine and placebo condition (lactose capsule). Naltrexone taken orally has a peak plasma concentration within 60 min and has been reported to produce sustained effects for up to 24 h from a single oral dose (11). To ensure a double-blind administration, one investigator, who did not conduct the exercise tests (P. J. O'Connor), distributed the capsules to a second investigator (D. B. Cook). The second investigator, who was unaware of what the capsules contained, administered all the capsules and conducted all of the exercise tests. The order in which the participants completed the conditions was randomized and counterbalanced.

Graded handgrip protocol. Each participant completed a minimum of three graded, dynamic handgrip exercise tests to fatigue. The exercise sessions were performed on 3 separate days. Before exercise the participants received either codeine, naltrexone, or placebo in a randomized, counterbalanced order. Exercise was performed with the dominant hand while the participants were in a supine position. The dominant hand and arm were extended laterally (60–90° angle) from the body and fully supported. Handgrip was performed at a rate of 30 contractions/min with the aid of a calibrated metronome. The first minute of exercise was done with no load. For each subsequent 1-min stage, the weight was increased by 1.13 kg until an exercise stage could not be completed. The participants were given 15 s of rest every minute while one of the investigators added 1.13 kg to the weight-support bar.

Pain and exertion assessment. Forearm muscle pain intensity was assessed by using a category scale with ratio properties. The pain intensity scale ranges from 0 (no pain at all) to 10 (extremely intense pain, almost unbearable). With this scale, if the subjective intensity increases above 10, the subject is free to choose any number larger in proportion to 10 that describes the proportionate growth of the sensation. Prior work with this instrument has provided evidence for both the validity and reliability of this tool for quantifying naturally occurring muscle pain during exercise (6).

The participants listened to a taped set of instructions, which informed them that they would be repeatedly asked to rate the intensity of the pain and exertion in their forearms and that they were to report aloud the number that corresponds to that intensity. Additionally, participants were instructed to remember to say the word "pain" when the pain in their forearm became just noticeable (pain threshold). An investigator recorded the minutes and seconds from a digital timer that was started at the beginning of exercise, and this quantified pain threshold.

Ratings of perceived exertion (RPE) were assessed during and after exercise by using Borg's 6–20 category scale (4) after explicit audiotaped and oral instructions (cf. 6). Our previously employed instructions were modified to obtain local ratings of forearm muscle exertion. A second scale for measuring perceived exertion [0–10 category-ratio scale; Borg, (3)] was employed to examine the relationship between ratings of pain and RPE by using scales that were both based on ratio-scaling methods.

During the graded exercise test forearm muscle pain and exertion ratings were obtained during the last 15 s of every 1-min exercise stage. At the point of fatigue (i.e., failure to maintain a handgrip contraction rate of 30 repetitions/min), the test was stopped, and the participants were asked to immediately stop contracting their forearm. Pain and exertion ratings were obtained every 15 s for 5 min during the recovery period to assess the abatement of pain and exertion perceptions.

Posttest information. Within 5 min after the completion of the exercise test, participants indicated in writing the reason they stopped exercising. Thereafter, they completed the short-form of the McGill Pain Questionnaire (MPQ) (15) to provide a multidimensional description of the forearm muscle pain experienced during the exercise test.

Primary statistical analyses. Pain ratings, RPE, and MSNA were analyzed by using a two-way [condition (codeine, naltrexone, placebo) × trials] repeated-measures ANOVA. The trials factor ranged from 5 different intensities for data obtained in association with exercise intensity [20, 40, 60, 80, and 100% of peak work (J)] to 20 time points associated with recovery (ratings obtained every 15 s for 5 min). When appropriate, eta^2 was used as a measure of the magnitude of an association among variables. Rough guidelines for the strength of association for a given eta^2 value are that 0.01 is considered a small effect, 0.09 medium, and 0.15 large. Pain threshold, peak pain, and baseline MSNA, arterial pressure, and heart rate among conditions were analyzed by using a one-way repeated-measures ANOVA. Significant main effects were followed up by simple contrast analysis to further delineate where the significant differences occurred. *T*-tests were used to analyze baseline MSNA, arterial pressure, heart rate, and selected pain threshold and peak variables between the control and placebo conditions. Pearson correlations were used to examine relationships between and among arterial pressure, MSNA, pain, and RPE. Effect sizes (d) were calculated according to the method described by Cohen (5) (mean 1 − mean 2/pooled SD) to provide a measure of the magnitude of the differences between selected variables. All table values are expressed as means ± SD, and all graphic values are means ± SE.

RESULTS

Pain threshold. Forearm muscle pain threshold data for codeine, naltrexone, and placebo conditions are presented in Table 1. One-way repeated-measures ANOVA revealed significant main effects for the amount of work completed at pain threshold and the perceived exertion using both the 0–10 and the 6–20 scales. Contrast analysis revealed that pain threshold was reported at a lower weight and RPE ratings were lower in the codeine condition compared with both the naltrexone and placebo conditions.

Peak exercise data. Peak exercise data for the codeine, naltrexone, and placebo conditions are shown in Table 2. There were no significant differences across conditions for any of the variables measured at peak exercise. Data based on verbal reports obtained postexercise indicated that six individuals, in at least one of three exercise tests, stopped exercising due in part to the pain they felt in their forearms. Overall, pain was a factor in the decision to stop exercising in 10 of the 36 exercise tests, whereas fatigue was indicated in 34 of

Table 1. *Measures of exercise intensity at pain threshold for codeine, naltrexone, and placebo conditions*

Variable	Codeine	Naltrexone	Placebo	*F* Value	*P* Value
Work, J	1.8 ± 1	2.2 ± 1	2.2 ± 1	5.1(2,22)	0.01
Perceived exertion (0–10)	2.7 ± 2	3.6 ± 2	3.8 ± 2	5.4(2,22)	0.01
Perceived exertion (6–20)	10.2 ± 3	11.4 ± 2	11.9 ± 3	4.3(2,22)	0.02

Values are means ± SD; n = 12 men.

Table 2. *Peak exercise data*

Variable	Codeine	Naltrexone	Placebo
Peak pain (0–10)	7.7 ± 3	8 ± 4	8.2 ± 4
Peak RPE (0–10)	9.9 ± 1	10.3 ± 2	11.4 ± 4
Peak RPE (6–20)	19.7 ± 1	19.3 ± 1	19.4 ± 1
Peak work, J	4.9 ± 1	4.8 ± 1	4.9 ± 1

Values are means ± SD; n = 12 men. RPE, rating of perceived exertion.

36 tests (both pain and fatigue were reported by some subjects).

Pain during exercise. Pain intensity ratings at 20, 40, 60, 80, and 100% of peak work (J) are shown in Fig. 1, *top*. Pain increased as a positively accelerating function of percent peak work (J) during maximal handgrip exercise [F (2, 22) = 55.7, $P < 0.001$]. There was no significant main effect for condition or interaction.

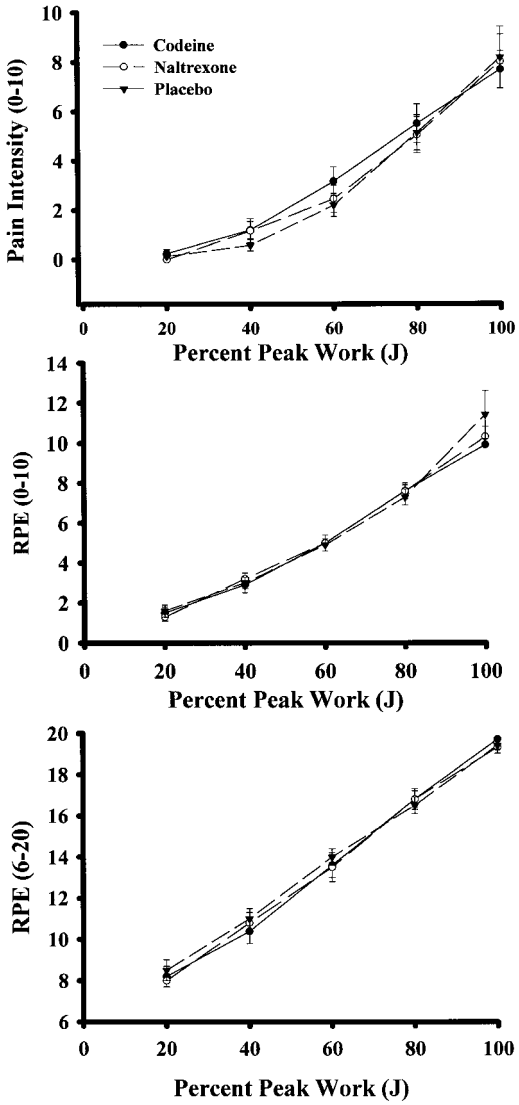


Fig. 1. Pain intensity (0–10 scale; *top*) and rating of perceived exertion [RPE; 0–10 (*middle*) and 6–20 scales (*bottom*)] at 20, 40, 60, 80, and 100% of peak work (J) during fatiguing handgrip for codeine, naltrexone, and placebo conditions (n = 12).

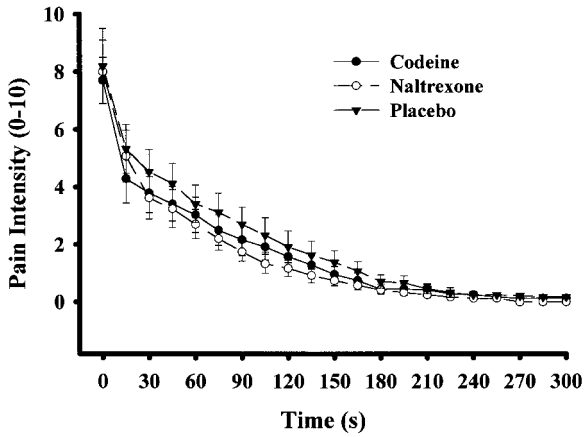


Fig. 2. Pain intensity ratings (0–10 scale) during recovery from fatiguing handgrip for codeine, naltrexone, and placebo conditions (n = 12).

Perceived exertion during exercise. Perceived exertion ratings at 20, 40, 60, 80, and 100% of peak work (J) are illustrated in Fig. 1, *middle* and *bottom*. Perceived exertion ratings using both the 0–10 and 6–20 scales increased as a function of percent peak work (J) [0–10 scale: F (2, 22) = 179.9, $P < 0.001$; 6–20 scale: F (2, 22) = 455.9, $P < 0.001$]. There were no significant main effects for condition or interactions for either scale.

Pain and exertion during recovery from handgrip exercise. Pain intensity ratings during recovery from maximal handgrip are shown in Fig. 2. Two-way repeated-measures ANOVA revealed no significant main effect for condition [F (2, 22) = 2.5, P = 0.10]; however, a significant interaction was detected [F (38, 418) = 1.5, P = 0.026, eta^2 = 0.12]. The SE bands in Fig. 2 show greater variability in the placebo condition. Inspection of individual responses revealed one influential subject exhibiting ratings that were >2 SD above the group mean at *trials 3*, *4*, and *6–14* in the placebo condition. There were no significant main effects or interactions for RPE during recovery (0–10 or 6–20 scales).

MPQ data. MPQ data for the codeine, naltrexone, and placebo conditions are presented in Table 3. There

Table 3. *McGill Pain Questionnaire data for codeine, naltrexone, and placebo conditions*

Variable	Codeine	Naltrexone	Placebo
Cramping pain	1.7 ± 1	1.6 ± 1	1.5 ± 1
Hot-burning pain	1.8 ± 1	1.8 ± 1	1.9 ± 1
Aching pain	1.9 ± 1	1.7 ± 1	1.8 ± 1
Tiring-exhausting pain	2.4 ± 1	2.2 ± 1	2.4 ± 1

Values are means ± SD and represent the average pain ratings based on scores of *1*) mild, *2*) moderate, and *3*) severe. The table represents the 4 verbal descriptors that were used most frequently and given the highest mean rating when the participants recalled the forearm muscle pain experienced during maximal handgrip exercise.

were no differences among the conditions for any of the verbal descriptors chosen to describe the forearm muscle pain, or its intensity, experienced during maximal handgrip.

Baseline cardiovascular and MSNA data. Preexercise baseline values for MSNA, heart rate, and mean arterial pressure are shown in Fig. 3. MSNA, heart rate, and mean arterial pressure were not significantly different across conditions before exercise.

MSNA, heart rate, and arterial pressure during exercise. MSNA burst frequency (bursts/30 s) and total activity (area/30 s) during handgrip for codeine, naltrexone, and placebo conditions are illustrated in Fig. 4. MSNA burst frequency increased during handgrip exercise [trial main effect: $F(4, 44) = 65.1, P < 0.001$]; however, there was no main effect for condition or interaction. Total MSNA increased during exercise [trial main effect: $F(4, 44) = 10.6, P < 0.001$], but there was no main effect for condition or interaction. A similar pattern of response, both physiologically and in terms of the ANOVA results, was observed for heart rate (peak values: codeine = 136 ± 16, naltrexone = 130 ± 10, placebo = 135 ± 14) and mean arterial pressure (peak values: codeine = 79 ± 14, naltrexone = 78 ± 12, placebo = 81 ± 9).

MSNA, heart rate, and arterial pressure during recovery from handgrip exercise. No significant main effects or interactions were observed during recovery from maximal handgrip for any of the MSNA, heart rate, or arterial pressure data.

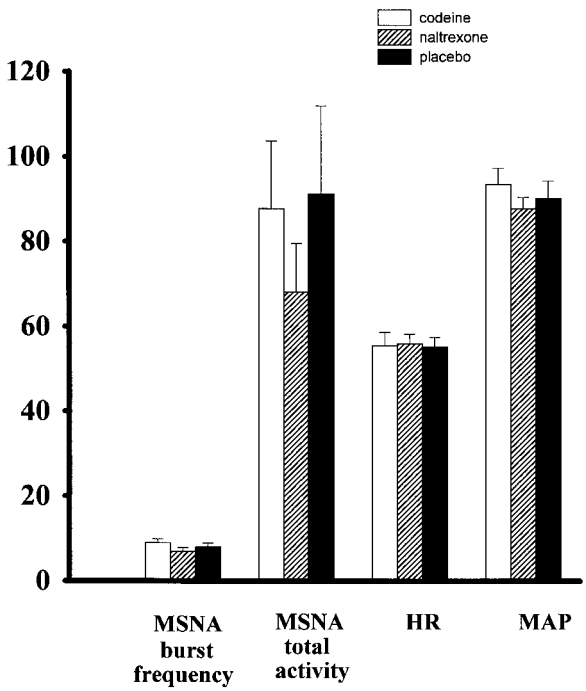


Fig. 3. Preexercise baseline values for muscle sympathetic nerve activity (MSNA; frequency: average bursts/30 s; total activity average area/30 s), heart rate (HR; beats/min), and mean arterial pressure (MAP; mmHg) for codeine, naltrexone, and placebo conditions ($n = 10$).

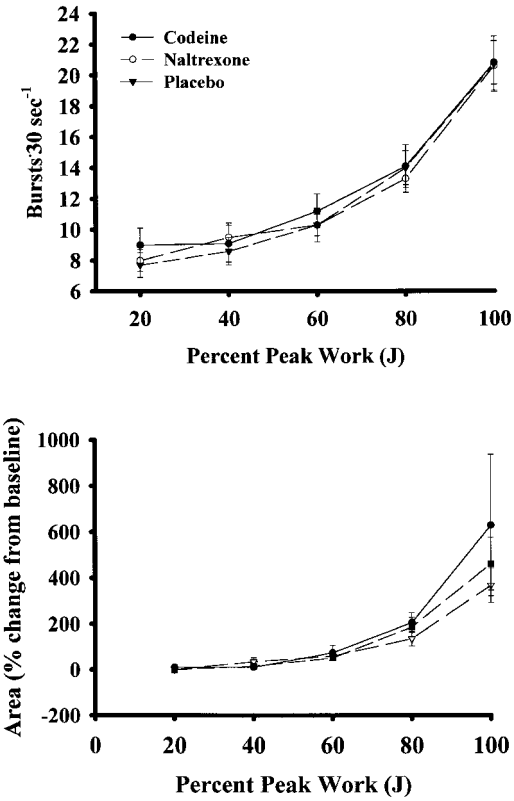


Fig. 4. MSNA [burst frequency (*top*) and total activity (*bottom*)] expressed as percent change from baseline at 20, 40, 60, 80, and 100% of peak work (J) during fatiguing handgrip for codeine, naltrexone, and placebo conditions ($n = 12$).

Selected relationships of interest. A strong negative relationship between forearm muscle pain ratings and systolic arterial pressure at 100% of peak exercise intensity in the placebo condition was observed ($r = -0.84$) and is illustrated in Fig. 5. Correlations between pain ratings and systolic arterial pressure during submaximal exercise ranged from −0.27 to 0.17. A significant negative correlation between pain and arterial pressure at 100% of peak exercise in the naltrexone condition was also observed ($r = -0.64$), whereas a significant positive correlation between pain and arterial pressure at 100% of peak exercise ($r = 0.74$) was observed in the codeine condition.

Nonsignificant bivariate correlations in the placebo condition indicated that pain was nonsignificantly related to exertion (0–10) during exercise (20% $r = 0.25$, 40% $r = -0.04$, 60% $r = 0.06$, 80%, $r = 0.14$, 100% $r = -0.02$). Relationships between pain and MSNA in the placebo condition also were low and nonsignificant (pain and burst frequency: 20% $r = -0.04$, 40% $r = -0.14$, 60% $r = -0.44$, 80% $r = 0.24$, 100% $r = 0.20$; pain and area: 20% $r = 0.12$, 40% $r = -0.22$, 60% $r = -0.21$, 80% $r = -0.08$, 100% $r = 0.24$).

DISCUSSION

The primary findings of the present investigation are that *1*) post-pain threshold, neither codeine nor nal-

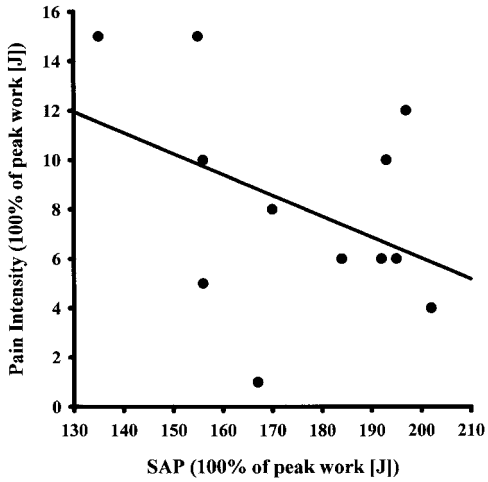


Fig. 5. Illustration of the relationship between pain and systolic arterial pressure (SAP) at 100% of peak work (J) in the placebo condition during fatiguing handgrip (n = 12).

trexone altered the perception of forearm muscle pain and *2*) neither drug had an effect on MSNA burst frequency or total activity.

Muscle pain. The lack of an effect for either codeine or naltrexone on muscle pain during exercise would seem to be at odds with what is known about type III and IV afferent fiber activity and the role that opioids play in regulating their activity. However, it is worth emphasizing that pain is a complex phenomenon that involves multiple systems working in parallel to regulate nociceptive activity. The simple act of inhibiting one of the modulatory systems may not result in an alteration in the subjective experience because of the redundancy of the system. Bradykinin, potassium, serotonin, and histamine, all of which have been shown to be released during exercise (16, 25, 30), act directly on type IV fibers, resulting in an increased firing rate (16). These algesics also sensitize the type III fibers that respond to the increased intramuscular pressure during exercise (16, 25). There are also a population of small afferent fibers reported to become active only during muscular contractions and that respond linearly to the force of the contraction (16). Combined, these algesic mechanisms increase the likelihood that the afferent nociceptive signal will be transmitted to supraspinal sites that are important in pain perception even when one possible mechanism is eliminated.

The timing of the administration of both drugs was based on reports that peak plasma concentrations of both codeine (and its *O*-demethylation to morphine) and naltrexone occur ~1 h after oral administration (20). We chose a dosage that is double that frequently prescribed for mild-to-moderate pain (28). Moreover, we did not want to exceed 60 mg because it has been reported that codeine administered at doses above this amount increases the likelihood of unwanted side effects, which in and of themselves could have altered pain responses (28). Other researchers consider the dose of naltrexone used in this study to be large (9, 11). Consequently, the naltrexone dose should have been adequate for the purposes of the present investigation. Thus the present results suggest that the endogenous opioid system, particularly the mu receptor-mediated portion, does not play a major role in the perception of muscle pain during exercise.

In accordance with our previous research examining leg muscle pain during maximal cycle ergometry (6), forearm muscle pain experienced during maximal handgrip exercise increased as a function of the exercise stimulus. Pain threshold in the placebo condition occurred on average at ~ 46% (6.6 kg) of the peak exercise intensity (14.3 kg), which is comparable to leg muscle pain threshold (~50%) observed during ramped maximal cycling exercise (6). Peak forearm muscle pain ratings averaged 8.2 on the 0–10 scale, a value identical to the average peak leg muscle pain reported in our previous cycle ergometry study (6). Moreover, the MPQ verbal descriptors most frequently used to describe forearm muscle pain were tiring-exhausting, aching, hot-burning, and cramping pain, which are also the same descriptors that were used to describe leg muscle pain during exercise. These results provide additional support for the reliability and validity of the 0–10 pain intensity scale used to quantify muscle pain during exercise. Moreover, these findings suggest that the intensity and quality of muscle pain experienced during graded or ramped maximal exercise tests are similar whether a large or small muscle mass is employed.

In the codeine condition, pain threshold occurred earlier and corresponded with lower exertion ratings than both the naltrexone and placebo conditions. The difference represented about one exercise stage (1.13 kg or 1 min of exercise) and corresponded to a moderate effect (d = ~0.50) for both comparisons. It is unclear why such an effect occurred. This effect, however, did not translate into altered pain ratings or performance during exercise or altered pain ratings during recovery. Therefore, simply altering someone's threshold for pain detection does not necessarily mean that the individual will experience changes in pain intensity above pain threshold. This is important because pain threshold is usually of little concern in sports. For example, with endurance athletes the greater concern is how long they can tolerate an intense bout of exercise.

There is a great deal of literature showing relationships between arterial pressure and pain (13, 21). It has been reported that both chronic hypertension and acute experimental increases in arterial pressure are associated with reduced pain sensitivity (13, 21). Experimental stimulation of arterial and cardiopulmonary baroreceptors (e.g., neck suction, physiological volume expansion, or pharmacological sympathetic stimulation) results in antinociceptive behavior in both animals and humans (7, 13, 21). Conversely, significant and parallel relationships between the degree of ischemic forearm muscle pain (produced during a sub-

maximal-effort tourniquet test) and arterial pressure have been reported (14).

In the present investigation, there was a significant negative relationship between pain ratings and systolic arterial pressure, but only at the highest exercise intensity (100% of peak exercise intensity) and only in the placebo and naltrexone conditions. It is unclear why this occurred only at the peak of exercise. This finding does suggest a link between elevated arterial pressure and pain inhibition and is in agreement with previous research demonstrating that arterial baroreceptor stimulation inhibits pain in rats and humans (7, 13, 21). However, given our sample size and lack of consistency across conditions, these results should be interpreted with caution. Nevertheless, this observation, were it found to generalize, would have potentially important clinical implications. For example, it might aid in our understanding and identification of those at risk for "silent" myocardial infarctions during exercise.

Perceived exertion. There was no effect of drug on perceived exertion during exercise or in recovery. This is not surprising given the lack of an effect of drug on the perception of pain and performance. In our previous work examining pain and exertion during exercise, we chose to employ the Borg 6–20 category scale for the measurement of perceived exertion and a 0–10 category-ratio scale to assess pain. The goal of that research was to determine the extent to which the constructs of pain and exertion could be differentiated during exercise. During the cycle ergometry protocol, perceived effort ratings were reported at low exercise intensities, but naturally occurring leg muscle pain typically did not occur until a moderate intensity of ~ 50% of peak power output was reached. Moreover, leg pain was only moderately related to exertion during exercise. From that study, it was concluded that pain and exertion were two separate, albeit related, constructs (6). However, the Borg 6–20 category scale is theoretically and empirically distinct from the 0–10 category-ratio scale, and the use of the two different scales may have made the original comparison less compelling. Therefore, in the present investigation we chose to add a measure of RPE and use Borg's 0–10 category-ratio scale (3). The pain and RPE 0–10 category-ratio scales are designed to overcome limitations of category scales (e.g., ceiling effects) by allowing users to choose a number above 10 when necessary. These scales are not only designed to have ratio properties (i.e., possessing a true 0 and unbounded), but they have been shown to perform similarly compared with ratio scaling methods (3, 6). The low correlations observed ($r = -0.04$ to 0.25) between pain (0–10) and RPE (0–10) in this experiment, and the occurrence of exertion before pain, again support the contention that pain and exertion are two separate constructs. The distinction between pain and exertion is potentially important because it allows researchers to determine the role of pain in various types of exercise performance. It also allows researchers to learn whether pain per se influences exertional perceptions.

MSNA. The results of the present experiment, while not directly comparable, appear to be in contrast to previous results obtained by Farrell et al. (8), who showed an augmentation of naloxone on MSNA during moderate-intensity, brief-duration, isometric handgrip. However, the results are in agreement with Ray and Pawelczyk (23), who showed no effect of naloxone on MSNA during moderate-intensity, brief-duration, isometric handgrip.

It is not immediately clear why different results were obtained in the earlier works by Ray and Pawelczyk (23) and Farrell et al. (8). However, one criticism of prior work attempting to examine the influence of the endogenous opioid system on MSNA responses to exercise has been the use of exercise stimuli that were of low intensity, short duration, and involved static contractions. Thus the exercise stimulus itself may not have been sufficient to stimulate the endogenous opioid system. Alternatively, it may be that the exercise stimulus used previously was not painful and thus negated the potential effects of naloxone. The strength of the present experimental design was that we employed painful, high-intensity, and dynamic exercise to fatigue while examining the MSNA responses after administration of both an opioid agonist and antagonist. Rhythmic exercise is different from isometric exercise used in our earlier studies because rhythmic exercise does not elicit continuous ischemia. Therefore, our results extend our previous findings to rhythmic exercise and strongly suggest that the endogenous opioid system does not modulate MSNA during handgrip.

No prior studies have adequately determined whether pain and MSNA are related. One study, in which muscle pain was poorly assessed and experienced by only 3 of 25 subjects, reported no relationship between pain and MSNA (32). In the present investigation, pain ratings during exercise were weakly or moderately related to MSNA burst frequency or total activity (pain and burst frequency: $r = -0.44$ to 0.24 and pain and total activity: $r = -0.21$ to 0.23). This lack of a strong relationship is not surprising given that nociceptive signals are modified at spinal and supraspinal sites that are independent of the sympathetic nervous system. Moreover, pain perception was not altered in the present study, which may have limited the potential for observing stronger relationships between MSNA and pain.

The present study could not definitively assess whether pain perception during exercise was altered by central mechanisms (i.e., higher brain systems). It is possible that "central command," associated with volitional effort, may have interacted with afferent feedback from the muscle to modulate pain perception. Future studies using postexercise muscle ischemia, which eliminates central command, may be useful in addressing this issue. In the present study, postexercise muscle ischemia was not assessed because it would have confounded our postexercise responses.

In summary, the results from this study indicate that the ingestion of either 60 mg of codeine or 50 mg

of naltrexone does not alter the perception of forearm muscle pain or exertion during maximal handgrip exercise to fatigue. Additionally, neither drug had an effect on MSNA when expressed in terms of burst frequency nor total integrated activity. We conclude that codeine and naltrexone, in practical doses, do not have an effect on naturally occurring muscle pain, exertion, or MSNA during exercise. Whether this finding generalizes to other opioid agonists and different modes of exercise remains to be tested.

Perspectives

The experience of muscle pain during exercise is a common phenomenon: the runner rounding the final curve and sprinting toward the finish of the 1,500 m at the Olympic Games, the 90-yr-old great-grandmother climbing a flight of stairs, and the patient with peripheral vascular disease simply walking to the grocery store all can experience intense muscle pain during exercise. These perceptions of pain provide strong motivation. Pain motivates the athlete, the great-grandmother, and the peripheral vascular disease patient all to slow down, presumably so they do not seriously injure themselves. By learning more about the mechanisms underlying this type of exercise-generated muscle pain, we may eventually be able to improve both well-being and athletic performance. We are surprised that muscle pain is one of the least studied types of pain and urge applied physiologists interested in muscle to consider including the perception of pain as a dependent measure in their future investigations involving exercise.

We acknowledge Keith M. Hume for invaluable services during the data collection stages of this project.

This investigation was supported by the Department of Exercise Science at the University of Georgia via the 1997 Louise E. Kindig Research Award (to D. B. Cook), National Institutes of Health Grants AR-44571 and HL-8503, and National Aeronautics and Space Administration Grant 9–1034 (to C. A. Ray).

REFERENCES

1. **Band CJ, Band PR, Deschamps M, Besner JG, and Coldman AJ.** Human pharmacokinetic study of immediate-release (codeine phosphate) and sustained-release (codeine contin) codeine. *J Clin Pharmacol* 34: 938–943, 1994.
2. **Boone JB, Sherraden T, Pierzchala K, Berger R, and Van Loon GR.** Plasma met-enkephalin and catecholamine responses to intense exercise in humans. *J Appl Physiol* 73: 388–392, 1992.
3. **Borg G.** A ratio scaling method for interindividual comparisons. In: *Reports from the Institute of Applied Psychology.* Stockholm: The University of Stockholm, vol. 27, 1972, p. 1–8.
4. **Borg G.** Subjective aspects of physical and mental load. *Ergonomics* 21: 215–220, 1978.
5. **Cohen J.** *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum, 1988.
6. **Cook DB, O'Connor PJ, Eubanks SA, Smith JC, and Lee M.** Naturally occurring muscle pain during exercise: assessment and experimental evidence. *Med Sci Sports Exerc* 29: 999–1012, 1997.
7. **Dworkin BR, Elbert T, Rau H, Birbaumer N, Pauli P, Droste C, and Brunia CHM.** Central effects of baroreceptor activation in humans: attenuation of skeletal reflexes and pain perception. *Proc Nat Acad Sci USA* 91: 6329–6333, 1994.
8. **Farrell PA, Ebert TJ, and Kampine JP.** Naloxone augments muscle sympathetic nerve activity during isometric exercise in humans. *Am J Physiol Endocrinol Metab* 260: E379–E388, 1991.
9. **Farrell PA, Gustafson AB, Garthwaite TL, Kalkoff RK, Cowley AW, and Morgan WP.** Influence of endogenous opioids on the response of selected hormones to exercise in humans. *J Appl Physiol* 61: 1051–1057, 1986.
10. **Hill JM and Kaufman MP.** Attenuation of reflex pressor and ventilatory responses to static muscular contraction by intrathecal opioids. *J Appl Physiol* 68: 2466–2472, 1990.
11. **Julien RM.** Narcotic and narcotic drugs. In: *A Primer of Drug Action.* New York: Freeman, 1995, p. 233–268.
12. **Kraemer WJ, Dziados JE, Marchitelli LJ, Gordon SE, Harmon EA, Mello R, Fleck SJ, Frykman PN, and Triplett NT.** Effects of different heavy-resistance exercise protocols on plasma β-endorphin concentrations. *J Appl Physiol* 74: 450–459, 1993.
13. **Lovick TA.** Integrated activity of cardiovascular and pain regulatory systems: role in adaptive behavioral responses. *Prog Neurobiol* 40: 631–644, 1993.
14. **Maixner W, Gracely RH, Zuniga JR, Humphrey CB, and Bloodworth GR.** Cardiovascular and sensory responses to forearm ischemia and dynamic hand exercise. *Am J Physiol Regulatory Integrative Comp Physiol* 259: R1156–R1163, 1990.
15. **Melzack R.** The short-form McGill Pain Questionnaire. *Pain* 30: 191–197, 1987.
16. **Mense S.** Group III and IV receptors in skeletal muscle: are they specific or polymodal? *Prog Brain Res* 113: 83–100, 1996.
17. **Mitchell JH, Kaufman MP, and Iwamoto GA.** The exercise pressor reflex: its cardiovascular effects, afferent mechanisms, and central pathways. *Ann Rev Physiol* 45: 229–242, 1983.
18. **Posner J.** A modified submaximal effort tourniquet test for evaluation of analgesics in healthy volunteers. *Pain* 19: 143–151, 1984.
19. **Potvin PJ and Schutz RW.** *Statistical Power for Repeated Measures ANOVA. Part B*. Vancouver, Canada: University of British Columbia, 1996.
20. **Quiding H, Anderson P, Bondesson U, Boreus LO, and Hynning PA.** Plasma concentrations of codeine and its metabolite, morphine, after single and repeated oral administration. *Eur J Clin Pharmacol* 30: 673–677, 1986.
21. **Randich A and Maixner W.** Interactions between cardiovascular and pain regulatory systems. *Neurosci Biobehav Rev* 8: 343–367, 1984.
22. **Ray CA, Hill GA, and Anderson EA.** Morphine does not alter muscle sympathetic nerve response to brief submaximal isometric handgrip in humans (Abstract). *FASEB J* 9: 8A 1995.
23. **Ray CA and Pawelczyk JA.** Naloxone does not affect the cardiovascular and sympathetic adjustments to static exercise in humans. *J Appl Physiol* 77: 231–235, 1994.
24. **Ray CA, Rea RF, Clary MP, and Mark AL.** Muscle sympathetic nerve responses to static leg exercise. *J Appl Physiol* 73: 1523–1529, 1992.
25. **Rotto DM, Schultz HD, Longhurst JC, and Kaufman MP.** Sensitization of group III afferents to static contraction by arachidonic acid. *J Appl Physiol* 68: 861–867, 1990.
26. **Schobel HP, Handwerker HO, Schmieder RE, Heusser K, Dominak P, and Luft FC.** Effect of naloxone on hemodynamic and sympathetic nerve responses to pain in normotensive vs. borderline hypertensive men. *J Auton Nerv Syst* 69: 49–55, 1998.
27. **Schwarz L and Kindermann W.** Changes in beta-endorphin levels in response to aerobic and anaerobic exercise. *Sports Med* 13: 25–36, 1992.
28. **Sindrup SH and Brosen K.** The pharmacogenetics of codeine hypoalgesia. *Pharmacogenetics* 5: 335–346, 1995.
29. **Stacher G, Steinringer H, Schneider S, Mittelbach G, Winklehner S, and Gaupmann G.** Experimental pain induced by electrical and thermal stimulation of the skin in healthy man: sensitivity to 75 and 150 mg diclofenac sodium compared with 60 mg codeine and placebo. *Br J Clin Pharmacol* 21: 35–43, 1986.
30. **Stebbins CL, Carretero OA, Mindroiu T, and Longhurst JC.** Bradykinin release from contracting skeletal muscle of the cat. *J Appl Physiol* 69*:* 1225–1230, 1990.
31. **Stein FH.** The control of pain in peripheral tissue by opiates. *Mech Dis* 332: 1685–1690, 1995.

32. **Victor RG, Seals DR, and Mark AL.** Differential control of heart rate and sympathetic nerve activity during dynamic exercise. *J Clin Invest* 79: 508–516, 1987.
33. **Yaksh TL and Malmberg AB.** Central pharmacology of nociceptive transmission. In: *Textbook of Pain*, edited by Wall PD and Melzack R. New York: Churchill Livingstone, 1994, p. 165–200.
34. **Zuckerman M and Lubin B.** *Manual for the MAACL-R: The Multiple Affect Adjective Checklist Revised.* San Diego, CA: Educational and Industrial Testing Service, 1985.

arthritis research&therapy

RESEARCH ARTICLE Open Access

Efficacy of different types of aerobic exercise in fibromyalgia syndrome: a systematic review and meta-analysis of randomised controlled trials

Winfried Häuser*[1,2], Petra Klose[3], Jost Langhorst[3], Babak Moradi[4], Mario Steinbach[4], Marcus Schiltenwolf[4] and Angela Busch[5]

Abstract

Introduction: The efficacy and the optimal type and volume of aerobic exercise (AE) in fibromyalgia syndrome (FMS) are not established. We therefore assessed the efficacy of different types and volumes of AE in FMS.

Methods: The Cochrane Library, EMBASE, MEDLINE, PsychInfo and SPORTDISCUS (through April 2009) and the reference sections of original studies and systematic reviews on AE in FMS were systematically reviewed. Randomised controlled trials (RCTs) of AE compared with controls (treatment as usual, attention placebo, active therapy) and head-to-head comparisons of different types of AE were included. Two authors independently extracted articles using predefined data fields, including study quality indicators.

Results: Twenty-eight RCTs comparing AE with controls and seven RCTs comparing different types of AE with a total of 2,494 patients were reviewed. Effects were summarised using standardised mean differences (95% confidence intervals) by random effect models. AE reduced pain (-0.31 (-0.46, -0.17); $P < 0.001$), fatigue (-0.22 (-0.38, -0.05); $P = 0.009$), depressed mood (-0.32 (-0.53, -0.12); $P = 0.002$) and limitations of health-related quality of life (HRQOL) (-0.40 (-0.60, -0.20); $P < 0.001$), and improved physical fitness (0.65 (0.38, 0.95); $P < 0.001$), post treatment. Pain was significantly reduced post treatment by land-based and water-based AE, exercises with slight to moderate intensity and frequency of two or three times per week. Positive effects on depressed mood, HRQOL and physical fitness could be maintained at follow-up. Continuing exercise was associated with positive outcomes at follow-up. Risks of bias analyses did not change the robustness of the results. Few studies reported a detailed exercise protocol, thus limiting subgroup analyses of different types of exercise.

Conclusions: An aerobic exercise programme for FMS patients should consist of land-based or water-based exercises with slight to moderate intensity two or three times per week for at least 4 weeks. The patient should be motivated to continue exercise after participating in an exercise programme.

Introduction

The key symptoms of fibromyalgia syndrome (FMS) are chronic widespread (both sides, above and below waist line, and axial skeletal) pain, fatigue, sleep disturbances and tenderness on palpation [1]. The estimated prevalence of FMS in western countries ranges from 2.2 to 6.6% [2]. Comorbidities with other functional somatic syndromes and mental disorders are common [3]. FMS is associated with high utilisation and costs of health services. Effective treatment options are therefore needed for medical and economic reasons [4].

Systematic reviews and evidence-based guidelines provide healthcare professionals and patients with a guide through the great variety of pharmacological and non-pharmacological treatment options in FMS. Three evidence-based guidelines available on the management gave different grades of recommendation for aerobic exercises (AE) (aerobic exercise with and without additional strength and flexibility training) in FMS. The

* Correspondence: whaeuser@klinikum-saarbruecken.de
[1] Department of Internal Medicine I, Klinikum Saarbrücken, Winterberg 1, D-66119 Saarbrücken, Germany
Full list of author information is available at the end of the article

American Pain Society [5] and the guidelines of the Association of the Scientific Medical Societies in Germany [6] gave the highest grade of recommendation for AE. The European League Against Rheumatism judged the published evidence for the efficacy of AE to be lacking [7]. Qualitative reviews on the efficacy of AE in FMS that searched the literature until December 2006 came to different conclusions on the short-term and long-term efficacy of AE in FMS [8-10].

More recently, Jones and Lipton reviewed over 70 FMS exercise studies and found similar results when protocols included yoga, tai chi and other movement-based therapies [11]. Two meta-analyses on exercise in FMS have been conducted. Busch and colleagues searched the literature until July 2005. Owing to significant clinical heterogeneity among the studies, only six studies with AE were meta-analysed. Moderate quality evidence was found that AE had positive effects on global well-being and physical function, but not on pain at post treatment [12]. The Ottawa Panel searched the literature until December 2006 and found most improvements for pain relief and increase of endurance at post treatment [13]. Outcomes at follow-up were not meta-analysed.

Not only the question of efficacy but also that of the dose and type of AE need to be clarified. The American Pain Society recommended encouraging patients to perform moderately intense AE (60 to 70% of age-adjusted predicted maximum heart rate (maxHR)) two or three times per week [5]. The evidence of this recommendation has not been tested by meta-analyses of head-to-head comparisons of different types and volumes of AE. Moreover, the question of whether continuing AE is required to maintain a symptom reduction had not been systematically addressed.

The aims of the present systematic review were to update the literature on AE in FMS and to assess whether AE has beneficial effects at post treatment and at follow-up on the key domains of FMS (pain, sleep, fatigue, depressed mood), compared with other therapies. In contrast to the Cochrane review [12], we intended to meta-analyse the outcomes of all randomised controlled trials (RCTs) available. Another aim was to asses which types, volumes and intensities of AE are effective by performing head-to-head comparisons of RCTs with different types and intensities of AE. The final aim was to assess whether ongoing exercise is necessary to maintain potential positive effects of AE.

Materials and methods

The present review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement [14] and the recommendations of the Cochrane Collaboration [15].

Protocol

Methods of analysis and inclusion criteria were specified in advance. We used the review protocol of our systematic review on multicomponent therapy in FMS [16].

Eligibility criteria

Types of studies

A RCT design comparing AE with a control group receiving no treatment, treatment as usual, attention control or any pharmacological or nonpharmacological therapy, or with head-to-head comparisons of different types or intensities of AE were included. Studies without randomisation were excluded.

Types of participants

Patients of any age diagnosed with FMS on recognised criteria were included.

Types of intervention

AE was assumed if the reported target heart rate of the training protocol was at least (on average) 40% of maxHR or if the training protocol included exercise involving at least one-sixth of the skeletal muscles (for example, walking, running, biking, aerobics, vibrations). At least 50% of the training session should consist of AE. In the case of mixed exercise, defined as a combination of AE with stretching and/or muscle strength [17], the length of AE should exceed the time with other types of exercise. Stretching during warm-up and cool-down periods was not defined as mixed exercise. No restrictions on frequency or duration of training were made.

We excluded studies or study arms in which AE was part of multicomponent therapy defined as a combination of AE with psychological therapy (structured education or relaxation therapy, cognitive-behavioural therapy) [16]. We excluded studies or study arms with balneotherapy (warm-water treatment without exercise).

Types of outcomes measures

Studies should assess at least one key domain of FMS (pain, sleep, fatigue, depressed mood and health-related quality of life (HRQOL)) (primary outcome measures). Secondary outcome measures were any measure of physical fitness.

Data sources and searches

The electronic bibliographic databases screened included the Cochrane Central Register of Controlled Trials (CENTRAL), EMBASE, MEDLINE, PsychInfo and SPORTDISCUS (through 31 March 2009). The search strategy for MEDLINE is detailed in Additional file 1. The search strategy was adapted for each database as necessary. No language restrictions were made. Only fully published papers were reviewed. In addition, reference sections of original studies, systematic reviews [8-10] and evidence-based guidelines on the management of FMS [4-6] were screened manually.

Study selection

The search was conducted by two authors (PK, JL). Two authors screened the titles and the abstracts of potentially eligible studies identified by the search strategy detailed above independently (PK, JL). The full-text articles were then examined independently by two authors to determine whether they met the selection criteria (MSc, JL). Discrepancies were rechecked and consensus was achieved by discussion. If needed, two other authors reviewed the data to reach a consensus (AB, WH).

Data collection process

Two authors independently extracted the data using standard extraction forms [16] (BM, MSc). Discrepancies were rechecked and consensus was achieved by discussion. If needed, a third author reviewed the data to reach a consensus (WH).

Based on our experiences of former systematic reviews in which none of the contacted authors provided these details on request, we did not ask for clarifications of study design in case of unclear randomisation, blinding or concealment of treatment allocation. We searched for further details of the study design in a Cochrane review [12].

When means or standard deviations (SDs) were missing, attempts were made to obtain these data through contacting 12 trial authors. Additional data were provided by four authors (see Tables 1 and 2). Where SDs were not available from the trial authors, they were calculated from *t* values, confidence intervals or standard errors when reported in articles [15]. If only the median was given, the median was used instead of the mean and a SD was substituted that was calculated as the mean of the SDs available for studies that used the same outcome scale.

Data items

The data for the study setting, participants, exclusion criteria, interventions, co-therapies, attendance rates, side effects reported and outcomes sought are presented in Tables 2 and 3.

When researchers reported more than one measure for an outcome, we used a predefined order of preference for analysis (details available on request).

If studies had two or more potential control groups, we used the following order to select for control group: treatment as usual, attention placebo, and active control to select the control group.

Risk of bias in individual studies

To ascertain the internal and external validity of the eligible RCTs, two pairs of reviewers (BM, WH; and MSc, Mst) working independently and with adequate reliability determined the adequacy of randomisation, concealment of allocation, blinding of outcome assessors and adequacy of data analysis (was intention-to-treat-analysis performed?) (internal validity). Furthermore we chose the item 'Were patients with mental disorders frequently associated with FMS (depressive and anxiety disorders) included in the studies?' as the marker of external validity.

Summary measures

Meta-analyses were conducted using RevMan Analyses software (RevMan 5.0.17) from the Cochrane collaboration [18]. Standardised mean differences (SMDs) were calculated by means and SDs or change scores for each intervention. The SMD used in Cochrane reviews is the effect size known as Hedge's (adjusted) *g* [15]. Examination of the combined results was performed by a random effects model (inverse variance method), because this model is more conservative than the fixed effects model and incorporates both within-study and between-study variance [19]. We used Cohen's categories to evaluate the magnitude of the effect size, calculated by the SMD: $g >$ 0.2 to 0.5, small effect size; $g >$ 0.5 to 0.8, medium effect size; $g >$ 0.8, large effect size [20].

Planned methods of analysis

Heterogeneity was tested using the I^2 statistic, with $I^2 >$ 50% indicating strong heterogeneity. τ^2 was used to determine how much heterogeneity was explained by subgroup differences [15].

Risk of bias across studies

Potential publication bias - that is, the association of publication probability with the statistical significance of study results - was investigated using visual assessment of the funnel plot (plots of effect estimates against its standard error) calculated by RevMan Analyses software. Publication bias may lead to asymmetrical funnel plots [15]. Moreover, we checked a potential small sample size bias by a sensitivity analysis of studies with very small (<25), small (25 to 50) and medium (>50) sample sizes.

Additional analyses

Subgroup analysis

The following subgroup analyses were pre-specified: types of AE (land-based, water-based and mixed; AE as monotherapy or combined with flexibility and/or strength), intensity of AE (very low intensity, <50% of maxHR; low intensity, 50 to 60% of maxHR; moderate intensity, 60 to 80% maxHR; intensity left up to patient), frequency of AE per week (1 time/week, 2 times/week, 3 times/week and >3 times/week), duration of the study (<7 weeks, 7 to 12 weeks, >12 weeks) and duration of total aerobic exercise (<1,000 minutes, 1,000 to 2,000 minutes, >2,000 minutes), and type of control group (attention placebo, treatment as usual, other active therapy). These subgroup analyses were also used to examine potential sources of clinical heterogeneity.

Table 1: Risk of bias (internal and external validity) of the randomised controlled trials' analysis

Author, year	Adequate randomisation	Adequate allocation concealment	Blinding of assessor	Intention-to-treat analysis	Inclusion of patients with mental disorders
Alentorn, 2008	0	0	+	-	+
Altan, 2004	0	0	0	-	-
Assis, 2006	+	+	-	+	-
Bircan, 2006	0	0	0	-	+
Buckelew, 2008	0	0	0	-	+
Da Costa, 2005	+	+	+	+	+
Ecvik, 2008	0	0	0	-	+
Etnier, 2009	0	0	0	-	+
Fontaine, 2007	0	0	0	-	+
Gowans, 2001	0	0	0	-	+
Gusi, 2006	0	0	0	-	-
Jentoft, 2001	0	0	-	+	+
Jones, 2008	+	+	+	+	-
King, 2002	+	0	+	+	+
Martin, 1996	0	0	+	+	+
McCain, 1988	0	0	+	-	+
Mengshoel, 1992	+	0	0	+	+
Meyer, 2000	0	0	0	-	+
Munguia, 2008	+	+	+	+	-
Nichols, 1994	0	0	0	-	+
Noregaard, 1997	0	0	0	+	+
Ramsay, 2000	0	0	+	+	0
Redondo, 2004	+	0	0	+	-
Richards, 2002	+	+	+	+	+
Rooks, 2007	+	+	+	+	0
Schachter, 2003	+	+	+	+	+
Sencan, 2004	0	0	0	+	0
Stephens, 2008	+	+	+	+	+
Tomas-Carus, 2008	0	+	+	+	0
Valim, 2003	0	0	+	+	+
Valkeinen, 2008	+	0	0	+	0
Van Santen, 2001	0	-	-	-	0
Van Santen, 2002	0	0	+	+	-
Vitorino, 2006	+	+	-	+	-
Wigers, 1996	0	0	+	+	0

+, yes; 0, unclear; -, no.

Table 2: Effect sizes of aerobic and mixed exercise on selected outcome variables

Outcome title	Number of study arms	Number of patients on aerobic exercise	Effect size[a]	Test for overall effect *P* value	Heterogeneity, *I2*; τ2 (%)
Post treatment					
01 Pain	29	567	-0.31 (-0.46, -0.17)	<0.001	26; 0.03
02 Fatigue	16	364	-0.22 (-0.38, -0.05)	0.009	9; 0.01
03 Sleep	9	184	0.01 (-0.19, 0.21)	0.92	0; 0
04 Depressed mood	19	456	-0.32 (-0.53, -0.12)	0.002	51; 0.10
05 HRQOL	25	526	-0.40 (-0.60, -0.20)	<0.001	63/0.15
06 Physical fitness	20	339	0.65 (0.38, 0.93)	<0.001	71/0.20
Latest follow-up					
01 Pain	9	187	-0.13 (-0.80, 0.54)	0.08	0/0
02 Fatigue	4	93	-0.23 (-0.62, 0.17)	0.26	42/0.07
03 Sleep	4	84	0.17 (-0.14, 0.47)	0.28	0/0
04 Depressed mood	8	151	-0.44 (-0.88, 0.01)	0.05	71/0.22
05 HRQOL	8	221	-0.27 (-0.48, -0.05)	0.02	14/0.01
06 Physical fitness	5	99	0.65 (0.35, 0.96)	<0.001	0/0

HRQOL: health-related quality of life. [a]Standardised mean difference (95% confidence interval).

Sensitivity analyses

The following sensitivity analyses were pre-specified: inadequate or unclear versus adequate sequence generation; inadequate or unclear allocation versus adequate concealment; intention-to-treat analysis, no versus yes; studies that provided medians of outcomes versus means of outcomes; and patients with mental disorders frequently associated with FMS excluded (yes or unclear). These sensitivity analyses were also used to examine potential sources of methodological heterogeneity.

Results

Study selection

The literature search produced 464 citations, of which 292 were double hits (study found in at least two data sources). By screening, 110 records were excluded: 23 evaluated AE, but not in FMS; 19 did not evaluate AE in

Table 3: Effect sizes of head-to-head comparisons of different types of aerobic exercise on selected outcome variables

Outcome title post treatment	Number of studies	Number of patients	Effect size[a]	Test for overall effect, *P* value	Heterogeneity, *I2*; τ2 (%)
Moderate intensity versus low intensity					
01 Pain	2	68	-0.08 (-1.41, 1.26)	0.91	78; 0.96
02 Depressed mood	2	68	-0.16 (-0.67, 0.13)	0.53	0; 0.
03 Physical fitness	2	68	0.25 (-0.26, 0.75)	0.34	0; 0
Land-based versus water based exercise					
01 Pain	9	187	-0.13 (-0.80, 0.54)	0.08	0/0
02 Depressed mood	8	151	-0.44 (-0.88, 0.01)	0.05	71/0.22

[a]Standardised mean difference (95% confidence interval).

FMS; 52 were review articles; and 18 were case reports or commentaries. Sixty of the full-text articles assessed for eligibility, and 25 full-text articles were excluded for the following reasons: two for publication of different outcomes of one trial in two publications [21,22]; six for lacking a control group [23-28]; three for lacking randomisation [29-31]; two because one could not conclude from the study protocol that the exercises performed met the predefined criteria of AE [32,33]; one because two different types of water-based exercise with similar intensity were compared [34]; one because the study did not assess a primary outcome measure [35]; and 10 because AE was combined with education or psychotherapy or pharmacotherapy [36-45]. Three RCTs comparing different intensities of AE [46-48], four RCTs comparing land-based with water-based exercise [49-52] and 28 RCTs with 29 study arms comparing AE with controls [53-80] were included in the qualitative and quantitative analyses (see Figure 1).

Study characteristics

Setting, referral and exclusion criteria (representativeness of study samples)

Fourteen studies each were conducted in North America, 13 studies in Europe and four studies each in South America (Brazil) and Asia (Turkey) (see Additional files 1, 2 and 3). Patients were recruited by register of hospitals, referral (general practitioner, rheumatologist, hospital departments), local self-help groups and newspaper advertisement. Thirty-two studies were conducted within the setting of a university, three within district hospitals. All studies were single-centre based. One study had two AE study arms.

Thirty-one studies excluded patients with internal diseases or with orthopaedic diseases precluding AE. Sixteen studies excluded patients with mental disorders including depression. Four studies excluded patients with unresolved litigation. No study reported comorbidities of the patients.

Participants

The median of the mean age of the participants was 45 years (13 to 59 years). One study included only children and adolescents. The median of the percentage of women was 100% (71 to 100%).

Interventions

AE was supervised by a trainer in 32 studies. AE included cycling, walking, aquatic jogging, games, dance and rhythmic or boxing movements. Aerobic intensity was reported in 27 studies as a target heart rate or percentage age-predicted maxHR determined by standard equations. Percentage maxHRs were usually progressive and ranged from 40 to 80% of the age-predicted maximum. The target heart rate of 21 studies was between low and moderate intensity (50 to 80%). Only one study prescribed a very low intensity (maxHR 30 to 50%), and three studies recommended that patients should exercise with a moderate intensity subjectively determined by the patient without measuring the heart rate. Three studies did not report the recommended intensity.

Sixteen studies reported the attrition rates, with a median of 67% (range 27 to 90%).

In 12 studies the controls received treatment as usual, and in 10 studies they received another active therapy (spa, hot packs, structured education, supervised relaxation, cognitive behavioural therapy, muscle strengthening, stretching). In six studies an attention control was used (nonstructured education, supervised recreational therapies, transcutaneous electrical neurostimulation or pharmacological placebo) (see Additional file 1).

Three studies compared different intensities of land-based AE, and four studies compared water-based AE with land-based AE (see Additional file 2).

A total 694/889 (78.1%) of the patients in the AE groups and 617/742 (83.1%) in the control groups completed therapy (z = -0.3, P = 0.7).

Fourteen studies performed follow-ups. The median of the latest follow-up was 26 (12 to 208) weeks. Five studies reported that the patients were motivated to continue exercise [51,56,70,71,75]. One study recommended no exercise until follow-up evaluation [61]. Two studies assessed the effects of continuing exercise on outcomes [25,80]. One study compared the outcomes of continuers of exercise versus noncontinuers at follow-up without mentioning whether continuing exercise had been recommended [80]. Two studies performed an uncontrolled follow-up [37,60].

Outcomes

There was a great variety of most outcomes measures (see Additional files 1, 2 and 3). Eleven studies reported on side effects. Five studies reported that no side effects occurred, and six studies reported an increase of symptoms leading to a drop out in some cases. Only six patients assigned to AE were designated to have an adverse event possibly related to exercise (metatarsal stress fracture, plantar fasciitis, ischialgia, transient knee pain).

Risk of bias within studies

Only two studies fulfilled all predefined criteria of internal and external validity (see Table 1).

Results of individual studies

The means, SDs, sample sizes and effect estimates of each study can be seen in the forest plots (see Additional files 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13).

Synthesis of results

Aerobic exercise patients versus controls

Data are reported as the SMD (95% confidence interval).

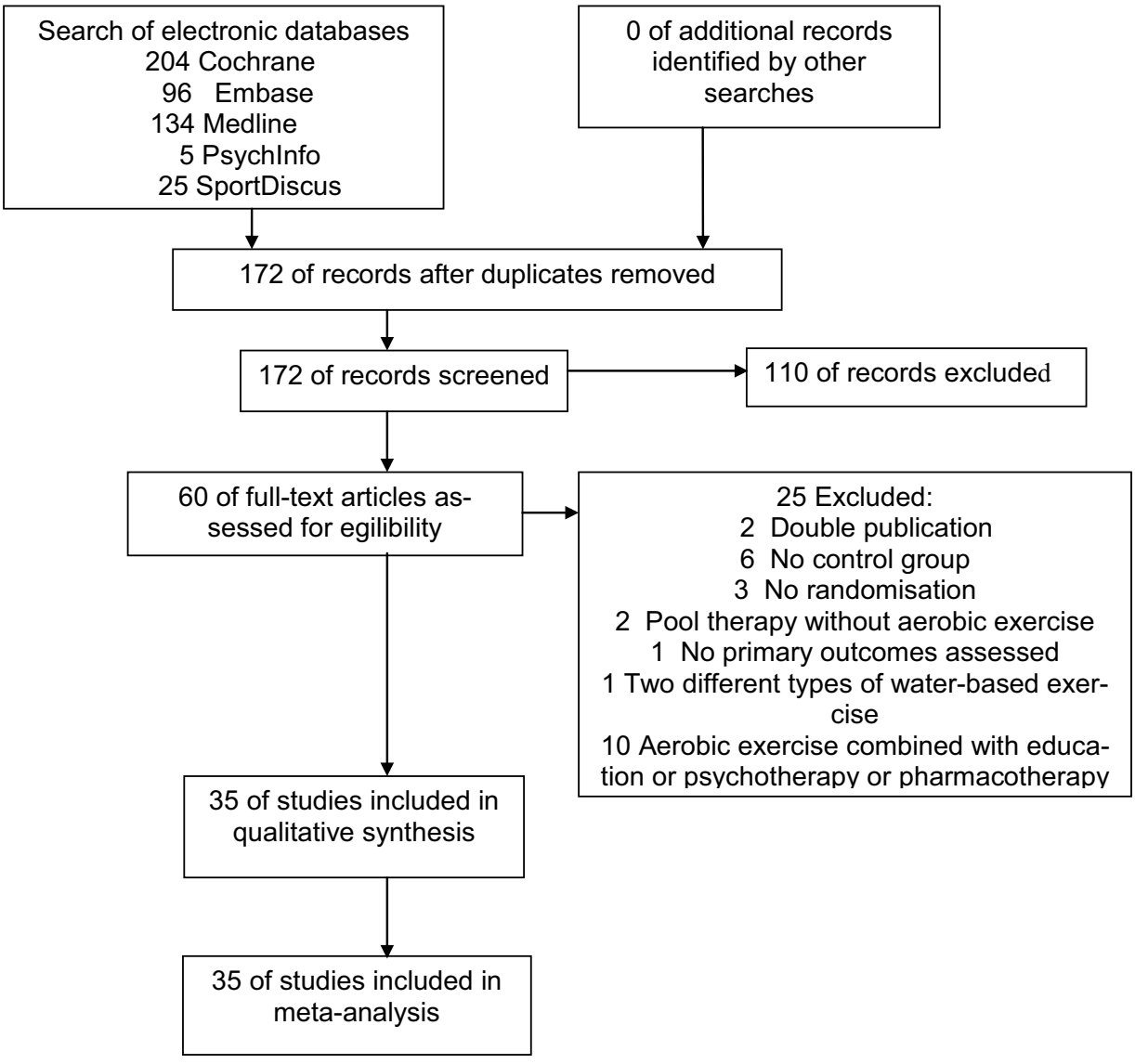


Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart. Schematic description of the results of the literature search.

At post treatment, AE reduced pain (-0.31 (-0.46, -0.17); $P < 0.001$), fatigue (-0.22 (-0.38, -0.05); $P = 0.006$), depressed mood (-0.32 (-0.53, -0.12); $P = 0.002$) and limitations of HRQOL (-0.40 (-0.60, -0.20); $P < 0.001$), and improved physical fitness (0.65 (0.38, 0.93); $P < 0.001$), compared with controls. The effect on sleep (0.01 (-0.19, 0.21); $P = 0.92$) was not significant. Based on Cohen's categories, the effects were small for pain, fatigue, depression and HRQOL, and were medium for physical fitness (see Table 4).

At latest follow-up, AE reduced depressed mood (-0.44 (-0.88, 0.01); $P = 0.05$) and limitations of HRQOL (-0.27 (-0.48, -0.05); $P = 0.01$), and improved physical fitness (0.65 (0.35, 0.96); $P < 0.001$), compared with controls. The effects were small for depressed mood and HRQOL, and were medium for physical fitness. The effects on pain (-0.13 (-0.80, 0.54); $P = 0.08$), fatigue (-0.23 (-0.62, 0.17); $P = 0.26$) and sleep (0.17 (-0.14, 0.47); $P = 0.26$) were not significant (see Table 4).

Land-based versus water-based aerobic exercise

There were no significant effects of water-based AE versus land-based AE on the outcomes pain and depressed mood at post treatment (see Table 2).

Moderate-intensity versus low-intensity aerobic exercise

There were no significant effects of moderate-intensity compared with low-intensity AE on the outcomes pain,

Table 4: Subgroup analysis for the effect size on pain at post treatment

Outcome title	Number of study arms	Number of patients on AE	Effect size[a]	Test for overall effect, *P* value	Heterogeneity, *I2*; τ2 (%)
Type of exercise					
Land-based	22	463	-0.29 (-0.46,-0.13)	0.0005	27; 0.03
Water-based	3	61	-0.67 (-1.04,-0.29)	0.0005	0; 0
Mixed	4	43	-0.03 (-0.45,0.39)	0.89	0; 0
Type of exercise					
AE only	12	273	-0.35 (-0.61,-0.09)	0.0008	48; 0.09
AE combined with other exercise	17	294	-0.28 (-0.45,-0.15)	0.001	0; 0
Duration of study					
<7 weeks	2	32	-1.16 (-1.86,-0.48)	0.001	36; 0.09
7 to 12 weeks	13	194	-0.24 (-0.50,-0.02)	0.03	16; 0.02
>12 weeks	12	338	-0.24 (-0.40,-0.08)	0.004	0; 0
Frequency of training/week					
1 time/week	2	37	-0.07 (-0.54,03.9)	0.48	Not applicable
2 times/week	5	127	-0.69 (-0.95,-0.27)	0.0004	35; 0.06
3 times/week	16	241	-0.35 (-0.62,-0.09)	0.009	48; 0.10
>3 times/ week	4	142	-0.13 (-0.38, 0.13)	0.33	2; 0
Total duration aerobic exercise[b]					
<1,000 minutes	10	175	-0.47 (-0.86,-0.08)	0.02	62; 0.19
1,000 to 2,000 minutes	9	175	-0.36 (-0.59,-0.13)	0.002	0; 0
>2,000 minutes	8	217	-0.15 (-0.34, 0.05)	0.15	0; 0
Intensity of AE[c]					
<50% maxHR	1	37	-0.09 (-0.54, 0.36)	Not applicable	Not applicable
Left up to patient	2	79	-0.42 (-0.77, -0.07)	0.02	0; 0
> 50% maxHR	21	367	-0.26 (-0.42,-0.11)	0.0007	0; 0
Type of control group					
Attention placebo	7	229	-0.27 (-0.62, 0.07)	0.12	67; 0.12
Therapy as usual	10	147	-0.47 (-0.71,-0.24)	<0.0001	0; 0
Active therapy	10	191	-0.27 (-0.49,-0.06)	0.01	0; 0

AE, aerobic exercise; maxHR, maximum of age-adjusted maximum heart rate. [a]Standardised mean difference (95% confidence interval). [b]If no precise duration of AE was given, 50% of the total exercise time was assumed for aerobic exercise. [c]Studies that did not report the intensity of training were excluded from analysis.

depressed mood and physical fitness at post treatment (see Table 3).

Effects of continuing exercise

One study found that continuers of exercise at follow-up reported less pain and depression than those who did not exercise [80]. One study found that exercising at follow-up was related to improvements in physical function and mood [37]. One study reported that pain returned close to the pretraining level during the subsequent de-training [61].

Risk of bias across studies

There was only substantial heterogeneity in the comparisons of depressed mood and HRQOL at post treatment and for depressed mood at latest follow-up (see Table 2). On visual inspection, the funnel plots of the outcomes post treatment were symmetrical and were thus not indicative for a publication bias (see Additional file 14). Studies with small sample sizes had no significant effect on pain at post treatment (see Table 5).

Additional analyses

Subgroup analysis

Subgroup analyses according to the types of AE, frequency, total time and intensity of AE and type of control groups did not change the significant effect of AE on pain at post treatment, except for a combination of water-based and land-based AE, total duration of AE >2,000 minutes, frequency of training 1 or >3 times/week and intensity <50% maxHR and attention placebo as control. Statistical heterogeneity of analysis for the effect size for pain was substantially increased in the case of a total duration of AE <1,000 minutes and attention placebo as control (see Table 4).

Sensitivity analysis

Sensitivity analyses according to potential risks of bias for the outcome pain at post treatment did not change the significant effect of AE on pain at post treatment, except for studies with sample size <25 and with only median of outcomes available. Statistical heterogeneity of analysis for the effect size for pain was substantially increased in the case of studies that included patients with mental disorders and with only the median of outcomes available (see Table 5).

Discussion

Summary of evidence

AE reduces pain, fatigue and depressed mood, and improves HRQOL and physical fitness, at post treatment. Positive effects of AE on depressed mood, HRQOL and physical fitness can be detected at latest follow-up. AE has no positive effect on sleep at post treatment, and on pain, fatigue and sleep at follow-up. Continuing exercise is necessary to maintain positive effects on pain.

The following statements are valid for pain reduction at post treatment. There is no evidence of a superiority of water-based over land-based exercise. AE with a slight to moderate intensity is effective. Low-intensity AE (<50% maxHR) is not effective. A frequency of AE of 2 to 3 times/week for at least 4 to 6 weeks is necessary for a reduction of symptoms. Combining AE with stretching or strengthening is no more effective than AE alone.

The evidence is applicable to the majority of patients in clinical practice except patients with internal and orthopaedic diseases that may prevent AE and male patients.

Limitations

Although every effort was made to obtain missing data (outcomes, study design) from the trial authors, it was not possible in every case to obtain these data; the included studies are therefore not represented fully in the meta-analyses. Only medians were available for three studies, but excluding these studies from analysis did not change the results.

The exercise protocol was insufficiently reported by some trials. The positive effects of the training can therefore possibly be attributed to other forms of exercise such as strength, stretching or relaxation, or in the case of pool-based exercise to the effects of warm water. Subgroup analyses did not, however, show a superiority of mixed exercise versus aerobic exercise nor a superiority of pool-based exercise versus land-based exercise.

The prescribed training intensity was either not assessed by heart rate telemetry or was not reported. No definitive conclusions on an effective intensity of AE are therefore possible.

The attendance rates during the study were inconsistently reported. If continuation of exercise until follow-up was recommended was inconsistently reported too. A subgroup analysis of studies with and without recommended exercise at follow-up was thus not possible.

Side effects were inconsistently reported. No definitive statement on the safety of AE in FMS is therefore possible.

The methodological quality of the studies varied. The positive effect on pain, however, was robust against potential methodological biases.

Given that formal blinding of participants and clinicians to the treatment arm is not possible in trials of exercise, we could have underestimated the extent to which clinicians' and participants' knowledge of group assignation influenced the true effect.

Males and adolescents were rarely included in the study populations. As no gender comparisons were reported, the evidence for the efficacy of AE in men and adolescents with FMS is limited.

Table 5: Sensitivity analysis for the effect size on pain at post treatment

Outcome title	Number of study arms	Number of patients on aerobic exercise	Effect size[a]	Test for overall effect, *P* value	Heterogeneity, *I2*; τ2 (%)
Adequate sequence generation					
Adequate	11	251	-0.20 (-0.38, -0.01)	0.04	0; 0
Unclear or nonadequate	18	348	-0.39 (-0.61, -0.18)	0.0004	39; 0.07
Allocation concealment					
Adequate	8	223	-0.24 (-0.47, -0.01)	0.04	24; 0.02
Unclear or nonadequate	19	344	-0.35 (-0.54, -0.16)	0.0002	28; 0.04
Blinding of assessor					
Yes	12	306	-0.20 (-0.36,-0.03)	0.02	0; 0
No or unclear	15	261	-0.41 (-0.66,-0.16)	0.001	38;0.08
ITT analysis					
Yes	13	315	-0.22 (-0.39, -0.06)	0.009	0; 0
No	14	252	-0.39 (-0.62, -0.16)	0.001	36; 0
Adequacy of outcomes for meta-analysis					
Yes (means)	24	517	-0.35 (-0.51,-0.19)	<0.0001	28;0.04
No (medians)	3	43	-0.05 {-0.43,-0.32)	0.78	0; 0
Sample size					
<25	3	30	-0.33 (-1.00,0.33)	0.33	34; 0.12
25 to 50	15	188	-0.41 (-0.70,-0.13)	0.005	46;0.11
>50	9	349	-0.23 (-0.39,-0.08)	0.004	0; 0
Patients with mental disorders included					
Yes	16	312	-0.43 (-0.78, -0.08)	0.02	73; 0.28
No or unclear	14	315	-0.40 (-0.61, -0.19)	0.0002	38; 0.05

ITT, intention to treat. [a]Standardised mean difference (95% confidence interval).

Agreements and disagreements with other systematic reviews

Our meta-analysis does not confirm the conclusion of a qualitative systematic review that the greatest effects occurred in exercise programmes that were of lower intensity than those of higher intensity [13]. Our data that AE reduces pain at post treatment are in line with the conclusion of the meta-analyses of the Ottawa Panel [13] and are in contrast to that of the Cochrane review [12]. Not only moderate-intensity AE as recommended by the American Pain Society [5], but also low-intensity AE seems to be effective in reducing pain.

Conclusions

Implications for clinical practice

The amount and intensity of initial AE should be adapted to the individual level of physical fitness. Patients should start at levels just below their capacity and gradually increase the duration and intensity until they are exercising with low to moderate intensity for 20 to 30 minutes 2 to 3 times/week [12]. It does not seem necessary to assess the heart rate during AE to find the optimum intensity. Patients should exercise with an intensity at which they are able to speak fluently with another person [17]. The choice of the type of AE should be left to the patient's

preferences and comorbidities and to the local offers of AE [11]. A training programme should last a minimum of 4 weeks. Patients should be educated that they may have some tolerable short-term increases in pain and fatigue but, if they exercise at an appropriate intensity, these symptoms should return to baseline levels within the first few weeks of exercise [12,17]. Patients should be motivated to continue exercise if they perceive a reduction of symptoms after the programme.

Because AE does not reduce sleeping disturbances, a combination of AE with medication effective for improving sleep - for example, tricyclic or dual antidepressants or pregabalin [81,82] - should be considered.

Implications for research

Four main questions need to be answered by future studies. By which methods (for example, education, booster sessions) can patients be motivated to continue exercise? Is aerobic and mixed exercise cost-effective [83]? Is the combination of AE with pharmacological therapy superior to AE or medication alone? Which sociodemographic and clinical variables predict a positive and negative treatment outcome?

Future studies on these topics should focus on larger sample sizes (multicentre studies including a sufficient number of men and adolescents and patients with mental and somatic comorbidities). Study quality could be improved by detailed reporting of demographic and clinical data of the study groups at baseline, exercise protocol and adherence to interventions (attendance rates, adherence to prescribed intensity assessed by heart rate telemetry), creation of a standardised protocol to report adverse events and use of standard outcome measures.

Additional material

Additional file 1 Search strategy for MEDLINE. The file contains the literature search strategy for the database MEDLINE.

Additional file 2 Main characteristics of studies with aerobic and mixed exercise in fibromyalgia syndrome. The file contains the main characteristics of studies with aerobic and mixed exercise in fibromyalgia syndrome including outcomes measures.

Additional file 3 Main characteristics of studies with head to head comparisons of different types of aerobic and mixed exercise in fibromyalgia syndrome. The file contains the main characteristics of studies with head-to-head comparisons of different types of aerobic and mixed exercise in fibromyalgia syndrome including outcome measures.

Additional file 4 Effect estimates (standardised mean differences) of aerobic exercise versus controls on pain at post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 5 Effect estimates (standardised mean differences) of aerobic exercise versus controls on fatigue and sleep at post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 6 Effect estimates (standardised mean differences) of aerobic exercise versus controls on depressed mood at post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 7 Effect estimates (standardised mean differences) of aerobic exercise versus controls on quality of life at post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 8 Effect estimates (standardised mean differences) of aerobic exercise versus controls on physical fitness at post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 9 Effect estimates (standardised mean differences) of aerobic exercise versus controls on pain and fatigue at latest follow-up. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 10 Effect estimates (standardised mean differences) of aerobic exercise versus controls on sleep and depressed mood at latest follow-up. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 11 Effect estimates (standardised mean differences) of aerobic exercise versus controls on quality of life and physical fitness at latest follow-up. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise groups is lower than in control group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 12 Effect estimates (standardised mean differences) of moderate versus low intensity on pain and depressed mood post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the exercise group with moderate intensity is lower than in exercise group with low intensity. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 13 Effect estimates (standardised mean differences) of water versus land-based aerobic exercise on pain and depressed mood post treatment. Forest plots show standardised mean differences (effect sizes) from the random effects model (inverse variance method). A negative effect indicates that the endpoint score of the outcome in the water-based exercise groups is lower than in the land-based exercise group in the study. The pooled (all studies together) effect size is weighted by the inverse variance of each study. IV, inverse variance (method); SD, standard deviation; Std. mean difference, standardised mean differences; random, random effects model; SD, standard deviation; total, number of patients; weight, relative weight (%) of the study in the calculation.

Additional file 14 Funnel plot of the comparisons of aerobic exercise versus controls on pain. Scatter plot of the intervention effect estimates (standardised mean differences (SMD)) from individual studies against their standard errors (SE) (on a reversed scale). Publication bias may lead to asymmetry in funnel plots on visual inspection.

Abbreviations

AE: aerobic and mixed exercise; FMS: fibromyalgia syndrome; HRQOL: health-related quality of life; maxHR: maximum heart rate; RCT: randomised controlled trial; SD: standard deviation; SMD: standardised mean difference.

Competing interests

WH received honoraria for educational lectures from Eli Lilly, Janssen-Cilag and Mundipharma, consulting fees from Eli-Lilly and Pfizer, and a congress travel grant from Eli-Lilly. MSc received consulting honoraria from Pfizer and MSD. None of these organisations financed this manuscript (including the article-processing charge). The other authors declare that they have no competing interests.

Authors' contributions

WH conceived the hypothesis of the manuscript, participated in the data collection, conducted the statistical analysis, wrote the first draft of the manuscript and had primary responsibility for the manuscript. PK, JL, BM, MSt, MSc and AB participated in the collection of the data and analysis of the studies (see Materials and methods). AB and MSc participated in the study design and the interpretation of the data. All authors critically reviewed, contributed and approved the final manuscript.

Authors' information

WH and MSc are vice-presidents of the German Interdisciplinary Association of Pain Therapy DIVS and were responsible for the development on the German interdisciplinary guideline on the classification, pathophysiology and management of FMS. AB is head of the Cochrane group on fibromyalgia.

Author Details

[1]Department of Internal Medicine I, Klinikum Saarbrücken, Winterberg 1, D-66119 Saarbrücken, Germany, [2]Department of Psychosomatic Medicine, Technische Universität München, Langestr. 3, D-81675 München, Germany, [3]Department of Internal Medicine V (Integrative Medicine), University of Duisburg-Essen, Kliniken Essen-Mitte, Am Deimelsberg 34a, D-45276 Essen, Germany, [4]Orthopaedic Clinic, University of Heidelberg, Schlierbacher Landstraße 200, D-69118 Heidelberg, Germany and [5]School of Physical Therapy, University of Saskatchewan, Saskatoon, 1121 College Drive, Saskatoon SK S7N OW3, Canada

Received: 4 November 2009 Revised: 8 February 2010
Accepted: 10 May 2010 Published: 10 May 2010

References

1. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, Tugwell P, Campbell SM, Abeles M, Clark P, Fam AG, Farber SJ, Fiechtner JJ, Franklin CM, Gatter RA, Hamaty D, Lessard J, Lichtbroun AS, Masi AT, McCain GA, Reynolds WJ, Romano TJ, Russel IJ, Sheon RP: **The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the multicenter criteria committee.** *Arthritis Rheum* 1990, **33:**160-172.
2. Branco JC, Bannwarth B, Failde I, Abello J Carbonell, Blotman F, Spaeth M, Saraiva F, Nacci F, Thomas E, Caubère JP, Le Lay K, Taieb C, Matucci-Cerinic M: **Prevalence of fibromyalgia: a survey in five European countries.** *Semin Arthritis Rheum* 2009 in press.
3. Henningsen P, Zimmermann T, Sattel H: **Medically unexplained physical symptoms, anxiety, and depression: a meta-analytic review.** *Psychosom Med* 2003, **65:**528-533.
4. Robinson RL, Jones ML: **In search of pharmacoeconomic evaluations for fibromyalgia treatments: a review.** *Expert Opin Pharmacother* 2006, **7:**1027-1039.
5. Burckhardt CS, Goldenberg D, Crofford L, Gerwin R, Gowans S, Jackson K, Kugel P, McCarberg W, Ruidn N, Schnaberg L, Taylor AG, Taylor J, Turk D: *Guideline for the Management of Fibromyalgia Syndrome. Pain in Adults and Children* APS Clinical Practice Guideline Series No. 4. Glenview, IL: American Pain Society; 2005.
6. Häuser W, Arnold B, Eich W, Felde E, Flügge C, Henningsen P, Herrmann M, Köllner V, Kühn E, Nutzinger D, Offenbächer M, Schiltenwolf M, Sommer C, Thieme K, Kopp I: **Management of fibromyalgia syndrome - an interdisciplinary evidence-based guideline.** *Ger Med Sci* 2008, **6:**Doc14.
7. Carville SF, Arendt-Nielsen S, Bliddal H, Blotman F, Branco JC, Buskila D, Da Silva JA, Danneskiold-Samsøe B, Dincer F, Henriksson C, Henriksson KG, Kosek E, Longley K, McCarthy GM, Perrot S, Puszczewicz M, Sarzi-Puttini P, Silman A, Späth M, Choy EH, European League Against Rheumatism: **EULAR evidence based recommendations for the management of fibromyalgia syndrome.** *Ann Rheum Dis* 2008, **67:**536-541.
8. van Koulil S, Effting M, Kraaimaat FW, van Lankveld W, van Helmond T, Cats H, van Riel PL, de Jong AJ, Haverman JF, Evers AW: **Cognitive-behavioural therapies and exercise programmes for patients with fibromyalgia: state of the art and future directions.** *Ann Rheum Dis* 2007, **66:**571-581.
9. Schiltenwolf M, Häuser W, Felde E, Flügge C, Häfner R, Settan M, Offenbächer M: **Physiotherapy, exercise and strength training and physical therapies in the treatment of fibromyalgia syndrome.** *Schmerz* 2008, **22:**303-312. [in German]
10. Gowans SE, de Hueck A: **Pool exercise for individuals with fibromyalgia.** *Curr Opin Rheumatol* 2007, **19:**168-173.
11. Jones KD, Lipton G: **Exercise interventions in fibromyalgia: clinical applications from the evidence.** *Rheum Dis Clin North Am* 2009, **35:**373-391.
12. Busch AJ, Barber KA, Overend TJ, Peloso PM, Schachter CL: **Exercise for treating fibromyalgia syndrome.** *Cochrane Database Syst Rev* 2007:CD003786.
13. Brosseau L, Wells GA, Tugwell P, Egan M, Wilson KG, Dubouloz CJ, Casimiro L, Robinson VA, McGowan J, Busch A, Poitras S, Moldofsky H, Harth M, Finestone HM, Nielson W, Haines-Wangda A, Russell-Doreleyers M, Lambert K, Marshall AD, Veilleux L, Ottawa Panel Members: **Ottawa Panel**

evidence-based clinical practice guidelines for aerobic fitness exercise in the management of fibromyalgia: part 1. *Phys Ther* 2008, **88:**857-871.
14. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group: **Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement.** *Ann Intern Med* 2009, **151:**264-269.
15. **Cochrane handbook for systematic reviews of intervention. Version 5.01** [http://www.cochrane.org/resources/handbook/handbook.pdf]
16. Häuser W, Bernardy K, Arnold B, Offenbächer M, Schiltenwolf M: **Efficacy of multicomponent treatment in fibromyalgia syndrome: a meta-analysis of randomized controlled clinical trials.** *Arthritis Rheum* 2009, **61:**216-224.
17. Rooks DS: **Talking to patients with fibromyalgia about physical activity and exercise.** *Curr Opin Rheumatol* 2008, **20:**208-212.
18. The Nordic Cochrane Centre: **Review Manager (RevMan) [computer program]. Version 5.0 for Windows.** Copenhagen: The Cochrane Collaboration; 2009.
19. Laird NM, Mosteller F: **Some statistical methods for combining experimental results.** *Int J Technol Assess Health Care* 1999, **6:**5-30.
20. Cohen J: *Statistical Power Analysis for the Behavoral Sciences* Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
21. Tomas-Carus P, Häkkinen A, Gusi N, Leal A, Häkkinen K, Ortega-Alonso A: **Aquatic training and detraining on fitness and quality of life in fibromyalgia.** *Med Sci Sports Exerc* 2007, **39:**1044-1050.
22. Munguía-Izquierdo D, Legaz-Arrese A: **Exercise in warm water decreases pain and improves cognitive function in middle-aged women with fibromyalgia.** *Clin Exp Rheumatol* 2007, **25:**823-830.
23. Dobkin PL, Abrahamowicz M, Fitzcharles MA, Dritsa M, da Costa D: **Maintenance of exercise in women with fibromyalgia.** *Arthritis Rheum* 2005, **53:**724-731.
24. Dobkin PL, Da Costa D, Abrahamowicz M, Dritsa M, Du Berger R, Fitzcharles MA, Lowensteyn I: **Adherence during an individualized home based 12-week exercise program in women with fibromyalgia.** *J Rheumatol* 2006, **33:**333-341.
25. Gowans SE, Dehueck A, Voss S, Silaj A, Abbey SE: **Six-month and one-year followup of 23 weeks of aerobic exercise for individuals with fibromyalgia.** *Arthritis Rheum* 2004, **51:**890-898.
26. Mannerkorpi K, Ahlmen M, Ekdahl C: **Six- and 24-month follow-up of pool exercise therapy and education for patients with fibromyalgia.** *Scand J Rheumatol* 2002, **31:**306-310.
27. Rooks DS, Silverman CB, Kantrowitz FG: **The effects of progressive strength training and aerobic exercise on muscle strength and cardiovascular fitness in women with fibromyalgia: a pilot study.** *Arthritis Rheum* 2002, **47:**22-28.
28. Dawson KA, Tiidus PM, Pierreynowski M, Crawford JP, Trotter J: **Evaluation of a community-based exercise program for diminishing symptoms of fibromyalgia.** *Physiother Canada* 2003, **55:**17-22.
29. Gandhi N, DePauw KP, Dolny DG, Freson T: **Effect of an exercise program on quality of life of women with fibromyalgia.** *Women Ther* 2002, **25:**91-103.
30. Meiworm L, Jakob E, Walker UA, Peter HH, Keul J: **Patients with fibromyalgia benefit from aerobic endurance exercise.** *Clin Rheumatol* 2000, **19:**253-257.
31. Salek AK, Khan MM, Ahmed SM, Rashid MI, Emran MA, Mamun MA: **Effect of aerobic exercise on patients with primary fibromyalgia syndrome.** *Mymensingh Med J* 2005, **14:**141-144.
32. Mannerkorpi K, Nyberg B, Ahlmen M, Ekdahl C: **Pool exercise combined with an education program for patients with fibromyalgia syndrome. A prospective, randomized study.** *J Rheumatol* 2000, **27:**2473-2481.
33. Ide MR, Belini MA, Caromano FA: **Effects of an aquatic versus non-aquatic respiratory exercise program on the respiratory muscle strength in healthy aged persons.** *Clinics (Sao Paulo)* 2005, **60:**151-158.
34. de Andrade SC, de Carvalho RF, Soares AS, de Abreu Freitas RP, de Medeiros Guerra LM, Vilar MJ: **Thalassotherapy for fibromyalgia: a randomized controlled trial comparing aquatic exercises in sea water and water pool.** *Rheumatol Int* 2008, **29:**147-152.
35. Verstappen FT, van Santen-Hoeuft HM, Bolwijn S: **Effect of a group activity program for fibromyalgia patients on physical fitness and well-being.** *J Musculoskeletal Pain* 1997, **5:**17-28.
36. Burckhardt CS, Mannerkorpi K, Hedenberg L, Bjelle A: **A randomized, controlled clinical trial of education and physical training for women with fibromyalgia.** *J Rheumatol* 1994, **21:**714-720.
37. Gowans SE, deHueck A, Voss S, Richardson M: **A randomized, controlled trial of exercise and education for individuals with fibromyalgia.** *Arthritis Care Res* 1999, **12:**120-128.
38. Hammond A, Freeman K: **Community patient education and exercise for people with fibromyalgia: a parallel group randomized controlled trial.** *Clin Rehabil* 2006, **20:**835-846.
39. Isomeri R, Mikkelson M, Latikka P, Kammonen K: **Effects of amitriptyline and cardiovascular fitness training on pain in patients with primary fibromyalgia.** *J Musculoskeletal Pain* 1993, **3/4:**253-256.
40. Gustafsson M, Ekholm J, Broman L: **Effects of a multiprofessional rehabilitation programme for patients with fibromyalgia syndrome.** *J Rehabil Med* 2002, **34:**119-127.
41. Cedraschi C, Desmeules J, Rapiti E, Baumgartner E, Cohen P, Finckh A, Allaz AF, Vischer TL: **Fibromyalgia: a randomised, controlled trial of a treatment programme based on self management.** *Ann Rheum Dis* 2004, **63:**290-296.
42. Keel PJ, Bodoky C, Gerhard U, Muller W: **Comparison of integrated group therapy and group relaxation training for fibromyalgia.** *Clin J Pain* 1998, **14:**232-238.
43. Lemstra M, Olszynski WP: **The effectiveness of multidisciplinary rehabilitation in the treatment of fibromyalgia: a randomized controlled trial.** *Clin J Pain* 2005, **21:**166-174.
44. Mason LW, Goolkasian P, McCain GA: **Evaluation of multimodal treatment program for fibromyalgia.** *J Behav Med* 1998, **21:**163-178.
45. Zijlstra TR, Laar MA van de, Bernelot HJ Moens, Taal E, Zakraoui L, Rasker JJ: **Spa treatment for primary fibromyalgia syndrome: a combination of thalassotherapy, exercise and patient education improves symptoms and quality of life.** *Rheumatology (Oxford)* 2005, **44:**539-546.
46. Meyer BB, Lemley KJ: **Utilizing exercise to affect the symptomology of fibromyalgia: a pilot study.** *Med Sci Sports Exerc* 2000, **32:**1691-1697.
47. Stephens S, Feldman BM, Bradley N, Schneiderman J, Wright V, Singh-Grewal D, Lefebvre A, Benseler SM, Cameron B, Laxer R, O'Brien C, Schneider R, Silverman E, Spiegel L, Stinson J, Tyrrell PN, Whitney K, Tse SM: **Feasibility and effectiveness of an aerobic exercise program in children with fibromyalgia: results of a randomized controlled pilot trial.** *Arthritis Rheum* 2008, **59:**1399-1406.
48. van Santen M, Bolwijn P, Landewé R, Verstappen F, Bakker C, Hidding A, Kemp D van Der, Houben H, Linden S van der: **High or low intensity aerobic fitness training in fibromyalgia: does it matter?** *J Rheumatol* 2002, **29:**582-587.
49. Assis MR, Silva LE, Alves AM, Pessanha AP, Valim V, Feldman D, Neto TL, Natour J: **A randomized controlled trial of deep water running: clinical effectiveness of aquatic exercise to treat fibromyalgia.** *Arthritis Rheum* 2006, **55:**57-65.
50. Evcik D, Yigit I, Pusak H, Kavuncu V: **Effectiveness of aquatic therapy in the treatment of fibromyalgia syndrome: a randomized controlled open study.** *Rheumatol Int* 2008, **28:**885-890.
51. Jentoft ES, Kvalvik AG, Mengshoel AM: **Effects of pool-based and land-based aerobic exercise on women with fibromyalgia/chronic widespread muscle pain.** *Arthritis Rheum* 2001, **45:**42-47.
52. de Melo Vitorino DF, de Carvalho BC, do Prado GF: **Hydrotherapy and conventional physiotherapy improbe sleep time and quality of life of fibromyalgia patients: randomised clinical trial.** *Sleep Med* 2006, **7:**293-296.
53. Alentorn-Geli E, Padilla J, Moras G, Lázaro Haro C, Fernández-Solà J: **Six weeks of whole-body vibration exercise improves pain and fatigue in women with fibromyalgia.** *J Altern Complement Med* 2008, **14:**975-981.
54. Altan L, Bingol U, Aykac M, Koc Z, Yurtkuran M: **Investigation of the effects of pool-based exercise on fibromyalgia syndrome.** *Rheumatol Int* 2004, **24:**272-277.
55. Bircan C, Karasel SA, Akgün B, El O, Alper S: **Effects of muscle strengthening versus aerobic exercise program in fibromyalgia.** *Rheumatol Int* 2008, **28:**527-532.
56. Buckelew SP, Conway R, Parker J, Deuser WE, Read J, Witty TE, Hewett JE, Minor M, Johnson JC, Van Male L, McIntosh MJ, Nigh M, Kay DR: **Biofeedback/relaxation training and exercise interventions for fibromyalgia: a prospective trial.** *Arthritis Care Res* 1998, **11:**196-209.
57. Da Costa D, Abrahamowicz M, Lowensteyn I, Bernatsky S, Dritsa M, Fitzcharles MA, Dobkin PL: **A randomized clinical trial of an individualized home-based exercise programe for women with fibromyalgia.** *Rheumatology (Oxford)* 2005, **44:**1422-1427.

58. Etnier JL, Karper WB, Gapin JI, Barella LA, Chang YK, Murphy KJ: **Exercise, fibromyalgia, and fibrofog: a pilot study.** *J Phys Act Health* 2009, **6:**239-246.
59. Fontaine KR, Haas S: **Effects of lifestyle physical activity on health status, pain, and function in adults with fibromyalgia syndrome.** *J Musculoskeletal Pain* 2007, **15:**3-9.
60. Gowans SE, deHueck A, Voss S, Silaj A, Abbey SE, Reynolds WJ: **Effect of a randomized, controlled trial of exercise on mood and physical function in individuals with fibromyalgia.** *Arthritis Rheum* 2001, **45:**519-529.
61. Gusi N, Tomas-Carus P, Hakkinen A, Hakkinen K, Ortega-Alonso A: **Exercise in waist-high warm water decreases pain and improves health-related quality of life and strength in the lower extremities in women with fibromyalgia.** *Arthritis Rheum* 2006, **55:**66-73.
62. Jones KD, Burckhardt CS, Deodhar AA, Deodhar AA, Perrin NA, Hanson GC, Bennett RM: **A six-month randomized controlled trial of exercise and pyridostigmine in the treatment of fibromyalgia.** *Arthritis Rheum* 2008, **58:**612-622.
63. King SJ, Wessel J, Bhambhani Y, Sholter D, Maksymowych W: **The effects of exercise and education, individually or combined, in women with fibromyalgia.** *J Rheumatol* 2002, **29:**2620-2627.
64. Martin L, Nutting A, MacIntosh BR, Edworthy SM, Butterwick D, Cook J: **An exercise program in the treatment of fibromyalgia.** *J Rheumatol* 1996, **23:**1050-1053.
65. McCain GA, Bell DA, Mai FM, Halliday PD: **A controlled study of the effects of a supervised cardiovascular fitness training program on the manifestations of primary fibromyalgia.** *Arthritis Rheum* 1988, **31:**1135-1141.
66. Mengshoel AM, Komnaes HB, Forre O: **The effects of 20 weeks of physical fitness training in female patients with fibromyalgia.** *Clin Exp Rheumatol* 1992, **10:**345-349.
67. Munguía-Izquierdo D, Legaz-Arrese A: **Assessment of the effects of aquatic therapy on global symptomatology in patients with fibromyalgia syndrome: a randomized controlled trial.** *Arch Phys Med Rehabil* 2008, **89:**2250-2257.
68. Nichols DS, Glenn TM: **Effects of aerobic exercise on pain perception, affect, and level of disability in individuals with fibromyalgia.** *Phys Ther* 1994, **74:**327-332.
69. Norregaard J, Lykkegaard JJ, Mehlsen J, Danneskiold-Samsooe B: **Exercise training in treatment of fibromyalgia.** *J Musculoskelet Pain* 1997, **5:**71-79.
70. Ramsay C, Moreland J, Ho M, Joyce S, Walker S, Pullar T: **An observer-blinded comparison of supervised and unsupervised aerobic exercise regimens in fibromyalgia.** *Rheumatology (Oxford)* 2000, **39:**501-515.
71. Redondo JR, Justo CM, Moraleda FV, Velayos YG, Puche JJ, Zubero JR, Hernández TG, Ortells LC, Pareja MA: **Long-term efficacy of therapy in patients with fibromyalgia: a physical exercise-based program and a cognitive-behavioral approach.** *Arthritis Rheum* 2004, **51:**184-192.
72. Richards SC, Scott DL: **Prescribed exercise in people with fibromyalgia: parallel group randomised controlled trial.** *BMJ* 2002, **325:**185.
73. Rooks DS, Gautam S, Romeling M, Cross ML, Stratigakis D, Evans B, Goldenberg DL, Iversen MD, Katz JN: **Group exercise, education, and combination self-management in women with fibromyalgia: a randomized trial.** *Arch Intern Med* 2007, **167:**2192-2200.
74. Schachter CL, Busch AJ, Peloso PM, Sheppard MS: **Effects of short versus long bouts of aerobic exercise in sedentary women with fibromyalgia: a randomized controlled trial.** *Phys Ther* 2003, **83:**340-358.
75. Şencan S, Ak S, Karan A, Müslümanoğlu L, Özcan E, Berker E: **A study to compare the therapeutic efficacy of aerobic exercise and paroxetine in fibromyalgia syndrome.** *J Back Musculoskeletal Rehabil* 2004, **17:**57-61.
76. Tomas-Carus P, Gusi N, Häkkinen A, Häkkinen K, Leal A, Ortega-Alonso A: **Eight months of physical training in warm water improves physical and mental health in women with fibromyalgia: a randomized controlled trial.** *J Rehabil Med* 2008, **40:**248-252.
77. Valim V, Oliveira L, Suda A, Silva L, de Assis M, Barros T Neto, Feldman D, Natour J: **Aerobic fitness effects in fibromyalgia.** *J Rheumatol* 2003, **30:**1060-1069.
78. Valkeinen H, Alén M, Häkkinen A, Hannonen P, Kukkonen-Harjula K, Häkkinen K: **Effects of concurrent strength and endurance training on physical fitness and symptoms in postmenopausal women with fibromyalgia: a randomized controlled trial.** *Arch Phys Med Rehabil* 2008, **89:**1660-1666.
79. van Santen M, Bolwijn P, Verstappen F, Bakker C, Hidding A, Houben H, Heijde D van der, Landewé R, Linden S van der: **A randomized clinical trial comparing fitness and biofeedback training versus basic treatment in patients with fibromyalgia.** *J Rheumatol* 2002, **29:**575-581.
80. Wigers SH, Stiles TC, Vogel PA: **Effects of aerobic exercise versus stress management treatment in fibromyalgia. A 4.5 year prospective study.** *Scand J Rheumatol* 1996, **25:**77-86.
81. Häuser W, Bernardy K, Üceyler N, Sommer C: **Treatment of fibromyalgia syndrome with antidepressants - a meta-analysis.** *JAMA* 2009, **301:**198-209.
82. Häuser W, Bernardy K, Üceyler N, Sommer C: **Treatment of fibromyalgia syndrome with gabapentin and pregabalin - a meta-analysis of randomized controlled trails.** *Pain* 2009, **145:**69-81.
83. Gusi N, Tomas-Carus P: **Cost-utility of an 8-month aquatic training for women with fibromyalgia: a randomized controlled trial.** *Arthritis Res Ther* 2008, **10:**R24.

doi: 10.1186/ar3002
Cite this article as: Häuser *et al.*, Efficacy of different types of aerobic exercise in fibromyalgia syndrome: a systematic review and meta-analysis of randomised controlled trials *Arthritis Research & Therapy* 2010, **12**:R79

J Rehabil Med 2009; 41: 347–352

ORIGINAL REPORT

PERSONAL CHARACTERISTICS INFLUENCING PATIENTS' ADHERENCE TO HOME EXERCISE DURING CHRONIC PAIN: A QUALITATIVE STUDY

Francesc Medina-Mirapeix, PT, PhD[1], Pilar Escolar-Reina, PT, PhD[1], Juan J. Gascón-Cánovas, MD, PhD[2], Joaquina Montilla-Herrador, PT, PhD[1] and Sean M. Collins, PT, ScD, CCS[3]

From the Departments of [1]Physical Therapy and [2]Public Health and Preventive Medicine, School of Medicine, University of Murcia, Murcia, Spain and [3]Department of Physical Therapy, University of Massachusetts, Lowell, USA

Objective: **To identify the beliefs and perceptions of patients with chronic neck and low back pain that influence adherence to home exercise during exacerbation and/or remission of pain.**

Design: **Qualitative study using a focus group technique.**

Subjects: **Thirty-four patients (23 women, age range 26–70 years) with chronic neck or low back pain who had participated in a home exercise programme.**

Methods: **Seven focus groups were formed. Participants were sampled purposefully from all patients with chronic neck or low back pain who attended for physiotherapy at 4 primary healthcare centres. Patients were interviewed about how they perceived their adherence to a home exercise programme during chronic pain. Data were analysed using a phenomenographic method.**

Results: **Several themes about patients' beliefs and perceptions were identified as factors related to adherence. These factors change when pain or disabilities appear, decrease or disappear for an extended period. Beliefs about illness and treatment are more likely when pain is present and when pain disappears for an extended period. However, patients consider perceptions about barriers, social support and physical environment when pain decreases.**

Conclusion: **These findings may represent an important potential for improving the adherence of patients with chronic pain to home exercise programmes.**

Key words: chronic pain, adherence, physical therapy, exercise, qualitative research.

J Rehabil Med 2009; 41: 347–352

Correspondence address: Francesc Medina-Mirapeix, Departamento de Fisioterapia. Facultad de Medicina, Universidad de Murcia, Campus de Espinardo, ES-30100, Murcia, Spain. E-mail: mirapeix@um.es

Submitted February 15, 2008; accepted November 20, 2008

INTRODUCTION

Exercise has been documented as an effective intervention for treatment of back and neck pain (1–4). Usually, exercises are taught and prescribed for home (5). However, research suggests that inadequate adherence to home exercise during the intervention period might attenuate the effectiveness of intervention (6, 7). It has also been suggested that recurrent cases of low back pain might be avoided if patients adhered to home exercise programmes after intervention (8, 9).

Several studies report that a lack of adherence to exercise is often a serious problem for patients with chronic pain. Estimates of what proportion of patients does not perform their exercises according to prescription vary depending on differences in the definition of adherence and measurement, but it is approximately ≥ 50% (6, 7, 10–12).

Research suggests that patients' personal characteristics influence adherence to home exercise programmes (13). Empirical studies have related patient's beliefs about seriousness, prognosis of illness and treatment efficacy with adherence (10, 14). Perceiving barriers to carrying out home exercise programmes has also been related to adherence (10, 12, 15). However, there are contradictory results from various studies concerning the association between adherence and perceptions such as pain or disability (10, 12, 16, 17).

It has been suggested that the relevance of pain in influencing adherence depends on the interaction between pain and patient's beliefs or other perceptions (17). Furthermore, several authors have suggested that patients might use different beliefs and perceptions to guide their adherence during periods of exacerbation or remission (10, 13, 17). However, this last issue has not been demonstrated for patients with neck or low back pain. Qualitative studies have not focused on this point of view in patients with neck or low back pain (15, 18) and quantitative studies that have investigated patients' perceptions, beliefs and adherence have not focused on its relative significance in exacerbations and remission of pain during the course of chronic pain. Therefore, although previous studies recognize the importance of patient's beliefs and perceptions on adherence to a home exercise programme, further work is needed to understand its importance during exacerbation and remission of pain.

The aim of this study was to explore patients' perceptions with the purpose of identifying those beliefs and perceptions that patients perceive to influence their adherence to a home exercise programme during exacerbation and remission of pain in the course of chronic pain. Physical therapists would benefit from a better understanding of such perceptions and their potential influence on adherence to their interventions as they attempt to maximize patient adherence.

METHODS

Qualitative methods provide a set of strategies for conducting a rigorous research study with the above aim (19). In order to describe completely the experience of adherence during exacerbation and remission of pain, a phenomenological study was undertaken. Phenomenology as a research approach aims to describe the experience of the everyday world as it appears, varied and complex (20). A focus group technique was used to obtain detailed data from patients' with experience of participating in home exercise programmes.

Participants

Four typical public primary healthcare centres in the region of Murcia, Spain, were selected. Murcia has a population of over 1 million and has a well-developed healthcare system that is mainly publicly operated. We selected these centres because patients with mechanical neck or low back pain referred to physical therapy intervention participate in both clinic visits and a home exercise programme during the period of intervention and afterwards.

Inclusion criteria for the study were: all patients with mechanical and chronic neck or low back pain who received and finished physical therapy treatment in the last 3 months. Exclusion criteria were: patients with mechanical neck and low back pain due to trauma, or patients with a physical or mental disability that precluded participation in focus groups (i.e. those who were deaf, blind, or had learning disability). Following research ethics committee approval, we identified patients from clinic records. A total of 94 patients were eligible for the study and a mixed purposive sampling strategy was used to select participants (21). Sampling was therefore dependent on the saturation of information.

Procedures

Recruitment. Stratification according to homogeneous and heterogeneous criteria was used to set up groups: homogeneous groups by centre (common experiences) and gender (to avoid diffidence in discussing health issues in the presence of the opposite gender), and heterogeneous groups by age and clinical condition (neck/back pain) to add variability of experiences to stimulate discussions. An invitation to join the study was sent by post to eligible patients and followed up with a telephone call. As patients declined to participate, we invited new patients to obtain a minimum group size of 4.

*Data collection. T*wo people, a moderator and an assistant, conducted all discussions in the public and neutral location of the city hall (i.e. not in the health centre). They used a topic guide initially derived from a literature review and later agreed upon by the research team. The topic guide was then reorganized after a preliminary analysis following the first focus group (Table I). During the interview dialogue the researcher posed questions such as: "What do you mean?", "Can you explain it more?", "How do you feel?", "What did you think?", "Please give an example". Audiotape was used for data collection during discussions. Videotape and field notes were used to record non-verbal language and incomplete or sarcastic expressions. Patients gave permission and were assured of confidentiality before the start of each session, prior to using these means of recording.

Table I. *Focus group interviewing guide*

Why did you go to the physical therapist?
How did you feel about having neck or low back pain before physiotherapy treatment?
What have you been told about chronic pain and its treatment?
Did you find easy your adherence to physiotherapist's instructions at the beginning of treatment? After your treatment, was it easier?
What kind of problems do you encounter for adherence when pain is not present?
Is there anything else you would like to say about your home programme or your pain?

Analysis

Interview transcripts were analysed in 5 steps: (*i*) overall impression of categories; (*ii*) independent generation of an initial code to label phrases; (*iii*) revision of categories and coding scheme as we accumulated data; (*iv*) elaboration and application of a final code scheme to the final data-set; and (*v*) exploration of the categories' relationships (22).

In the first step, transcript and observational notes were read to gain a sense of entirety, to identify significant phrases and to obtain tentative ideas about categories and relationships. It was agreed to define 2 kinds of categories to code information (23): (*i*) substantive categories, which help understand the experiences of patients; and (*ii*) themes or organizational categories, which gather substantive categories in logical areas according to study objectives. The agreement was reached using concepts that participants used or theoretical terms employed in the literature for substantive categories. The differences in the initial coding schemes generated independently by each of 4 researchers were resolved by discussion. This step was iterative, allowing emerging categories as the groups progressed. When saturation was being reached, it was implied when no new major themes arose by the end of the seventh focus group. The defined categories were presented to the physical therapists who treated the participants, as an external audit of the initial results before applying codes to the final data-set of phrases. Subsequently, a final coding scheme was elaborated by 2 researchers (PER and FMM) and confirmed for consistency through blind review of 2 transcripts. Disagreements between the 2 researchers were resolved by discussion. Codes were then applied to the final data-set, and category relationships within and among patients were explored.

Describing. Using a phenomenological method, a synthesis of the transformed meaning units was described, thus explaining the inner core of the phenomenon. Finally, taking into account all results, the essence of the investigated phenomenon was described (24).

RESULTS

There were 34 participants in this study (22 of these had chronic neck pain, and 23 were women). Their mean age was 48 years and age range 25–70 years. All participants were included in home exercise programmes by physiotherapists. Since their inclusion they had experienced periods of exacerbation and remission of pain. Most patients expressed beliefs and perceptions in some form to report problems with adherence to home exercise programmes.

The patients' experience was expressed in 5 themes: beliefs about illness and adherence, and perceptions in relation to barriers, support social and physical environment. Patients balanced these beliefs and perceptions to decide adherence to their home exercise programme. We classified them according to the emergent taxonomy shown in Table II, which identified variation in categories of beliefs and perceptions that concern patients under our 3 pre-established conditions: perception of presence, decrease or absence of pain or disabilities. From these themes and classification, an essential structure emerged. Identifier, for example Interviewed Person (IP), and demographic characteristics are given for the quotes below.

When pain or disabilities appear

When patients perceive pain or disabilities associated with pain, they report deciding whether they should adhere to home exercise programme recommendations. At the moment of

Table II. *Beliefs and perceptions associated with adherence to home exercise*

When pain or disabilities appear
Beliefs about illness
Prognosis expectations
Beliefs about adherence
Outcome expectations with exercises
When pain or disabilities decrease
Perceived barriers
Lack of time for exercises
Tiredness
Forgetting
Adverse effects of exercises
Comorbidities
Perceptions of support social
Incentives from family
Interactions with people exercising
Perceptions of physical environment
Entertainment
Recreational centres
Beliefs about adherence
Self-efficacy
When pain disappears for long time
Beliefs about illness
Vulnerability to relapse
Beliefs about adherence
Distance between adherence and its benefits or costs

perception, patients report doing exercise regardless of other considerations.

IP 3: *"When the pain bothers me or my usual activities, I remember the exercises, and leave what I am doing, and do the exercises that the therapist gave me" [Male, 53 years].*

Some patients report that their beliefs can interfere in the decision that occurs between perception of pain and adherence to exercise recommendations. Patients report that these beliefs are related to illness and treatment.

1. Beliefs about illness. Prognosis expectations are the beliefs that patients regard in this phase. Those patients who believe their problem is chronic and immutable tend to have a resigned attitude toward their pain, and consequently decide not to adhere to exercise recommendations. However, patients with optimistic *prognosis expectations* do not associate this optimism to adherence.

IP 16: *"I have my problem since so many years and nobody could help me. Because that often I don't do advice of the brochure that physical therapist gave me" [Female, 63 years].*

2. Beliefs about adherence. Patients also assess the credibility of the treatment offered. If patients doubt the effectiveness of the recommended advice, or if its rationale is not clear, they are less likely to adhere. Conversely, when patients believe that treatment is effective they report having high *outcome expectations* and consequently adhere to recommendations.

IP 10: *"Exercising was for the pain, I saw myself with disabilities and I hoped to get better with this treatment" [Male, 55 years].*

When pain or disabilities decrease

When the pain or disabilities associated with pain decrease, patients report perceiving that the home exercise programme requires some degree of alteration to their lifestyle. They report deciding about whether to adhere to recommendations once again. Patients initially prioritize to complete daily routine activities and discontinue exercises. Additionally, it is a positive reinforcement for the patients that symptoms take time to reappear after this decision.

IP 20: *"When I feel better, I forget the exercises and do other things; besides that, I don't have pain again" [Male, 35 years].*

In spite of the initial prioritization, patients try to maintain some degree of exercise. However, patients report that several perceptions – related to barriers, social support and physical environment – and beliefs about ability to adhere, interfere with their intention and then they do not give priority to their home exercise programme.

1. Perception of barriers. Perceived barriers are associated with low or no adherence. Common barriers usually include lack of time to fit exercises into a daily routine, tiredness, forgetting to exercise, adverse effects of exercises and symptoms associated with comorbidities.

IP 5: *"After work I arrive home at 9:00 o'clock at night, have dinner, sit down and put my feet up to watch TV" [Male, 45 years].*

IP 22: *"I had to stop using the bicycle because my knee was swelling. She also recommended that I walk, but I cannot do that either" [Male, 65 years].*

IP 32: *"I have another problem. Then, the days I feel good I can do exercises and the day I don't feel good I can't" [Female, 46 years].*

Patients report that a lack of time to fit exercises into their daily routine leads to barriers such as forgetting to exercise or tiredness.

IP 13: *"Being in the house I usually forget to do exercises because I am doing other things I am very busy and when I finish I want to sit or lay down because I am tired" [Female, 44 years].*

2. Perceptions of social support. Patients perceive that social support from family by means of incentives and reminders is helpful to adherence at times. Nevertheless, they recognize that this kind of support has less influence on adherence than social support from interactions with people exercising.

IP 8: "In the clinic I had to comply, after, in my house, nobody was watching me or telling me what to do, sometimes my wife told me to do exercises and then I did them, but generally I did not" [Male, 61 years].

IP 32: *"Exercise is different in the clinic than in my house, because in the clinic I was in front of other people and at home I am alone" [Female, 46 years].*

3. Perceptions of physical environment. Patients report having effective resources from the physical environment to overcome perceived barriers. These resources include using entertainment, such as television at home, and attendance at recreational cen-

tres. Some patients even feel that attending recreational centres is fundamental for adherence in post-treatment periods.

IP 28: *"I exercised every day when I woke up in the morning. I turned on the TV and I did the exercises while I watched" [Female, 52 years].*

IP 14: *"If I don't go to a gym or a recreation centre then I don't exercise, and if I do it's boring, unless I turn on the TV, a record or have a partner" [Female, 45 years].*

4. Beliefs about adherence. Patients' self-efficacy to overcome the common barriers to do exercises is a belief that has a strong influence on adherence. Low self-efficacy is associated with low adherence, and high self-efficacy is associated with high adherence.

IP 14: *"I can't do the exercises. I know it depends on my will-power to have a routine, but when I wake up I go directly to do the things I have to do, and I don't stop to exercise… A woman who tells herself to take care of herself or that she needs to exercise, she finds the time" [Female, 45 years].*

When pain disappears for long time

When pain or disabilities are absent, patients tend to make decisions regarding adherence, and this usually results in low or no adherence. Patient's beliefs about illness and benefits/costs of adherence influence their decision.

IP 11: *"I exercised in my house because I could move my arm better. I did them for a long time until I saw my arm didn't have pain and my hand was no longer asleep. Since then I have not done the exercises" [Male, 49 years].*

1. Beliefs about illness. Patient's beliefs about vulnerability to relapse influence this decision in relation to adherence. Nevertheless, patients report feeling no vulnerability to relapses as a consequence of not undertaking the prescribed home exercise programme. Only a few believe they might have a relapse.

IP 3: *"I exercise because I am afraid that I will have the pain again [Male, 53 years]".*

2. Beliefs about adherence. Even when beliefs about vulnerability to relapse are present, they are not strong enough to promote adherence. Thus, when patients believe that continuing exercises might prevent relapses, they face a conflict between knowing that they should perform (i.e. adherence to exercises and other advice) and at the same time feeling it is difficult to adhere. Most patients attenuate or stop exercising because relapse might be a long time away and they prefer exercising only if pain reappears. Only a few prefer initially to continue exercising. This decision is highly influenced by fear of relapse.

IP 30: *"After a time being good I stopped exercises. I am not doing well but when I feel pain again I will probably restart the exercises" [Female, 37 years].*

The essential structure

The essential meaning of patients' experiences was desire to live without pain and without exercise programmes that alter their lifestyle. If either of these factors disturbs their lifestyle patients decide about adherence to exercise programmes.

Conditions for adherence were different in subjects under conditions of exacerbation and remission of pain. During exacerbation of pain, conditions for adherence were to have high expectations about the prognosis of illness and outcomes of exercises. When pain decreased, essential conditions were related self-efficacy to overcome perceived barriers and to having social and environment support.

Lack of these conditions gave way to feelings of worse pain management and difficulty in accepting adherence to home programme. It also led, especially when pain had disappeared for a long time, to feelings of guilt about subsequent relapses. However, patients who had these feelings were not discouraged and trusted themselves or their capacity to carry out exercise programmes.

DISCUSSION

We examined the beliefs and perceptions of patients that influence their adherence to home exercise programmes during periods of exacerbation and remission of pain during chronic pain. The study provides evidence on several issues. First, patients relate adherence to perceptions of pain itself or disabilities associated with pain. Secondly, these perceptions interact with other patients' perceptions or beliefs to decide adherence to a home exercise programme. Thirdly, these perceptions and beliefs change over periods of pain and disability exacerbation and remission, and between patients.

Regarding the first issue, our participants associated positively perceptions of pain or disabilities and adherence, but for a limited time. Previous studies have reported contradictory relevance of perception of pain itself or disabilities associated with pain (10, 12, 16, 17). Our finding may explain apparent discrepancies between studies with back pain patients, resulting from variance created by the measurement of adherence at different points of time across the spectrum of chronic pain or disability. For example, studies that measured both disability and adherence at the same time found significant relationships (10), while studies that measured initial disability and follow-up adherence found no relationships (12).

Perceptions of pain or disability could be relevant because they can contribute to a patient's belief of a more severe condition (14) or vulnerability to further problems as a consequence of not undertaking the home exercise programme (17). Both beliefs, severity and vulnerability, are related to adherence to physical therapy activities in empirical studies with a variety of musculoskeletal conditions (10, 25, 26). However, these beliefs were not explicitly identified as in our taxonomy, during the period of pain.

Our taxonomy included patients' beliefs and perceptions that interact with perception of or not of pain in different periods of chronic pain. The taxonomy's distinction in periods of pain exacerbation, remission and disappearance suggests that there is a dynamic influence between pain perceptions and other perceptions or beliefs. The dynamic influence of determinants of adherence according to another determinant, such as pain, is a central component of social cognitive theory (27).

This study suggests that only beliefs about illness prognosis or outcome expectations are able to interact negatively with

such perceptions when pain or disability appears. Thus, when patients believed their complaints would continue or their exercises would not help them, lower adherence was reported. These finding regarding illness prognosis support empirical studies in physiotherapy and medical research (10, 28). On the other hand, the relevance of outcome expectations is mirrored in another study that also found that sport's injury patients with lower outcomes expectations were less adherent (25). This is reinforced by another study that related high levels of adherence to beliefs about the effectiveness of rehabilitation (26).

Perceived barriers to exercise, such as comorbidity, adverse effects, or lack of time to fit recommendations into a daily routine, have been strongly associated with adherence (10, 29, 30). Lack of time is a very consistent barrier identified between studies (13, 15). For some subjects, reporting lack of time may be a more socially acceptable excuse (10, 15) and may reflect a lack of interest in their commitment to compliance or could be the reflection of poor behaviour skills, such as time management (31). Thus, lack of time may not be a true determinant of adherence, but a perceived determinant. Furthermore, our participants also reported that, at times, the perception of a lack of time appears to be related to other barriers such as tiredness and forgetting. Traditionally, forgetting has been considered a determinant of non-intentional adherence (17). However, our finding suggests that sometimes this non-intentional adherence could be due to a previous and implicit prioritization between the exercises and other activities that leads to a lack of evocation of the reminder for exercising.

While perceived barriers are negatively associated with adherence when pain decreases, perceptions of social support and physical environment and belief of self-efficacy are positively associated. Self-efficacy has been related to adherence to home exercise programmes (17). Social support from family and from social interactions has been related to adherence to clinic programmes of physiotherapy (32), but not to home programmes. In addition, the type of social support might be more relevant for home exercise programme adherence and has not been studied. Patients in this study perceived that social interactions impact on adherence to a greater extent than does social support from their family.

Prevention of relapse is not something our patients wish to avoid when pain disappears for an extended time. Besides, prevention as the desired outcome is not strong enough to promote adherence because the relapse may be a long time away. According to social cognitive theory (27), it is likely that a patient's balance between distal desired outcomes and proximal costs influences adherence behaviour. Balance between costs and benefits of treatment has been identified in qualitative studies of other conditions (33, 34) and in psychological models such as the theory of planned behaviour (35).

Recommendations for practice and research

Most factors identified in this study have clear implications for patient management in physical therapy as well as other instances in healthcare providing self-management therapies. The predominant emergent view is that large improvements could be made in designing therapeutic encounters in order to maximize adherence. First, it is a problem that patients often do not communicate their beliefs about treatment, particularly regarding adherence to home exercise programmes when pain or disabilities decrease or disappear. Patients probably lack basic background knowledge about why it is important to follow the exercises even without pain. Thus, it is not odd that patients use their symptoms and disabilities to decide whether they should adhere to the home exercise programme. This knowledge and belief can be addressed by the therapists in the ordinary clinical situation in order to improve adherence. In this respect, this study provides potential support for enhancing the impact of educational interventions by targeting them to address factors that emerge in each period of chronic pain. Therapists should first establish patient's prognostic expectations and their perceived credibility of treatment, and only later reinforce positive factors and offer balance between perceived barriers or other problems with knowledge and beliefs of benefits.

The results of this study also have potential implications for patterns of delivery of physiotherapy. Patients usually stop adhering to home exercise programmes at the end or after a period of treatment, when pain decreases or disappears. For these patients, improving adherence might be an unrealistic aim if there is no physiotherapy follow-up intervention, such as programme adjustment or reinforcement of schedules.

This study focused on home exercise programmes and did not address other common home interventions, such as activities for self-management of pain (e.g. heat, rest) and self-care of back or neck (e.g. rest position, posture) (36, 37). Perhaps patients perceive that different beliefs and perceptions influence their adherence. Because this study was limited to home exercise, future research should explore other home activities.

Focus group studies have some potential disadvantages. They involve relatively small numbers of people; therefore, findings may not be representative of the general population in terms of opinions voiced. However, this qualitative study was designed to highlight the phenomenon being studied, and not to measure variables. Future research should provide more comprehensive and sensitive measurement of factors related to non-adherence during different periods of chronic pain.

In conclusion, this study has provided a deeper understanding of patients' beliefs and perceptions and their relationship with adherence to home exercise programmes during periods of exacerbation and remission of chronic pain. Knowledge of patients' priorities regarding the most important beliefs and perceptions that have high potential for adherence to home exercise may be helpful in improving the quality of care of patients with neck or low back pain. Adherence is usually a reasoned response in relation to a person's beliefs and perceptions. Managing adherence successfully can be a difficult task that cannot be accomplished simply by informing or instructing patients about home exercise. Overcoming negative perceptions and beliefs will require comprehension that the significance of each specific determinant of adherence must be considered in other determinants and techniques of continuing education, such as programme adjustment or reinforcement schedules.

ACKNOWLEDGEMENT

This study was supported by a research grant from the Ministry of Health and Consumer Affairs (PI030317) Spain.

REFERENCES

1. Bigos SJ, Bowyer R, Braen RG, Brown K, Deyo R, Haldeman S, et al., editors. Acute low back problems in adults. Clinical practice guidelines. Rockville, MD: United States Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1994.
2. Sarig-Bahat H. Evidence for exercise therapy for mechanical neck disorders. Man Ther 2003; 8: 10–12.
3. Kool J, de Bie R, Oesch P, Knüsel O, van den Brandt P, Bachmann S. Exercise reduces sick leave in patients with non-acute non-specific low back pain: a meta-analysis. J Rehabil Med 2004; 36: 49–62.
4. Medina-Mirapeix F, Meseguer-Henarejos AB, Montilla-Herrador J, Escolar-Reina P, Collins SM, Saturno-Hernández PJ. Reliability and validity of patient reports for physical therapy quality assessment: an empirical analysis regarding the use of exercises for neck pain in Spain. J Rehabil Med 2006; 38: 354–359.
5. Richardson C, Jull G, Hodges P, Hides J, editors. Therapeutic exercise for spinal segmental stabilization in low back pain. London: Churchill Livingstone; 1999.
6. Kolt GS, McEvoy JF. Adherence to rehabilitation in patients with low back pain. Man Ther 2003; 8: 110–116.
7. Iversen MD, Fossel AH, Katz JN. Enhancing function in older adults with chronic low back pain: a pilot study of endurance training. Arch Phys Med Rehabil 2003; 84: 1324–1331.
8. Linton SJ, Hellsing AL, Bergström G. Exercise for workers with musculoskeletal pain: does enhancing compliance decrease pain? J Occup Rehabil 1996; 6: 177–189.
9. Middleton A. Chronic low back pain: patient compliance with physiotherapy advice and exercise, perceived barriers and motivation. Phys Ther Rev 2004; 9: 153–160.
10. Sluijs EM, Kok GJ, van der Zee J. Correlates of exercise compliance in physical therapy. Phys Ther 1993; 73: 771–782.
11. Friedrich M, Gittler G, Halberstadt Y, Cermak T, Heiller I. Combined exercise and motivation program: effect on the compliance and level of disability of patients with chronic low back pain: a randomized controlled trial. Arch Phys Med Rehabil 1998; 79: 475–487.
12. Alexandre NM, Nordin M, Hiebert R, Campello M. Predictors of compliance with short-term treatment among patients with back pain. Rev Pan Salud Pública 2002; 12: 86–95.
13. Bassett SF. The assessment of patient adherence to physiotherapy rehabilitation. NZ J Physiother 2003; 31: 60–66.
14. Mayo NE. Patient compliance: practical implications for physical therapists. A review of the literature. Phys Ther 1978; 58: 1083–1090.
15. Dean SG, Smith JA, Payne S, Weinman J. Managing time: an interpretative phenomenological analysis of patients' and physiotherapists' perceptions of adherence to therapeutic exercise for low back pain. Disabil Rehabil 2005; 27: 625–636.
16. Jones SL, Jones PK, Katz J. Compliance for low back patients in the emergency department. A randomized trial. Spine 1988; 13: 553–556.
17. Turk DC, Rudy TE. Neglected topics in the treatment of chronic pain patients-relapse, noncompliance and adherence enhancement. Pain 1991; 44: 5–28.
18. Cook F, Hassenkamp AM. Active rehabilitation for chronic low back pain. The patient's perspective. Physiotherapy 2000; 86: 61–68.
19. Charmaz K. Grounded theory. In: Smith JA, Harre R, van Langenhove L, editors. Rethinking methods in psychology. London: Sage Publications, Inc; 1995, p. 27–49.
20. Giorgi A. Some theoretical and practical issues regarding the psychological phenomenological method. Saybrook Review 1989; 7: 71–85.
21. Patton MQ, editor. Qualitative evaluation and research methods, 3rd edn. Thousand Oaks, CA: Sage Publications, Inc; 2002.
22. Miles MB, Huberman AM, editors. Qualitative data analysis: an expanded sourcebook. Thousand Oaks, CA: Sage Publications, Inc; 1994.
23. Maxwell JA, editor. Qualitative research design: an interactive approach, 2nd edn. Thousand Oaks, CA: Sage Publications, Inc; 2005.
24. Giorgi A. One type of analysis of descriptive data: procedures involved in following scientification on emotional method. Methods 1989; 1: 39–61.
25. Taylor AH, May S. Threat and coping appraisal as determinants of compliance with sports injury rehabilitation: an application of Protection Motivation Theory. J Sports Sci 1996; 14: 471–482.
26. Brewer BW, Cornelius AE, Van Raalte JL, Petitpas AJ, Sklar JH, Pohlman MH, et al. Protection motivation theory and adherence to sport injury rehabilitation revisited. The Sport Psychologist 2003; 17: 95–103.
27. Bandura A, editor. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall; cop. 1986.
28. Gerber KE. Compliance in the chronically ill: an introduction to the problem. In: Gerber KE, Nehemkis AM editors. Compliance: the dilemma of the chronically ill. New York: Springer; 1986, p. 12–23.
29. Beazley DL, Paris DR, Partington JT. Correlates of compliance among chiropractic patients treated for low back pain. J Manipulative Physiol Ther 1986; 9: 125–129.
30. Campbell R, Evans M, Tucker M, Quilty B, Dieppe P, Donovan JL. Why don't patients do their exercises? Understanding noncompliance with physiotherapy in patients with osteoarthritis of the knee. J Epidemiol Community Health 2001; 55: 132–138.
31. Buckworth J, Dishman RK, editor. Exercise psychology. Champaign, IL: Human Kinetics; 2002.
32. Byerly PN, Worrell T, Gahimer J, Domholdt E. Rehabilitation compliance in athletic training environment. J Athl Train 1994; 29: 352–355.
33. Benson J, Britten N. Patients' decisions about whether or not to take antihypertensive drugs: qualitative study. BMJ 2002; 325: 873–877.
34. Donovan JL, Blake DR. Patient non-compliance: deviance or reasoned decision-making? Soc Sci Med 1992; 34: 507–513.
35. Ajzen I, editor. Attitudes, personality and behaviour. Buckingham: Open University Press; 1988.
36. Kerssens JJ, Sluijs EM, Verhaak PFM, Knibbe HJ, Hermans IMJ. Educating patient educators: enhancing instructional effectiveness in physical therapy for low back pain patients. Patient Educ Couns 1999; 378: 165–176.
37. Escolar Reina P. Adherencia a la fisioterapia en los pacientes con algias vertebrales: magnitud y factores explicativos [Adherence to physical therapy in patient with back pain. Magnitude and explanatory factors]. PhD Thesis. Murcia: Universidad de Murcia; 2007 (in Spanish).

Systematic review

Systematic review of management of chronic pain after surgery

V. Wylde[1], J. Dennis[1], A. D. Beswick[1], J. Bruce[3], C. Eccleston[4,5], N. Howells[2], T. J. Peters[1] and R. Gooberman-Hill[1]

[1]Musculoskeletal Research Unit, School of Clinical Sciences, University of Bristol, and [2]Avon Orthopaedic Centre, Department of Trauma and Orthopaedics, North Bristol NHS Trust, Bristol, [3]Warwick Clinical Trials Unit, University of Warwick, Warwick, and [4]Centre for Pain Research, University of Bath, Bath, UK, and [5]Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium

Correspondence to: Dr V. Wylde, Musculoskeletal Research Unit, Learning and Research Building, Southmead Hospital, Bristol BS10 5NB, UK (e-mail: v.wylde@bristol.ac.uk)

Background: **Pain present for at least 3 months after a surgical procedure is considered chronic postsurgical pain (CPSP) and affects 10–50 per cent of patients. Interventions for CPSP may focus on the underlying condition that indicated surgery, the aetiology of new-onset pain or be multifactorial in recognition of the diverse causes of this pain. The aim of this systematic review was to identify RCTs of interventions for the management of CPSP, and synthesize data across treatment type to estimate their effectiveness and safety.**

Methods: **MEDLINE, Embase, PsycINFO, CINAHL and the Cochrane Library were searched from inception to March 2016. Trials of pain interventions received by patients at 3 months or more after surgery were included. Risk of bias was assessed using the Cochrane risk-of-bias tool.**

Results: **Some 66 trials with data from 3149 participants were included. Most trials included patients with chronic pain after spinal surgery (25 trials) or phantom limb pain (21 trials). Interventions were predominantly pharmacological, including antiepileptics, capsaicin, epidural steroid injections, local anaesthetic, neurotoxins, *N*-methyl-D-aspartate receptor antagonists and opioids. Other interventions included acupuncture, exercise, postamputation limb liner, spinal cord stimulation, further surgery, laser therapy, magnetic stimulation, mindfulness-based stress reduction, mirror therapy and sensory discrimination training. Opportunities for meta-analysis were limited by heterogeneity. For all interventions, there was insufficient evidence to draw conclusions on effectiveness.**

Conclusion: **There is a need for more evidence about interventions for CPSP. High-quality trials of multimodal interventions matched to pain characteristics are needed to provide robust evidence to guide management of CPSP.**

Paper accepted 21 April 2017

Published online 6 July 2017 in Wiley Online Library (www.bjs.co.uk). **DOI:** 10.1002/bjs.10601

Introduction

Pain present for at least 3 months after a surgical procedure is described as chronic postsurgical pain (CPSP)[1]. CPSP affects between 10 and 50 per cent of patients after common operations such as mastectomy, cardiac surgery, hysterectomy, hernia repair, joint replacement, back surgery and also more minor procedures[2–8]. In a European survey[9] of surgical patients, the prevalence of moderate to severe CPSP at 12 months after operation was 11·8 per cent. Chronic pain is associated with poor general health, disability, depression[9–12] and social withdrawal, and increases the risk of further co-morbidities[13]. CPSP has been defined previously as pain that develops after surgery[5], and a proposed update to the definition includes the possibility that CPSP is pain that increases in intensity after surgery[14]. This update allows for the possibility that pain among patients who undergo surgery to relieve pain is also, appropriately, included in the definition.

Risk factors for CPSP may be genetic, psychosocial, or related to preoperative or acute postoperative pain severity[2,15]. However, certain surgical procedure-related factors are key for the development of chronic pain[16]. Surgical procedures lasting longer than 3 h may increase the risk of postoperative pain[5]. A major surgical factor in the development of chronic pain is nerve injury, and patients undergoing thoracic, breast and hernia surgery are at particular risk of neuropathic pain[8]. Inflammation resulting from intraoperative tissue injury can contribute

to central sensitization and further pain[2]. Inadequate preventive analgesia may also contribute[17].

Knowledge of determinants and predictors of CPSP can guide the development of interventions and help target care. Possible forms of management for CPSP may focus on the underlying condition that needed surgery, on the aetiology of the pain, or be multifactorial in recognition of the diverse causes of postoperative pain. Although some forms of management may have limited applicability outside of the specific condition for which they were intended, others may be transferrable, regardless of the surgical procedure.

There are systematic reviews of pharmacological and other interventions for the management of chronic pain, defined generally, or specific to the presumed mechanisms (such as neuropathic pain[17] and cancer pain[18]). A number of Cochrane reviews[17,19–29] have included studies evaluating interventions for CPSP, although this was not the primary focus of these reviews. It is rare for any review to focus specifically on chronic pain in the postoperative context. Exceptions include reviews that have focused on interventions for chronic pain after particular surgical procedures, including phantom limb pain after amputation[30] and knee replacement[31]. The aim of the present review was to identify RCTs of interventions for the management of CPSP and to synthesize data across treatment type to provide an estimate of their effectiveness and safety. In keeping with recommended practice, a systematic review is a key step toward the development of future trials to evaluate interventions for CPSP[32].

Methods

The protocol was registered in the international prospective register of systematic reviews (PROSPERO; http://www.crd.york.ac.uk/PROSPERO/) on 15 January 2015 (registration number 15957). The review was conducted in accordance with PRISMA guidelines[33].

Eligibility criteria

Published articles describing RCTs involving any intervention that aimed to provide management of CPSP were included. Eligible studies reflected PICO criteria[34]: patients aged 18 years or more and at least 90 per cent of study participants reporting CPSP; interventions for pain received by patients at a minimum of 3 months after surgery; comparison arm of placebo, usual care or an alternative pain management intervention; and outcomes were pain reported using any data collection tool(s).

Information sources and searches

MEDLINE, Embase, PsycINFO, CINAHL and the Cochrane Library were searched from inception to 23 March 2016. The search strategies were modified for different bibliographic databases (*Appendix S1*, supporting information). No language restrictions were applied. Reference lists were checked and registers inspected; grey literature (literature not formally published as journal articles) was sought in OpenGrey (http://www.greynet.org/opengreyrepository.html), a database of grey literature, on 30 March 2016. A minimum sample size was not specified in the protocol to ensure inclusion of all treatments of potential interest to clinicians and researchers working in a range of surgical specialties.

Published conference abstracts were followed up to obtain any full publications, but otherwise excluded. After completion of data extraction, relevant systematic reviews were identified from the Cochrane Database of Systematic Reviews, and included studies were reviewed to identify any studies missed in the initial searches because eligibility was not apparent from the title and abstract.

Study selection

All records identified in the search were imported into EndNote X7 (Thomson Reuters, New York, New York, USA). Abstracts or full-text articles were screened to remove obviously irrelevant reports. Reasons for excluding studies were recorded as free text in EndNote X7. This was performed by one author who was over-inclusive if eligibility was not clear. A sample of 10 per cent was double screened by a second author, which identified one eligible study that had been missed. The final selection of studies was then performed in duplicate by two authors. When there was insufficient information to determine eligibility, study author e-mail addresses were obtained and supplementary information was requested.

Data collection

Data from included studies were extracted using standard forms by one author and checked by a second author. Study setting, participant demographics, methodology, recruitment, duration, treatment characteristics, length of follow-up, outcomes, tools used to measure outcomes, and information for the risk-of-bias assessment were recorded. Authors of studies were contacted where necessary for clarification and to provide missing or incomplete data.

Outcomes

In accordance with GRADE guidelines[35] and Cochrane guidance, the total number of outcomes planned to be included in this review was limited to seven (2 primary and 5 secondary). The primary clinical effectiveness outcome was pain intensity and the primary harm outcome was serious adverse events. These reflect recommendations from the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT)[36] and the Cochrane Pain, Palliative and Supportive Care Review Group (PaPaS)[37]. Studies were required to report the primary outcome of pain intensity to be eligible for inclusion in the review. The first secondary outcome was the presence or absence of neuropathic pain, which is particularly relevant to chronic pain after a surgical intervention[8]. The other four secondary outcomes reflected the IMMPACT core outcome domains for chronic pain clinical trials: physical functioning, emotional functioning, participants' ratings of global improvement and satisfaction with treatment, and participant disposition[36]. No limits were placed on the tools used to measure these outcomes.

Risk of bias in individual studies

Risk of bias was assessed using the Cochrane risk-of-bias tool[38]. Two authors assessed the risk of bias independently across the six domains of the tool for each study. Results are reported through graphical representation of bias judgements grouped by intervention.

Statistical analysis

In the protocol, meta-analyses were planned using RevMan 5 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark) with general guidance from the Cochrane Handbook[38]. For dichotomous (binary) data, the odds ratio with 95 per cent c.i. would be used. For continuous data, if outcomes were measured identically across studies, an overall mean difference and 95 per cent c.i. would be calculated. If continuous outcomes were measured differently across studies, overall standardized mean differences and 95 per cent c.i. would be calculated. For data from crossover trials, the generic inverse-variance method in RevMan 5 would be used.

Opportunities for meta-analysis were limited by heterogeneity between studies. Even when multiple studies for a particular intervention were identified, variation in modes of administration, comparator groups and/or format of outcome data precluded pooling. Thus, the majority of results are reported narratively, with results of meta-analyses described only for gabapentin and capsaicin. Planned subgroup analysis of pharmacotherapy, physical/self-management and multidisciplinary interventions was not possible owing to clinical and methodological heterogeneity. Results for pain outcome at final follow-up within individual studies are presented as reported by investigators (*Tables S1* and *S2*, supporting information).

Results

Included trials

Searches identified 17 029 articles, of which 660 were considered potentially relevant after initial screening. Author e-mail addresses were traced for 57 of 78 studies that contained insufficient information to determine eligibility, and further data were requested. Replies were received for 16 studies, and only one was eligible for inclusion. The remaining articles were assumed to be ineligible as the abstract or full text made no reference to patients having CPSP. After evaluation of full-text articles, 66 trials[39–104] with data from 3149 participants were included (*Fig. 1*).

An overall summary of trial characteristics is provided in *Table 1* and characteristics of individual trials are shown in *Tables S1* and *S2* (supporting information). *Table 2* summarizes studies according to the index surgery and intervention. Individual components of risk-of-bias assessment are provided in *Appendix S2* (supporting information).

Trial design

Trials were generally small, ranging in size from three to 250 participants (median 38). The study with three participants was a pilot trial, but a total of 18 trials recruited fewer than 20 participants. Sample size calculations were reported in 34 of 66 studies; of these, 13 failed to recruit or retain sufficient numbers of participants to meet their calculation. Authors did not always state dates of recruitment; publication dates ranged from 1989 to 2016. There has been an increase over time in the number of published trials in this field, from ten trials published before 2001 to 23 published between 2011 and 2015.

Interventions

The primary method of reporting results was grouped according to treatment type. The majority of studies evaluated pharmacological interventions, and so studies were grouped as primarily pharmacological, or as primarily physical, surgical, psychological and other (*Table 1*).

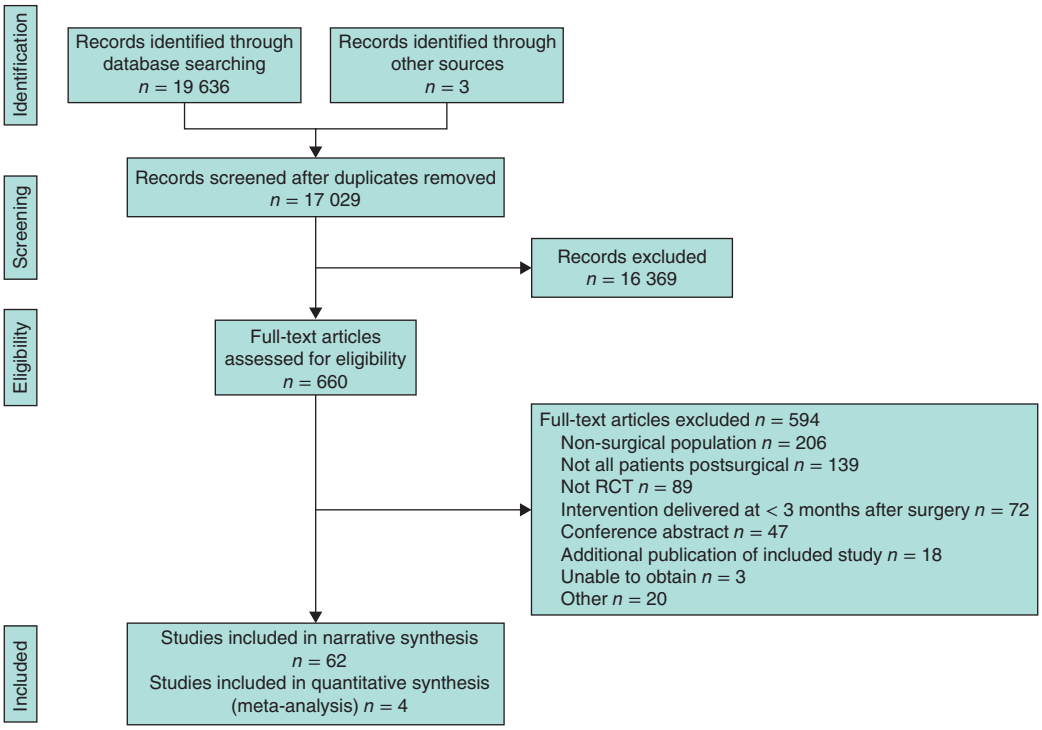


Fig. 1 Flow diagram showing selection of articles for review

Outcome measurements

A visual analogue scale (VAS) was used to assess pain intensity in 37 of 66 studies (56 per cent). Validated pain-specific tools were used infrequently; the most common was the McGill Pain Questionnaire, which was used in eight studies. Only one trial referenced IMMPACT criteria[36] for patient outcomes. In terms of presentation of the primary outcome (pain intensity) as the percentage of patients reporting 30 or 50 per cent improvement, a format preferred by both IMMPACT[105] and the Cochrane PaPaS Group[106], only a minority of authors (6 of 31 trials published after 2008) were compliant. Serious adverse events and the secondary outcomes in this review were reported inconsistently and there was no opportunity to summarize these outcomes; therefore, only pain outcomes are presented in *Tables S1* and *S2* (supporting information).

Pharmacological interventions

Antidepressants

Four trials, including data from 177 participants, evaluated the effect of antidepressants on chronic pain after amputation[39,40] or breast surgery[41,42]. Risk of bias was evident in two studies owing to incomplete outcome data[40,42], and a change to the definition of responder and potential funder bias[40]. Amitriptyline was evaluated in three trials[39,40,42], with some issues suggesting risk of bias, and venlafaxine in one trial[41] with a low risk of bias. There was no evidence that a 4–6-week course of antidepressants reduced pain intensity compared with placebo, except in one trial[42] involving 20 patients which found that patients reported lower breast scar pain intensity after 4 weeks of 100 mg/day amitriptyline compared with placebo. However, this trial also found evidence that amitriptyline resulted in more adverse events than placebo.

Antiepileptics

Eight trials including data from 338 participants evaluated the effects of antiepileptic medications on CPSP. The largest number of studies for any one technology was for gabapentin (6 studies, 293 participants with pain after amputation[43,44], breast cancer surgery[45], sternotomy[46] and spinal surgery[47,48]). Four trials were at risk of bias owing to issues relating to blinding[48], blinding and randomization[45], incomplete outcome data[44], and blinding and single-authored article[47]. Meta-analysis using the generic inverse-variance method was possible for the primary outcome for one subgroup only (gabapentin *versus* placebo for 6 weeks) involving two crossover trials with 43 patients[43,44]. This demonstrated a within-person mean difference in pain intensity measured on a VAS and numerical

Table 1 Overall summary of trial characteristics

Trial design	Parallel design (41), crossover study (25)
No. of arms	Two arms (53), three arms (12), five arms (1)
Countries	USA (21), Germany (8), Denmark (7), Iran (3), China (2), Egypt (2), Finland (2), France (2), Italy (2), Korea (2), Sweden (2), Turkey (2), Belgium (1), Canada (1), Israel (1), Mozambique (1), The Netherlands (1), Norway (1), Serbia (1), Spain (1), Switzerland (1), UK (1), international multisite (1)
Surgery types	Spinal surgery (25), amputation (21), breast cancer surgery (8), inguinal hernia repair (3), neck dissection for cancer (2), knee replacement (1), sternotomy (1), abdominal surgery (1), shoulder surgery (1), various surgical procedures (3)
Interventions	
Pharmacological	Antidepressants as analgesics (4), antiepileptics (8), capsaicin (3), epidural steroid injections and associated interventions (11), local anaesthetic (11), neurotoxins (3), *N*-methyl-D-aspartate receptor antagonist (7), opioids (6), calcitonin (1), naloxone as an adjuvant to morphine (1)
Physical, surgical, psychological and other pain management	Acupuncture/dry needling (2), exercise (4), limb cover/liner for patients who had undergone amputation (2), spinal cord stimulation (5), further surgery (2), laser therapy (1), magnetic stimulation (1), mindfulness-based stress reduction (1), mirror therapy for amputation (1), sensory discrimination training (1), joint manipulation (1), combined package of hot packs, ultrasound treatment and transcutaneous electrical nerve stimulation (1)
Comparator interventions	Active treatment (31), placebo or sham (31), usual care (2), waiting list (1), no treatment (1)

Values in parentheses are number of studies.

rating scale score of −1·12 (95 per cent c.i. −1·89 to −0·36; $I^2 = 53$ per cent), favouring gabapentin (*Fig. S1*, supporting information). Similar results are reported in a Cochrane review[30] and no studies additional to those included in the previous review were identified. The two trials that compared gabapentin with non-steroidal anti-inflammatory drugs (NSAIDs) found that gabapentin provided more effective pain relief after 1 month[46] and 6 months[48] of treatment. An 8-week course of gabapentin was found to be superior to a stellate ganglion block using bupivacaine in a trial involving 60 patients after mastectomy[45]. The addition of 1 month of oral gabapentin to standard epidural corticosteroids was found to result in lower pain at 6 months compared with epidural corticosteroids alone after spinal surgery[47].

One trial[49] with a low risk of bias found no differences in pain relief between levetiracetam and placebo over 4 weeks of treatment. Pain relief after taking pregabalin for 7 weeks for chronic pain after abdominal surgery was compared with placebo in a study of 13 patients[50]; results favouring the treatment group must be interpreted with caution as the trial was terminated early by the industry sponsor.

Capsaicin

Three trials including data from 174 participants evaluated capsaicin for relief of chronic pain after inguinal hernia repair[51], mastectomy[52], and diverse procedures for cancer[53]. All studies were at risk of bias because of issues relating to blinding of a preparation with a burning sensation and erythema. One study[52] was also at risk of bias owing to selective reporting. The trial[51] assessing a single 60-min application of a capsaicin patch (8 per cent) found no evidence of pain relief compared with placebo after 3 months. Two trials[52,53] of low-dose (0·075 per cent) capsaicin topical cream applied four times daily for 6–8 weeks reported some evidence of reduced pain intensity compared with placebo. Meta-analysis suggested a modest positive effect of capsaicin topical cream on the proportion of patients reporting pain improvement (odds ratio 2·64, 95 per cent c.i. 1·02 to 6·86; $I^2 = 0$ per cent) (*Fig. S2*, supporting information), although caution is warranted owing to risk of bias and, as a previous Cochrane review[107] advised, the total number of events was too few to be reliable. In both trials, a commonly reported side-effect was local skin reaction.

Epidural injections and associated interventions

Eleven trials including data from 886 participants evaluated epidural injections and associated interventions after spinal surgery[54–64]. Risk of bias was evident in ten studies relating to allocation concealment[54–56,64], blinding[54,58,61], incomplete outcome data[54,57,61,63], selective reporting[56,57,59,62] and single-authored article[63]. Two trials[55,56] comparing epidural injections with, and without steroids found no difference in pain relief between groups. The addition of steroids to 3-monthly morphine epidural injections was not found to influence pain intensity after 6 months[59]. A three-arm trial[63] involving 206 patients evaluated epidural injections of 1 mg indomethacin, 2 mg indomethacin and 80 mg methylprednisolone, and found that all treatments resulted in a similar pain reduction. Two trials evaluating epidural injections of steroids (prednisolone acetate) *versus* saline alone produced contrasting results: one[57] noted no benefit at 120 days and the other[58] reported a reduction in pain at 18 months after multiple epidural injections. The addition of hyaluronidase to an epidural steroid injection was found to lead to lower pain intensity at 4 weeks[62] and 12 months[60], and a combination of hyaluronidase and triamcinolone provided more effective pain relief for 12 weeks than either agent alone[61]. Two trials reported that adding percutaneous adhesiolysis to an epidural injection

Table 2 Summary of included studies according to the index surgery and intervention

	Amputation	Spinal surgery	Breast cancer	Abdominal surgery	Hernia repair	Knee replacement	Neck dissection	Sternotomy	Shoulder surgery	Mixed
Pharmacological interventions										
Antidepressants	2[39,40]		2[41,42]							
Antiepileptics	2[43,44]	2[47,48]	2[45,49]	1[50]				1[46]		
Capsaicin			1[52]		1[51]					1[53]
Epidural injection and associated interventions		11[54–64]								
Local anaesthetic	4[70–73]	2[67,68]	2[45,74]		2[65,66]					1[69]
Neurotoxins	1[77]					1[75]	1[76]			
NMDA receptor antagonist	6[78–83]									1[69]
Opioids	4[40,70,71,84]	1[59]	1[85]							
Other	1[78]									1[86]
Physical, surgical, psychological and other interventions										
Acupuncture/dry needling							1[88]		1[87]	
Exercise	1[91]	3[89,90,92]								
Limb cover/liner	2[93,94]									
Spinal cord stimulation		5[95–99]								
Surgery		2[90,95]								
Other	2[102,104]	3[92,100,103]	1[101]							

Nine trials[40,45,59,69–71,90,92,95] were included twice or more as they evaluated interventions which fall into different categories. NMDA, *N*-methyl-D-aspartate.

led to better pain relief at 6 months[64] and 12 months[54] after treatment.

Local anaesthetics

Eleven trials including data from 324 participants assessed the effectiveness of local anaesthetics in providing pain relief after inguinal hernia repair[65,66], spinal surgery[67–69], amputation[70–73] and breast cancer surgery[45,74]. Risk of bias was evident in three studies, and concerned incomplete outcome data[73], random sequence generation and blinding[45], and early trial termination[65]. Interventions assessed included lidocaine block[65], repeated epidural nerve blocks[68], stellate ganglion block[45,74], bupivacaine[72], intravenous lidocaine[67,69,70], ropivacaine[73], oral mexiletine[71] and lidocaine patch[66].

Five trials evaluated local anaesthetic nerve blocks. No difference in pain intensity was found after ultrasound-guided lidocaine nerve block compared with placebo block[65] or after repeated epidural sympathetic nerve block compared with saline blocks[68]. A trial[45] of stellate ganglion blocks for pain after breast surgery found that they were inferior to gabapentin; another trial[74] noted that pain relief at 8 weeks was improved with ultrasound guidance compared with unguided blocks. One trial[72] found that injections of bupivacaine into contralateral painful muscle sites that mirror phantom limb pains were more effective at providing pain relief than placebo saline injections.

Five trials evaluated systemic administration of intravenous local anaesthetic or oral mexiletine. Two trials[67,69] found that intravenous lidocaine did not reduce pain intensity compared with saline, and one[70] reported that it reduced stump pain, but not phantom limb pain. A pilot trial[73] of three patients reported that ropivacaine reduced phantom limb pain after 12 weeks, although no statistical tests were performed on this small patient sample. An 8-week course of oral mexiletine was found to have no effect on pain intensity compared with placebo in a trial with 60 patients[71].

A single trial[66] of 21 patients evaluated lidocaine patches (5 per cent); applied for 2 weeks, they were found to produce similar results to placebo patches.

Neurotoxins

Three trials including data from 91 participants evaluated botulinum toxin A injections for chronic pain after knee replacement[75], neck dissection[76] and lower limb amputation[77]. One study[77] had evidence of bias relating to incomplete outcome data. In patients with knee replacement treated with botulinum toxin A, pain intensity was reduced compared with that in the placebo group after 2 and 7 months, with no increase in adverse events[75]. In a dose-finding study[76] involving patients with chronic pain after neck dissection, a lower dose of botulinum A toxin was associated with reduced pain intensity. There was little evidence that botulinum A was more effective than lidocaine/Depo-Medrol® (Pharmacia & Upjohn, New York, New York, USA) injection after 6 months in a pilot trial of patients with phantom limb pain[77].

N-methyl-D-aspartate receptor antagonists

Seven trials, including data from 122 participants, evaluated *N*-methyl-D-aspartate (NMDA) receptor antagonists. One study[78] had evidence of risk of bias relating to blinding and incomplete outcome data. No differences in pain relief after 3–5 weeks of memantine compared with placebo were found in four trials[79–82] involving patients with pain after amputation. Three studies, involving 11–20 patients each, evaluated ketamine alone or in conjunction with calcitonin with placebo for patients with phantom limb pain[78,83] or pain after diverse procedures[69]. All trials provided evidence that the intervention reduced pain intensity, although follow-up was short (80 min to 48 h).

Opioids

Six trials including data from 297 participants evaluated opioids for chronic pain after amputation[40,70,71,84], breast surgery[85] and spinal surgery[59]. Opioids evaluated included tramadol[40], oral morphine[71,84,85], morphine infusion[70] and epidural morphine[59]. In three trials, risk of bias was noted relating to blinding[84], incomplete outcome data[40] and selective reporting[59]. Compared with placebo, oral morphine was found to provide better pain relief at 4–6 weeks[70,71,84], although a common side-effect was constipation. Trials evaluating 4 weeks of tramadol compared with placebo[40], 6 weeks of morphine compared with gabapentin with, or without NSAIDs[85], and 3-monthly injections of epidural morphine and steroids compared with steroids alone[59], found no differences in pain intensity between treatment groups.

Other pharmacological interventions

One trial[78], with risk of bias relating to blinding and incomplete outcome data, evaluated intravenous calcitonin for phantom limb pain in 20 patients, and found no effect up to 48 h after infusion compared with saline. Another trial[86], with no clear evidence of risk of bias, evaluated low doses of oral/or intravenous naloxone as a supplement in 12 patients whose severe CPSP was already managed by continuous intrathecal morphine administration. No evidence of an effect on pain relief was found after two 3-week sessions on differing doses of the drug across a 9-week period

Physical, surgical, psychological and other interventions

Acupuncture/dry needling

No evidence of differences in pain relief was found in a trial of 20 patients comparing dry needling and physiotherapy with physiotherapy alone for pain after shoulder surgery[87]. Another trial[88] involving 70 patients with chronic pain after neck dissection reported that acupuncture resulted in better pain relief than usual care after 42 days. Neither participants nor assessors were blinded and this may have introduced bias.

Exercise

Four trials[89–92] involving 323 participants evaluated exercise interventions, often as a component of a broader package of care. Two studies were at risk of bias owing to lack of blinding[90,92]. No evidence of differences in pain relief was found in trials comparing 3 months of exercise with, and without hyperextension exercises after lumbar surgery[89], and 3 weeks of exercise combined with a cognitive intervention compared with lumbar fusion after disc herniation surgery[90]. A 4-week training programme of progressive muscle relaxation, mental imagery and phantom exercises was found to be more effective at relieving phantom limb pain than a general exercise programme[91]. A trial[92] of treatment of pain after laminectomy found that 8 weeks of low-tech exercises (McKenzie-type and spinal stabilization training exercise) or high-tech exercises (cardiovascular, isotonic and isokinetic exercises) resulted in a reduction in pain-related disability compared with no treatment.

Limb cover/lining

One trial[93] with 57 patients reported that non-invasive limb covering for 12 weeks compared with sham limb covering did not reduce phantom limb pain. Another trial[94] involving 30 patients, which was at risk of bias owing to incomplete outcome data, found evidence that a stump liner worn by amputees for 2 weeks reduced pain compared with a placebo liner.

Spinal cord stimulation

Five trials including 260 participants assessed the impact of spinal cord stimulation (SCS) on chronic pain after spinal surgery. Two studies[95,96] had a risk of bias related to blinding, and another[97] owing to blinding and commercial interests. Two trials found that patients who received SCS for 6 months reported better pain relief than those who had conventional management (100 patients)[97] or reoperation (60)[95]. Subcutaneous stimulation as adjunct therapy to SCS was noted to provide better relief of back pain, but not leg pain, compared with sham treatment in a trial of 20 participants[96]. Burst SCS was found to be more effective at providing pain relief after 1 week than tonic or placebo SCS[98]. In a trial involving 15 patients[99], there was no difference in pain after 2 weeks on stimulation with 1000- *versus* 500-Hz bursts.

Surgery

No evidence of differences in pain relief was found when lumbar fusion was compared with exercise for chronic pain after disc herniation surgery in a trial of 60 participants[90]. In a trial[95] involving 60 patients with failed back surgery syndrome, less pain relief after 6 months was reported by patients who had reoperation than was reported by patients who had SCS.

Other interventions

No evidence of differences in pain relief were found in a trial of cutaneous magnetic stimulation for 24 h compared with sham treatment after spinal surgery in 17 patients[100]. Four weeks of laser therapy compared with placebo laser therapy was found to reduce mastectomy pain at 12 weeks in a trial of 61 participants[101]. An unblinded trial that included ten patients[102], which was at risk of bias, found that 2 weeks of sensory discrimination training led to a reduction in phantom limb pain at 3 months compared with comprehensive psychophysiological assessment. One trial[103] with 40 patients, reported that 8 weeks of mindfulness-based stress reduction following spinal surgery led to better pain relief after 12 weeks compared with that in the waiting list control group. However, the trial was at risk of bias because of lack of blinding and incomplete outcome data. In a three-arm trial[104] that included 22 patients with phantom limb pain, mirror therapy was found to reduce pain intensity after 4 weeks compared with sham mirror therapy and mental visualization. The study was at risk of bias owing to lack of blinding. A trial[92] involving patients with pain after laminectomy found that an 8-week course of joint manipulation did not reduce pain-related disability compared with no treatment. The same trial also found no difference between a combined package of hot packs, ultrasound treatment and transcutaneous electrical nerve stimulation, and no treatment.

Discussion

The best evidence to guide the implementation of effective interventions comes from their evaluation in high-quality randomized trials, and ultimately in systematic reviews and meta-analyses. Given the prevalence and impact of CPSP, it is imperative to establish robust methods for its management. This systematic review aimed to provide a comprehensive evaluation of the evidence base for the management of CPSP. Although some of the interventions identified were procedure-specific, others had wider applicability to other types of CPSP. However, owing to heterogeneity in the interventions and trial design, pooling of data in meta-analysis was rarely possible or warranted. Of the 66 included trials, most evaluated pharmacological interventions. For all interventions, there was insufficient evidence to draw conclusions on effectiveness or harm.

There are few systematic reviews in the field of CPSP, with existing reviews focusing on predictors[108], characteristics[8] and prevention[109–112]. The previous reviews that have evaluated treatments have been procedure-specific, focusing on chronic pain after total knee replacement[31] and phantom limb pain[30]. This contrasts with other areas of pain research in which numerous systematic reviews[21,22,113–116] of treatments have been published. Typically, the focus of a review is on a defined condition (such as fibromyalgia, back pain) or a presumed mechanism of chronic pain (for example neuropathic pain). Patients with CPSP are, of course, embedded within broader trials investigating chronic pain, but it has not previously been possible to identify these patients.

This review highlighted some difficulties with conducting a broad systematic review of CPSP. First, there was heterogeneity in the definition of CPSP within research studies; some trials included only patients with neuropathic pain and there was variability across studies in key eligibility criteria, such as duration and severity of pain. Second, one-third of the studies included in the review evaluated interventions for phantom limb pain. Although previous reviews of CPSP have also included amputation[109,111], the commonality in the aetiology of phantom limb pain and other forms of CPSP could be questioned. However, phantom limb pain was included as the aim of this review was to provide a broad overview of interventions for chronic pain in the surgical context. The identification of interventions that show effectiveness in one well studied surgical model could provide directions for the evaluation of interventions for CPSP in other surgical areas.

It is important to acknowledge the limitations of this review when interpreting the results. Searches yielded a large volume of literature and therefore initial eligibility screening was performed in duplicate for only 10 per cent of the studies; this may have increased the risk of eligible studies being discarded[117]. However, the final selection of studies was undertaken by two reviewers in accordance with guidance from the Cochrane Handbook[38]. Given the hidden nature of patients with CPSP included within other trials, relevant studies were often difficult to identify from titles and abstracts, and required investigation of the full text to establish whether or not patients were likely to have CPSP. Although the search terms identified a large volume of literature, search of relevant Cochrane reviews identified three other relevant studies that were not identified in the initial searches. This highlights the difficulty of conducting such a systematic review owing to limited reference to the

patient sample by conventional means – there is no medical subject heading (MeSH) for CPSP, so indexers and even study authors did not necessarily use CPSP as a descriptor or keyword. Limits were not placed on the tools used to assess secondary outcomes and the resulting heterogeneity precluded their inclusion in analysis. Adverse events were found to be poorly and inconsistently reported. This has previously been described as a common issue in chronic pain trials[30], and poor reporting precluded conclusions about intervention safety in this review. Opportunities for meta-analysis were limited because of variability in the identified interventions, and conclusions are predominantly based on narrative synthesis. However, this review has produced a comprehensive overview of the evidence for management of CPSP, and the findings have a number of methodological and clinical implications.

Only three trials included patients with CPSP after various surgical procedures; the remainder focused on one surgery type. Of these, the majority of trials were conducted to evaluate treatments for phantom limb pain and failed back surgery syndrome, which is likely to reflect the historical recognition of these pain conditions[118,119]. Although an encouraging temporal increase in the number of trials conducted was identified, the paucity of research into the management of CPSP, particularly after operations other than amputation and spinal surgery, highlights the need for further research. The majority of trials in this review evaluated pharmacological interventions, reflecting the commonplace role of these therapies in the management of chronic pain. There was insufficient evidence to evaluate the effectiveness of any treatment modality in reducing CPSP. It has previously been proposed that commonly prescribed pharmacological treatments are insufficient to treat chronic non-cancer pain when used in isolation[120]. Given the complex and multifactorial nature of CPSP, an individualized and multimodal model of care may be required, as recommended more widely for chronic non-cancer pain[120].

Similar to a previous review of interventions for phantom limb pain[30], the present analysis identified the need for more methodological rigour in the reporting and conduct of randomized trials in this field. This need is highlighted by the unclear or high risk of bias rating assigned to many aspects of the included trials. Frequently encountered issues included lack of transparency, as shown by lack of preregistration of trials or publication of trial protocols, failure to report conduct/results according to CONSORT standards[121–123], and limited and variable assessment of pain and adverse events. IMMPACT recommendations[36] suggest the use of a comprehensive approach to pain assessment in clinical trials addressing chronic pain. Many of the trials included in this review were conducted before publication of the IMMPACT recommendations in 2003. However, the trials conducted and published after the IMMPACT guidance generally limited their outcome assessment to pain intensity. Inconsistent reporting of the secondary outcomes of interest precluded their analysis, highlighting the need for standardization of outcomes assessment. For many included trials, sample sizes were small and duration of follow-up was short, limiting the conclusions that can be drawn about the therapeutic benefit of interventions in the context of chronic pain.

For many included trials, threats to both internal and external validity existed. Reports of trials did not always include a sample size calculation. The inclusion criteria did not specify an *a priori* sample size, owing to the heterogeneity of the definition of CPSP and the range of potential interventions. Such a broad approach allowed this review to meet the intended aim of comprehensiveness, and to identify and present all trials within this complex and evolving field, including those in which events led to early trial termination or lower recruitment than planned. In keeping with recommendations in the Cochrane risk-of-bias tool, sample size was not considered to present a risk of bias *per se*[124], although small studies do not improve the precision of estimates. The relatively small sample sizes in some of the studies that met the inclusion criteria, as well as the high risk of bias among many of the largest trials, impacted on both results and generalizability.

In addition to issues of bias in trial conduct and reporting, the authors were initially keen to report on the quality of the evidence, potentially using GRADE. However, this was not possible because of the inability to estimate effects of treatments: all findings would have been downgraded for quality owing to the absence of evidence for synthesis. However, as this field develops and more trials emerge, it would be expected that new reviews will report effect estimates and examine the quality of the evidence.

This review highlights the need for more evidence about interventions for CPSP, and a focus not on the presumed pathological mechanism or location of pain, but on the relationship of pain to surgery. Many patients experience CPSP and it is imperative that evidence-based interventions are offered to these individuals to improve postoperative outcome. Trials to date have focused on pharmacological interventions, and no trials have been conducted to evaluate multimodal interventions matched to pain characteristics for the management of CPSP. Given the complexity of pain that extends or emerges after surgery, individualized interventions should be developed and evaluated. High-quality trials of these interventions

are needed to provide a robust evidence base to guide the management of CPSP.

Acknowledgements

This article summarizes independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research programme (RP-PG-0613-20001). The views expressed in this article are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health.
Disclosure: The authors declare no conflict of interest.

References

1 International Association for the Study of Pain. Classification of chronic pain. Descriptions of chronic pain syndromes and definitions of pain terms. Prepared by the International Association for the Study of Pain, Subcommittee on Taxonomy. *Pain Suppl* 1986; **3**: S1–S226.
2 Kehlet H, Jensen TS, Woolf CJ. Persistent postsurgical pain: risk factors and prevention. *Lancet* 2006; **367**: 1618–1625.
3 Brandsborg B, Nikolajsen L, Hansen CT, Kehlet H, Jensen TS. Risk factors for chronic pain after hysterectomy: a nationwide questionnaire and database study. *Anesthesiology* 2007; **106**: 1003–1012.
4 Thomson S. Failed back surgery syndrome – definition, epidemiology and demographics. *Br J Pain* 2013; **7**: 56–59.
5 Macrae WA. Chronic post-surgical pain: 10 years on. *Br J Anaesth* 2008; **101**: 77–86.
6 Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ Open* 2012; **2**: e000435.
7 Hoofwijk DM, Fiddelers AA, Peters ML, Stessel B, Kessels AG, Joosten EA *et al.* Prevalence and predictive factors of chronic postsurgical pain and poor global recovery 1 year after outpatient surgery. *Clin J Pain* 2015; **31**: 1017–1025.
8 Haroutiunian S, Nikolajsen L, Finnerup NB, Jensen TS. The neuropathic component in persistent postsurgical pain: a systematic literature review. *Pain* 2013; **154**: 95–102.
9 Fletcher D, Stamer UM, Pogatzki-Zahn E, Zaslansky R, Tanase NV, Perruchoud C *et al.* Chronic postsurgical pain in Europe: an observational study. *Eur J Anaesthesiol* 2015; **32**: 725–734.
10 Smith BH, Elliott AM, Chambers WA, Smith WC, Hannaford PC, Penny K. The impact of chronic pain in the community. *Fam Pract* 2001; **18**: 292–299.
11 Leadley RM, Armstrong N, Reid KJ, Allen A, Misso KV, Kleijnen J. Healthy aging in relation to chronic pain and quality of life in Europe. *Pain Pract* 2014; **14**: 547–558.
12 Magni G, Caldieron C, Rigatti-Luchini S, Merskey H. Chronic musculoskeletal pain and depressive symptoms in the general population. An analysis of the 1st National Health and Nutrition Examination Survey data. *Pain* 1990; **43**: 299–307.
13 Hawton A, Green C, Dickens AP, Richards SH, Taylor RS, Edwards R *et al.* The impact of social isolation on the health status and health-related quality of life of older people. *Qual Life Res* 2011; **20**: 57–67.
14 Werner MU, Kongsgaard UE. I. Defining persistent post-surgical pain: is an update required? *Br J Anaesth* 2014; **113**: 1–4.
15 Hinrichs-Rocker A, Schulz K, Jarvinen I, Lefering R, Simanski C, Neugebauer EA. Psychosocial predictors and correlates for chronic post-surgical pain (CPSP) – a systematic review. *Eur J Pain* 2009; **13**: 719–730.
16 Perkins FM, Kehlet H. Chronic pain as an outcome of surgery. A review of predictive factors. *Anesthesiology* 2000; **93**: 1123–1133.
17 Chaparro LE, Wiffen PJ, Moore RA, Gilron I. Combination pharmacotherapy for the treatment of neuropathic pain in adults. *Cochrane Database Syst Rev* 2012; (7)CD008943.
18 Schmidt-Hansen M, Bennett MI, Arnold S, Bromham N, Hilgart JS. Oxycodone for cancer-related pain. *Cochrane Database Syst Rev* 2015; (2)CD003870.
19 Wiffen PJ, Derry S, Moore RA, Lunn MPT. Levetiracetam for neuropathic pain in adults. *Cochrane Database Syst Rev* 2014; (7)CD010943.
20 Oosterhuis T, Costa LOP, Maher CG, De Vet HCW, Van Tulder MW, Ostelo RWJG. Rehabilitation after lumbar disc surgery – an updated Cochrane review. *Physiotherapy (United Kingdom)* 2015; **101**: eS1158–eS1159.
21 Moore RA, Wiffen PJ, Derry S, Toelle T, Rice AS. Gabapentin for chronic neuropathic pain and fibromyalgia in adults. *Cochrane Database Syst Rev* 2014; (4)CD007938.
22 McNicol ED, Midbari A, Eisenberg E. Opioids for neuropathic pain. *Cochrane Database Syst Rev* 2013; (8)CD006146.
23 Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database Syst Rev* 2005; (3)CD000335.
24 Challapalli V, Tremont-Lukats IW, McNicol ED, Lau J, Carr DB. Systemic administration of local anesthetic agents to relieve neuropathic pain. *Cochrane Database Syst Rev* 2005; (4)CD003345.
25 Derry S, Moore RA. Topical capsaicin (low concentration) for chronic neuropathic pain in adults. *Cochrane Database Syst Rev* 2012; (9)CD010111.
26 Derry S, Wiffen PJ, Moore RA, Quinlan J. Topical lidocaine for neuropathic pain in adults. *Cochrane Database Syst Rev* 2014; (7)CD010958.
27 Gallagher HC, Gallagher RM, Butler M, Buggy DJ, Henman MC. Venlafaxine for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015; (8)CD011091.

28 Staal JB, de Bie R, de Vet HC, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low-back pain. *Cochrane Database Syst Rev* 2008; (3)CD001824.
29 Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain. *Cochrane Database Syst Rev* 2011; (2)CD008112.
30 Alviar MJ, Hale T, Dungca M. Pharmacologic interventions for treating phantom limb pain. *Cochrane Database Syst Rev* 2016; (10)CD006380.
31 Beswick AD, Wylde V, Gooberman-Hill R. Interventions for the prediction and management of chronic postsurgical pain after total knee replacement: systematic review of randomised controlled trials. *BMJ Open* 2015; **5**: e007387.
32 Rosenthal R, Bucher HC, Dwan K. The use of systematic reviews when designing and reporting surgical trials. *Ann Surg* 2017; **265**: e35–e36.
33 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535.
34 Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995; **123**: A12–A13.
35 Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R *et al.* GRADE guidelines: 12. Preparing summary of findings tables – binary outcomes. *J Clin Epidemiol* 2013; **66**: 158–172.
36 Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB *et al.* Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003; **106**: 337–345.
37 Cochrane Pain, Palliative and Supportive Care Review Group. *PaPaS Author and Referee Guidance: Supplementary Information with Evidence About Pain Studies and Their Outcomes Over the Past Two Decades*; 2011. http://papas.cochrane.org/sites/papas.cochrane.org/files/uploads/PaPaSAuthorandRefereeGuidance.pdf [accessed 19 May 2017].
38 Higgins J, Green SE. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). The Cochrane Collaboration; 2011. www.cochrane-handbook.org [accessed 19 May 2017].
39 Robinson LR, Czerniecki JM, Ehde DM, Edwards WT, Judish DA, Goldberg ML *et al.* Trial of amitriptyline for relief of pain in amputees: results of a randomized controlled study. *Arch Phys Med Rehabil* 2004; **85**: 1–6.
40 Wilder-Smith CH, Hill LT, Laurent S. Postamputation pain and sensory changes in treatment-naive patients: characteristics and responses to treatment with tramadol, amitriptyline, and placebo. *Anesthesiology* 2005; **103**: 619–628.
41 Tasmuth T, Härtel B, Kalso E. Venlafaxine in neuropathic pain following treatment of breast cancer. *Eur J Pain* 2002; **6**: 17–24.
42 Kalso E, Tasmuth T, Neuvonen Pertti J. Amitriptyline effectively relieves neuropathic pain following treatment of breast cancer. *Pain* 1995; **64**: 293–302.
43 Smith DG, Ehde DM, Hanley MA, Campbell KM, Jensen MP, Hoffman AJ *et al.* Efficacy of gabapentin in treating chronic phantom limb and residual limb pain. *J Rehabil Res Dev* 2005; **42**: 645–654.
44 Bone M, Critchley P, Buggy DJ. Gabapentin in postamputation phantom limb pain: a randomized, double-blind, placebo-controlled, cross-over study. *Reg Anesth Pain Med* 2002; **27**: 481–486.
45 Hoseinzade H, Mahmoodpoor A, Agamohammadi D, Sanaie S. Comparing the effect of stellate ganglion block and gabapentin on the post mastectomy pain syndrome. *Rawal Med J* 2008; **33**: 21–24.
46 Biyik I, Gülcüler M, Karabiga M, Ergene O, Tayyar N. Efficacy of gabapentin *versus* diclofenac in the treatment of chest pain and paresthesia in patients with sternotomy. *Anadolu Kardiyol Derg* 2009; **9**: 390–396.
47 Zencirci B. Analgesic efficacy of oral gabapentin added to standard epidural corticosteroids in patients with failed back surgery. *Clin Pharmacol* 2010; **2**: 207–211.
48 Khosravi MB, Azemati S, Sahmeddini MA. Gabapentin *versus* naproxen in the management of failed back surgery syndrome; a randomized controlled trial. *Acta Anaesthesiol Belg* 2014; **65**: 31–37.
49 Vilholm OJ, Cold S, Rasmussen L, Sindrup SH. Effect of levetiracetam on the postmastectomy pain syndrome. *Eur J Neurol* 2008; **15**: 851–857.
50 Silverman A, Samuels Q, Gikas H, Nawras A. Pregabalin for the treatment of abdominal adhesion pain: a randomized, double-blind, placebo-controlled trial. *Am J Ther* 2012; **19**: 419–428.
51 Bischoff JM, Ringsted TK, Petersen M, Sommer C, Uçeyler N, Werner MU. A capsaicin (8%) patch in the treatment of severe persistent inguinal postherniorrhaphy pain: a randomized, double-blind, placebo-controlled trial. *PLoS ONE* 2014; **9**: e109144.
52 Watson CPN, Evans RJ. The postmastectomy pain syndrome and topical capsaicin: a randomized trial. *Pain* 1992; **51**: 375–379.
53 Ellison N, Loprinzi CL, Kugler J, Hatfield AK, Miser A, Sloan JA *et al.* Phase III placebo-controlled trial of capsaicin cream in the management of surgical neuropathic pain in cancer patients. *J Clin Oncol* 1997; **15**: 2974–2980.
54 Manchikanti L, Singh V, Cash KA, Pampati V, Datta S. A comparative effectiveness evaluation of percutaneous adhesiolysis and epidural steroid injections in managing lumbar post surgery syndrome: a randomized, equivalence controlled trial. *Pain Physician* 2009; **12**: E355–E368.
55 Manchikanti L, Malla Y, Cash KA, McManus CD, Pampati V. Fluoroscopic cervical interlaminar epidural injections in managing chronic pain of cervical postsurgery syndrome: preliminary results of a randomized, double-blind, active control trial. *Pain Physician* 2012; **15**: 13–25.

56 Manchikanti L, Singh V, Cash KA, Pampati V, Datta S. Fluoroscopic caudal epidural injections in managing post lumbar surgery syndrome: two-year results of a randomized, double-blind, active-control trial. *Int J Med Sci* 2012; **9**: 582–591.

57 Meadeb J, Rozenberg S, Duquesnoy B, Kuntz JL, Le Loët X, Sebert JL *et al.* Forceful sacrococcygeal injections in the treatment of postdiscectomy sciatica. A controlled study *versus* glucocorticoid injections. *Joint Bone Spine* 2001; **68**: 43–49.

58 Revel M, Auleley GR, Alaoui S, Nguyen M, Duruoz T, Eck-Michaud S *et al.* Forceful epidural injections for the treatment of lumbosciatic pain with post-operative lumbar spinal fibrosis. *Rev Rhum Engl Ed* 1996; **63**: 270–277.

59 Rocco AG, Frank E, Kaul AF, Lipson SJ, Gallo JP. Epidural steroids, epidural morphine and epidural steroids combined with morphine in the treatment of post-laminectomy syndrome. *Pain* 1989; **36**: 297–303.

60 Yousef AA, EL-Deen AS, Al-Deeb AE. The role of adding hyaluronidase to fluoroscopically guided caudal steroid and hypertonic saline injection in patients with failed back surgery syndrome: a prospective, double-blinded, randomized study. *Pain Pract* 2010; **10**: 548–553.

61 Kim SB, Lee KW, Lee JH, Kim MA, An BW. The effect of hyaluronidase in interlaminar lumbar epidural injection for failed back surgery syndrome. *Ann Rehabil Med* 2012; **36**: 466–473.

62 Rahimzadeh P, Sharma V, Imani F, Faiz HR, Ghodraty MR, Nikzad-Jamnani AR *et al.* Adjuvant hyaluronidase to epidural steroid improves the quality of analgesia in failed back surgery syndrome: a prospective randomized clinical trial. *Pain Physician* 2014; **17**: E75–E82.

63 Aldrete JA. Epidural injections of indomethacin for postlaminectomy syndrome: a preliminary report. *Anesth Analg* 2003; **96**: 463–468.

64 Chun-jing H, Hao-xiong N, jia-xiang N. The application of percutaneous lysis of epidural adhesions in patients with failed back surgery syndrome. *Acta Cir Bras* 2012; **27**: 357–362.

65 Bischoff JM, Koscielniak-Nielsen ZJ, Kehlet H, Werner MU. Ultrasound-guided ilioinguinal/iliohypogastric nerve blocks for persistent inguinal postherniorrhaphy pain: a randomized, double-blind, placebo-controlled, crossover trial. *Anesth Analg* 2012; **114**: 1323–1329.

66 Bischoff JM, Petersen M, Uçeyler N, Sommer C, Kehlet H, Werner MU. Lidocaine patch (5%) in treatment of persistent inguinal postherniorrhaphy pain: a randomized, double-blind, placebo-controlled, crossover trial. *Anesthesiology* 2013; **119**: 1444–1452.

67 Park CH, Jung SH, Han CG. Effect of intravenous lidocaine on the neuropathic pain of failed back surgery syndrome. *Korean J Pain* 2012; **25**: 94–98.

68 Fredman B, Zohar E, Ben Nun M, Iraqi R, Jedeikin R, Gepstein R. The effect of repeated epidural sympathetic nerve block on 'failed back surgery syndrome' associated chronic low back pain. *J Clin Anesth* 1999; **11**: 46–51.

69 Kvarnström A, Karlsten R, Quiding H, Emanuelsson BM, Gordh T. The effectiveness of intravenous ketamine and lidocaine on peripheral neuropathic pain. *Acta Anaesthesiol Scand* 2003; **47**: 868–877.

70 Wu CL, Tella P, Staats PS, Vaslav R, Kazim DA, Wesselmann U *et al.* Analgesic effects of intravenous lidocaine and morphine on postamputation pain: a randomized double-blind, active placebo-controlled, crossover trial. *Anesthesiology* 2002; **96**: 841–848.

71 Wu CL, Agarwal S, Tella PK, Klick B, Clark MR, Haythornthwaite JA *et al.* Morphine *versus* mexiletine for treatment of postamputation pain: a randomized, placebo-controlled, crossover trial. *Anesthesiology* 2008; **109**: 289–296.

72 Casale R, Ceccherelli F, Labeeb AA, Biella GE. Phantom limb pain relief by contralateral myofascial injection with local anaesthetic in a placebo-controlled study: preliminary results. *J Rehabil Med* 2009; **41**: 418–422.

73 Ilfeld BM, Moeller-Bertram T, Hanling SR, Tokarz K, Mariano ER, Loland VJ *et al.* Treating intractable phantom limb pain with ambulatory continuous peripheral nerve blocks: a pilot study. *Pain Med* 2013; **14**: 935–942.

74 Liu CJ, Cai HF, Liu DC, Liu YQ, Sun ZG, Li N *et al.* [Efficacy of ultrasound-stellate ganglion block in breast cancer with postoperative neuropathic pain.] *Chin J Contemp Neurol Neurosurg* 2013; **13**: 872–875.

75 Singh JA, Mahowald ML, Noorbaloochi S. Intraarticular botulinum toxin A for refractory painful total knee arthroplasty: a randomized controlled trial. *J Rheumatol* 2010; **37**: 2377–2386 [Erratum in: *J Rheumatol* 2011; **38**: 1534].

76 Wittekindt C, Liu WC, Preuss SF, Guntinas-Lichius O. Botulinum toxin A for neuropathic pain after neck dissection: a dose-finding study. *Laryngoscope* 2006; **116**: 1168–1171.

77 Wu H, Sultana R, Taylor KB, Szabo A. A prospective randomized double-blinded pilot study to examine the effect of botulinum toxin type A injection *versus* Lidocaine/Depomedrol injection on residual and phantom limb pain: initial report. *Clin J Pain* 2012; **28**: 108–112.

78 Eichenberger U, Neff F, Sveticic G, Björgo S, Petersen-Felix S, Arendt-Nielsen L *et al.* Chronic phantom limb pain: the effects of calcitonin, ketamine, and their combination on pain and sensory thresholds. *Anesth Analg* 2008; **106**: 1265–1273.

79 Maier C, Dertwinkel R, Mansourian N, Hosbach I, Schwenkreis P, Senne I *et al.* Efficacy of the NMDA-receptor antagonist memantine in patients with chronic phantom limb pain – results of a randomized double-blinded, placebo-controlled trial. *Pain* 2003; **103**: 277–283.

80 Wiech K, Kiefer RT, Töpfner S, Preissl H, Braun C, Unertl K *et al.* A placebo-controlled randomized crossover trial of the N-methyl-D-aspartic acid receptor antagonist, memantine, in patients with chronic phantom limb pain. *Anesth Analg* 2004; **98**: 408–413, table of contents.

81 Schwenkreis P, Maier C, Pleger B, Mansourian N, Dertwinkel R, Malin JP *et al.* NMDA-mediated mechanisms in cortical excitability changes after limb amputation. *Acta Neurol Scand* 2003; **108**: 179–184.

82 Nikolajsen L, Gottrup H, Kristensen AG, Jensen TS. Memantine (a *N*-methyl-D-aspartate receptor antagonist) in the treatment of neuropathic pain after amputation or surgery: a randomized, double-blinded, cross-over study. *Anesth Analg* 2000; **91**: 960–966.

83 Nikolajsen L, Hansen CL, Nielsen J, Keller J, Arendt-Nielsen L, Jensen TS. The effect of ketamine on phantom pain: a central neuropathic disorder maintained by peripheral input. *Pain* 1996; **67**: 69–77.

84 Huse E, Larbig W, Flor H, Birbaumer N. The effect of opioids on phantom limb pain and cortical reorganization. *Pain* 2001; **90**: 47–55.

85 Patarica-Huber E, Boskov N, Pjevic M. Multimodal approach to therapy-related neuropathic pain in breast cancer. *J BUON* 2011; **16**: 40–45.

86 Block L, Lundborg C, Bjersing J, Dahm P, Hansson E, Biber B. Ultralow dose of naloxone as an adjuvant to intrathecal morphine infusion improves perceived quality of sleep but fails to alter persistent pain: a randomized, double-blind, controlled study. *Clin J Pain* 2015; **31**: 968–975.

87 Arias-Buría JL, Valero-Alcaide R, Cleland JA, Salom-Moreno J, Ortega-Santiago R, Atín-Arratibel MA *et al.* Inclusion of trigger point dry needling in a multimodal physical therapy program for postoperative shoulder pain: a randomized clinical trial. *J Manipulative Physiol Ther* 2015; **38**: 179–187.

88 Pfister DG, Cassileth BR, Deng GE, Yeung KS, Lee JS, Garrity D *et al.* Acupuncture for pain and dysfunction after neck dissection: results of a randomized controlled trial. *J Clin Oncol* 2010; **28**: 2565–2570.

89 Manniche C, Asmussen K, Lauritsen B, Vinterberg H, Karbo H, Abildstrup S *et al.* Intensive dynamic back exercises with or without hyperextension in chronic back pain after surgery for lumbar disc protrusion. A clinical trial. *Spine (Phila Pa 1976)* 1993; **18**: 560–567.

90 Brox JI, Reikerås O, Nygaard Ø, Sørensen R, Indahl A, Holm I *et al.* Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. *Pain* 2006; **122**: 145–155.

91 Brunelli S, Morone G, Iosa M, Ciotti C, De Giorgi R, Foti C *et al.* Efficacy of progressive muscle relaxation, mental imagery, and phantom exercise training on phantom limb: a randomized controlled trial. *Arch Phys Med Rehabil* 2015; **96**: 181–187.

92 Timm KE. A randomized-control study of active and passive treatments for chronic low back pain following L5 laminectomy. *J Orthop Sports Phys Ther* 1994; **20**: 276–286.

93 Hsiao AF, York R, Hsiao I, Hansen E, Hays RD, Ives J *et al.* A randomized controlled study to evaluate the efficacy of noninvasive limb cover for chronic phantom limb pain among veteran amputees. *Arch Phys Med Rehabil* 2012; **93**: 617–622.

94 Kern U, Altkemper B, Kohl M. Management of phantom pain with a textile, electromagnetically-acting stump liner: a randomized, double-blind, crossover study. *J Pain Symptom Manage* 2006; **32**: 352–360.

95 North RB, Kidd DH, Farrokhi F, Piantadosi SA. Spinal cord stimulation *versus* repeated lumbosacral spine surgery for chronic pain: a randomized, controlled trial. *Neurosurgery* 2005; **56**: 98–106.

96 Van Gorp EJ, Teernstra OPM, Gültuna I, Hamm-Faber T, Bürger K, Schapendonk R *et al.* Subcutaneous stimulation as ADD-ON therapy to spinal cord stimulation is effective in treating low back pain in patients with failed back surgery syndrome: a multicenter randomized controlled trial. *Neuromodulation* 2016; **19**: 171–178.

97 Kumar K, Taylor RS, Jacques L, Eldabe S, Meglio M, Molet J *et al.* Spinal cord stimulation *versus* conventional medical management for neuropathic pain: a multicentre randomised controlled trial in patients with failed back surgery syndrome. *Pain* 2007; **132**: 179–188.

98 Schu S, Slotty PJ, Bara G, Knop M, Edgar D, Vesper J. A prospective, randomised, double-blind, placebo-controlled study to examine the effectiveness of burst spinal cord stimulation patterns for the treatment of failed back surgery syndrome. *Neuromodulation* 2014; **17**: 443–450.

99 Van Havenbergh T, Vancamp T, Van Looy P, Vanneste S, De Ridder D. Spinal cord stimulation for the treatment of chronic back pain patients: 500-Hz *vs.* 1000-Hz burst stimulation. *Neuromodulation* 2015; **18**: 9–12.

100 Weintraub MI, Steinberg RB, Cole S. The role of cutaneous magnetic stimulation in failed back syndrome. *Semin Integr Med* 2005; **3**: 101–103.

101 Ebid AA, El-Sodany AM. Long-term effect of pulsed high-intensity laser therapy in the treatment of post-mastectomy pain syndrome: a double blind, placebo-control, randomized study. *Lasers Med Sci* 2015; **30**: 1747–1755.

102 Flor H, Denke C, Schaefer M, Grüsser S. Effect of sensory discrimination training on cortical reorganisation and phantom limb pain. *Lancet* 2001; **357**: 1763–1764.

103 Esmer G, Blum J, Rulf J, Pier J. Mindfulness-based stress reduction for failed back surgery syndrome: a randomized controlled trial. *J Am Osteopath Assoc* 2010; **110**: 646–652.

104 Chan BL, Witt R, Charrow AP, Magee A, Howard R, Pasquina PF *et al.* Mirror therapy for phantom limb pain. *N Engl J Med* 2007; **357**: 2206–2207.

105 Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT *et al.* Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008; **9**: 105–121.

106 Moore A, Eccleston C, Derry S, Wiffen P, Bell RF, Straube S *et al.* 'Evidence' in chronic pain – establishing

best practice in the reporting of systematic reviews. *Pain* 2010; **150**: 386–389.
107 Derry S, Lloyd R, Moore RA, McQuay HJ. Topical capsaicin for chronic neuropathic pain in adults. *Cochrane Database Syst Rev* 2009; (4)CD007393.
108 Masselin-Dubois A, Attal N, Fletcher D, Jayr C, Albi A, Fermanian J *et al.* Are psychological predictors of chronic postsurgical pain dependent on the surgical model? A comparison of total knee arthroplasty and breast surgery for cancer. *J Pain* 2013; **14**: 854–864.
109 Chaparro LE, Smith SA, Moore RA, Wiffen PJ, Gilron I. Pharmacotherapy for the prevention of chronic pain after surgery in adults. *Cochrane Database Syst Rev* 2013; (7)CD008307.
110 Humble SR, Dalton AJ, Li L. A systematic review of therapeutic interventions to reduce acute and chronic post-surgical pain after amputation, thoracotomy or mastectomy. *Eur J Pain* 2015; **19**: 451–465.
111 McNicol ED, Schumann R, Haroutounian S. A systematic review and meta-analysis of ketamine for the prevention of persistent post-surgical pain. *Acta Anaesthesiol Scand* 2014; **58**: 1199–1213.
112 Clarke H, Bonin RP, Orser BA, Englesakis M, Wijeysundera DN, Katz J. The prevention of chronic postsurgical pain using gabapentin and pregabalin: a combined systematic review and meta-analysis. *Anesth Analg* 2012; **115**: 428–442.
113 Moore RA, Chi C-C, Wiffen PJ, Derry S, Rice ASC. Oral nonsteroidal anti-inflammatory drugs for neuropathic pain. *Cochrane Database Syst Rev* 2015; (10)CD010902.
114 Moore RA, Derry S, Aldington D, Cole P, Wiffen PJ. Amitriptyline for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015; (7)CD008242.
115 Finnerup NB, Attal N, Haroutounian S, McNicol E, Baron R, Dworkin RH *et al.* Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. *Lancet Neurol* 2015; **14**: 162–173.
116 Wiffen PJ, Derry S, Moore RA, Stannard C, Aldington D, Cole P *et al.* Buprenorphine for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015; (9)CD011603.
117 Edwards P, Clarke M, DiGuiseppi C, Pratap S, Roberts I, Wentz R. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat Med* 2002; **21**: 1635–1640.
118 Pool JL. Posterior cordotomy for relief of phantom limb pain. *Ann Surg* 1946; **124**: 386–391.
119 Blume H, Richardson R, Rojas C. Epidural nerve stimulation of the lower spinal cord and cauda equina for the relief of intractable pain in failed low back surgery. *Appl Neurophysiol* 1982; **45**: 456–460.
120 Turk DC, Wilson HD, Cahana A. Treatment of chronic non-cancer pain. *Lancet* 2011; **377**: 2226–2235.
121 Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c332.
122 Moher D, Schulz KF, Altman DG; CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001; **134**: 657–662.
123 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I *et al.* Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996; **276**: 637–639.
124 Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD *et al.*; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Appendix S1 Search terms (Word document)

Appendix S2 Risk of bias by intervention type (Word document)

Table S1 Characteristics of included studies evaluating pharmacological interventions (Word document)

Table S2 Characteristics of included studies evaluating physical, surgical, psychological and other interventions (Word document)

Fig. S1 Forest plot showing trials of gabapentin *versus* placebo for treatment of chronic phantom limb pain (Word document)

Fig. S2 Forest plot for trials of low-dose capsaicin *versus* placebo for treatment of chronic postsurgical pain after cancer surgery (Word document)

OPEN ACCESS

Should exercises be painful in the management of chronic musculoskeletal pain? A systematic review and meta-analysis

Benjamin E Smith,[1,2] Paul Hendrick,[3] Toby O Smith,[4] Marcus Bateman,[1] Fiona Moffatt,[3] Michael S Rathleff,[5,6] James Selfe,[7] Pip Logan[2]

[1]Department of Physiotherapy, Derby Teaching Hospitals NHS Foundation Trust, Derby, UK
[2]Division of Rehabilitation and Ageing, School of Medicine, University of Nottingham, Nottingham, UK
[3]Division of Physiotherapy and Rehabilitation Sciences, School of Health Sciences, University of Nottingham, Nottingham University Hospitals (City Campus), Nottingham, UK
[4]University of East Anglia, Norwich, UK
[5]Research Unit for General Practice in Aalborg, Department of Clinical Medicine at Aalborg University, Aalborg, Denmark
[6]Department of Occupational Therapy and Physiotherapy, Department of Clinical Medicine, Aalborg University Hospital, Aalborg, Denmark
[7]Manchester Metropolitan University, Manchester, UK

Correspondence to
Benjamin E Smith, Department of Physiotherapy, London Road Community Hospital, Derby Teaching Hospitals NHS Foundation Trust, Derby DE1 2QY, UK;
benjamin.smith3@nhs.net

Accepted 1 April 2017
Published Online First 8 June 2017

ABSTRACT

Background Chronic musculoskeletal disorders are a prevalent and costly global health issue. A new form of exercise therapy focused on loading and resistance programmes that temporarily aggravates a patient's pain has been proposed. The object of this review was to compare the effect of exercises where pain is allowed/encouraged compared with non-painful exercises on pain, function or disability in patients with chronic musculoskeletal pain within randomised controlled trials.

Methods Two authors independently selected studies and appraised risk of bias. Methodological quality was evaluated using the Cochrane risk of bias tool, and the Grading of Recommendations Assessment system was used to evaluate the quality of evidence.

Results The literature search identified 9081 potentially eligible studies. Nine papers (from seven trials) with 385 participants met the inclusion criteria. There was short- term significant difference in pain, with moderate quality evidence for a small effect size of −0.27 (−0.54 to −0.05) in favour of painful exercises. For pain in the medium and long term, and function and disability in the short, medium and long term, there was no significant difference.

Conclusion Protocols using painful exercises offer a small but significant benefit over pain-free exercises in the short term, with moderate quality of evidence. In the medium and long term there is no clear superiority of one treatment over another. Pain during therapeutic exercise for chronic musculoskeletal pain need not be a barrier to successful outcomes. Further research is warranted to fully evaluate the effectiveness of loading and resistance programmes into pain for chronic musculoskeletal disorders.

PROSPERO registration CRD42016038882.

BACKGROUND

Musculoskeletal disorders are one of the most prevalent and costly disorders globally.[1 2] Low back pain is considered the leading cause of years lived with disability worldwide, ahead of conditions such as depression, diabetes, cardiovascular disease and cancer, with a global point prevalence of 9.4%.[3 4] Neck pain and other musculoskeletal pain ranks fourth and sixth in terms of years lived with disability, with a global point prevalence of 5% and 8%, respectively.[5 6] In the UK, an estimated one in four people suffer from chronic musculoskeletal disorders,[7] with an estimated economic consequence of 8.8 million working days lost.[8]

Previous systematic reviews have assessed the effectiveness of various interventions for musculoskeletal disorders, including pharmaceutical therapies,[9–12] psychological-based therapies[13–16] and physical-based therapies, including manual therapy[17–19] and exercise.[16 20–24] These have all presented poor to moderate results in terms of effectiveness at improving pain and function, and have identified limitations in the quality of included trials when drawing conclusions.

There is a high level of uncertainty and lack of sufficient level 1 evidence on which to base treatment for people with musculoskeletal disorders. A systematic review of self-management interventions for chronic musculoskeletal pain concluded that strong evidence existed that changes in the psychological factors, self-efficacy and depression were predictors of outcomes, irrespective of the intervention delivered, and strong evidence existed that positive changes in patients' pain catastrophising and physical activity were mediating factors.[25] Experimental studies have also demonstrated that stimulus context and the emotional response to pain affect the experience of pain,[26–28] and have led to the development of desensitisation interventions for chronic musculoskeletal disorders.[29–31]

It has been proposed that modern treatment therapies for chronic musculoskeletal pain and disorders should be designed around loading and resistance programmes targeting movements and activities that can temporarily reproduce and aggravate patients' pain and symptoms.[31–33] Pain does not correlate with tissue damage,[34] and psychological factors such as catastrophising and fear avoidance behaviours play an important role in the shaping of the physiological responses to pain, and therefore the development and maintenance of chronic pain.[35] It is thought that such an exercise programme could facilitate the reconceptualisation of pain by addressing fear avoidance and catastrophising beliefs within a framework of 'hurt not equalling harm'.[36 37] Through this, proponents support the prescription of exercises into pain for chronic musculoskeletal pain and disorders.[31 37 38] We define 'exercise into pain' as a therapeutic exercise where pain is encouraged or allowed.

No previous systematic reviews have evaluated the effectiveness of exercises into pain for chronic musculoskeletal pain. Therefore the object of this review was to compare the effect of exercises into pain compared with non-painful exercises on pain, function or disability in patients with chronic musculoskeletal pain within randomised controlled

To cite: Smith BE, Hendrick P, Smith TO, *et al*. *Br J Sports Med* 2017;**51**:1679–1687.

Table 1 Search strategy

1	Randomised controlled trials as
2	Topic/
3	randomised controlled trial.pt
4	controlled clinical trial.pt
5	or/1-3
6	Exp Pain
7	Exp Musculoskeletal Disease
8	Exp Musculoskeletal Pain
9	Or/5-7
10	Rehabilitation
11	Bone
12	Joint
13	Muscle
14	Exp Exercise therapy
15	Physiotherapy
16	Physical therapy
17	Physical-therapy
18	Exp Exercise Or/9-17
19	(exercise adj7 pain$).af
20	High load
21	Loaded$
22	Resistance$
23	Eccentric$
24	Concentric$
25	Weight loaded
26	Weight-loaded
27	Weight resistance
28	Weight-resistance
29	High-load
30	Heavy load
31	Heavy-load
32	Direction$ preference
33	Directional-preference
34	Or/19-33
35	4 and 8 and 18 and 34 (limited to English)

trials (RCTs), specifically exercises that were prescribed with instructions for patients to experience pain, or where patients were told it was acceptable and safe to experience pain, and to compare any difference in contextual factors and prescription parameters of the prescribed exercise intervention.

METHODS

This systematic review followed the recommendations of the PRISMA statement,[39] and was registered with the International Prospective Register of Systematic Reviews (PROSPERO; http://www.crd.york.ac.uk/prospero/, reference CRD42016038882).

Search strategy

An electronic database search was conducted on titles and abstract from inception to October 2016 on the following databases: the Allied and Complimentary Medicine Database, the Cumulative Index to Nursing and Allied Health Literature, the Cochrane Library, Embase, Medline, SPORTDiscus and Web of Science. For the keywords and keywords search strategy used, please see table 1. The database searches were accompanied by hand searches of the reference list of included articles, and the grey literature and ongoing trials were searched using the following databases: Open Grey, WHO International Clinical Trials Registry Platform, ClinicalTrials.gov and the bjsports-2016-097383 portfolio.

For inclusion, the studies had to meet the following criteria: adults recruited from the general population with any musculoskeletal pain or disorder greater than 3 months; participants with pain suggestive of non-musculoskeletal pain, for example, headache, migraine, bowel/stomach pain, cancer, fibromyalgia, chest pain, and breathing difficulties were excluded. Studies had to have a primary treatment arm of therapeutic exercises that was advised to be purposively painful, or where pain was allowed or tolerated. The comparison group had to use therapeutic exercises that were pain-free. Included studies were required to report pain, disability or function. Studies had to be full RCT published in English. Studies that were not randomised or quasi-random were excluded.

Study selection

One reviewer (BES) undertook the searches. Titles and abstracts were screened by one reviewer (BES), with potential eligible papers retrieved and independently screened by two reviewers (BES and PH). Initial inclusion agreement was 81%, and using Cohen's statistic method the kappa agreement was $k=0.47$, which is considered 'fair to moderate' agreement.[40–42] All initial disagreements were due to intervention criteria, specifically the levels of pain during the therapeutic exercises in each intervention arm,[43–50] and were resolved through consensus. Three trials needed further information with regard to their control exercise to ascertain if they met the inclusion criteria, and all three were contacted.[51–53] All three responded with further information, and after discussion there was consensus to include two of the three trials.[51 52]

Data extraction

The following data were extracted from the included articles: trial design, participant information, intervention and control exercise, setting, follow-up periods and outcome data.[54] The data were independently extracted and transcribed to a standard table by one reviewer (BES), and then 25% of the data were independently checked by a second reviewer (PH). Effectiveness was judged in the short term (≤3 months from randomisation), medium term (>3 and<12 months) and long term (≥12 months), as recommended by the 2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group.[55]

Quality assessment

Each included study was appraised independently by two reviewers (BES and PH) for methodological quality using the Cochrane risk of bias tool for randomised clinical trials.[56] The tool was originally developed in 2008, and updated in 2011, and is based on seven key bias domains[57]: sequence generation and allocation concealment (both within the domain of selection bias or allocation bias), blinding of participants and personnel (performance bias), blinding of outcome assessors (detection bias), incomplete outcome data (attrition bias) and selective reporting (reporting bias).[56] For each domain the reviewers judged the risk of bias as 'high', 'low' or 'unclear'. Percentage agreement between the two reviewers for the individual risk of bias domains for the Cochrane risk of bias tool was 86%, with a kappa of $\kappa=0.76$, which is considered 'substantial or good',[40–42] and disagreements were resolved through consensus.

We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system to rate the overall

quality of the body of evidence in each pooled analysis.[58] We did not evaluate the publication bias domain in this review as it is not recommended to assess funnel plot asymmetry with a meta-analysis of fewer than 10 trials.[59] A GRADE profile was completed for each pooled estimate. Where only single trials were available, evidence from studies with <400 participants was downgraded for inconsistency and imprecision and rated as low-quality evidence. Three reviewers assessed these factors for each outcome and agreed by consensus (BES, PH and TOS).

The quality of evidence was defined as the following: (1) high quality—further research is unlikely to change our confidence in the estimate of effect; the Cochrane risk of bias tool identified no risks of bias and all domains in the GRADE classification were fulfilled; (2) moderate quality—further research is likely to have an important impact on our confidence in the estimate of effect, and one of the domains in the GRADE classification was not fulfilled; (3) low quality—further research is likely to have an important impact on our confidence and is likely to change the estimate; two of the domains were not fulfilled in the GRADE classification; and (4) very low quality—we are uncertain about the estimate; three of the domains in the GRADE classification were not fulfilled.[60 61]

Statistical analysis

Clinical heterogeneity was assessed through visual examination of the data extraction table on details related to participant characteristics, intervention, study design and process in the included studies. Based on this assessment, the reviewers judged there to be low clinical heterogeneity and accordingly it was appropriate to perform a meta-analysis where feasible. The primary outcome was a measure of pain, disability or function. As pain scores were reported on different scales, we used the standardised mean difference (SMD).[62] We a priori defined effect size interpretation as 0.2 for a 'small' effect size, 0.5 for a 'medium' effect size and 0.8 for a 'large' effect size, as suggested by Cohen (1988).[63] If data were not available, the associated corresponding author was contacted. Failing this, the mean and SD were estimated, assuming normal distribution, from medians and IQRs.[64] Statistical between-study heterogeneity was assessed with the $I^2$2 statistic. We considered 0%–25% as low, 26%–74% moderate and 75% and over as high statistical heterogeneity.[65] When outcomes presented with low statistical heterogeneity, data were pooled using a fixed-effects model.[66] When analyses presented with moderate or high statistical heterogeneity, a DerSimonian and Laird random-effects model was adopted.[67]

All data analyses were performed using the OpenMetaAnalyst software.[68]

Sensitivity analysis

A sensitivity analysis was performed for the primary and secondary analyses using only trials that presented with a low risk of bias.[56] In addition we carried out a sensitivity analysis to assess the impact of studies where mean and SD were estimated from medians and IQRs, and outcome measures of pain were pooled scores set within pain domains from patient-reported outcome measures, for example, the Shoulder Pain and Disability Index (SPADI).[69]

RESULTS

Study identification

The search results are presented in figure 1. The database search produced 9081 results, with no additional findings from reference list searches or unpublished searches. After duplicates were removed, 37 papers were appropriate for full-text review.

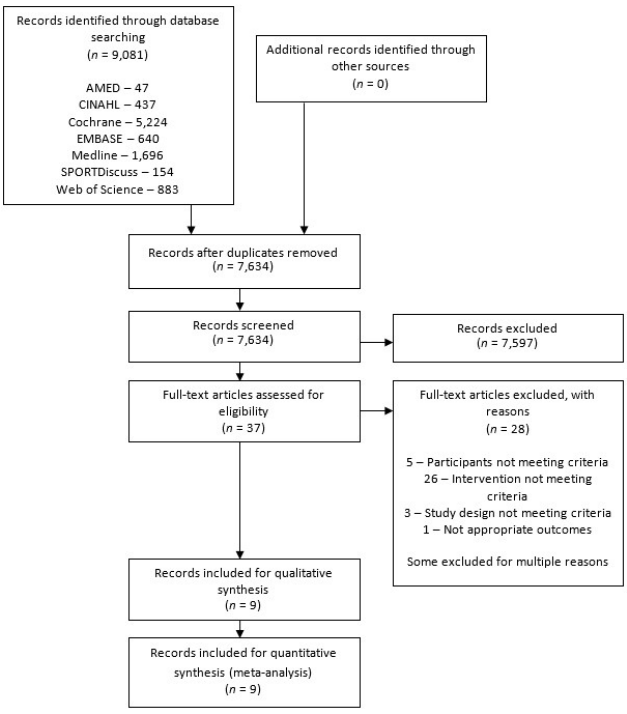


Figure 1 PRISMA 2009 flow diagram.

After full-text review, 28 articles were excluded, 5 were due to participants not meeting the criteria, 26 because the intervention did not meet the criteria, 3 because of study design not meeting criteria, and 1 due to inappropriate outcome measures. Some articles were excluded for multiple reasons. Therefore nine articles were included in the final review. Of the included articles, there were two occurrences of the same trial reporting different time points over two publications.[43 70–72]

Characteristics of included trials

A summary of the characteristics and main findings of the included trials can be found in table 2.

The two occurrences of the same trial reporting different time points over two articles were analysed as single trials to prevent multiplicity in analyses.[43 70–72] All trials investigated home-based exercises, had a roughly even composition of women and men (46% women), with similar mean ages of participants (mean age 47, range 19–83). One trial included low back pain,[43 72] three included shoulder pain,[47 52 70 71] two included Achilles pain[73 74] and one included plantar heel pain.[51]

Three trials used a Visual Analogue Scale to measure pain,[43 70–72 74] two trials used the SPADI,[47 52] one used the Knee Injury and Osteoarthritis Outcome Score (KOOS),[73] and one used the Foot Function Index (FFI) including pain at worse and pain on first step on a numerical rating scale (0–10).[51]

Where pain outcomes were included within patient-reported outcome measures, these data were extracted.[47 52 73] Two trials that used the SPADI had insufficient data in the publication to complete a meta-analysis for pain,[47 52] and both were contacted and asked to supply pain domain data. Littlewood *et al*[52] replied and provided all the available data; however, Maenhout *et al*[47] did not respond. One trial reported outcomes in medians and IQRs,[74] and was contacted and asked for further data. They

Table 2 Characteristics of included trials

Study characteristics	Participant characteristics	Intervention and setting	Outcome data/results
Aasa *et al* (2015)[43] Michaelson *et al* (2016)[72] 2 groups: 1. High-load lifting exercise 2. Low-load motor control exercises	70 patients recruited from occupational healthcare services in Sweden (mean age 42, 56% female); inclusion criteria included: (a) adults with low back pain >3 months' duration and (b) with or without leg pain	Physiotherapy clinic, sports centre and home setting 1. n=35; group exercises based at a sports centre (5 participants in each group), with pain up to 50 mm Visual Analogue Scale acceptable, such that the pain subsided after each set of exercises; 12 treatment sessions over an 8-week period (weeks 1–4, 2 sessions per week; weeks 5–8, 1 session per week); 60 min in duration; no home exercises 2. n=35; pain-free individual exercises at a physiotherapy centre; 12 treatment sessions over an 8-week period (weeks 1–4, 2 sessions per week; weeks 5–8, 1 session per week); 20–30 min in duration; exercises involved improving control around joint neutral positions; in supine, four-point kneeling, sitting, and/or standing positions; Plus home exercises, 10 repetitions 2–3× a day	Main outcome assessed at baseline, 2-month and 12-month follow-up was 7 day average pain on a Visual Analogue Scale (0–100 mm) and Roland-Morris Disability Questionnaire (0–24) Group 1: mean pain at baseline 43 (SD 24), 2 months 22 (SD 21), 12 months 24 (SD 27) and 24 months 27 (SD 27) Group 2: mean pain at baseline 47 (SD 28), 2 months 30 (SD 26), 12 months 25 (SD 22) and 24 months 30 (SD 29) Group 1: mean disability at baseline 7.2 (SD 4.3), 2 months 3.8 (SD 4.0), 12 months 3.6 (SD 4.2) and 24 months 3.8 (SD 3.9) Group 2: mean disability at baseline 7.1 (SD 3.9), 2 months 3.6 (SD 4.2), 12 months 3.3 (SD 3.6) and 24 months 3.6 (SD 3.7) Both groups had significant improvements in their pain and disability levels; no significant between-group difference for pain at any follow-up (2 months p=0.71; 12 months p=0.94; 24 months p=0.89); no significant between-group difference for disability at any follow-up (2 months p=0.77; 12 months p=0.74; 24 months p=0.99)
Holmgren *et al* (2012)[70] Hallgren *et al* (2014)[71] 2 groups: 1. Specific exercises group 2. Control exercise group Patients were given the option at 3 months of continuing to have an arthroscopic subacromial decompression.	97 patients recruited from the waiting list for an arthroscopic subacromial decompression from a university hospital in Sweden (mean age 52, 37% female); inclusion criteria included (a) adults with lateral shoulder pain >6 months, (b) failed 3 months of previous primary care, (c) signs of impingement symptoms and (d) positive Neer's impingement test of a subacromial anaesthetic injection	Physiotherapy and home setting 1. n=51; eccentric rotator cuff exercises and concentric/eccentric scapula exercises; recommendation of 5/10 numerical rating scale for pain during exercises, such that the pain subsided by the next exercise session; 7 physiotherapy appointment, weekly first 2 weeks, alternative weeks thereafter; exercises to be performed at home once or twice a day for 12 weeks 2. n=46; pain-free upper limb and neck exercises; 7 physiotherapy appointment, weekly first 2 weeks, alternative weeks thereafter; exercises to be performed at home once or twice a day for 12 weeks	Main outcome of Constant-Murley score (C-M) (0–100), along with shoulder assessment scores and pain scores taken at baseline, 3 months and 12 months, including pain at rest measured on Visual Analogue Scale (0–100 mm) Group 1: mean C-M at baseline 48 (SD 15), 3 months 72 (SD 19) and 12 months 83 (SD 14) Group 2: mean C-M at baseline 43 (SD 15), 3 months 52 (SD 23) and 12 months 76 (SD 18) Group 1: mean pain at rest at baseline 15 (SD 19), 3 months 10 (SD 14) and 12 months 2 (SD 6) Group 2: mean pain at rest at baseline 20 (SD 21), 3 months 20 (SD 25) and 12 months 4 (SD 13) Both groups had significant improvements in all outcomes at 3 months and 1 year follow-up. Significantly more patients in the control group decided to have surgery (63%) than those in the specific exercise group (24%; p<0.0001).
Littlewood *et al* (2015)[52] 2 groups: 1. Self-managed exercises 2. Usual physiotherapy	86 patients recruited from UK, NHS physiotherapy waiting list (mean age 55, 50% female); inclusion criteria included (a) adults with shoulder pain >3 months, (b) maintained shoulder ROM and (c) pain with resisted movements	Physiotherapy and home setting 1. n=42; single shoulder exercise guided by the symptomatic response, requiring pain to be produced during exercise, such that the pain subsided after the exercises; typically involving a weighted shoulder abduction exercise of 3 sets of 10–15 repetitions; pragmatic approach to number of follow-ups, timings of appointments and point of discharge; that is, the treating physiotherapist and patient will determine these factors 2. n=44; usual physiotherapy,* including advice, stretching, exercise, manual therapy, massage, strapping, acupuncture, electrotherapy, corticosteroid injection at the discretion of the treating physiotherapist; pragmatic approach to number of follow-ups, timings of appointments and point of discharge; that is, the treating physiotherapist and patient will determine these factors	Main outcome of the Shoulder Pain and Disability Index (SPADI) (0–100) at baseline, 3, 6 and 12 months Group 1: mean at baseline 49.1 (SD 18.3), 3 months 32.4 (SD 20.2), 6 months 16.6 (SD 19.7) and 12 months 14.2 (SD 20.0) Group 2: mean at baseline 49.0 (SD 18.0), 3 months 30.7 (SD 19.7), 6 months 24.0 (SD 19.7) and 12 months 21.4 (SD 25.4) Statistically significant and clinically important within group changes for SPADI from baseline to all three follow-up points. There were no statistically significant differences between the groups across all the outcomes at 3, 6 or 12 months, (p=0.75, 0.19 and 0.32, respectively).

Continued

Table 2 Continued

Study characteristics	Participant characteristics	Intervention and setting	Outcome data/results
Maenhout *et al* (2013)[47] 2 groups: 1. Traditional rotator cuff training with heavy load eccentric training 2. Traditional rotator cuff training	61 patients recruited from a shoulder surgeon's clinic in Belgium (mean age 39.8, 41% female); inclusion criteria included (a) adults with >3 months of shoulder pain, (b) painful arc, (c) 2 out of 3 impingement tests, (d) pain on palpation of rotator cuff tendons	Physiotherapy and home setting 1. n=31; the same exercises as group 2, plus a heavy loaded eccentric exercise of abduction within the scapular plane; 3 sets of 15 repetitions, such that the patient experiences pain on the last set, up to 5/10 Visual Analogue Scale, such that the pain subsided by the following morning. 2. n=30; pain-free, traditional rotator cuff exercises of internal and external rotation with a resisted rubber band; performed once a day, with 3 sets of 10 repetitions; both groups had exercise prescription and monitoring through 9 physiotherapy appointments over 12 weeks	Main outcome of the SPADI (0–100) at baseline, 6 weeks and 12 weeks Group 1: mean at baseline 44.3 (SD 11.5), 6 weeks 17.7 (SD 12.0) and 12 weeks 14.5 (SD 11.7). Group 2: mean at baseline 42.0 (SD 11.0), 6 weeks 25.4 (SD 11.9) and 12 weeks 17.0 (SD 11.4) In both groups pain and function, measured with the SPADI score, improved significantly over time ($p>0.001$). When comparing between groups, improvement of the SPADI score was not significantly different.
Nørregaard *et al* (2007)[73] 2 groups: 1. Eccentric exercises 2. Stretching exercises	45 patients recruited from a clinic of sports medicine in Denmark (mean age 42, 49% female); inclusion criteria included (a) adults with Achilles pain >3 months, (b) local thickening >2 mm on ultrasound, (c) diffuse posterior ankle pain	Sports medicine clinic and home setting 1. n=21; information leaflet with home exercise programme on; to be performed twice a day, for 12 weeks; 1 follow-up appointment at 3 months; 3 sets of 15 repetitions of eccentric calf exercises, with knee straight and semi-flexed; patients told to expect pain during the exercises, but to avoid increasing daily pain or morning stiffness 2. n=24; information leaflet with home exercise programme on; to be performed twice a day, for 12 weeks; 1 follow-up appointment at 3 months; pain-free standing stretches for gastrocnemius and soleus; 5 repetitions of 30 s each	Outcome measures were tenderness on palpation, ultrasound and pain, as measured by the Knee Injury and Osteoarthritis Outcome Score (KOOS) (0–4) and patient's global assessment; follow-up was at baseline, 3, 6, 9, 12 weeks and 1 year Group 1: mean pain domain from KOOS at baseline 1.6 (SD 0.6), 3 weeks 0.1 (SD 0.1), 6 weeks 0.3 (SD 0.1), 9 weeks 0.4 (SD 0.2), 12 weeks 0.4 (SD 0.2) and 1 year score was 1.0 (SD 0.2) Group 2: mean pain domain from KOOS at baseline 1.6 (SD 0.6), 3 weeks 0.2 (SD 0.1), 6 weeks 0.3 (SD 0.1), 9 weeks 0.3 (SD 0.2), 12 weeks 0.4 (SD 0.2) and 1 year score was 0.7 (SD 0.2) There were significant improvements in all dimensions of the KOOS compared with baseline, with no differences between group differences.
Rathleff *et al* (2015)[51] 2 groups: 1. High-load strengthening exercises 2. Stretching exercises	48 patients recruited from a university hospital, regional hospital and private clinic in Denmark (mean age 46, 66% female); inclusion criteria included (a) adults with plantar fasciitis >3 months, (b) pain on palpation, (c) local thickening >4 mm on ultrasound	Home based exercises 1. n=24; information leaflet, heel inserts and a prescription of a high-load strength programme; consisting of single calf raises with a towel rolled up under the toes for maximum toe extension, activating the windlass mechanism; each calf raises was 3 s up, 2 s pause, 3 s down; weight was added in rucksacks, starting at 12 repetition maximum for three sets, and slowly progressed over 3 months; patients were advised to perform the exercise every other day; exercises were allowed to be painful, with no postincrease in pain. 2. n=24; information leaflet, heel inserts and a prescription of pain-free* plantar-specific stretches; patients were asked to stretch the plantar fascia in a cross-legged position by extending their toes, hold for 10 s, 10 times, 3× a day for 3 months	Primary outcome was Foot Function Index at 1, 3, 6 and 12 months, including pain at worse and pain on first step on a numerical rating scale (0–10). Mean scores for group 1 pain at worse at baseline was 7.9 (SD 1.7), 1 month 6.1 (95% CI 5.1 to 7.2), 3 months 3.5 (95% CI 2.3 to 4.7), 6 months 2.5 (95% CI 1.4 to 3.6) and 12 months 2.9 (95% CI 1.7 to 4.0). Mean scores for group 2 pain at worse at baseline was 7.5 (SD 1.6), 1 month 6.1 (95% CI 5.2 to 7.1), 3 months 6.1 (95% CI 4.4 to 7.7), 6 months 3.4 (95% CI 2.0 to 4.7) and 12 months 1.8 (95% CI 0.7 to 3.0). At 3 months group 1 had significantly lower pain scores than group 2 ($p<0.05$). At months 1, 6 and 12, there was no significant difference between groups.
Silbernagel *et al* (2001)[74] 2 groups: 1. Eccentric exercises 2. Regular concentric/ eccentric exercises	40 patients recruited from mailings to hospitals, clinics and sports clubs in Sweden (mean age 45, 23% female); inclusion criteria included (a) adults with Achilles pain >3 months	Clinic and home setting 1. n=22; progressive eccentric exercise programme to be performed 2× a day, plus three sets of six different stretching exercises, 20 s each, as well as balance, toe/heel walking exercises; weekly physiotherapy contact for 12 weeks; pain was allowed during the exercises up to 5/10 Visual Analogue Scale, such that the pain subsided by the following morning with no morning stiffness 2. n=18. 3 x a day of regular concentric and eccentric calf strengthening, plus two sets of the stretching exercises from group 1. Physiotherapy contacts 3–5 x during the 12 weeks. Exercises must be pain-free.	Outcomes of pain on palpation (Visual Analogue Scale) (0–100 mm) taken at baseline, 6 weeks, 3 and 6 months. Other outcomes included pain on walking and pain on stairs (yes/no), various objective measures, plus a non-validated functional questionnaire Median±IQR scores for pain on palpation for group 1 at baseline was 49±26.2, 6 weeks 40±27.5, 3 months 35±24.8 and 6 months 21±20. Median±IQR scores for group 2 at baseline was 27±21.5, 6 weeks 20±20, 3 months 31±26 and 6 months 9±17.5. There was a significant decrease in pain on palpation in both treatment groups; no significant differences between groups were seen.

*Information not within publication, authors contacted for clarification.
ROM, range of motion.

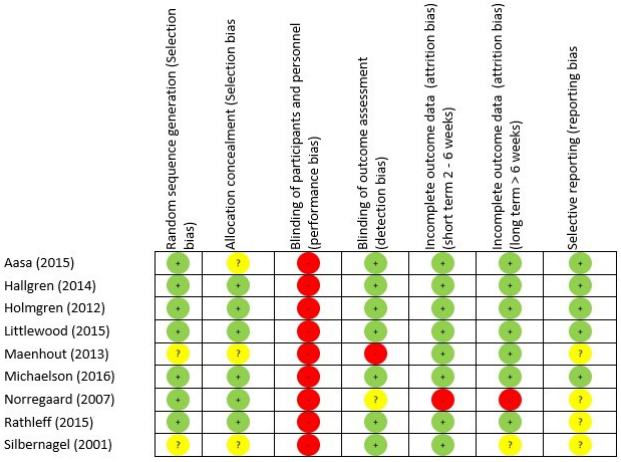


Figure 2 Risk of bias summary.

were unable to supply this, so the mean and SD were estimated assuming normal distribution.[64]

All seven trials recorded short-term follow-up of pain, four trials recorded medium-term follow-up of pain,[47 51 52 74] and five trials recorded long-term follow-up for pain.[43 51 52 70–73]

Trial quality and bias

The two papers reporting long-term outcomes for the trials that reported different time points made reference to the short-term outcome papers with regard to design parameters; therefore, trial quality and bias were assessed accordingly.[43 70–72]

No trial had greater than three 'high risk' of bias scores for a domain (figure 2).

The greatest risk of bias was with the blinding of participants and personnel (100%) (figure 3). The greatest amount of uncertainty was with regard to selective reporting bias, as many of the trials failed to include trials register details, or protocol details (44%).[47 51 73 74] Other common areas of bias with the included trials were with attrition bias, one trial failed to adequately describe attrition,[43] and two trials had large dropout rates[52 73]; however, Littlewood *et al*[52] received a 'low risk' score as their participant attrition was balanced across the intervention and control groups,[75] and an intention-to-treat analysis was performed. The risk of bias assessment tool highlights common trial write-up errors, with a number of papers failing to give an appropriate level of detail to adequately assess selection bias risk (33%).[43 47 74]

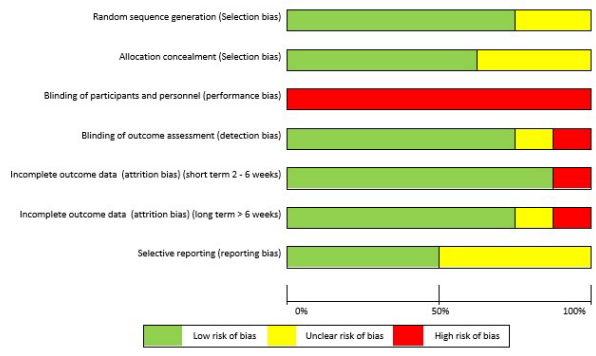


Figure 3 Risk of bias graph.

Narrative synthesis of disability and function outcomes

Of the seven trials, six reported some form of patient-reported outcome measure of disability or function. One reported Roland-Morris Disability Questionnaire,[43 72] one reported Constant-Murley and the Disabilities of the Arm Shoulder and Hand score,[70 71] two reported the SPADI,[47 52] one reported the KOOS,[73] and one reported the FFI.[51] With the exception of Rathleff *et al*,[51] there was clinically significant improvements in all outcomes, with no clear superiority. At 3-month follow-up for Rathleff *et al*,[51] the intervention group had a statistically significant lower FFI than the control group (p=0.016). At 1, 6 and 12 months, there were no differences between groups (p>0.34).

Contextual factors

With regard to the parameters of pain in the exercise intervention the participants were advised to adhere to, each trial gave different instructions, the key differences being if pain was allowed[43 51 72 74] or recommended.[47 52 70 71 73] In addition other differences were if an acceptable level of pain measured on a pain scale was advised,[47 70 71 74] and a time frame for the pain to subside by, for instance, if the pain had to subside immediately,[43 51 52 72] by the next session[70 71] or by the next day.[47 73 74] Clinically significant improvements in patient-reported outcome measures were reported across all interventions and control exercises, and all time points. It is not clear from the data if one approach was superior to the others.

Meta-analysis of pain

Short-term results

Six trials with 385 participants reported post-treatment effect on pain. Combining the results of these trials demonstrated significant benefit (SMD) of exercises into pain compared with pain-free exercises for musculoskeletal pain in the short term, with a small effect size of −0.28 (95% CI −0.49 to −0.08; figure 4). Statistical heterogeneity was negligible, I^2=0%. The quality of evidence (GRADE) was rated as 'low quality' due to trial design and low participant numbers (table 3).

For sensitivity analysis in the short term, we repeated the meta-analysis, removing two trials that used a patient-reported outcome measures index and had high dropout rates,[52 73] and the Silbernagel *et al*[74] trial where the mean and SD were estimated from medians and IQRs. The results of the data synthesis produced very similar results, with a small effect size of −0.27 (95% CI −0.54 to −0.05), with low statistical heterogeneity of I^2=22%. The quality of evidence (GRADE) was rated as 'moderate quality' due to low participant numbers (table 3).

Medium-term results

In the medium-term follow-up, meta-analysis demonstrated significant benefit (SMD) for exercises into pain compared with pain-free exercises for musculoskeletal pain, with a medium effect size of −0.59 (95% CI −1.03 to −0.15) (see figure 5). The statistical heterogeneity was moderate, I^2=50%. The quality of evidence (GRADE) was rated as 'low quality' due to trial design and low participant numbers (table 3).

Sensitivity analysis was not possible for medium-term results as two trials were excluded, one for using a patient-reported outcome measures index,[51] and one due to means and SD being estimated from medians and IQRs.[74] The one remaining trial showed no significant difference in the medium term.[51] The quality of evidence (GRADE) was rated as 'low quality' due to it being only from a single trial (table 3).

Smith BE, *et al*. *Br J Sports Med* 2017;**51**:1679–1687. doi:10.1136/bjsports-2016-097383

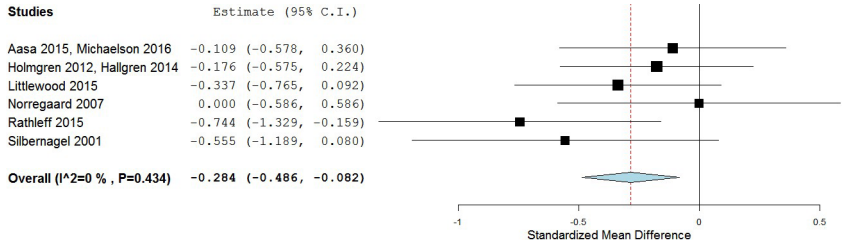


Figure 4 Forest plot of exercises into pain versus pain-free exercises—short term. Negative values favour painful intervention, whereas positive favour pain-free.

Long-term results

In the long term follow-up, meta-analysis demonstrated no statistical difference between exercises into pain and pain-free exercises, with an effect size of 0.01 (95% CI −0.39 to 0.41) (figure 6). The statistical heterogeneity was high, I^2=70%. The quality of evidence (GRADE) was rated as 'very low quality' due to trial design, heterogeneity and low participant numbers (table 3).

For sensitivity analysis in the long term, we repeated the meta-analysis, removing the two trials that used a patient-reported outcome measures index.[52 73] The results of the data synthesis found no statistical difference between exercises into pain and pain-free exercises, with an effect size of 0.13 (95% CI −0.14 to 0.40). The statistical heterogeneity was negligible, I^2=0%. The quality of evidence (GRADE) was rated as 'moderate quality' due to low participant numbers (table 3).

DISCUSSION

Summary of main findings

There was a significant short-term benefit for exercises into pain over pain-free exercises for patient-reported outcomes of pain, with a small effect size and moderate quality of evidence. There appears to be no difference at medium-term or long term follow-up, with the quality of the evidence rated as moderate to low.

Clinical and research implications

Traditionally, healthcare practitioners have been reluctant to encourage patients to continue with exercise into pain when they are treating chronic musculoskeletal pain,[76] with some research suggesting clinicians' fear being the primary deterrent.[77] The results of our systematic review show that there does not appear to be a scientific basis for this fear in relation to outcome measures of pain, and also potentially function and disability. This is an important point when considering what advice is given on any short-term exacerbations of musculoskeletal pain during physical activity or exercise by healthcare practitioners, particularly when physical inactivity is one of the 10 leading risk factors for death worldwide,[78] and when an estimated €1.9 billion a year in healthcare and €9.4 billion a year in economic costs in the UK are attributable to physical inactivity.[79]

A theoretical rationale for a positive response to exercises into pain is the positive impact on the central nervous system.[31 37] Specifically, the exercise addresses psychological factors such as fear avoidance, kinesiophobia and catastrophising, and is set within a framework of 'hurt not equalling harm', thus, in time, reducing the overall sensitivity on the central nervous system, with a modified pain output.[31 37] The exercise-induced endogenous analgesia effect

Table 3 GRADE summary of findings table

Summary of results				Quality of the evidence (GRADE)			
Follow-up	Number of participants(trials)	SMD (95% CI)	Design	Inconsistency	Indirectness	Imprecision	Quality
Short term	385 (6 trials)	−0.28 (−0.49 to −0.08)	Limitations*	No inconsistency	No indirectness	Imprecision†	**Low** ⊕⊕◯◯
Medium term	173 (3 trials)	−0.59 (−1.03 to −0.15)	Limitations*	No inconsistency	No indirectness	Imprecision†	**Low** ⊕⊕◯◯
Long term	345 (5 trials)	0.01 (−0.39 to 0.41)	Limitations*	Inconsistency‡	No indirectness	Imprecision†	**Very low** ⊕◯◯◯
		Sensitivity analysis					
Short term	215 (3 trials)	−0.27 (−0.54 to −0.05)	No limitations	No inconsistency	No indirectness	Imprecision†	**Moderate** ⊕⊕⊕◯
Medium term	40 (1 trials)	−0.32 (−0.95 to 0.31)	No limitations	Inconsistency§	No indirectness	Imprecision†	**Low** ⊕⊕◯◯
Long term	215 (3 trials)	0.13 (−0.14 to 0.40)	No limitations	No inconsistency	No indirectness	Imprecision†	**Moderate** ⊕⊕⊕◯

*Lack of blinding of participants and personnel, attrition bias, unable to adequately assess selection bias risk.
†<400 participants for each outcome.
‡Large statistical heterogeneity; I^2=70%.
§Only single trial available, <400 participants therefore downgraded for inconsistency and imprecision.
Short term, ≤3 months; medium term, >3 and <12 months; long term, ≥12 months.
High quality: further research is unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect.
Very low quality: we are uncertain about the estimate.
GRADE, Grading of Recommendations Assessment, Development and Evaluation; SMD, standardised mean difference.

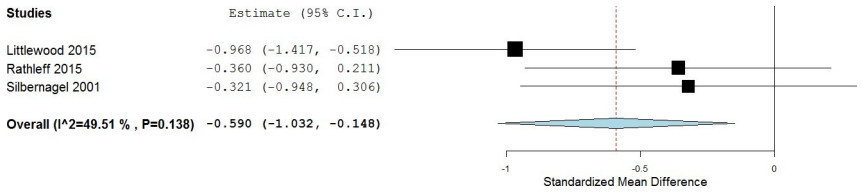


Figure 5 Forest plot of exercises into pain versus pain-free exercises—medium term. Negative values favour painful intervention, whereas positive favour pain-free.

is thought to occur due to a release of endogenous opioids and activation of spinal inhibitory mechanisms.[80–84] However, a recent systematic review has established that no firm conclusions could be reached about pain modulation during exercise therapy for chronic musculoskeletal pain.[85] Indeed one experimental study has shown a dysfunction of endogenous analgesia in patients with musculoskeletal pain,[86] and therefore exercising non-painful body parts with patients with chronic musculoskeletal pain has been recommended.[87] However, it is worth noting that empirical data within this field are greatly lacking, and this systematic review shows that painful exercises may even improve the clinical outcomes. Additionally, exercise prescription in the included trials was primarily based on strength and conditioning principles, with the exception of Littlewood *et al*,[52] suggesting a tissue-focused approach, and therefore could still have been giving a 'hurt is harm' message to the majority of participants.

Significant improvements in patient-reported pain can be achieved with a range of contextual factors, such as varying degrees of pain experiences and postrecovery time for therapeutic exercise. In addition to the aspect of pain, an important difference between the intervention arm and the control arm is the higher loads, or levels of resistance, employed with the exercises into pain, and it is unknown if the difference in responses can be attributable to these two elements of the different exercise programmes. Research has shown a 'dose response' to exercise for musculoskeletal pain—the more incremental exercise (with appropriate recovery period) a person does the greater his/her improvements in pain[88–90]; the short-term benefits of exercises into pain over pain-free exercises could be explained by this dose effect, or response to load/resistance. However to our knowledge the optimal 'dose' of therapeutic exercise for musculoskeletal pain has not been established. Furthermore, little is known if it is possible or appropriate to identify individuals most suitable to exercise interventions.

Our review only investigated patient-reported outcome measures of pain and function/disability. It has been hypothesised that exercise therapy, where it has been advised that the experience of pain is safe and allowed, may address other patient-reported outcome measures—fear avoidance, self-efficacy and catastrophising beliefs[37 38]—and therefore may lead to improvements in function, quality of life and disability, despite pain levels. Unfortunately none of the trials included in this review recorded the level of pain patients actually experienced during their exercise programme, preventing any detailed attempt to fully explain any mechanisms of effect. This aspect of exercise prescription clearly warrants further investigation in relation to chronic musculoskeletal pain. Any future trials should consider the role of pain with exercises and clearly define the parameters employed to ensure translation of findings into practice and further evaluation of optimal 'dosage'.

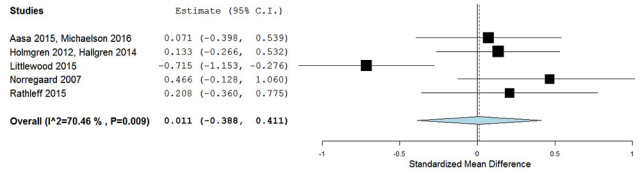


Figure 6 Forest plot of exercises into pain versus pain-free exercises—long term. Negative values favour painful intervention, whereas positive favour pain-free. AMED, Allied and Complimentary Medicine Database; CINAHL, Cumulative Index to Nursing and Allied Health Literature.

Strengths and limitations of included trials

We chose not to perform subgroup analyses by anatomical region and/or tissue structures. The labelling of musculoskeletal structures as sources of pain has been debated for many years, with polarising opinions.[91 92] However, the diagnostic labelling of patients into tissue-specific pathology characteristically suffers from poor reliability and validity.[93–98] A strength of this review is that despite the trials including subjects suffering from musculoskeletal pain at different body locations, there exists low statistical heterogeneity at short-term follow-up and for the sensitivity analyses carried out.

The overall quality of the included papers can be considered relativity high, with only three domains in the Cochrane risk of bias tool (disregarding blinding of participants) demonstrating clear risk of bias across all domains for all trials. However taking into account other factors assessed with the GRADE analysis, the quality of the evidence was rated as moderate to low. Therefore our results can be considered to have moderate to low internal validity, with future research likely to alter our conclusions.

The main source of bias within the included trials were blinding; no trial blinded the participants. Knowledge of group assignment may affect participants' behaviour, for example with patient-reported outcome measures such as pain scales or compliance with therapy interventions.[99] However, it is accepted that blinding in physiotherapy and physical intervention trials is difficult to achieve.[24]

Another limitation of the included trials is the high level of attrition suffered by some of the trials in both treatment arms. For example Littlewood *et al*[52] suffered from 51% dropout at 12-month follow-up. A high level of attrition can overestimate the treatment effect size and could bias the results of our meta-analysis. However, we minimised the risk of bias on our results by conducting a sensitivity analysis on trials with a large dropout, identified using the Cochrane risk of bias tool and assessed level of evidence using the GRADE classification.

Limitations of this review

For pragmatic reasons one reviewer screened titles and abstracts. An extensive literature search was carried out, with two reviewers independently screening full texts for inclusion, and a sample of the data extraction independently verified. Additionally an attempt was made to retrieve unpublished trials; however, it may be that not all trials were retrieved, particularly considering we did not search for papers published in languages other than English and US spelling was used in the search terms. This review excluded trials where participants had a diagnosis of more widespread pain disorders like fibromyalgia.

CONCLUSION

The results of this systematic review indicates that protocols using exercises into pain offer a small but significant benefit over pain-free exercises in the short term, with moderate quality of the evidence for outcomes of pain in chronic musculoskeletal pain in adults. There appears to be no difference at medium-term or long-term follow-up, with moderate to low quality of evidence, demonstrating pain need not be ruled out or avoided in adults with chronic musculoskeletal pain.

What are the findings?

- Protocols using exercises into pain for chronic musculoskeletal pain offer a small but significant benefit over pain-free exercises in the short term.
- Adults with musculoskeletal pain can achieve significant improvements in patient-reported outcomes with varying degrees of pain experiences and postrecovery time with therapeutic exercise.
- Pain during therapeutic exercise for chronic musculoskeletal pain need not be a barrier to successful outcomes.
- Protocols using exercises into pain typically have higher loads and dose of exercise.

Correction notice This paper has been amended since it was published Online First. The authors have noticed that figure 4 was a duplication of figure 6. The correct figure 4 has now been uploaded.

Contributors BES was responsible for conception and design, publication screening, acquisition of data, analysis and interpretation, and drafting and revising the manuscript. PH was responsible for conception and design, publication screening, acquisition of data, data interpretation, and reviewing and revising the manuscript. TOS was responsible for conception and design, data interpretation, and reviewing and revising the manuscript. All authors were involved in interpretation, reviewing revisions to the manuscript and final approval of the version to be published. All have read and approved the final version.

Funding This report is an independent research arising from a Clinical Doctoral Research Fellowship, Benjamin E Smith, ICA-CDRF-2015-01-002, supported by the National Institute for Health Research (NIHR) and Health Education England (HEE).

Disclaimer The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR, HEE or the Department of Health.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Lawrence RC, Helmick CG, Arnett FC, *et al*. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998;41:778–99.
2 Murray CJ, Vos T, Lozano R, *et al*. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the global burden of disease study 2010. *Lancet* 2012;380:2197–223.
3 Hoy D, March L, Brooks P, *et al*. The global burden of low back pain: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:968–74.
4 Vos T, Flaxman AD, Naghavi M, *et al*. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010. *Lancet* 2012;380:2163–96.
5 Hoy D, March L, Woolf A, *et al*. The global burden of neck pain: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1309–15.
6 Smith E, Hoy DG, Cross M, *et al*. The global burden of other musculoskeletal disorders: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1462–9.
7 Department of Health. *The musculoskeletal services framework- A joint responsibilty: doing itdifferently–*. London: DH Publications Orderline, 2006:1–72.
8 HSE. Work-related musculoskeletal disorder statistics, great britain 2016. 2016:1–20 www.hse.gov.uk/statistics.
9 Mason L, Moore RA, Edwards JE, *et al*. Topical NSAIDs for chronic musculoskeletal pain: systematic review and meta-analysis. *BMC Musculoskelet Disord* 2004;5:28.
10 Noble M, Treadwell JR, Tregear SJ, *et al*. Long-term opioid management for chronic noncancer pain. *Cochrane Database Syst Rev* 2010;1.
11 Roelofs PD, Deyo RA, Koes BW, *et al*. Non-steroidal anti-inflammatory drugs for low back pain. *Cochrane Database Syst Rev* 2008 1:CD000396.
12 Chou R, Huffman LH. Medications for acute and chronic low back pain: a review of the evidence for an american Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med* 2007;147:505–14.
13 Ravenek MJ, Hughes ID, Ivanovich N, *et al*. A systematic review of multidisciplinary outcomes in the management of chronic low back pain. *Work* 2010;35:349–67.
14 Scascighini L, Toma V, Dober-Spielmann S, *et al*. Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes. *Rheumatology* 2008;47:670–8.
15 van Geen JW, Edelaar MJ, Janssen M, *et al*. The long-term effect of multidisciplinary back training: a systematic review. *Spine* 2007;32:249–55.
16 van Middelkoop M, Rubinstein SM, Kuijpers T, *et al*. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *Eur Spine J* 2011;20:19–39.
17 Rubinstein SM, van Middelkoop M, Assendelft WJ, *et al*. Spinal manipulative therapy for chronic low-back pain: an update of a Cochrane review. *Spine* 2011;36:E825–46.
18 Miller J, Gross A, D'Sylva J, *et al*. Manual therapy and exercise for neck pain: a systematic review. *Man Ther* 2010;15:334–54.
19 Furlan AD, Imamura M, Dryden T, *et al*. Massage for low-back pain. *Cochrane Database Syst Rev* 2008: (4):CD001929.
20 Hall A, Maher C, Latimer J, *et al*. The effectiveness of Tai Chi for chronic musculoskeletal pain conditions: a systematic review and meta-analysis. *Arthritis Rheum* 2009;61:717–24.
21 Slade SC, Keating JL. Unloaded movement facilitation exercise compared to no exercise or alternative therapy on outcomes for people with nonspecific chronic low back pain: a systematic review. *J Manipulative Physiol Ther* 2007;30:301–11.
22 Hendrick P, Te Wake AM, Tikkisetty AS, *et al*. The effectiveness of walking as an intervention for low back pain: a systematic review. *Eur Spine J* 2010;19:1613–20.
23 Hall J, Swinkels A, Briddon J, *et al*. Does aquatic exercise relieve pain in adults with neurologic or musculoskeletal disease? A systematic review and meta-analysis of randomized controlled trials. *Arch Phys Med Rehabil* 2008;89:873–83.
24 Smith BE, Littlewood C, May S. An update of stabilisation exercises for low back pain: a systematic review with meta-analysis. *BMC Musculoskelet Disord* 2014;15:416.
25 Miles CL, Pincus T, Carnes D, *et al*. Can we identify how programmes aimed at promoting self-management in musculoskeletal pain work and who benefits? A systematic review of sub-group analysis within RCTs. *Eur J Pain* 2011;15.
26 Moseley GL, Arntz A. The context of a noxious stimulus affects the pain it evokes. *Pain* 2007;133:64–71.
27 Harvie DS, Broecker M, Smith RT, *et al*. Bogus visual feedback alters onset of movement-evoked pain in people with neck pain. *Psychol Sci* 2015;26:385–92.
28 Lobanov OV, Zeidan F, McHaffie JG, *et al*. From cue to meaning: brain mechanisms supporting the construction of expectations of pain. *Pain* 2014;155:129–36.
29 Nijs J, Lluch Girbés E, Lundberg M, *et al*. Exercise therapy for chronic musculoskeletal pain: innovation by altering pain memories. *Man Ther* 2015;20:216–20.
30 Meeus M, Nijs J, Van Wilgen P, *et al*. Moving on to movement in patients with chronic joint pain. *Pain Clin Updat* 2016;24:1–8.
31 Nijs J, Roussel N, Paul van Wilgen C, *et al*. Thinking beyond muscles and joints: therapists' and patients' attitudes and beliefs regarding chronic musculoskeletal pain are key to applying effective treatment. *Man Ther* 2013;18:96–102.
32 Zusman M. Associative memory for movement-evoked chronic back pain and its extinction with musculoskeletal physiotherapy. *Phys Ther Rev* 2008;13:57–68.
33 Zusman M. Mechanisms of musculoskeletal physiotherapy. *Phys Ther Rev* 2004;9:39–49.
34 Moseley GL. Reconceptualising pain according to modern pain science. *Phys Ther Rev* 2007;12:169–78.
35 Quartana PJ, Campbell CM, Edwards RR. Pain catastrophizing: a critical review. *Expert Rev Neurother* 2009;9:745–58.
36 Moseley GL. Joining forces – Combining Cognition-Targeted Motor Control Training with Group or Individual Pain Physiology Education: A Successful Treatment For Chronic Low Back Pain. *J Man Manip Ther* 2003;11:88–94.
37 Littlewood C, Malliaras P, Bateman M, *et al*. The central nervous system--an additional consideration in 'rotator cuff tendinopathy' and a potential basis for understanding response to loaded therapeutic exercise. *Man Ther* 2013;18:468–72.
38 Smith BE, Hendrick P, Logan P. Patellofemoral pain: challenging current practice - A case report. *Man Ther* 2016;22:216–9.
39 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and Meta-Analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
40 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.

41 Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. New York: John Wiley, 1981.
42 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
43 Aasa B, Berglund L, Michaelson P, *et al*. Individualized low-load motor control exercises and education versus a high-load lifting exercise and education to improve activity, pain intensity, and physical performance in patients with low back pain: a randomized controlled trial. *J Orthop Sports Phys Ther* 2015;45:77–85.
44 Beyer R, Kongsgaard M, Hougs Kjær B, *et al*. Heavy Slow Resistance Versus Eccentric training as treatment for Achilles Tendinopathy. *Am J Sports Med* 2015; 43:1704–11.
45 Geraets JJ, Goossens ME, de Groot IJ, *et al*. Effectiveness of a graded exercise therapy program for patients with chronic shoulder complaints. *Aust J Physiother* 2005;51:87–94.
46 Harts CC, Helmhout PH, de Bie RA, *et al*. A high-intensity lumbar extensor strengthening program is little better than a low-intensity program or a waiting list control group for chronic low back pain: a randomised clinical trial. *Aust J Physiother* 2008;54:23–31.
47 Maenhout AG, Mahieu NN, De Muynck M, *et al*. Does adding heavy load eccentric training to rehabilitation of patients with unilateral subacromial impingement result in better outcome? A randomized, clinical trial. *Knee Surg Sports Traumatol Arthrosc* 2013;21:1158–67.
48 Østerås B, Østerås H, Torstensen TA, *et al*. Dose-response effects of medical exercise therapy in patients with patellofemoral pain syndrome: a randomised controlled clinical trial. *Physiotherapy* 2013;99:126–31.
49 Schenk R, Dionne C, Simon C, *et al*. Effectiveness of mechanical diagnosis and therapy in patients with back pain who meet a clinical prediction rule for spinal manipulation. *J Man Manip Ther* 2012;20:43–9.
50 van der Plas A, de Jonge S, de Vos RJ, *et al*. A 5-year follow-up study of Alfredson's heel-drop exercise programme in chronic midportion Achilles tendinopathy. *Br J Sports Med* 2012;46:214–8.
51 Rathleff MS, Mølgaard CM, Fredberg U, *et al*. High-load strength training improves outcome in patients with plantar fasciitis: a randomized controlled trial with 12-month follow-up. *Scand J Med Sci Sports* 2015;25:e292–300.
52 Littlewood C, Bateman M, Brown K, *et al*. A self-managed single exercise programme versus usual physiotherapy treatment for rotator cuff tendinopathy: a randomised controlled trial (the SELF study). *Clin Rehabil* 2016;30:686–96.
53 Østerås B, Østerås H, Torstensen TA, *et al*. Long-term effects of medical exercise therapy in patients with patellofemoral pain syndrome: results from a single-blinded randomized controlled trial with 12 months follow-up. *Physiotherapy* 2013;99:311–6.
54 Higgins J, Deeks J. Cochrane Handbook: General Methods For Cochrane Reviews: Ch 7: Selecting studies and collecting data. In: Higgins PTJ, Green S, eds. *Cochrane handbook for: systematic reviews of interventions*. USA: Wiley-Blackwell, 2011:151–86.
55 Furlan AD, Pennick V, Bombardier C, *et al*. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009;34:1929–41.
56 Higgins JP, Altman DG, Gøtzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
57 Clijsen R, Fuchs J, Taeymans J. Effectiveness of exercise therapy in treatment of patients with patellofemoral pain syndrome: systematic review and meta-analysis. *Phys Ther* 2014;94:1697–708.
58 Atkins D, Best D, Briss PA, *et al*. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
59 Sterne JA, Egger M, Moher D. Cochrane handbook: General methods for cochrane reviews: Ch 10: Addressing reporting biases. In: Higgins PTJ, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. USA: Wiley-Blackwell, 2011:297–334.
60 Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Chinese J. Evidence-Based Med* 2009;9:8–11.
61 Schünemann H, Brożek J, Guyatt G, *et al*. GRADE handbook for grading quality of evidence and strength of recommendations. *GRADE Work. Gr* 2013.
62 Hedges L V, Olkin I. Statistical methods for meta-analysis. *Phytochemistry* 1985;72:369.
63 Cohen J. Statistical power analysis for the behavioral sciences. *Stat Power Anal Behav Sci* 1988.567.
64 Hozo SP, Djulbegovic B, Hozo I, *et al*. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol* 2005;5:13.
65 Higgins JP, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
66 Yusuf S, Peto R, Lewis J, *et al*. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis* 1985;27:335–71.
67 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
68 Wallace BC, Dahabreh IJ, Trikalinos TA, *et al*. Closing the gap between Methodologists and End-Users: r as a computational Back-End. *Wiley Interdiscip Rev Comput* 2012;49:1–15.
69 Roach KE, Budiman-Mak E, Songsiridej N, *et al*. Development of a shoulder pain and disability index. *Arthritis Care Res* 1991;4:143–9.
70 Holmgren T, Björnsson Hallgren H, Öberg B, *et al*. Effect of specific exercise strategy on need for surgery in patients with subacromial impingement syndrome: randomised controlled study. *BMJ* 2012;344:e787.
71 Hallgren HC, Holmgren T, Oberg B, *et al*. A specific exercise strategy reduced the need for surgery in subacromial pain patients. *Br J Sports Med* 2014;48:1431–6.
72 Michaelson P, Holmberg D, Aasa B, *et al*. High load lifting exercise and low load motor control exercises as interventions for patients with mechanical low back pain: a randomized controlled trial with 24-month follow-up. *J Rehabil Med* 2016;48:456–63.
73 Nørregaard J, Larsen CC, Bieler T, *et al*. Eccentric exercise in treatment of Achilles tendinopathy. *Scand J Med Sci Sports* 2007;17:133–8.
74 Silbernagel KG, Thomeé R, Thomeé P, *et al*. Eccentric overload training for patients with chronic Achilles tendon pain--a randomised controlled study with reliability testing of the evaluation methods. *Scand J Med Sci Sports* 2001;11:197–206.
75 Higgins JPT, Altman DG, Jonathan AC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated march 2011). The cochrane collaboration, 2011, 2008. www.cochrane-www.handbook.org.
76 Littlewood C, Lowe A, Moore J. Rotator cuff disorders: a survey of Current UK physiotherapy practice. *Shoulder Elbow* 2012;4:64–71.
77 Littlewood C, Mawson S, May S, *et al*. Understanding the barriers and enablers to implementation of a self-managed exercise intervention: a qualitative study. *Physiotherapy* 2015;101:279–85.
78 WHO. Global health risks: mortality and burden of disease attributable to selected Major risks. *Bull World Health Organ* 2009;87:646.
79 ISCA/Cebr. The economic cost of physical inactivity in Europe. *ISCA / Cebr Rep* Published Online First: 2015.
80 Koltyn KF. Exercise-induced hypoalgesia and intensity of exercise. *Sports Med* 2002;32:477–87.
81 Koltyn KF. Analgesia following exercise: a review. *Sports Med* 2000;29:85–98.
82 Koltyn KF, Arbogast RW. Perception of pain after resistance exercise. *Br J Sports Med* 1998;32:20–4.
83 Ray CA, Carter JR. Central modulation of exercise-induced muscle pain in humans. *J Physiol* 2007;585:287–94.
84 Millan MJ. Descending control of pain. *Prog Neurobiol* 2002;66:355–474.
85 Fuentes C JP, Armijo-Olivo S, Magee DJ, *et al*. Effects of exercise therapy on endogenous pain-relieving peptides in musculoskeletal pain: a systematic review. *Clin J Pain* 2011;27:365–74.
86 Van Oosterwijck J, Nijs J, Meeus M, *et al*. Lack of endogenous pain inhibition during exercise in people with chronic whiplash associated disorders: an experimental study. *J Pain* 2012;13:242–54.
87 Nijs J, Kosek E, Van Oosterwijck J, *et al*. Dysfunctional endogenous analgesia during exercise in patients with chronic pain: to exercise or not to exercise? *Pain Physician* 2012;15:ES205–13.
88 Littlewood C, Malliaras P, Chance-Larsen K. Therapeutic exercise for rotator cuff tendinopathy: a systematic review of contextual factors and prescription parameters. *Int J Rehabil Res* 2015;38:95–106.
89 Østerås B, Østerås H, Torstensen TA, *et al*. Dose-response effects of medical exercise therapy in patients with patellofemoral pain syndrome: a randomised controlled clinical trial. *Physiotherapy* 2013;99:126–31.
90 Rathleff MS, Roos EM, Olesen JL, *et al*. Exercise during school hours when added to patient education improves outcome for 2 years in adolescent patellofemoral pain: a cluster randomised trial. *Br J Sports Med* 2015;49:1–7.
91 Cools AM, Michener LA. Shoulder pain: can one label satisfy everyone and everything? *Br J Sports Med* 2017;51:416–7.
92 Sloan TJ, Walsh DA. Explanatory and diagnostic labels and perceived prognosis in chronic low back pain. *Spine* 2010;35:E1120–5.
93 Smith BE, Thacker D, Crewesmith A, *et al*. Special tests for assessing meniscal tears within the knee: a systematic review and meta-analysis. *Evid Based Med* 2015;20:88–97.
94 May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review. *Aust J Physiother* 2006;52:91–102.
95 May S, Chance-Larsen K, Littlewood C, *et al*. Reliability of physical examination tests used in the assessment of patients with shoulder problems: a systematic review. *Physiotherapy* 2010;96:179–90.
96 Seffinger MA, Najm WI, Mishra SI, *et al*. Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. *Spine* 2004;29:E413–25.
97 Smith TO, Davies L, Donell ST. The reliability and validity of assessing medio-lateral patellar position: a systematic review. *Man Ther* 2009;14:355–62.
98 Smith TO, Hunt NJ, Donell ST. The reliability and validity of the Q-angle: a systematic review. *Knee Surg Sports Traumatol Arthrosc* 2008;16:1068–79.
99 Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002;359:696–700.

BMJ Open

What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? A systematic review

Jonathan Room,[1,2] Erin Hannink,[1] Helen Dawes,[1] Karen Barker[3]

To cite: Room J, Hannink E, Dawes H, *et al*. What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? A systematic review. *BMJ Open* 2017;**7**:e019221. doi:10.1136/bmjopen-2017-019221

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-019221).

Received 18 August 2017
Revised 27 October 2017
Accepted 30 October 2017

[1]Physiotherapy Research Unit, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[2]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[3]Centre for Movement and Occupational Rehabilitation Sciences (MOReS), Oxford Brookes University, Oxford, UK

Correspondence to
Jonathan Room;
jonathan.room@ouh.nhs.uk

ABSTRACT

Objectives To conduct a systematic review of interventions used to improve exercise adherence in older people, to assess the effectiveness of these interventions and to evaluate the behavioural change techniques underpinning them using the Behaviour Change Technique Taxonomy (BCTT).

Design Systematic review.

Methods A search was conducted on AMED, BNI, CINAHL, EMBASE, MEDLINE and PsychINFO databases. Randomised controlled trials that used an intervention to aid exercise adherence and an exercise adherence outcome for older people were included. Data were extracted with the use of a preprepared standardised form. Risk of bias was assessed with the Cochrane Collaboration's tool for assessing risk of bias. Interventions were classified according to the BCTT.

Results Eleven studies were included in the review. Risk of bias was moderate to high. Interventions were classified into the following categories: comparison of behaviour, feedback and monitoring, social support, natural consequences, identity and goals and planning. Four studies reported a positive adherence outcome following their intervention. Three of these interventions were categorised in the feedback and monitoring category. Four studies used behavioural approaches within their study. These were social learning theory, socioemotional selectivity theory, cognitive behavioural therapy and self-efficacy. Seven studies did not report a behavioural approach.

Conclusions Interventions in the feedback and monitoring category showed positive outcomes, although there is insufficient evidence to recommend their use currently. There is need for better reporting, use and the development of theoretically derived interventions in the field of exercise adherence for older people. Robust measures of adherence, in order to adequately test these interventions would also be of use.

PROSPERO registration number CRD42015020884.

Strengths and limitations of this study

- This systematic review adds to the evidence for exercise adherence interventions specifically for older people.
- Uses a predefined behaviour change taxonomy allowing the categorisation and evaluation of interventions.
- Studies included were of moderate to high risk of bias.

INTRODUCTION

Exercise is an effective treatment option for a variety of conditions[1] and in a number of chronic conditions its effectiveness may be comparable to drug interventions.[2] This type of therapeutic exercise is defined as a subset of physical activity that is structured and planned, with the aim of maintaining or improving one or more aspects of physical fitness, in this way it differs from physical activity which is defined as any bodily movement generated by skeletal muscle.[3] Prescribed exercise is a common treatment option used by health professionals such as physiotherapists.[4] No definitive figure exists regarding the number of exercise programmes prescribed in a given year. However to give some indication as to the magnitude of this number, in 2014 there were 23 006 physiotherapists in the UK.[5] A UK survey of organisations offering outpatient physiotherapy reported that of the 54% of organisations to respond 1 480 893 new patients were seen in a year.[6] It is known from surveys of practice that exercise is a commonly used treatment modality across a range of conditions.[7–11] It is therefore reasonable to assume that a significant number of exercise programmes are being prescribed yearly.

Adherence to exercise is known to be variable. In their seminal paper, Sluijs *et al*[12] reported that 22% of patients were non-compliant, with 41% being partially compliant. Similar figures have been demonstrated subsequently.[13] It is known that exercise adherence can affect treatment outcomes, with factors such as pain, physical function,

1. Medline; exp AGED/;
2. Medline; older.ti,ab;.
3. Medline; (Older AND NEAR AND Adult).ti,ab;
4. Medline; elderly.ti,ab;
5. Medline; adherence.ti,ab;
6. Medline; exp PATIENT COMPLIANCE/;
7. Medline; exp PATIENT PARTICIPATION/;
8. Medline; attendance.ti,ab;
9. Medline; ((Change OR changes OR Changing) AND NEAR AND (Behaviour OR Behavior)).ti,ab;
10. Medline; ((Modify OR Modifies OR Modifying) AND NEAR AND (Behaviour OR Behavior)).ti,ab;
11. Medline; adhering.ti,ab;
12. Medline; complying.ti,ab;
13. Medline; exp MOTIVATION/;
14. Medline; concordance.ti,ab;
15. Medline; co-operation.ti,ab;
16. Medline; engagement.ti,ab;
17. Medline; exp EXERCISE/;
18. Medline; exp REHABILITATION/;
19. Medline; (older ADJ5 patient*)
20. Medline; 1 OR 2 OR 3 OR 4 OR 19;
21. Medline; 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16;
22. Medline; 17 OR 18;
23. Medline; 20 AND 21 AND 22;

Figure 1 An example of search terms from Medline.

physical performance and self-perceived effect of exercise being higher in those with better adherence.[14 15] Therefore, low levels of adherence may limit the effectiveness of prescribed exercise. This makes adherence an important consideration for those who prescribe exercise.

Adherence is defined by WHO as the '*extent to which a person's behaviour corresponds with agreed recommendations*

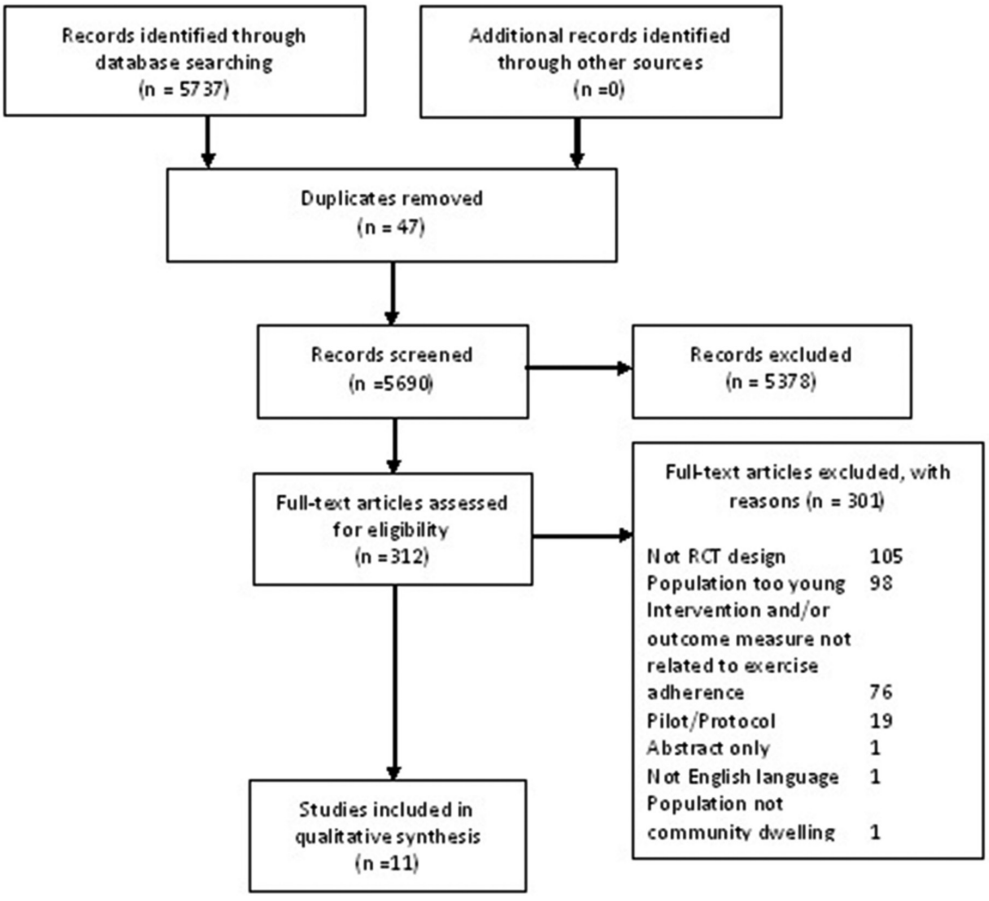


Figure 2 Flow diagram of study selection. RCT, randomised controlled trial.

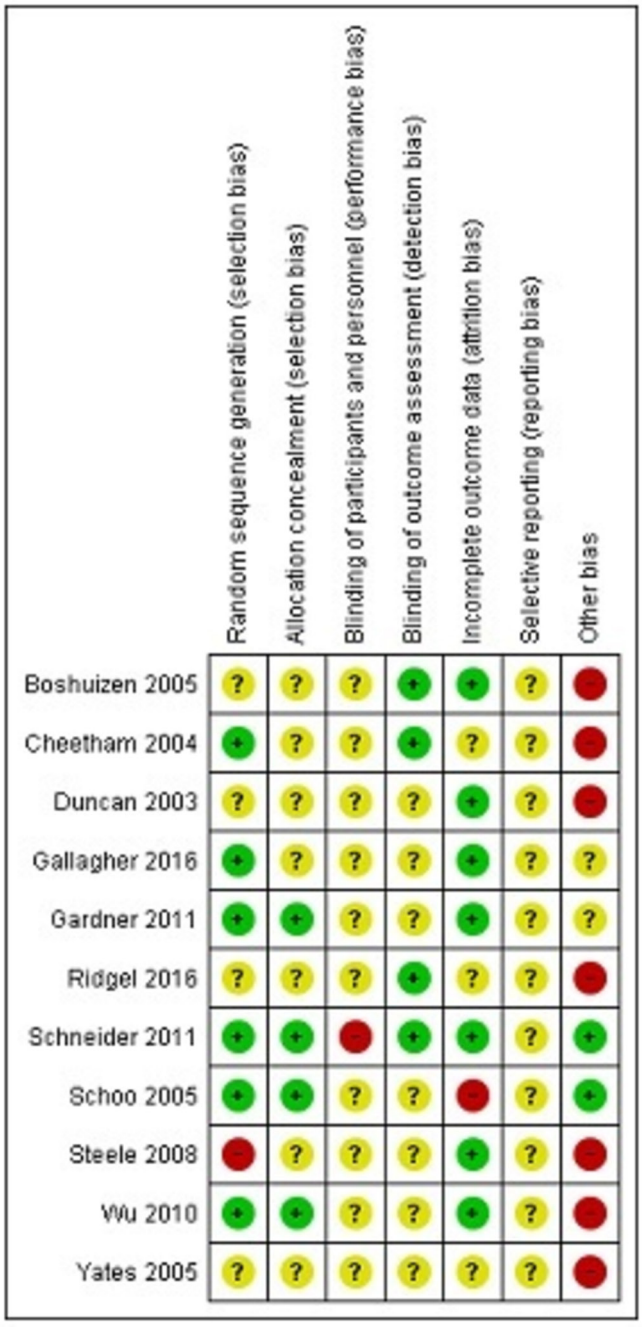


Figure 3 Risk of bias assessment.

from a healthcare provider'.[16] Adherence to medical treatments, in particular medication is well reviewed.[17–19] Considering exercise adherence, previous literature has focused largely on factors relating to adherence.[20–25] A previous review[22] reported that adherence to treatment given in physiotherapy including prescribed exercises was influenced by low baseline levels of physical activity, low in-treatment adherence, low self-efficacy, depression, anxiety, helplessness, poor social support, greater number of perceived barriers to exercise and increased pain levels during exercise. Adherence to home-based physical therapies has been linked to several factors including intention to engage in home-based physical therapy, self-motivation, self-efficacy, previous adherence to exercise-related behaviours and also social support.[20]

Adhering to exercise is important for all populations, however, there are several factors that make it crucial for older people. Exercise adherence in this population is affected by health status,[21] and it is known that older people are more likely to have long-term conditions (LTCs) or multiple LTCs,[26] for which prescribed exercise is a treatment option.[1] Exercise engagement is known to be poor in older people following discharge from hospital,[27] or discharge from physiotherapy.[28] This is a critical consideration because treatment outcomes in this population are linked to compliance with interventions.[29] There are a number of factors that have been identified as affecting exercise adherence in older people, including low self-efficacy, low motivation, depression, lack of interest, fear of falling, health status, physical ability, low expectations, socioeconomic status and exercise programme characteristics.[21 27 28] Programme design was also a factor noted by Farrance *et al*[30] in a mixed-method systematic review of community-based exercise interventions for older people. They also reported six key themes related to adherence, these being social connectedness, participant perceived benefits, programme design, empowering/energising effects, instructor and individual behaviour. While it is important to understand the role of these personal factors and programme characteristics, it is also crucial to establish if there is anything clinicians can do to enhance adherence to prescribed exercise in older people.

Exercise adherence interventions aim to increase the likelihood that people will follow prescribed exercise, in this way they fulfil the definition by the National Institute for Health and Care Excellence[31] of a behaviour change intervention, *'…sets of techniques, used together, which aim to change the health behaviours of individuals, communities or whole populations'*. Many previous behavioural interventions have been designed using what Martin Eccles calls the ISLAGIATT principle, *'it seemed like a good idea at the time'*.[32] This lack of theoretical underpinning could potentially limit the effectiveness of interventions. For this reason and so that interventions can be described and categorised, it is important to review the theories or approaches that underpin exercise adherence interventions. One way this can be achieved is through using a method to categorise behavioural approaches, such as the Behaviour Change Technique Taxonomy (BCTT) developed by Michie *et al*.[33]

A Cochrane review exploring interventions to improve exercise adherence in those aged 18 years and over with chronic musculoskeletal pain[34] reported that interventions such as self-management techniques and supervised as well as individualised exercise might improve adherence. More recently, Peek *et al*[35] reviewed adherence to self-management strategies prescribed by physiotherapists. They found that interventions using activity monitoring and feedback systems, written instructions and behavioural exercise programmes with booster sessions may be effective in promoting adherence. Although both

these reviews were undertaken on adult populations, they did not breakdown the population further and, there remains a need to consider interventions specific to older populations. Disease-specific reviews that are relevant to older people have taken place, in particular considering arthritis. Ezzat *et al*[36] reported limited evidence for exercise adherence interventions in an arthritis population. Nicolson *et al*[37]concluded that booster sessions, and behavioural graded exercise could improve adherence for those with osteoarthritis, in addition to motivational approaches for those with chronic low back pain. However, given that exercise is prescribed for a breadth of conditions,[1] there is need to consider a broader, non-disease-specific review for older people to draw evidence from a wider population.

Other approaches that have shown potential in adherence include peer delivered programmes and arthritis self-management programmes. Burton *et al*[38] reviewed the effectiveness of peers delivering programmes, or motivating older people to increase physical activity, finding that involving peers in exercise programmes can promote adherence. Williamson *et al*[39] reviewed behavioural physical activity interventions in those with lower limb osteoarthritis. They report that self-management programmes for those with osteoarthritis demonstrate a small but significant improvement in short-term physical activity. Although both these examples focus on physical activity, rather than exercise, there may be some crossover, and there remains a need to review interventions in the field of therapeutic exercise. While we know there is no clear guidance regarding approaches for therapists to optimise adherence to prescribed exercise, there are studies that consider older patients and adherence,[40–42] but no evidence synthesis as yet. Therefore, the aim of this review is threefold to:

- Establish what interventions have been described in the literature to improve adherence to prescribed exercise in older people.
- Determine to what extent these interventions are effective at improving exercise adherence.
- Describe any underlying behavioural techniques or theory behind these interventions.

METHODS

The steps taken in the design and conduct of this review have been done so with consideration of the Preferred Reporting Items for Systematic reviews and Meta-analysis (PRISMA).[43 44] This review has been registered with PROSPERO, registration number CRD42015020884 available at http://www.crd.york.ac.uk/PROSPERO.

Data sources and searches

The following electronic databases were searched from inception up to May 2017 AMED, BNI, CINAHL, EMBASE, Medline and PsycINFO. Additionally, the reference lists of papers included in the systematic review were screened.

Search terms

Search terms were developed by JR. The terms were expanded at two consensus meetings attended by healthcare researchers. The list was reviewed by a healthcare librarian and further changes made. The terms make use of both subject headings and free text search terms. Figure 1 is an example of the search from Medline; Medical Subject Headings terms are shown in bold.

Study selection

All databases were searched by JR, once studies were returned, titles and abstracts were screened and full texts were retrieved if the study was potentially relevant. A second reviewer EH also independently screened the title and abstracts of the studies retrieved by the Embase database. This comprised 1179 hits which was 20.55% of all the studies retrieved. JR and EH compared results, any disagreements were resolved by discussion. A third reviewer KB was available if agreement could not be reached. Once full texts had been retrieved, JR and EH independently assessed the studies against the inclusion criteria. After reviewing all full texts, results were compared. Where disagreement occurred this was resolved through discussion, KB was available if agreement could not be reached.

Eligibility criteria

Studies were included if they met the following inclusion criteria:

- Including a population that had a mean age of 65 years or older.
- Including a population that is community dwelling.
- Randomised controlled trials (RCTs).
- Studies including intervention(s) aiming to improve adherence, compliance, concordance to or engagement with exercise, compared with either no adherence, compliance, concordance or engagement intervention; another adherence, compliance, concordance or engagement intervention or an intervention which does not aim to improve adherence, compliance, concordance or engagement.
- A comparator group which was also undertaking the exercise programme. Where a no intervention control group occurred, there needed to be a least two active intervention groups to offer a comparison.
- Published in English.
- Peer reviewed.

Studies were excluded for the following reasons:

- Studies including a population with a diagnosis of dementia or cognitive impairment.
- Any study design that was not an RCT.
- Protocols, feasibility and pilot studies including pilot RCTs.

Data extraction

Two reviewers independently extracted information from the included studies using separate, standardised preprepared forms. Data were extracted about study design,

participants, setting, type and dose of intervention, underlying theory behind the intervention, the comparator arm, the method of assessment, outcome measures used and study findings.

Quality assessment

Two reviewers assessed study quality independently. One reviewer was blinded to author, journal, publication date and affiliations. The Cochrane Collaboration's tool for assessing risk of bias was used.[45] Each study was reviewed for the following items: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, incomplete data, selective outcome reporting and other sources of bias. Each item was graded as low risk of bias, high risk of bias or uncertain risk. For sequence generation, if an appropriate method of randomly generating group allocation, to allow for comparable groups was described, this was scored as low risk of bias. If a non-random method was described and/or groups were significantly different in baseline characteristics, it was scored as high risk. If the description was not clear, it was marked as uncertain risk. For allocation concealment, where a method of concealing assignment, that is, the participant or investigator could not predict assignment, such as opaque sealed envelopes this was scored as low risk. If a method enabling participants or investigators to be able to predict assignment was used, this was graded as high risk. Where information was not clear, this was scored as uncertain risk. For blinding of participants and personnel, if where possible studies reported blinding of participants and personnel this was scored as low risk. If it was possible to blind participants and personnel but this was not done, this was graded as high risk. It is acknowledged that blinding of participants and personnel is very difficult in exercise and rehabilitation studies, therefore if the reviewers felt that blinding was not possible it was scored as uncertain risk. For blinding of outcome assessors, studies where there was specific mention of steps to blind either outcome assessors, or those handling the data if outcomes were self-report questionnaires or surveys filled out at home, this was scored as low risk. If outcome assessors were not blinded this was rated as high risk. Where unclear this was graded as uncertain risk. For incomplete data, if there was little or no incomplete data, and if appropriate measures were taken to deal with missing data, this was scored as low risk. If there was a large amount of missing data, or no appropriate steps to manage missing data this was marked as high risk. Where it was unclear, this was scored as uncertain risk. For selective outcome reporting, if a study protocol was available and all outcomes described were reported this was scored as low risk. If not all outcomes were described this was scored as high risk. Where no study protocol was available this was rated as uncertain risk. For the final domain, other sources of bias, studies were scored as low risk if the reviewers felt there were no other sources of bias that could affect the results. As high risk if there were other potential sources of bias, such as small sample size, where unclear this was scored as uncertain risk. Disagreement between reviewers was resolved through discussion. If consensus was not met a third reviewer was available. If it was necessary authors were contacted for further information.

Data synthesis

The heterogeneous nature of the interventions and the different outcomes used for measuring exercise adherence prevented the use of meta-analysis. Therefore, the interventions are classified according to the behaviour change techniques that they employ, as described in the predefined BCTT.[33] This taxonomy categorises behaviour change techniques by the active ingredients they use. Interventions from included studies were grouped into categories according to the techniques that they employed. Study interventions were reviewed and compared against the definitions used to define each technique. All techniques in the taxonomy fall within 1 of 16 categories. The interventions were placed into categories according to the techniques that were identified during the process of reviewing and comparing against technique definitions.

RESULTS

A total of 5737 papers were identified through database searches, after screening the title and abstract and removing duplicates 5425 were removed. The full text was retrieved for the remaining 312 papers, 301 papers were removed at this stage as they did not meet the inclusion criteria. This left 11 studies which were included in the review.[40–42 46–53] A flow chart of this process can be seen in figure 2.

Risk of bias

Eleven studies were assessed using the Cochrane Collaboration's tool for assessing risk of bias. No studies were rated as low risk of bias, two as moderate risk of bias and nine as high risk of bias. The scores for each domain can be seen in figure 3. The most common area where risk of bias was observed was in the relatively small sample sizes, and the lack of sample size justification of most studies.

Types of intervention

Several types of intervention were identified, these are categorised according to the BCTT developed by Michie *et al*.[33] For a full description of the studies see table 1.

Comparison of behaviour

One study with high risk of bias[41] compared exercise instruction given in audio and video format in addition to written instructions, they found no significant difference in mean exercise adherence between groups at 1–4 weeks (P=0.690) and 5–8 weeks (P=0.538).

Feedback and monitoring

One study[47] with a high risk of bias, provided heart failure participants with individual graphic feedback, related to their exercise goal. They found a significant difference

Table 1 study characteristics

Author	Population	Adherence intervention	Behavioural theory	Measure of adherence	Control or comparison	Results	Conclusion
Boshuizen *et al*[48]	Frail elderly Mean age years (SD) High guidance 80.0 (6.7), dropouts 80.8 (5.3) Medium guidance 79.3 (7.0), dropouts 79.9 (5.9), control 77.2 (6.5), dropouts 75.2 (10.5) Gender Male=5.6%	Guidance and supervision 1) Two supervised sessions and one unsupervised a week 2) One supervised session and two unsupervised session a week	None given	Percentage of exercise sessions (supervised+unsupervised), taken from physical therapists records and self-report diaries. At each location different physical therapists collected outcomes from those that lead the training	Asked to remain habitually active	Percentage of exercise sessions undertaken High guidance 79 (range 57–100) Medium guidance 72 (range 20–93) (No significant difference between groups)	No significant difference in the number of exercise sessions completed between the groups
Cheetham *et al*[52]	Intermittent claudication Mean age 67 years Gender Male=67%	Weekly exercise and motivation class	None given	Self-reported compliance at 6 months—asked whether they walked <3 times, three times or >3 times a week Data compiled by blinded personnel	Exercise advice: verbal and written	Average frequency of 30 min walks to near pain undertaken Supervised exercise <3 times 2 three times 8 >3 times 19 Advice <3 times 9 three times 11 >3 times 9	A larger number of people in the exercise class group reported to be walking either three times a week or >3 times a week (P=0.012)
Duncan and Pozehl [47]	Heart failure Mean age 66.4 years Gender Male=87.5%	Individualised graphic feedback on exercise goals, participation and problem solving	Social learning theory	Exercise diaries—number of sessions completed	Exercise programme without adherence intervention	Exercise sessions completed 12 weeks Control 59.3 (SD 11.1) Intervention 62.3 (SD 6.4) 24 weeks Control 41.2 (SD 9.7) Intervention 59.6 (SD 10.6)* *Significant difference (P<0.01)	The adherence intervention can increase exercise sessions completed after finishing a supervised exercise programme in patients with heart failure
Gallagher [49]	Physical therapy patients with low back, hip or knee symptoms Mean age years (SD) 69.3 (6.87) Gender Male=28.3%	Printed messages and magnets underpinned by socioemotional selectivity theory 1) Emotional and meaningful message 2) Factual and information message	Socioemotional selectivity theory	Self-reported adherence to their exercise programme (used to calculated adherence score)	One message group compared with the other	Average adherence score % Emotional 60% (SD 34.4%) Fact 55.3% (SD 34.0%) (No significant difference between groups)	No significant difference found in participants' adherence between the message groups
Gardner *et al*[50]	Intermittent claudication Mean age years (SD) Control 65 (10) Supervised ex 66 (12) Home ex 65 (11) Gender Male=47.9%	Supervised vs unsupervised exercise 1) Home exercise (no supervision) for 12 weeks+step activity monitor 2) Supervised exercise for 12 weeks+step activity monitor	None given	Total exercise sessions, using step activity monitor and exercise log book	Encouraged to walk more on their own	Total exercise sessions completed % Supervised group 84.8 (SD 20.9) Home group 82.5 (SD 27.7) (No significant difference between groups)	The relatively high adherence rate in home-based exercise was similar to that found with the supervised exercise group

Continued

Table 1 Continued

Author	Population	Adherence intervention	Behavioural theory	Measure of adherence	Control or comparison	Results	Conclusion
Ridgel *et al*[53]	Parkinson's disease and depression Mean age years (SD) 70.2 (7.9) Gender Male=63.3%	Psychoeducation, peer education/support, group exercise (Enhanced EXerCisE thErapy (EXCEED) group)	None given	Number of exercise sessions attended, recorded by a research assistant. Those performing outcome measures were blinded to group assignment	Self-guided psychoeducation and exercise (SGE). No group interactions or peer education	Number of exercise sessions attended at 12 weeks EXCEED=20.7 (SD 8.1) SGE=22.0 (SD 8.0)	Both groups attended a similar number of exercise sessions
Schneider *et al*[42]	Older adults who engage in aerobic or strengthening exercise <3 or more days a week Mean age years (SD) 71.8 (5.1) Gender Male=24.1%	Cognitive behavioural therapy (CBT) 1) CBT group 2) Attention-control education group	CBT	Time spent exercising in the past month. Exercise behaviour was assessed by a research assistant blinded to group allocation	Control group—no CBT, no education group	Time spent exercising Strengthening exercises (h) 3 months to 6 months CBT 1.0 (SD 0.8) 1.0 (0.7) Education 1.1 (SD 0.8) 1.0 (0.7) Control 1.3 (SD 1.3) 1.2 (1.3) 9 months 12 months CBT 1.0 (1.2) 0.9 (1.0) Education 1.2 (2.3) 1.2 (2.4) Control 1.0 (1.0) 1.1 (1.1) (No significant difference between groups)	No significant difference with time spent exercising between groups
Schoo *et al*[41]	Osteoarthritis of the hip and/or knee Mean age Years (SD) Brochure 71.1 (6.83) Audio 70.9 (7.23) Video 69.2 (6.36) Gender Male=33%	Exercise programme instruction method: 1) Brochure+audio tape 2) Brochure+video tape	None given	Home exercise log sheets	Brochure-only group	Home exercise adherence (median) 1–4 weeks: Brochure 93% Video 92% Audio 89% 5–8 weeks: Brochure 89.5% Video 81.5% Audio 87% (No significant differences)	Audio and video tapes given in addition to an exercise brochure, did not show an increase in adherence compared with the brochure only group
Steele *et al*[46]	Chronic lung disease Mean age 67 years Gender Male=92.5%	Weekly phone calls and one home visit over 3 months Consisting of dealing with queries about exercise adherence, problem solving, exercise maintenance, recommendations about health problems, encouragement, evaluated home safety, assistance in establishing an individualised exercise routine. Receiving a digital pedometer and exercise handbook	None given	Exercise diary—total minutes of exercise	Continued care from referring provider. Recommendation for continuation of the exercise programme. Invited to attend the lung club group sessions.	Minutes of exercise Pre-intervention Control 14 (SD 14) Intervention 21 (SD 19)* Postintervention Control 28 (SD 21) Intervention 30 (SD 32) 20 weeks Control 16 (SD 19) Intervention 32 (SD 46)* 1 year Control 22 (SD 25) Intervention 33 (SD 36) *Significant differences (P<0.05)	The adherence intervention gave limited improvement in the short term regarding self-reported maintenance of exercise after pulmonary rehabilitation in highly sedentary chronic lung disease patients. No long-term benefit was found

Continued

Table 1 Continued

Author	Population	Adherence intervention	Behavioural theory	Measure of adherence	Control or comparison	Results	Conclusion
Wu *et al*[40]	People at risk of falling Mean age Years (SD) Tele ex 76.1 (7.9) Comm ex 74.1 (6.9) Home ex 75.9 (6.3) Gender Male=15.6%	Method of delivering exercise programme: 1) Instructor lead video call at home 2) Instructor lead community-based group	None given	Log sheets -Number of sessions -Time exercising	Home exercise with a digital versatile disc (DVD)	Total time exercising (h) Tele 30 (SD 12) Comm 31 (SD 12) Home 17 (SD 17) Attendance rate (%) Tele 69 (SD 27) Comm 72 (SD 27) Home 38 (SD 46) (Tele and comm significantly higher for time exercising and attendance rate (both P<0.01)	Compared with home exercise, tele ex (video conferencing) and comm ex (community class) were better for total time spent exercising and number of exercise sessions completed
Yates *et al*[51]	Postcardiac rehabilitation Mean age years (SD) 66.7 (9.4) Gender Male=69%	Booster sessions, structured education and counselling given 1) Over the phone 2) In clinic	Self-efficacy theory	Considered adherent if they had performed exercise ≥3 times a week	Usual care—one telephone call at 4–6 weeks	Adherence rate at 3 months Control 50% Clinic 70% Phone 75% Adherence rate at 6 months Control 50% Clinic 40% Phone 63% (No significant difference were found between groups)	Adherence to the recommended exercise programme was greater in the two treatment groups compared with usual care, but differences were not significant

*Significant (P<0.05) difference between control and intervention.

between number of exercise sessions completed between a group that received the intervention and a control group at 24 weeks (P<0.01).

Another study[40] with a high risk of bias, compared adherence with a Tai Chi exercise programme delivered through an interactive telecommunication approach, or a class in a community centre, compared with exercising at home with a digital versatile disc (DVD) for people at risk of falling. They found the telecommunication and community-based groups had significantly higher results for time exercising and attendance rate compared with home exercise (P<0.01).

A further study[52] with a high risk of bias, compared a weekly exercise and motivation classes lasting 6 months against written and verbal exercise advice, for those with intermittent claudication. Participants were advised to walk at least three times a week to near maximal pain. At 6 months, there were more participants in the intervention group who reported to be walking either three times a week, or more than three times a week, in comparison to the advice group (P<0.012).

Social support

A study[46] with a high risk of bias, tested an adherence intervention for participants with chronic lung disease. The intervention included weekly phone calls and one home visit over a 3-month period. The phone calls and visit included dealing with queries about exercise adherence and exercise maintenance, problem solving, discussion and recommendations about health problems and encouragement. With home visits that evaluated home safety and helped establish an individualised exercise routine. This study found a short-term difference in minutes of exercise undertaken, between the intervention and a control at 20 weeks (P<0.05). Although this difference was absent at 1 year follow-up.

A second study with a high risk of bias[48] looked at guidance and supervision for the frail elderly, testing the difference between a high guidance group and a medium guidance group. They found no difference between groups for percentage of exercise sessions undertaken.

A third study with a high risk of bias[53] compared a group that received psychoeducation, peer support and group exercise, with a group that undertook self-guided psychoeducation and exercise. They found that both groups attended a similar number of exercise sessions at 12 weeks.

A study[50] with a moderate risk of bias, investigated supervised exercise versus home-based exercise with no supervision. They found no significant difference between groups with regard to total exercise sessions completed (P=0.712).

Natural consequences

One study[49] with moderate risk of bias provided two different types of adherence messages based on Socio-emotional Selectivity Theory,[54] one message emphasised emotionally meaning reasons to exercise, for example, spending time with loved ones. The other message emphasised knowledge-related goals, for example, stronger muscles. No significant difference was found in an average adherence score between the two groups 2 weeks after discharge from physical therapy (P=0.03).

Identity

One study[42] with a high risk of bias, compared cognitive behavioural therapy (CBT), an attention control education group and a control group. The primary emphasis of the CBT was to teach older people to recognise and modify their thoughts or interpretations about exercise. They found no significant difference in time spent exercising between groups at 3, 6, 9 and 12 months.

Goals and planning

A study[51] with a high risk of bias, compared a structured educational counselling booster session, given over the phone, or face to face, compared with usual care. During the booster sessions, participant's individualised goals were used as a basis for intervening. Where participants were progressing towards goal achievement they received praise, and were encouraged to attribute their accomplishment to their own ability. In addition, discussion of factors inhibiting achievement of goals took place. They found no significant differences in adherence rates between groups at 3 and 6 months.

Behavioural theories

Four studies used behavioural theories to justify their chosen intervention. Duncan and Pozehl[47] delivered an intervention which offered individual graphic feedback, related to the exercise goal. This was underpinned by Social Learning Theory, a theory in which Bandura suggested that people can learn through observation of others, their behaviour and the outcomes of their behaviour.[55] Gallagher's[49] intervention used two different types of adherence messages, one message emphasised emotionally meaning reasons to exercise, while the other message emphasised knowledge related goals. This was based on Socioemotional Selectivity theory,[54] a theory that posits that time effects the pursuit of social goals. Social motives can fall into those that deal with the acquisition of knowledge, or those that relate to regulation of emotion. Once time is perceived as limited emotional goals take priority over knowledge acquisition. Schneider *et al*[42] used a CBT intervention. CBT works on the principle that thoughts, emotions, physical feelings, situations and actions are connected, CBT aims to help people break down any negative thought cycles.[56] Finally, Yates *et al*[51] used booster sessions delivered over the phone or face to face. Bandura's self-efficacy[57] was used to inform their intervention. Self-efficacy refers to the magnitude of a person's belief in their ability to undertake a task and achieve a desired goal. Seven studies did not cite a specific behavioural theory to justify their intervention. Of the studies which reported a behavioural theory, one reported a significant improvement in exercise adherence, this was Duncan and Pozehl.[47]

DISCUSSION

This review investigated interventions tested in RCTs to improve exercise adherence in older adults. Interventions were categorised using the BCTT.[33] Interventions categorised in the feedback and monitoring group demonstrated positive results for exercise adherence, although risk of bias limits generalisability of these results. The inconclusive results mirror similar results to adherence prompting interventions in other populations. Peek *et al*[35] investigated interventions to support adherence to physiotherapy prescribed self-management strategies, they found that although some interventions had a positive impact on adherence, there was insufficient data to recommend their use clinically. Another review by McLean *et al*[58] investigated interventions to improve adherence to musculoskeletal physiotherapy treatment, they found moderate evidence that a motivational cognitive behavioural programme is effective at enhancing attendance to clinic sessions which were exercise based, but conflicting evidence that adherence approaches improve short-term exercise adherence, and strong evidence that adherence interventions were not effective at enhancing long-term exercise adherence. Although it has previously been found that there is evidence that interventions can improve exercise adherence in disease-specific populations which are relevant to older people, such as osteoarthritis and rheumatoid arthritis,[36] back pain and hip and knee osteoarthritis.[37]

This review provides a synthesis of evidence specifically for older patients, without considering a specific condition. Four papers reported positive results. Three of these interventions were categorised in the feedback and monitoring category. Namely, exercise delivered by telecommunication or in a community class setting,[40] supervised exercise and motivation classes[52] and graphic feedback delivered by a healthcare professional.[47] Interestingly, the method of feedback or monitoring differed across these studies, yet they all demonstrated positive results. Although the limitations of these studies reduce their generalisability, there may be scope for further investigation in this area. It may be that monitoring and feedback interventions can help to overcome some of the barriers to exercise adherence in older people, such as low self-efficacy and motivation,[27] or help to facilitate exercise adherence, for example, it has been previously reported that adherence is generally better in programmes with supervision.[21] Peek *et al*[35] reported that activity monitoring and feedback systems may help to promote adherence. This is in line with evidence from other populations.[59 60] It has been reported that feedback may improve adherence to an exercise programme, for adults with borderline hypertension.[59] Feedback is also of use in areas such as self-care in those with diabetes.[60] One of the common factors that these interventions possess is that the number of contacts with healthcare professionals is greater than the control. This is reflective of work in other areas where number of contacts can affect behaviour change, such as with using exercise advice to treat young adults with prehypertension and hypertension.[61]

Prescribed exercise is a prominent treatment option, which is likely to be used further as people live longer,[62] with more likelihood of LTCs.[26] Strategies to promote adherence should therefore remain an important factor for those who prescribe exercise. An area in which exercise adherence research could move forwards would be to consider the theory that underpins interventions. Measuring adherence is essentially measuring behaviour change in participants, that is, the participant's behaviour corresponding to recommendation from a healthcare provider,[16] in this case following an exercise programme. It is interesting that seven of the studies included did not appear to have used any behavioural theory. This could have potential impact on the effectiveness of interventions. An important aspect in developing complex interventions, as outlined by the MRC's guidance is using the best available evidence and appropriate theory.[63] If adherence interventions lack theoretical underpinning, then the chances of successfully changing people's behaviour may be limited. It may also affect the ability to appropriately categorise and replicate interventions.

Even where behavioural approaches are considered, there may still be room for further consideration. Michie *et al*[64] developed a framework for behaviour change interventions, The Behaviour Change Wheel. This model posits that the three crucial components to behaviour change are capability, opportunity and motivation. Interventions may need to target one, two or even all three components to facilitate change. Approaches targeting only one area may not result in the desired change in behaviour. For example, giving information may target capability, while having no effect on opportunity or motivation. Well-developed interventions underpinned by appropriate theory, are likely to maximise the potential for behaviour change, in this case adherence to prescribed exercise.

One of the challenges to research in the field of exercise adherence is measuring adherence itself. It has previously been reported that numerous methods for reporting exercise adherence exist, however, on the whole there is a lack of measures with reported validity and reliability.[65] This is in line with the results of this review. The papers included used a diverse range of adherence outcome measures. Robust outcome measures would offer greater confidence in the effects of interventions, in addition to making the comparison of interventions and meta-analysis more straightforward. A further consideration in the area of exercise adherence interventions is that of contextual equivalence of intervention and control groups. Bishop *et al*[66] reviewed the contextual effects and behaviour change techniques of both control and target interventions, in trials from a Cochrane review of physical activity. They conclude that a broad range of control interventions are used in this field. This in turn may influence effect size, due to the different behaviour change techniques that are included within the numerous different control interventions. It is important that future work considers

the contextual equivalence of control and intervention groups in the area of exercise adherence, such as considered in the review by Nicolson *et al*.[37]

Strengths and limitations

This review systematically searched the literature with clear inclusion and exclusion criteria using an appropriate risk of bias assessment tool. It also used a predefined BCTT allowing the categorisation and evaluation of interventions. Limitations of this review include the moderate to high risk of bias of the studies, in particular due to small sample sizes leading to underpowered studies. Also, it was not possible to perform meta-analyses due to the heterogeneous nature of the interventions and measurements of adherence, it is also known that there is a lack of well-developed measures of adherence for therapeutic exercise,[65] making it more challenging to capture the effect of adherence interventions. Another consideration is that although it was beyond the scope of this review to analyse health outcomes, adherence promoting interventions do need to be considered in the context of these health outcome results. For it is appropriate to ensure that intervening to promote adherence also offers an improvement in health outcome, or at least causes no harm. Finally, only papers published in English were considered for this review. It is possible that there are studies published in languages other than English that would have changed the results of the review.

Future research

Interventions that focused on feedback and monitoring demonstrated significant results. However, these types of intervention need to be tested in appropriately powered trials. Second, there is need for the development of adherence interventions underpinned by appropriate theory. Finally, there is need for robust adherence measures that are valid and reliable to be developed, in order to adequately assess the effectiveness of interventions.

CONCLUSION

This review provides an overview of interventions to improve exercise adherence in older people. Interventions grouped in the feedback and monitoring category of the BCTT demonstrated positive effects on exercise adherence, although risk of bias limits the generalisability of these approaches. There is need for better reporting, use and the development of theoretically derived interventions in the field of exercise adherence for older people. Robust measures of adherence, in order to adequately test these interventions would also be of use.

Acknowledgements The authors acknowledge Mary Boulton who provided supervision to JR through the study.

Contributors JR conceived the study and was responsible for study design and search strategy, JR and EH were responsible for data extraction, quality assessment and data analysis. KB and HD provided methodological advice. JR drafted the manuscript, this was revised with input from EH, KB and HD. All authors approved the final version.

Funding This work was supported by the Centre for Movement and Occupational Rehabilitation Science (MOReS) Oxford Brookes University, and the Physiotherapy Research Unit, Nuffield Orthopaedic Centre, Oxford .

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement There are no additional data available for this review.

REFERENCES

1. Hoffmann TC, Maher CG, Briffa T, *et al*. Prescribing exercise interventions for patients with chronic conditions. *CMAJ* 2016;188:510–8.
2. Naci H, Ioannidis JP. Comparative effectiveness of exercise and drug interventions on mortality outcomes: metaepidemiological study. *BMJ* 2013;347:f5577.
3. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;100:126–31.
4. Chartered Society of Physiotherapy. What is physiotherapy? 2013. http://www.csp.org.uk/your-health/what-physiotherapy (accessed 18 Jun 2015).
5. Chartered Society of Physiotherapy. NHS in England reports rise in physio numbers. 2015 http://www.csp.org.uk/news/2015/04/09/nhs-england-reports-rise-physio-numbers
6. Chartered Society of Physiotherapy. Physiotherapy outpatient services survey 2012. *London* 2013.
7. Artz N, Dixon S, Wylde V, *et al*. Physiotherapy provision following discharge after total hip and total knee replacement: a survey of current practice at high-volume NHS hospitals in England and wales. *Musculoskeletal Care* 2013;11:31–8.
8. Rushton A, Wright C, Heap A, *et al*. Survey of current physiotherapy practice for patients undergoing lumbar spinal fusion in the United Kingdom. *Spine* 2014;39:E1380–7.
9. Palmer S, Cramp F, Lewis R, *et al*. Diagnosis, management and assessment of adults with joint hypermobility syndrome: a UK-wide survey of physiotherapy practice. *Musculoskeletal Care* 2015;13:101–11.
10. Grieve R, Palmer S. Physiotherapy for plantar fasciitis: a UK-wide survey of current practice. *Physiotherapy* 2017;103:193–200.
11. Smith TO, Chester R, Clark A, *et al*. A national survey of the physiotherapy management of patients following first-time patellar dislocation. *Physiotherapy* 2011;97:327–38.
12. Sluijs EM, Kok GJ, van der Zee J. Correlates of exercise compliance in physical therapy. *Phys Ther* 1993;73:771–82. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=8234458&retmode=ref&cmd=prlinks%5Cnpapers2://publication/uuid/B42C2D96-9561-42D1-89E0-C5621F8BF246
13. Alexandre NM, Nordin M, Hiebert R, *et al*. Predictors of compliance with short-term treatment among patients with back pain. *Rev Panam Salud Publica* 2002;12:86–95.
14. Pisters MF, Veenhof C, Schellevis FG, *et al*. Exercise adherence improving long-term patient outcome in patients with osteoarthritis of the hip and/or knee. *Arthritis Care Res* 2010;62:1087–94.
15. van Gool CH, Penninx BW, Kempen GI, *et al*. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. *Arthritis Rheum* 2005;53:24–32.
16. World Health Organisation. *Adherence to long-term therapies*. Geneva 2003.
17. van Dulmen S, Sluijs E, van Dijk L, *et al*. Patient adherence to medical treatment: a review of reviews. *BMC Health Serv Res* 2007;7:55.
18. Nieuwlaat R, Wilczynski N, Navarro T, *et al*. Interventions for enhancing medication adherence (Review). *Cochrane Database Syst Rev* 2014;11:CD000011.
19. Schroeder K, Fahey T, Ebrahim S. Interventions for improving adherence to treatment in patients with high blood pressure

in ambulatory settings. *Cochrane Database Syst Rev* 2004;3:CD004804.
20. Essery R, Geraghty AW, Kirby S, *et al*. Predictors of adherence to home-based physical therapies: a systematic review. *Disabil Rehabil* 2017;39:519–34.
21. Picorelli AM, Pereira LS, Pereira DS, *et al*. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. *J Physiother* 2014;60:151–6.
22. Jack K, McLean SM, Moffett JK, *et al*. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. *Man Ther* 2010;15:220–8.
23. Resnick B, D'Adamo C, Shardell M, *et al*. Adherence to an exercise intervention among older women post hip fracture. *J Clin Sport Psychol* 2008;2:41–56. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2859720&tool=pmcentrez&rendertype=abstract
24. Slovinec D'Angelo ME, Pelletier LG, Reid RD, *et al*. The roles of self-efficacy and motivation in the prediction of short- and long-term adherence to exercise among patients with coronary heart disease. *Health Psychol* 2014;33:1344–53.
25. Palazzo C, Klinger E, Dorner V, *et al*. Barriers to home-based exercise program adherence with chronic low back pain: Patient expectations regarding new technologies. *Ann Phys Rehabil Med* 2016;59:107–13.
26. Department of Health. Long term conditions compendium of information. *Leeds* 2012 http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4100717%5Cnhttp://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publications
27. Hill AM, Hoffmann T, McPhail S, *et al*. Factors associated with older patients' engagement in exercise after hospital discharge. *Arch Phys Med Rehabil* 2011;92:1395–403.
28. Forkan R, Pumper B, Smyth N, *et al*. Exercise adherence following physical therapy intervention in older adults with impaired balance. *Phys Ther* 2006;86:401–10 http://www.ncbi.nlm.nih.gov/pubmed/16506876
29. Fairhall N, Sherrington C, Cameron ID, *et al*. A multifactorial intervention for frail older people is more than twice as effective among those who are compliant: complier average causal effect analysis of a randomised trial. *J Physiother* 2017;63:40–4.
30. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. *Prev Med* 2016;87:155–66.
31. NICE. *Behaviour change: individual approaches*. London 2014.
32. Michie S, Atkins L, West R. The behaviour change wheel: a guide to designing interventions. *Silverback Publishing* 2014.
33. Michie S, Richardson M, Johnston M, *et al*. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;46:81–95.
34. Jordan JL, Holden MA, Mason EE, *et al*. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database Syst Rev* 2010;1:CD005956.
35. Peek K, Sanson-Fisher R, Mackenzie L, *et al*. Interventions to aid patient adherence to physiotherapist prescribed self-management strategies: a systematic review. *Physiotherapy* 2016;102:127–35.
36. Ezzat AM, MacPherson K, Leese J, *et al*. The effects of interventions to increase exercise adherence in people with arthritis: a systematic review. *Musculoskeletal Care* 2015;13:1–18.
37. Nicolson PJA, Bennell KL, Dobson FL, *et al*. Interventions to increase adherence to therapeutic exercise in older adults with low back pain and/or hip/knee osteoarthritis: a systematic review and meta-analysis. *Br J Sports Med* 2017;51:791–9.
38. Burton E, Farrier K, Hill KD, *et al*. Effectiveness of peers in delivering programs or motivating older people to increase their participation in physical activity: Systematic review and meta-analysis. *J Sports Sci* 2017;25:1–13.
39. Williamson W, Kluzek S, Roberts N, *et al*. Behavioural physical activity interventions in participants with lower-limb osteoarthritis: a systematic review with meta-analysis. *BMJ Open* 2015;5:e007642.
40. Wu G, Keyes L, Callas P, *et al*. Comparison of telecommunication, community, and home-based Tai Chi exercise programs on compliance and effectiveness in elders at risk for falls. *Arch Phys Med Rehabil* 2010;91:849–56.
41. Schoo AMM, Morris ME, Bui QM. The effects of mode of exercise instruction on compliance with a home exercise program in older adults with osteoarthritis. *Physiotherapy* 2005;91:79–86.
42. Schneider JK, Cook JH, Luke DA. Unexpected effects of cognitive-behavioural therapy on self-reported exercise behaviour and functional outcomes in older adults. *Age Ageing* 2011;40:163–8.
43. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Reprinted from Annals of Internal Medicine). *Phys Ther* 2009;89:873–80.
44. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–e34.
45. Higgins JP, Altman DG. Assessing risk of bias in included studies. In: Higgins JP, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley-Blackwell, 2008:187–241.
46. Steele BG, Belza B, Cain KC, *et al*. A randomized clinical trial of an activity and exercise adherence intervention in chronic pulmonary disease. *Arch Phys Med Rehabil* 2008;89:404–12.
47. Duncan K, Pozehl B. Effects of an exercise adherence intervention on outcomes in patients with heart failure. *Rehabil Nurs* 2003;28:117–22. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emed6&AN=12875144%5Cnhttp://sfx.scholarsportal.info/uhn?sid=OVID:embase&id=pmid:12875144&id=&issn=0278-4807&isbn=&volume=28&issue=4&spage=117&pages=117-122&date=2003&title=Rehabilita
48. Boshuizen HC, Stemmerik L, Westhoff MH, *et al*. The effects of physical therapists' guidance on improvement in a strength-training program for the frail elderly. *J Aging Phys Act* 2005;13:5–22.
49. Gallagher KM. Helping older adults sustain their physical therapy gains. *J Geriatr Phys Ther* 2016;39:20–9.
50. Gardner AW, Parker DE, Montgomery PS, *et al*. Efficacy of quantified home-based exercise and supervised exercise in patients with intermittent claudication: a randomized controlled trial. *Circulation* 2011;123:491–8.
51. Yates BC, Anderson T, Hertzog M, *et al*. Effectiveness of follow-up booster sessions in improving physical status after cardiac rehabilitation: health, behavioral, and clinical outcomes. *Appl Nurs Res* 2005;18:59–62.
52. Cheetham DR, Burgess L, Ellis M, *et al*. Does supervised exercise offer adjuvant benefit over exercise advice alone for the treatment of intermittent claudication? A randomised trial. *Eur J Vasc Endovasc Surg* 2004;27:17–23.
53. Ridgel AL, Walter BL, Tatsuoka C, *et al*. Enhanced Exercise Therapy in Parkinson's disease: a comparative effectiveness trial. *J Sci Med Sport* 2016;19:12–17.
54. Carstensen LL, Isaacowitz DM, Charles ST. Taking time seriously. A theory of socioemotional selectivity. *Am Psychol* 1999;54:165–81.
55. Bandura A. Social learning theory. *Soc Learn Theory* 1971;1:46.
56. NHS Choices. Cognitive Behavioural Therapy (CBT). 2016 http://www.nhs.uk/Conditions/Cognitive-behavioural-therapy/Pages/Introduction.aspx (accessed 15 Jan 2017).
57. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977;84:191–215.
58. McLean SM, Burton M, Bradley L, *et al*. Interventions for enhancing adherence with physiotherapy: a systematic review. *Man Ther* 2010;15:514–21.
59. Shakudo M, Takegami M, Shibata A, *et al*. Effect of feedback in promoting adherence to an exercise programme: a randomized controlled trial. *J Eval Clin Pract* 2011;17:7–11.
60. Polonsky WH, Fisher L. When does personalized feedback make a difference? A narrative review of recent findings and their implications for promoting better diabetes self-care. *Curr Diab Rep* 2015;15:50.
61. Williamson W, Foster C, Reid H, *et al*. Will exercise advice be sufficient for treatment of young adults with prehypertension and hypertension? A systematic review and meta-analysis. *Hypertension* 2016;68:78–87.
62. Suzman R, Beard J. *Global Health and Aging*, US Department of State. 2011:1–32. [Online]. papers2://publication/uuid/150A301A-0A61-489C-80D4-3C889BE672E2
63. Craig P, Dieppe P, Macintyre S, *et al*. *Developing and evaluating complex interventions: New Guidance*. London, 2008. http://eprints.ucl.ac.uk/103060/ (accessed 30 Jul 2014).
64. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;6:42.
65. Bollen JC, Dean SG, Siegert RJ, *et al*. A systematic review of measures of self-reported adherence to unsupervised home-based rehabilitation exercise programmes, and their psychometric properties. *BMJ Open* 2014;4:e005044.
66. Bishop FL, Fenge-Davies AL, Kirby S, *et al*. Context effects and behaviour change techniques in randomised trials: a systematic review using the example of trials to increase adherence to physical activity in musculoskeletal pain. *Psychol Health* 2015;30:104–21.

Cochrane Library

Cochrane Database of Systematic Reviews

Interventions for preventing falls in older people in care facilities and hospitals (Review)

Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, Kerse N

Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, Kerse N.
Interventions for preventing falls in older people in care facilities and hospitals.
Cochrane Database of Systematic Reviews 2018, Issue 9. Art. No.: CD005465.
DOI: 10.1002/14651858.CD005465.pub4.

www.cochranelibrary.com

TABLE OF CONTENTS

[Intervention Review]

Interventions for preventing falls in older people in care facilities and hospitals

Ian D Cameron[1], Suzanne M Dyer[2], Claire E Panagoda[3], Geoffrey R Murray[4], Keith D Hill[5], Robert G Cumming[6], Ngaire Kerse[7]

[1]John Walsh Centre for Rehabilitation Research, Kolling Institute, Northern Sydney Local Health District, St Leonards, Australia. [2]DHATR Consulting, Parkside, Australia. [3]John Walsh Centre for Rehabilitation Research, Kolling Institute, Northern Sydney Local Health District, St Leonards, Australia. [4]Aged Care, Rehabilitation and Palliative Care, Illawarra Shoalhaven Local Health District, Warrawong, Australia. [5]School of Physiotherapy and Exercise Science, Faculty of Health Sciences, Curtin University, Perth, Australia. [6]School of Public Health, Sydney Medical School, University of Sydney, Sydney, Australia. [7]Department of General Practice and Primary Health Care, University of Auckland, Auckland, New Zealand

Contact address: Ian D Cameron, John Walsh Centre for Rehabilitation Research, Kolling Institute, Northern Sydney Local Health District, St Leonards, NSW, 2065, Australia. ian.cameron@sydney.edu.au.

Editorial group: Cochrane Bone, Joint and Muscle Trauma Group.
Publication status and date: Edited (no change to conclusions), published in Issue 9, 2018.

Citation: Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, Kerse N. Interventions for preventing falls in older people in care facilities and hospitals. *Cochrane Database of Systematic Reviews* 2018, Issue 9. Art. No.: CD005465. DOI: 10.1002/14651858.CD005465.pub4.

ABSTRACT

Background

Falls in care facilities and hospitals are common events that cause considerable morbidity and mortality for older people. This is an update of a review first published in 2010 and updated in 2012.

Objectives

To assess the effects of interventions designed to reduce the incidence of falls in older people in care facilities and hospitals.

Search methods

We searched the Cochrane Bone, Joint and Muscle Trauma Group Specialised Register (August 2017); Cochrane Central Register of Controlled Trials (2017, Issue 8); and MEDLINE, Embase, CINAHL and trial registers to August 2017.

Selection criteria

Randomised controlled trials of interventions for preventing falls in older people in residential or nursing care facilities, or hospitals.

Data collection and analysis

One review author screened abstracts; two review authors screened full-text articles for inclusion. Two review authors independently performed study selection, 'Risk of bias' assessment and data extraction. We calculated rate ratios (RaR) with 95% confidence intervals (CIs) for rate of falls and risk ratios (RRs) and 95% CIs for outcomes such as risk of falling (number of people falling). We pooled results where appropriate. We used GRADE to assess the quality of evidence.

Main results

Thirty-five new trials (77,869 participants) were included in this update. Overall, we included 95 trials (138,164 participants), 71 (40,374 participants; mean age 84 years; 75% women) in care facilities and 24 (97,790 participants; mean age 78 years; 52% women) in hospitals. The majority of trials were at high risk of bias in one or more domains, mostly relating to lack of blinding. With few exceptions, the quality of evidence for individual interventions in either setting was generally rated as low or very low. Risk of fracture and adverse events were generally poorly reported and, where reported, the evidence was very low-quality, which means that we are uncertain of the estimates. Only the falls outcomes for the main comparisons are reported here.

Care facilities

Seventeen trials compared exercise with control (typically usual care alone). We are uncertain of the effect of exercise on rate of falls (RaR 0.93, 95% CI 0.72 to 1.20; 2002 participants, 10 studies; I^2 = 76%; very low-quality evidence). Exercise may make little or no difference to the risk of falling (RR 1.02, 95% CI 0.88 to 1.18; 2090 participants, 10 studies; I^2 = 23%; low-quality evidence).

There is low-quality evidence that general medication review (tested in 12 trials) may make little or no difference to the rate of falls (RaR 0.93, 95% CI 0.64 to 1.35; 2409 participants, 6 studies; I^2 = 93%) or the risk of falling (RR 0.93, 95% CI 0.80 to 1.09; 5139 participants, 6 studies; I^2 = 48%).

There is moderate-quality evidence that vitamin D supplementation (4512 participants, 4 studies) probably reduces the rate of falls (RaR 0.72, 95% CI 0.55 to 0.95; I^2 = 62%), but probably makes little or no difference to the risk of falling (RR 0.92, 95% CI 0.76 to 1.12; I^2 = 42%). The population included in these studies had low vitamin D levels.

Multifactorial interventions were tested in 13 trials. We are uncertain of the effect of multifactorial interventions on the rate of falls (RaR 0.88, 95% CI 0.66 to 1.18; 3439 participants, 10 studies; I^2 = 84%; very low-quality evidence). They may make little or no difference to the risk of falling (RR 0.92, 95% CI 0.81 to 1.05; 3153 participants, 9 studies; I^2 = 42%; low-quality evidence).

Hospitals

Three trials tested the effect of additional physiotherapy (supervised exercises) in rehabilitation wards (subacute setting). The very low-quality evidence means we are uncertain of the effect of additional physiotherapy on the rate of falls (RaR 0.59, 95% CI 0.26 to 1.34; 215 participants, 2 studies; I^2 = 0%), or whether it reduces the risk of falling (RR 0.36, 95% CI 0.14 to 0.93; 83 participants, 2 studies; I^2 = 0%).

We are uncertain of the effects of bed and chair sensor alarms in hospitals, tested in two trials (28,649 participants) on rate of falls (RaR 0.60, 95% CI 0.27 to 1.34; I^2 = 0%; very low-quality evidence) or risk of falling (RR 0.93, 95% CI 0.38 to 2.24; I^2 = 0%; very low-quality evidence).

Multifactorial interventions in hospitals may reduce rate of falls in hospitals (RaR 0.80, 95% CI 0.64 to 1.01; 44,664 participants, 5 studies; I^2 = 52%). A subgroup analysis by setting suggests the reduction may be more likely in a subacute setting (RaR 0.67, 95% CI 0.54 to 0.83; 3747 participants, 2 studies; I^2 = 0%; low-quality evidence). We are uncertain of the effect of multifactorial interventions on the risk of falling (RR 0.82, 95% CI 0.62 to 1.09; 39,889 participants; 3 studies; I^2 = 0%; very low-quality evidence).

Authors' conclusions

In care facilities: we are uncertain of the effect of exercise on rate of falls and it may make little or no difference to the risk of falling. General medication review may make little or no difference to the rate of falls or risk of falling. Vitamin D supplementation probably reduces the rate of falls but not risk of falling. We are uncertain of the effect of multifactorial interventions on the rate of falls; they may make little or no difference to the risk of falling.

In hospitals: we are uncertain of the effect of additional physiotherapy on the rate of falls or whether it reduces the risk of falling. We are uncertain of the effect of providing bed sensor alarms on the rate of falls or risk of falling. Multifactorial interventions may reduce rate of falls, although subgroup analysis suggests this may apply mostly to a subacute setting; we are uncertain of the effect of these interventions on risk of falling.

PLAIN LANGUAGE SUMMARY

Interventions for preventing falls in older people in care facilities and hospitals

Review question

How effective are interventions designed to reduce falls in older people in care facilities and hospitals?

Background

Falls by older people in care facilities, such as nursing homes, and hospitals are common events that may cause loss of independence, injuries, and sometimes death as a result of injury. Effective interventions to prevent falls are therefore important. Many types of interventions are in use. These include exercise, medication interventions that include vitamin D supplementation and reviews of the drugs that people are taking, environment or assistive technologies including bed or chair alarms or the use of special (low/low) beds, social environment interventions that target staff members and changes in the organisational system, and knowledge interventions. A special type of intervention is the multifactorial intervention, where the selection of single interventions such as exercise and vitamin D supplementation is based on an assessment of a person's risk factors for falling. Falls are reported in two ways in our review. One outcome is rate of falls, which is the number of falls. The other outcome is risk of falling, which is the number of people who had one or more falls.

Search date

We searched the healthcare literature for reports of randomised controlled trials relevant to this review up to August 2017.

Study characteristics

This review included 95 randomised controlled trials involving 138,164 participants. Seventy-one trials (40,374 participants) were in care facilities, and 24 (97,790 participants) in hospitals. On average, participants were 84 years old in care facilities and 78 years old in hospitals. In care facilities, 75% were women and in hospitals, 52% were women.

Quality of the evidence

The majority of trials were at high risk of bias, mostly relating to lack of blinding. With few exceptions, the quality of evidence for individual interventions in either setting was generally rated as low or very low. Risk of fracture and adverse events were generally poorly reported and, where reported, the evidence was very low quality, which means that we are uncertain of the estimates.

Key results

There was evidence, often from single studies, for a wide range of interventions used for preventing falls in both settings. However, in the following we summarise only the falls outcomes for four key interventions in care facilities and three key interventions in hospitals.

Care facilities

We are uncertain of the effect of exercise on the rate of falls (very low-quality evidence) and it may make little or no difference to the risk of falling (low-quality evidence).

General medication review may make little or no difference to the rate of falls (low-quality evidence) or the risk of falling (low-quality evidence).

Prescription of vitamin D probably reduces the rate of falls (moderate-quality evidence) but probably makes little or no difference to the risk of falling (moderate-quality evidence). The population included in these studies appeared to have low vitamin D levels.

We are uncertain of the effect of multifactorial interventions on the rate of falls (very low-quality evidence). They may make little or no difference to the risk of falling (low-quality evidence).

Hospitals

We are uncertain whether physiotherapy aimed specifically at reducing falls in addition to usual rehabilitation in the ward has an effect on the rate of falls or reduces the risk of falling (very low-quality evidence).

We are uncertain of the effect of bed alarms on the rate of falls or risk of falling (very low-quality evidence).

Multifactorial interventions may reduce the rate of falls, although this is more likely in a rehabilitation or geriatric ward setting (low-quality evidence). We are uncertain of the effect of these interventions on risk of falling.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [Explanation]

Exercise compared with usual care for falls prevention in care facilities

Population and setting: older (≥ 65 years) residents of care facilities
Intervention: exercise
Comparison: usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Usual care	Corresponding risk Exercise				
Rate of falls Length of follow-up: 3 to 12 months	Low-risk population[1]		**RaR 0.93** (0.72 to 1.20)	**2002** (10 studies)	+ooo VERY LOW[7]	These results were heterogeneous: subgroup analysis by type of exercise did not explain the heterogeneity Four additional trials (N = 130) with data not suitable for pooling reported a reduction in the rate of falls
	1000 per 1000 py	**930** (720 to 1200) **per 1000 py**				
	High-risk population[2]					
	3500 per 1000 py	**3255** (2520 to 4200) **per 1000 py**				
Risk of falling Length of follow-up: 3 to 12 months	Low-risk population[3]		**RR 1.02** (0.88 to 1.18)	**2090** (10 studies)	++oo LOW[8]	1 additional trial (2 comparisons, N = 110) reported no significant difference in the risk of falling
	250 per 1000	**255** (220 to 295) **per 1000**				
	Moderate-risk population[4]					
	500 per 1000	**510** (440 to 590) **per 1000**				
	High-risk population[5]					

	700 per 1000	**714** (616 to 826) **per 1000**				
Risk of fracture Length of follow-up: 6 months	Average risk population[6]		**RR 0.88** (0.25 to 3.14)	**183** 1 study	+ooo VERY LOW[9]	This outcome poorly reported.
	42 per 1000	**37** (11 to 132) **per 1000**				
Adverse events Length of follow-up: 4 to 12 months	See comment	See comment	Not estimable.	**1032** (4 studies)	+ooo VERY LOW [10]	1 serious adverse event reported (death due to a ruptured abdominal aortic aneurysm one week after the follow-up tests, association could not definitely be ruled out) in 1 trial (183 participants) Three trials reported no differences in adverse events: • 1 trial (639 participants) reporting aches and pains, P = 0.75 • 1 trial (194 participants) reported no statistical difference in severe soreness (10 exercise versus 11 control), severe bruises (2 versus 1), severe fatigue (4 versus 1) • 1 trial reported no adverse events

*Illustrative risks for the control group were derived from all or subgroups of trials in care facilities reporting the outcome. The exact basis for the **assumed risk** for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Low risk was based on the mean control risk of the 17 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.07, rounded to 1.0 per person year; thus 1000 per 1000 person years.
[2] High risk was based on the mean control risk of the 18 (top third) trials with the highest rate of falls. The mean rate of falls = 3.69, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[3] Low risk was based on the mean control risk of the 20 trials with the lowest risk of falling. The mean risk of falling = 0.268, rounded to 0.25; thus 250 per 1000 people.
[4] Moderate risk was based on the mean control risk of the 20 trials reporting a moderate risk of falling, not described as high-risk populations. The mean risk of falling = 0.539, rounded to 0.5; thus 500 per 1000 people.
[5] High risk was based on the mean control risk of the 13 trials reporting a high risk of falling, including populations with a description as a high-risk population. The mean risk of falling = 0.680, rounded to 0.7; thus 700 per 1000 people.
[6] Risk based on the median control risk of fracture of the trials reporting this outcome. Median risk = 0.042; thus 42 per 1000.
[7] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of bias for blinding (not feasible), baseline imbalance, attrition bias and high or unclear risk of bias in method of ascertaining falls), one level for inconsistency (considerable heterogeneity I^2 = 76%) and one level for publication bias (suspected based on asymmetry of funnel plots).
[8] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of bias based on blinding (not feasible), baseline imbalance and high or unclear risk of selection bias) and one level for publication bias (strongly suspected based on asymmetry of funnel plots).
[9] The quality of the evidence was downgraded two levels for imprecision (extremely wide confidence intervals that include the possibility of both important benefit and harm) and one level for publication bias (strongly suspected based on asymmetry of funnel plots).
[10] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of bias for selection bias, baseline imbalance and selective reporting), two levels for imprecision (inadequate power to assess rare adverse events) and two levels for 'other reasons' (publication bias strongly suspected based on asymmetry of funnel plots and adverse events unlikely to have been recorded systematically).

BACKGROUND

Description of the condition

Studies of falls in nursing facilities show considerable variation in falls incidence rates but a "middle of the road" figure provided in a review of incidence rates is 1.7 falls per person-year, compared with 0.65 falls per person-year for older people living in the community (Rubenstein 2006). In a study conducted in 40 Canadian residential care facilities, 62% of participants fell over a one-year period, with a falls rate of 2.51 falls per person per year (Kennedy 2015). It should be noted, however, that routine recording of falls incidents in standard reporting systems is likely to under-estimate the incidence of falls (Hill 2010; Sutton 1994). In a prospective one-year study in 528 nursing homes in Bavaria, Germany, about 75% of falls occurred in the residents' rooms or in bathrooms; 41% occurred during transfers and 36% when walking (Becker 2012). The fall rate was higher in men (2.8 falls per person year) than women (1.49 falls per person year), and falls were less common in people requiring the lowest and highest levels of care. Lord 2003 also found that fall rates were lower in frailer people who were unable to rise from a chair or stand unaided. In this group, increased age, male sex, higher care classifications, incontinence, psychoactive medication use, previous falls and slow reaction times were associated with increased falls. Systematic reviews have shown that in nursing homes, falls history, walking aid use, moderate disability, cognitive impairment, wandering, Parkinson's disease, dizziness, use of sedatives, antipsychotics, antidepressants and total number of medications used are associated with an increased risk of falling (Deandrea 2013; Muir 2012). In residents with dementia, age, use of psychotropic drugs, fair or poor general health, gait impairment and trunk restraint use are associated with an increased number of falls (Kropelin 2013).

In hospital settings, a falls incidence of 5.71 falls per 1000 bed days has been found in 16 US general medical surgical and speciality units (Shorr 2012), 6.45 falls per 1000 bed days in 24 Australian medical and surgical wards (Barker 2016), 10.9 falls per 1000 bed days in eight Australian rehabilitation/geriatric units (Hill 2015) and 17.1 falls per 1000 bed days in psychogeriatric wards (Nyberg 1997). In elderly care wards in an UK district general hospital in 2004, the reported rate was as high as 18.0 falls per 1000 bed days (Healey 2004). A similar rate has been reported in some high-risk wards in Australia (Barker 2016).

Systematic reviews have shown that risk factors for falls in hospital inpatients are falls history, age, cognitive impairment, sedative and antidepressant use, gait instability, agitated confusion and urinary incontinence (Deandrea 2013; Oliver 2004). For older patients in rehabilitation hospital settings, risk factors include carpet flooring, vertigo, being an amputee, confusion, cognitive impairment, stroke, sleep disturbance, anticonvulsants, tranquillisers, antihypertensive medications, previous falls and need for transfer assistance (Vieira 2011).

There is considerable mortality and morbidity associated with falls in care facilities and hospitals. A study in 24 Australian medical and surgical wards reported a fall injury rate of 2.36 per 1000 bed days (Barker 2016). A study in both these settings reported an incidence of 533 per 1000 person years for all injuries, 20 per 1000 person years for hip fracture, and 270 per 1000 person years for head injuries, for which 13% (14/107) required medical attention (Nurmi 2002). Overall, men were 1.5 times more likely to be injured than women. Older people who sustain a hip fracture while in hospital have been shown to have poor outcomes compared with people sustaining similar fractures in the community (Murray 2007). Falls have been reported to be the most common cause of death from an external cause in residents of care facilities (Ibrahim 2015).

Description of the intervention

The majority of falls are caused by complex combinations of factors operating at the time of each fall event. Interventions may target risk factors in participants or target staff and clinicians with the aim of improving clinical practice or the organisation of care. In some studies, single interventions have been evaluated while in others, interventions with more than one component have been evaluated. Delivery of multiple-component interventions may be based on individual assessment of risk (a multifactorial intervention) or the same components are provided to all participants (a multiple intervention). A taxonomy has been developed to describe and classify types of intervention (Lamb 2007; Lamb 2011). Key intervention categories include exercise, medication (drug target) interventions which include interventions targeting vitamin D and medication reviews, environment or assistive technologies including bed/chair alarms or the use of low/low beds, social environment interventions which target staff members and changes in the organisational system, knowledge interventions and multifactorial interventions.

The majority of randomised controlled trials considered within this review provide a comparison with 'usual care' in the care facilities and hospitals involved. Typically, 'usual care' will include standard practices for managing commonly known, potentially modifiable, risk factors for falls and, moreover, the components of usual care will vary both over time and between settings.

Why it is important to do this review

A systematic review is required to summarise evidence of the impact of purposeful interventions designed to prevent falls, in addition to the unknown impact of routine (and probably variable) care in care facilities and hospitals. Despite routine activities attempting to reduce falls, falls are common in these settings and they result in considerable mortality and morbidity. Results will inform healthcare professionals, researchers, policy makers, informal

care givers and consumers. This review is an update of a Cochrane Review first published in 2010 (Cameron 2010), and previously updated in 2012 (Cameron 2012).

OBJECTIVES

To assess the effects of interventions designed to reduce the incidence of falls in older people in care facilities and hospitals.

METHODS

Criteria for considering studies for this review

Types of studies

We considered for inclusion all randomised trials, including quasi-randomised trials (for example, alternation), cluster-randomised trials and trials in which treatment allocation was inadequately concealed.

Types of participants

We included trials of interventions to prevent falls in older people, of either sex, in care facilities or hospitals. We considered trials for inclusion if the majority of participants were over 65 years or the mean age was over 65 years, and the majority were living in care facilities or were patients in hospital. We excluded trials conducted in places of residence that do not provide residential health-related care or rehabilitative services, for example retirement villages or sheltered housing. Trials with participants resident in the community and in care facilities were included either in this review or in the Cochrane Review of interventions for preventing falls in older people living in the community (Gillespie 2012), depending on the proportion of participants in each setting. Inclusion in either review was determined by discussion between the authors of both reviews. Trials recording falls in both settings may be included in both reviews.

We subdivided care facilities based on level of care provided. We defined high-level care facilities as "establishments that are primarily engaged in providing inpatient nursing and rehabilitative services for long-term care patients. The care is generally provided for an extended period of time to individuals requiring nursing care. These establishments have a permanent core staff of registered or licensed practical nurses that, along with other staff, provide nursing care in combination with personal care" (OECD 2011). We defined intermediate-care facilities as "institutions which provide health-related care and services to individuals who do not require the degree of care which hospitals or skilled nursing facilities provide, but because of their physical or mental condition require care and services above the level of room and board" (NLM 2012). Some facilities provided both these levels of care. For cluster-randomised trials, the classification of the level of care was based on the description of the facility. For individually-randomised trials where the level of care provided by the facility was clearly described, this description informed the classification. Where the inclusion/exclusion criteria of a trial selected patients who required high or intermediate level of care from a mixed-care facility, the classification was based upon the care needs of the individual participants. For trials in hospitals, participants included staff or in-patients. We excluded interventions that took place in emergency departments, outpatient departments or where hospital services were provided in community settings. We subdivided hospitals into those providing acute, and those providing subacute care. We defined subacute care as "medical and skilled nursing services provided to patients who are not in an acute phase of an illness but who require a level of care higher than that provided in a long-term care setting" (NLM 2012).

Studies recruiting participants post-stroke were excluded as interventions to prevent falls in this population are reviewed in a separate Cochrane Review *Interventions for preventing falls in people after stroke* (Verheyden 2013).

Types of interventions

Any intervention designed to reduce falls in older people compared with any other intervention, usual care or placebo. We grouped interventions using the fall-prevention classification system (taxonomy) developed by the Prevention of Falls Network Europe (ProFaNE) (Lamb 2011). Interventions have been grouped by combination (single, multiple, or multifactorial), and then by the type of intervention (descriptors). Full details are available in the ProFaNE taxonomy manual (Lamb 2007). The possible intervention descriptors are: exercises, medication (drug target, i.e. withdrawal, dose reduction or increase, substitution, provision), surgery, management of urinary incontinence, fluid or nutrition therapy, psychological interventions, environment/assistive technology, social environment, interventions to increase knowledge, other interventions.

Types of outcome measures

We included only trials that reported raw data or statistics relating to rate or number of falls, or number of participants sustaining at least one fall during follow-up (fallers). Trials that reported only those participants who had more than one fall were included. Trials that reported only specific types of fall (e.g. injurious falls) were not included. Trials that focused on intermediate outcomes such as improved balance or strength, and did not report falls or falling as an outcome, were excluded.

Primary outcomes

- Rate of falls (falls per unit of person time that falls were monitored)
- Number of fallers (risk of falling)

Secondary outcomes

- Number of participants sustaining fall-related fractures
- Complications of the interventions
- Economic outcomes

Search methods for identification of studies

Electronic searches

We searched the Cochrane Bone, Joint and Muscle Trauma Group Specialised Register (to 3 August 2017), the Cochrane Central Register of Controlled Trials (CENTRAL) (2017, Issue 8), MEDLINE (including Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily, Ovid MEDLINE and Versions) (1946 to 3 August 2017), Embase (1980 to 2017 Week 31), and CINAHL (1982 to 3 August 2017). We also searched ongoing trial registers via the World Health Organization's ICTRP Search Portal (3 August 2017) and ClinicalTrials.gov (3 August 2017). We did not apply any language restrictions.

For this update, the search results were limited from 2012 onwards. The search update process was run in two stages: the first search was run in February 2016 and a second top-up search was run in August 2017. Details of the search strategies used for previous versions of the review are given in Cameron 2012.

In MEDLINE (OvidSP), subject-specific search terms were combined with the sensitivity- and precision-maximising version of the Cochrane Highly Sensitive Search Strategy for identifying randomised trials in MEDLINE (Lefebvre 2011). We modified this strategy for use in CENTRAL, Embase, and CINAHL (*see* Appendix 1 for all strategies).

Searching other resources

We also checked reference lists of articles and further trials were identified by contact with researchers in the field. For the first version of this review, we identified trials in care facilities and hospitals included in Gillespie 2003.

Data collection and analysis

Data collection and analysis were carried out according to methods stated in the published protocol (Cameron 2005), and subsequently amended to concur with updated methods in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011a) as described in Differences between protocol and review. Data collection and analysis were carried out according to methods stated in the published protocol (Cameron 2005), which were based on the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011a).

Selection of studies

From the title, abstract, or descriptors, one review author screened all abstracts to identify potentially relevant trials for full review. Two review authors screened potentially relevant abstracts. From the full text, two review authors independently assessed potentially eligible trials for inclusion and resolved disagreement by discussion, or by adjudication with a third review author. Full-text review was undertaken using Covidence. Disagreement was resolved by discussion and consensus or third party adjudication when necessary. We contacted trial authors for additional information if necessary to assess eligibility.

Data extraction and management

Pairs of review authors independently extracted data using a pretested data extraction form for studies included to 2012. For this update,again pairs of review authors independently extracted data from the identified studies using Covidence. Multiple reports from the same study were linked as a single study in Covidence and evidence from all reports were reviewed in undertaking data extraction. Where data were unclear authors were contacted whenever possible for clarification. Disagreement was resolved by discussion and consensus or third party adjudication when necessary.

Assessment of risk of bias in included studies

Pairs of review authors independently assessed risk of bias for each included study based on recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b). Assessors were not blinded to author and source institution. Review authors did not assess their own trials. Disagreement was resolved by consensus, or by third party adjudication.

We assessed risk of bias for the following domains: sequence generation (selection bias); allocation concealment (selection bias); blinding of participants and personnel (performance bias); blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), and selective reporting (reporting bias). Since all the outcomes collected in our review are susceptible to the same risk of bias, we have not assessed outcomes for risk of detection bias or completeness of outcome data separately. Additionally, we assessed bias in the recall of falls due to less reliable methods of ascertainment (Hannan 2010), and bias resulting from major imbalances in key baseline characteristics (e.g. age, gender, previous falls, medical status, dependency, cognitive function). Assessors rated the risk of bias as low, high or unclear for each domain.

We established additional criteria within currently existing domains for assessing the additional risks of bias associated with cluster randomisation (Section 16.3.2; Higgins 2011b). Thus 'recruitment bias' was considered as a component of selection bias under allocation concealment; 'baseline imbalance' resulting from small numbers of clusters was considered in bias resulting from major imbalances in key characteristics; risk of bias resulting from 'loss of clusters' was considered under incomplete outcome data; and 'incorrect analysis' that failed to take into account the effect of clustering and that could not be satisfactorily remedied was considered under selective outcome reporting. We did not assess the risk of bias relating to the 'comparability with individually-randomised trials' as a separate item as it is impossible to establish suitable criteria for an individual trial out of context. The potential for differences in effects between cluster- and individually-randomised trials was considered in our assessment of the quality of the evidence and in our Discussion.

Our criteria for 'Risk of bias' assessments are shown in Appendix 2.

Measures of treatment effect

We have reported the treatment effect for rate of falls as a rate ratio (RaR) and 95% confidence interval (CI). For number of fallers and number of participants sustaining fall-related fractures we have reported a risk ratio (RR) and 95% CI. We used results reported at discharge from hospital for trials that continued to monitor falls after discharge.

Rate of falls

The rate of falls is the total number of falls per unit of person time that falls were monitored (e.g. falls per person year). The rate ratio compares the rate of falls in any two groups during each trial.

We used a rate ratio (for example, incidence rate ratio or hazard ratio for all falls) and 95% CI if these were reported in the paper. If both adjusted and unadjusted rate ratios were reported, we used the unadjusted estimate, unless the adjustment was for clustering. If a rate ratio was not reported but appropriate raw data were available, we used Excel to calculate a rate ratio and 95% CI. We used the reported rate of falls (falls per person year) in each group and the total number of falls for participants contributing data, or we calculated the rate of falls in each group from the total number of falls and the actual total length of time falls were monitored (person years) for participants contributing data. In cases where data were only available for people who had completed the study, or where the trial authors had stated there were no losses to follow-up, we assumed that these participants had been followed up for the maximum possible period. Where there were no falls in one arm of a study, and a low total number of falls and/or participants (e.g. Beck 2016; Cadore 2014), the rate of falls cannot be determined. Such data were therefore not pooled, however the omission of these data from the pooled analysis is considered unlikely to change any estimate of effect.

Risk of falling

For number of fallers, a dichotomous outcome, we used a risk ratio as the treatment effect. The risk ratio compares the number of people who fell once or more (fallers) in the intervention and control arms of each trial.

We used a reported estimate of risk (hazard ratio for first fall, risk ratio (relative risk), or odds ratio) and 95% CI if available. If both adjusted and unadjusted estimates were reported we used the unadjusted estimate, unless the adjustment was for clustering. If an odds ratio was reported, or there was no effect estimate and 95% CI, and appropriate data were available, we calculated a risk ratio and 95% CI using the *csi* command in Stata or in Review Manager. For the calculations, we used the number of participants contributing data in each group if this was known; if not reported, we used the number randomised to each group.

Secondary outcomes

For the number of participants sustaining one or more fall-related fractures, we used a risk ratio as described in 'Risk of falling' above.

Unit of analysis issues

For trials that were cluster randomised, for example by care facility or ward, we performed adjustments for clustering (Higgins 2011c), if this was not done in the published report. We used intra-cluster correlation coefficients reported by Dyer 2004 (falls per person year 0.100, number of residents falling 0.071, and residents sustaining a fracture 0.026).

For trials with multiple intervention groups, we either combined the groups or included only one pair-wise comparison (intervention versus control) in any analysis in order to avoid the same group of participants being included twice.

For trials that excluded the intervention period from the falls outcomes, we did not pool the outcomes data with other studies.

Dealing with missing data

Only the available data were used in the analyses; we did not impute missing data.

Assessment of heterogeneity

We assessed heterogeneity within a pooled group of trials using a combination of visual inspection of the graph along with consideration of the Chi² test (with statistical significance set at $P < 0.10$), and the I^2 statistic (Higgins 2003). We based our interpretation of the I^2 results on that suggested by Higgins 2011a: 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial

heterogeneity; and 75% to 100% may represent very substantial ('considerable') heterogeneity.

Assessment of reporting biases

To explore the possibility of publication and other reporting biases, we constructed funnel plots for analyses that contained more than 10 studies.

Data synthesis

We classified interventions into those taking place in care facilities and those taking place in hospitals, and pooled these separately because participant characteristics and the environment warrants different types of interventions in the different settings, possibly implemented by people with different skill mixes.

Within these categories, we grouped the results of trials with comparable interventions and participant characteristics, and compiled forest plots using the generic inverse variance method in Review Manager. This method enabled pooling of the adjusted and unadjusted treatment effect estimates (rate ratios or risk ratios) that were reported in the paper, or we calculated from data presented in the paper (*see* Measures of treatment effect). Where the total number of patients, rather than admissions, could not be determined, we did not pool these data with other studies. Where the reported trial outcomes did not include falls during the intervention period, we did not pool these data with those of other trials. Where appropriate, we pooled results of comparable studies using both fixed-effect and random-effects models. We chose the model to report by careful consideration of the extent of heterogeneity and whether it can be explained by factors such as the number and size of included studies, or the level of care provided. We used 95% CIs throughout. We considered, on a case by case basis, not pooling data where there was considerable heterogeneity (I^2 statistic value of greater than 75%) that could not be explained by the diversity of methodological or clinical features among trials. Where it was inappropriate to pool data, we still presented trial data in the analyses or tables for illustrative purposes and reported these in the text.

Subgroup analysis and investigation of heterogeneity

We minimised heterogeneity as much as possible by grouping trials as described previously (using ProFaNE categories of interventions). We categorised broad interventions further by grouping subtypes of interventions according to ProFaNE (e.g. for exercise interventions). We explored heterogeneity by carrying out subgroup analyses based on level of care and level of cognition at enrolment in care facilities and hospitals where possible. We subdivided the care facilities into high, intermediate or mixed levels of care. The levels of care of the facilities reflect the levels of dependence of the participants. In hospitals, the level of care was subdivided by acute versus subacute or mixed levels of care. We also carried out subgroup analyses by stratification of intervention types according to ProFaNE (e.g. for exercise types, medication target interventions), and type of fracture. Subgroup analyses based upon the individual components of the multifactorial interventions was precluded by the study design and reporting. Data were inadequate for conducting a subgroup analysis by level of frailty of the participants in trials of exercise in care facilities.

We grouped trials by level of cognition into those that included only participants with cognitive impairment versus those with no cognitive impairment, or a mixed sample at enrolment.

We used the random-effects model to pool data in all subgroup analyses testing for subgroup differences due to the high risk of false-positive results when comparing subgroups in a fixed-effect model (Higgins 2011d). We used the test for subgroup differences available in Review Manager to determine whether there was evidence for a difference in treatment effect between subgroups.

Sensitivity analysis

Where there was substantial statistical heterogeneity we carried out a post-hoc sensitivity analysis to explore the effect of removing trials from the analysis if visual inspection of the graph showed poorly overlapping confidence intervals. Where there was considered to be significant statistical heterogeneity for rate of falls but not risk of falling, sensitivity analyses were carried out to determine the likely effects of using random-effects versus fixed-effect meta-analyses for the risk of falling (e.g. for exercise versus usual care in care facilities and multifactorial interventions in care facilities). We conducted post-hoc sensitivity analyses for exercise in care facilities, excluding trials with 20 participants or less in each arm of the trial to explore the possibility of small-trial effects, due to the observed asymmetry in the Funnel plots. We conducted a sensitivity analysis for exercise compared to usual care in care facilities including Cadore 2014, which had zero falls in the intervention arm, using one fall in the intervention arm to examine the likely effect of omitting this trial from the analysis. We also conducted a sensitivity analysis excluding one trial with a known non-normal distribution of falls in the intervention arm from the analysis of general medication review in care facilities for the rate of falls outcomes.

Sensitivity analyses according to study quality were not possible as most studies were at potential risk of bias.

Economic issues

We have noted the results from any economic evaluations (cost-effectiveness analysis, cost-utility analysis) incorporated in included studies. We also extracted from each trial reporting a cost analysis, cost description or analytic model, the type of resource use reported (e.g. delivering the intervention, hospital admissions, medication use) and the cost of the items for each group.

Assessing the quality of the evidence and 'Summary of findings' tables

For each comparison, we used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the quality of the body of evidence (Schünemann 2011) for each outcome listed in Types of outcome measures. For all comparisons where there were two or more trials, GRADE assessment was performed independently by two review authors and disagreement was resolved by discussion, or by adjudication with a third review author. We adopted a different approach for single trial comparisons, where we started with the assumption that the quality of evidence was likely to be very low. This reflected assumptions of downgrading at a minimum for serious risk of bias (typically performance and detection bias), for serious indirectness (trial being conducted was a single trial or setting), and for serious imprecision (failure to meet the 200 to 300 events optimal size criteria) (Guyatt 2011). Where these assumptions did not hold, we performed GRADE assessment as above. The quality rating 'high' is reserved for a body of evidence based on randomised controlled trials. We 'downgraded' the quality rating to 'moderate', 'low' or 'very low' depending on the presence and extent of five factors: study limitations, inconsistency of effect, imprecision, indirectness or publication bias. We used the GRADE approach to assess quality of evidence related to the primary and secondary outcomes listed in the Types of outcome measures. We prepared a 'Summary of findings' table for each of the main categories of interventions, for listed outcomes.

We selected the following comparisons for presentation in 'Summary of findings' tables as these are the most common falls prevention activities considered and applied in clinical settings. In care facilities: exercise, vitamin D supplementation, medication review and multifactorial interventions; in hospitals: exercise, bed alarms and multifactorial interventions.

RESULTS

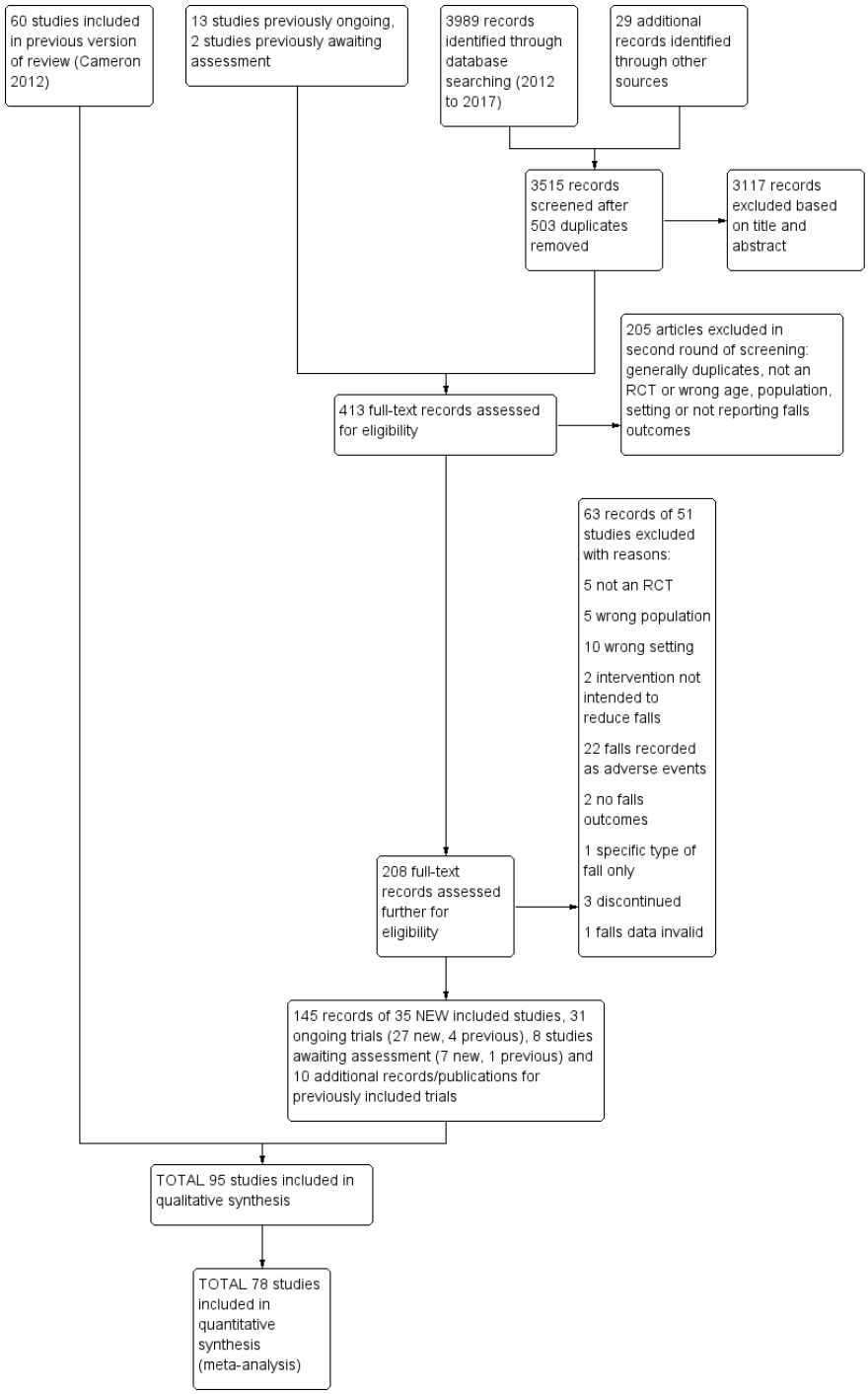
Description of studies

Results of the search

For this update we screened a total of 3989 records from the following databases: Cochrane Bone, Joint and Muscle Trauma Group Specialised Register (0 records); CENTRAL (127), MEDLINE (1104), Embase (1211), CINAHL (314) the WHO ICTRP (450) and Clinicaltrials.gov (783). We also found 29 potentially eligible studies from other sources. After removal of 503 duplicates, 3515 citations were screened for inclusion.

Screening of the search update identified a total of 413 records for potential inclusion, for which full-text reports were obtained. Thirty-five new trials were included in this update, 27 new ongoing trials identified and seven new studies await classification. In addition, a new subgroup analysis (Stenvall 2012) from the Stenvall 2007 trial and a cost-effectiveness analysis (Haines 2013) of Haines 2011 have been added. A flow diagram summarising the study selection process is shown in Figure 1.

Figure 1. Study flow diagram



Overall, there are now 95 included trials, 105 excluded studies, eight studies awaiting classification and 31 ongoing trials.
Due to the review size, not all links to references have been inserted in the text but can be viewed in Table 1.

Included studies

Thirty-five additional trials have been included in this update, 28 trials in care facilities and seven in hospitals (*see* Table 1). This review now contains 95 trials with 138,164 participants. Details of individual trials are provided in the Characteristics of included studies, and are briefly outlined below.

Design

Participants were individually randomised in 53 studies, whereas 42 studies used a cluster-randomised design (*see* Table 1).

Settings

The included trials were carried out in 23 countries (*see* Table 1). Of the 71 studies (40,374 participants) in care facilities, 17 were in high-level care facilities, 17 were in intermediate-level care facilities and 37 were in facilities with mixed levels of care, or combinations of facilities that included both high and intermediate levels of care. Of the 24 studies (97,790 participants) in hospital settings, 10 were in an acute hospital setting, 12 were in subacute settings, and 2 were in both acute and subacute care settings (*see* Table 1).
Van Gaal 2011a and Van Gaal 2011b have been included as two separate trials although reported in the same paper as the participants were randomised separately in two settings (nursing homes and hospitals) and results are reported by setting.

Participants

The mean age of participants was 83.5 years in care facilities and 77.6 years in hospitals. In care facilities, 75.3% were women and in hospitals, 51.6% were women.
All participants were women in seven trials (Bischoff 2003; Chapuy 2002; Faber 2006; Irez 2011; Jarvis 2007; Kovacs 2012; Sihvonen 2004). Ten studies specifically recruited participants with cognitive impairment (Buettner 2002; Chenoweth 2009; Klages 2011; Kovacs 2013; Mador 2004; Neyens 2009; Shaw 2003; Toulotte 2003; Van de Ven 2014; Whitney 2017). Exceptionally, Stenvall 2007 only recruited people with a proximal femoral (hip) fracture.

Interventions

Using ProFaNE taxonomy, all studies were categorised by intervention and grouped by combination (single, multiple, or multifactorial) (*see* Appendix 3). The first column of Appendix 3 shows the intervention classification (single, multiple, or multifactorial) and setting type (care facility or hospital). The components of included 'Exercises' interventions, 'Environmental/assistive technology' and 'Medication (drug target)' interventions are shown in Appendix 4, Appendix 5 and Appendix 6 respectively.
In care facilities, 54 trials tested the effect of a single intervention only, three trials tested both single and multiple interventions (Huang 2016; Imaoka 2016; Sambrook 2012), one trial tested a multiple intervention only (Schnelle 2003), and 13 trials tested a multifactorial intervention. In hospitals, 18 trials tested the effect of a single intervention and six tested a multifactorial intervention. Seven studies tested the effect of two interventions (Faber 2006; Haines 2011; Huang 2016; Nowalk 2001; Sambrook 2012; Saravanakumar 2014; Tuunainen 2013), and one tested three interventions (Imaoka 2016) in comparison with usual care. Donald 2000 was a 2 x 2 factorial study of supervised exercises and flooring types that has been classified as two single interventions.
In general, included studies compared an active falls prevention intervention with a control group comprising 'usual care', that typically would have included standard falls prevention activities. Often, however, standard practice in terms of falls prevention activities was not clearly described. Additional descriptions of the control groups provided for individual trials are provided in the Characteristics of included studies table, the 'Summary of findings' tables available for some comparisons, and the analyses headings and/or footnotes. A general description of the control arms for the main intervention categories is also given below.
In care facilities, 17 trials of exercise provided a comparison with usual care, defined as no exercise, no change in previous lifestyle or exercise type or level unlikely to change physical performance and nine trials provided a comparison of two different exercise programmes (see Table 2). Trials of medication target interventions in care facilities more often provided a comparison with placebo (see Table 3). Trials of vitamin D supplementation in care facilities provided estimates of effect compared with usual care or placebo. In hospitals, multifactorial interventions were generally compared with a control group consisting of standard falls prevention activities. Whether or not the control arm included some of the multifactorial intervention components was not always clearly reported. Additional detail is provided in the description of individual studies in the results text and within the Characteristics of included studies table.

Outcomes

The source of data used for calculating outcomes for each trial for

generic inverse variance analysis is shown in Appendix 7. Seventeen trials met our inclusion criteria but did not report data that could be included in pooled analyses. Reported results from these trials are presented in the text or additional tables. Raw data for rate of falls and number of fallers when reported or when they could be calculated are shown in Appendix 8. Twenty-four trials reported data on fractures suitable for use in pooled analyses, other reported fractures data is presented in the text. Twenty-nine trials clearly reported data on adverse events, but in many of these it was not clear if adverse-event data were recorded systematically; for the majority of trials, this outcome was not reported.

Excluded studies

Overall there were 105 excluded studies (*see* Characteristics of excluded studies for details). Of the 51 newly excluded studies (see Figure 1): five were excluded as they were not randomised; five were conducted in the wrong population (e.g. including participants post stroke); 10 were conducted in the wrong setting (in most of these, the majority of participants were living in the community); two studies of flooring interventions were excluded as the intent was to reduce fall injuries, rather than falls (Drahota 2013; NCT01618786); 22 studies were excluded as they measured falls as a potential adverse outcome of the intervention; two did not report falls outcomes; one study was excluded as it reported a specific type of falls only (Sahota 2014); three trials were discontinued and one had invalid falls data (DeSure 2013).

Of the 54 studies excluded in the previous version of this review: 21 trials were excluded because the intervention they tested was not designed to reduce falls, rather falls were measured as a potential adverse outcome of an intervention with a different aim; in 11 trials the majority of participants were living in the community; eight excluded trials did not provide sufficient data on falls or fallers; seven included participants post stroke and seven were not randomised (Cameron 2012). Of note is that four trials that had been excluded in Cameron 2012 because they included participants with post-stroke hemiplegia, have now either been retracted (Sato 2000; Sato 2005a; Sato 2005b; see Retraction Watch) or, for Sato 2011, likely to be retracted in future because of serious concerns about research misconduct as revealed in Bolland 2016.

Studies awaiting classification

Three studies await publication of full reports containing falls data (*see* Characteristics of studies awaiting classification). One of these is a study of whole body vibration in care facilities (Tallon 2013), another is likely to be an additional conference abstract of an already included study (Frohnhofen 2013), and the third is a thesis for which no study publication has been identified (MacRitchie 2001). Five newly published studies were identified in the top-up search and await full assessment (Dever 2016; Hewitt 2014; Raymond 2017; Van der Linden 2017; Wylie 2017).

Ongoing studies

We are aware of 31 ongoing studies, 14 set in care facilities and 17 in hospitals (*see* Characteristics of ongoing studies for details). The ongoing studies in care facilities include five exercise trials in care facilities (two of whole body vibration), one trial of a multiple intervention of exercise and nutrition, one of nutrition, three of medication review, one of vitamin D supplementation, three of service model changes, and one of a telesurveillance system; two trials are likely to have been completed, one of whole body vibration (JPRN-UMIN000000555) and one of vitamin D supplementation (JPRN-UMIN000008361). The ongoing studies in hospitals include three trials of medication review, four of exercise, one of an education intervention, five social environment interventions including one of student training, one psychological intervention, one of a sensor technology, one educational intervention, and one multifactorial intervention; five trials are likely to be completed, three of medication review (ISRCTN42003273; NCT01876095; NCT02570945), one of exercise (Hassett 2016), and one of telesurveillance (NCT01561872).

Risk of bias in included studies

Details of 'Risk of bias' assessment for nine items for each trial are shown in the Characteristics of included studies. Summary results for these items are shown in Figure 2, Figure 3 and Table 4.

Figure 2. 'Risk of bias' summary: review authors' judgements about each methodological quality item for each included study.

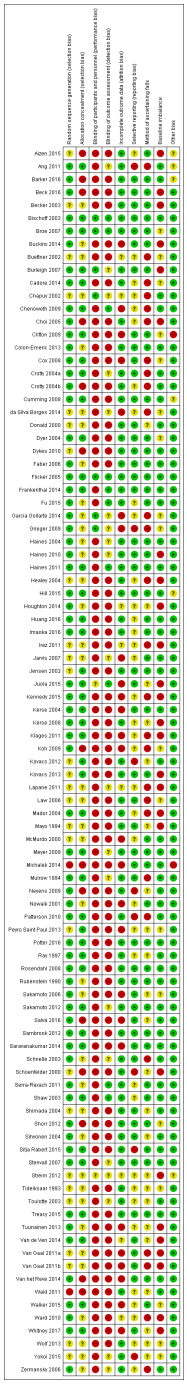
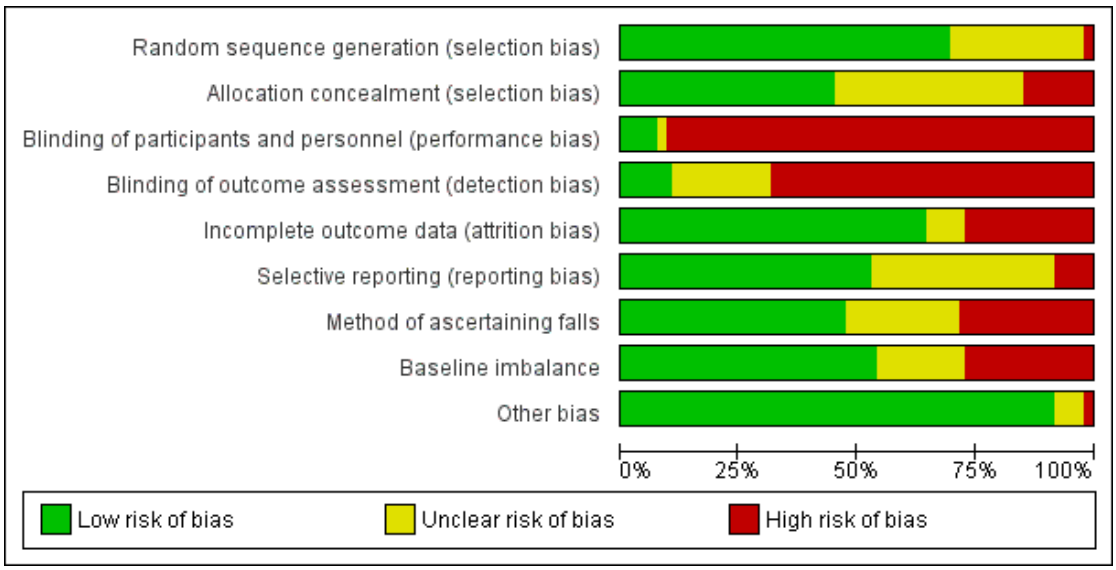


Figure 3. 'Risk of bias' graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.



The majority of included studies were considered at high risk of bias for at least one domain. In particular, there was a high risk of performance bias for the majority of studies due to lack of blinding. Only three trials were considered at low risk of bias for all or the majority of domains (Bischoff 2003; Broe 2007; Flicker 2005), these all examined vitamin D supplementation in comparison with placebo. However, for many other types of interventions, blinding was generally not feasible (e.g. exercise, bed alarms). The risk of bias was often unclear, in particular for risk of selection bias due to allocation concealment. Potential bias varied within comparison groups and it is difficult to judge whether any bias would result in an over- or under-estimation of treatment effect.

Allocation

Under half of included studies (39 in all) were considered at low risk of selection bias; this often reflected lack of clarity on the methods for allocation concealment. We assessed risk of bias in sequence generation as low in 66 trials, high in two trials that described inappropriate methods (Michalek 2014; Wald 2011), and unclear in the remaining 27 trials, usually because of a lack of reporting of methods. We judged methods for concealment of allocation prior to group assignment to carry low risk of bias in 43 trials, high in 14 trials and to be unclear in the remaining 38 trials, again typically due to lack of reporting. Barker 2016, a cluster-randomised trial, is an example of a trial at high risk of selection bias due to lack of allocation concealment: although the initial cluster allocation was concealed, the subsequent recruitment of participants into the study (i.e. admission to the ward) was not.

Blinding

Blinding of participants and personnel was uncommon and indeed blinding of these was not feasible for many intervention types (e.g. exercise, multifactorial interventions). In all, 86 trials were at high risk of performance bias, with just seven trials being at low risk and the remaining two trials being judged at unclear risk of bias. The likelihood of detection bias in relation to the ascertainment of falls by outcome assessors was also high in 65 trials, generally as falls were ascertained by staff who were not blinded (e.g. Barker 2016). Risk of bias was low in 10 trials, most commonly in vitamin D trials where administration of a placebo was possible (e.g. Flicker 2005) and unclear in 20 trials.

Incomplete outcome data

The risk of attrition bias due to incomplete outcome data was assessed as high in 26 trials (the high risk of attrition in some trials

is likely to be related to longer periods of follow-up; e.g. 12 months for Juola 2015 and 16 months for Kennedy 2015). Risk of bias was low in 61 trials, where there was no loss to follow-up (this occurred more frequently in a hospital setting: e.g. Barker 2016; Hill 2015) or losses were balanced between groups (e.g. Cadore 2014; Kerse 2008). Risk of bias was unclear in eight trials, which generally reflected unclear reporting (e.g. Van de Ven 2014).

Selective reporting

Reporting bias was judged as unclear in 37 trials, generally as no protocol was identified (e.g. Healey 2004), and low risk in 50 trials where results were reported according to the protocol (e.g. Potter 2016), or all expected falls outcomes were reported (e.g. Law 2006). Eight trials were at high risk, usually where outcomes mentioned in the protocol or methods were not reported (e.g. Ang 2011).

Other potential sources of bias

The method of ascertaining falls was judged to be at a low risk of bias for 45 trials, at high risk of bias for 27 trials, generally where falls were poorly defined (e.g. Healey 2004), and at unclear risk for 23 trials when methods were not reported (e.g. Sakamoto 2006). The risk of bias relating to imbalance in baseline characteristics was considered to be low in 51 trials, high in 26 trials, and unclear in 18 trials. Risk of baseline imbalance usually occurred in small trials (e.g. Buckinx 2014) or cluster-randomised trials (e.g. Becker 2003; Choi 2005; Van Gaal 2011a; Van Gaal 2011b; Whitney 2017). Two trials were considered to be a high risk of other bias, this was due to the author being employed by the company producing the intervention (Clifton 2009), or the individual randomisation being to one of two clusters, hence the trial was not truly individually randomised (Michalek 2014). There was a low risk of other bias in 87 trials and unclear risk in six trials due to unusual study design (stepped-wedge trial in Aizen 2015; Hill 2015; and including a non-randomised patient preference arm in Streim 2012) or ongoing falls prevention activities (Aizen 2015; Ang 2011; Barker 2016; Cumming 2008).

Cluster-randomised trials

There were a large number of included cluster-randomised trials (44%, 42/95), many of which had a large number of participants (e.g. Barker 2016; Shorr 2012). Risk of bias particular to cluster-randomised trials were considered within other domains (*see* Assessment of risk of bias in included studies). However, it is worth noting that some of these trials contained a small number of clusters and hence were more prone to baseline imbalance (e.g. Choi 2005; Van Gaal 2011a; Van Gaal 2011b), and in some cases prediction of allocation concealment (e.g. Choi 2005; Koh 2009). Loss of whole clusters could also lead to a high risk of attrition bias (e.g. Cox 2008).

Effects of interventions

See: **Summary of findings for the main comparison** Summary of findings: Exercise compared with usual care in care facilities; **Summary of findings 2** Summary of findings: General medication review compared with usual care in care facilities; **Summary of findings 3** Summary of findings: Vitamin D supplementation in care facilities; **Summary of findings 4** Summary of findings: Multifactorial interventions compared with usual care in care facilities; **Summary of findings 5** Summary of findings: Additional exercise plus physiotherapy compared with usual physiotherapy in hospitals; **Summary of findings 6** Summary of findings: Bed alarms compared with usual care in hospitals; **Summary of findings 7** Summary of findings: Multifactorial interventions compared with usual care in hospitals

We present results by setting (care facilities or hospitals), combination (single, multiple, or multifactorial) and intervention type (categorised according to ProFaNE, Lamb 2011) in Appendix 3.

Care facilities: single interventions

Single interventions consist of one major category of intervention only and are delivered to all participants in the group.

Exercise

Twenty-five trials (2848 participants) investigated exercise as a single intervention (*see* Table 2), four trials (986 participants) were cluster randomised (Choi 2005; Kerse 2008; Rosendahl 2008; Yokoi 2015), and the remaining 22 trials (1862 participants) were individually randomised. However, many of these trials were small (median 60 participants, range 16 to 682; *see* Table 1). The types of exercise are shown in Table 2. The control arm of the different trials also varied. Four trials included three arms (Faber 2006; Nowalk 2001; Saravanakumar 2014; Tuunainen 2013). One was a cross-over trial (Toulotte 2003). The trials are categorised below, both according to the ProFaNE exercise category (see Appendix 4) and the comparator arm of the trial. A summary of the evidence from exercise versus usual care for falls prevention in care facilities is provided in Summary of findings for the main comparison.

Only two trials reported on the impact of exercise interventions on fractures (Rosendahl 2008, Sitja Rabert 2015). Nine trials reported on adverse events, while 16 trials did not report adverse-event data. In seven trials, the reported data were incomplete and not suitable for pooling with other studies (Buettner 2002; Cadore 2014; da Silva Borges 2014; Imaoka 2016; Nowalk 2001; Serra-Rexach 2011; Toulotte 2003); *see* Analysis 1.2 and Analysis 4.2). Falls data from Imaoka 2016 excluded the intervention period and thus are not presented in the forest plot.

Exercise versus usual care

Seventeen trials (2406 participants) compared an exercise intervention with usual care, defined as no exercise, no change in previous lifestyle or exercise type or level unlikely to change physical performance (e.g. seated flexibility exercise programme). Four trials (986 participants) of exercise in comparison with usual care were cluster randomised (Choi 2005; Kerse 2008; Rosendahl 2008; Yokoi 2015), the remaining 13 trials (1420 participants) were individually randomised. Faber 2006, included two exercise intervention arms, we combined the results from the two intervention groups in these analyses. As there is considerable clinical heterogeneity within these studies, we undertook analyses to explore heterogeneity, which are reported below.

Rate of falls

Ten trials (2002 participants) reporting on the impact of exercise in comparison with usual care in care facilities on the rate of falls had considerable statistical heterogeneity (I^2 = 76%, heterogeneity P < 0.0001). Nevertheless, as these trials were considered clinically similar in terms of the intervention, comparator, patient group and outcomes, these trials were pooled with a random effects meta-analysis (Analysis 1.1: Rate ratio (RaR) = 0.93, 95% confidence interval (CI) 0.72 to 1.20). We are uncertain whether exercise reduces the rate of falls in care facilities as the quality of the evidence was assessed as very low (Summary of findings for the main comparison).

In a subgroup analysis by broad types of exercise, there was no evidence of a difference between subgroups (Analysis 2.1: test for subgroup differences P = 1.00).

To explore further the heterogeneity in these findings, we carried out a post-hoc subgroup analysis by level of care (high or intermediate levels of care, or mixed levels). There was evidence of a difference between these subgroups that partially explained the heterogeneity (Analysis 3.1: test for subgroup differences Chi^2 = 6.39, I^2 = 69%, 2 df, P = 0.04). In studies of facilities providing mixed levels of care, the heterogeneity was no longer evident (I^2 = 0%, P = 0.41) and there was no evidence of an effect (Analysis 3.1.3 RaR: 1.08, 95% CI 0.92 to 1.28, 3 trials, 477 participants: I^2 = 0%). However, heterogeneity remained considerable for trials in a high or intermediate level of care (I^2 = 78%, P = 0.001).

Four additional trials (130 participants) reported outcomes on rate of falls with data not suitable for pooling (Analysis 1.2); all reported a reduction in falls.

Risk of falling

Pooled data from 10 trials (2090 participants) indicated exercise may make little or no difference to the risk of falling (risk ratio (RR) with random-effects RR 1.02, 95% CI 0.88 to 1.18: I^2 = 23%; Analysis 1.3; low-quality evidence, Summary of findings for the main comparison).

There were no subgroup differences in post-hoc analyses for number of fallers between different levels of care (Analysis 3.2; test for subgroup differences P = 0.56) or types of exercise (Analysis 2.2; test for subgroup differences P = 0.71).

Faber 2006 carried out a post-hoc subgroup analysis and found that the intervention in frail participants may increase risk of falling (hazard ratio (HR) 2.95, 95% CI 1.64 to 5.32; 115 participants), while in the pre-frail subgroup there was no strong evidence for a reduction in the risk of falling (HR 0.62, 95% CI 0.29 to 1.33; 105 participants) (test for subgroup difference P $\leq$ 0.10). Other trials did not provide data suitable for a post-hoc subgroup analysis of the effectiveness of the intervention according to the frailty of the participants.

Nowalk 2001 (N = 110) reported that there was no significant difference in the risk of falling between "Fit NB Free" individually-tailored combination exercises, or the "Living and Learning/Tai Chi" in comparison with usual routine activities; data were not suitable for pooling (Analysis 1.2).

Risk of fracture

One trial of functional exercises (Rosendahl 2008, 183 participants) found no strong evidence for a reduction in the risk of hip fracture (Analysis 1.4.1: RR 0.16, 95% CI 0.01 to 2.81; 3 fractures) or total fractures (Analysis 1.4.2: RR 0.88, 95% CI 0.25 to 3.14; 10 fractures). We are uncertain whether exercise reduces the risk of fracture as the quality of the evidence was assessed as very low (Summary of findings for the main comparison).

Adverse events

Two trials (833 participants) of exercise compared with usual care reported the rates of adverse event outcomes including aches, pains, fatigue, soreness and bruises. Kerse 2008 (639 participants) reported no differences in the level of adverse outcomes on negative binomial regression adjusted for clustering (aches and pains at six months exercise 46.7, 95% CI 39.3 to 54.9 versus usual care 51.1, 95% CI 43.8 to 58.4, P = 0.75). Mulrow 1994 (194 participants) found no difference in the proportion of participants reporting severe soreness (Analysis 1.7.1: RR 0.91, 95% CI 0.40 to 2.04), severe bruises (Analysis 1.7.2: RR 2.00, 95% CI 0.18 to 21.69) or severe fatigue (Analysis 1.7.3: RR 4.00, 95% CI 0.46 to 35.14); there were no injuries during the therapy sessions. One trial (16 participants) reported that there were no adverse events (Schoenfelder 2000). One trial (183 participants) reported a death due to a ruptured abdominal aortic aneurysm one week after the follow-up tests of the exercise intervention for which association could not definitely be excluded by geriatric review (Rosendahl 2008). We are uncertain of the effects of exercise on adverse events as the quality of the evidence has been assessed as very low; Summary of findings for the main comparison).

Sensitivity analysis

As a sensitivity analysis, the pooled analysis of rate of falls was conducted with a fixed-effect model. This made little difference to the estimate of effect (RaR 1.01, 95% CI 0.91 to 1.13). The pooled analysis of the risk of falling with a fixed-effect model also made little difference to the estimate of effect (RR 1.04, 95% CI 0.92 to 1.18). We also conducted a sensitivity analysis including Cadore 2014, which had zero falls in the intervention arm, calculated using one fall in lieu of zero in this arm. This had little impact on the effect estimate (RaR 0.85, 95% CI 0.63 to 1.13; I^2 = 81%). To further explore the heterogeneity in the results, outcomes for all trials excluding two trials (Schoenfelder 2000; Sihvonen 2004) with 20 participants or less in each arm of the trial were pooled (this chosen threshold was arbitrary but considered indicative of 'very small' trials). This did not reduce the heterogeneity for rate of falls (Analysis 1.5: I^2 = 70%), or change the overall pooled estimate of rate of falls (Analysis 1.5: RaR 0.91, 95% CI 0.72 to 1.15) or risk of falling (Analysis 1.6: RR 1.04, 95% CI 0.89 to 1.21; I^2 = 25%).

Funnel plots testing for publication bias

We constructed funnel plots of trials of exercise versus usual care for both the rate of falls and risk of falling outcomes. The funnel plots appeared asymmetrical for both rate of falls and risk of falling (Figure 4 and Figure 5), which may indicate publication bias or lower methodological quality leading to spuriously inflated effects in the smaller trials. In addition to the trials included in the funnel plots, there were four other trials reporting a reduction in the rate of falls.

Figure 4. Funnel plot of comparison: 1 Care facilities: Exercise vs usual care (grouped by level of care), outcome: 1.1 Rate of falls. NB four additional trials with data unsuitable for pooling reported a reduction in the rate of falls.

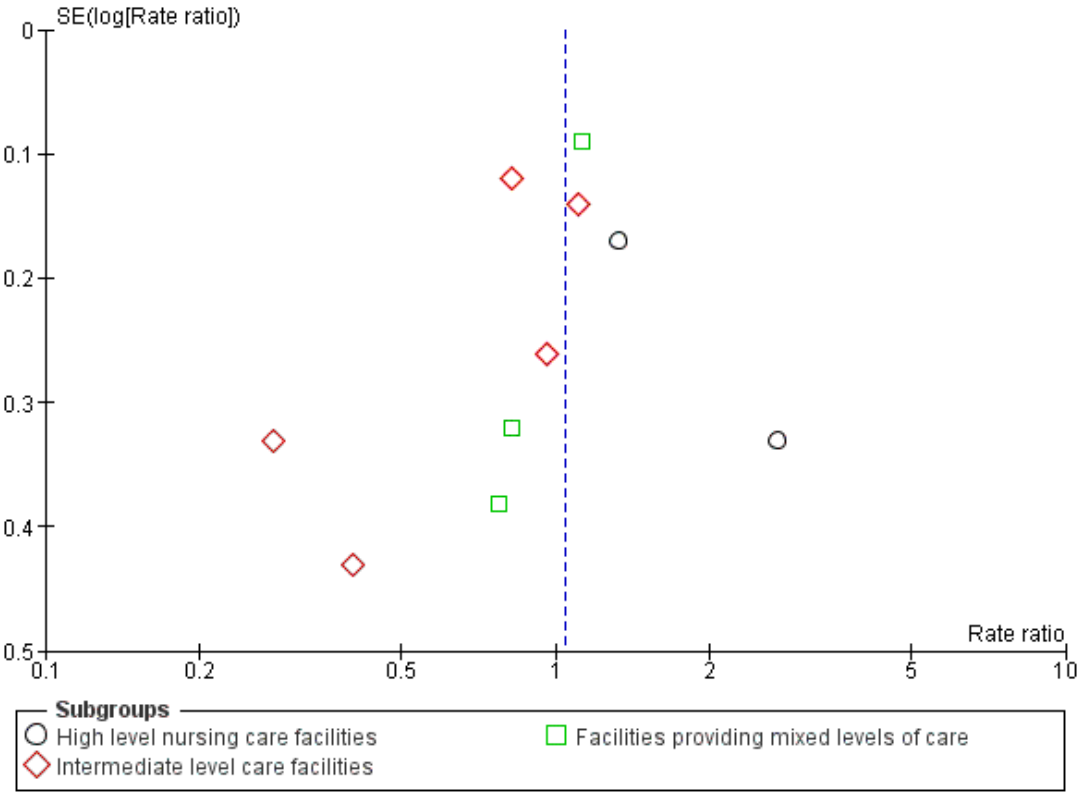
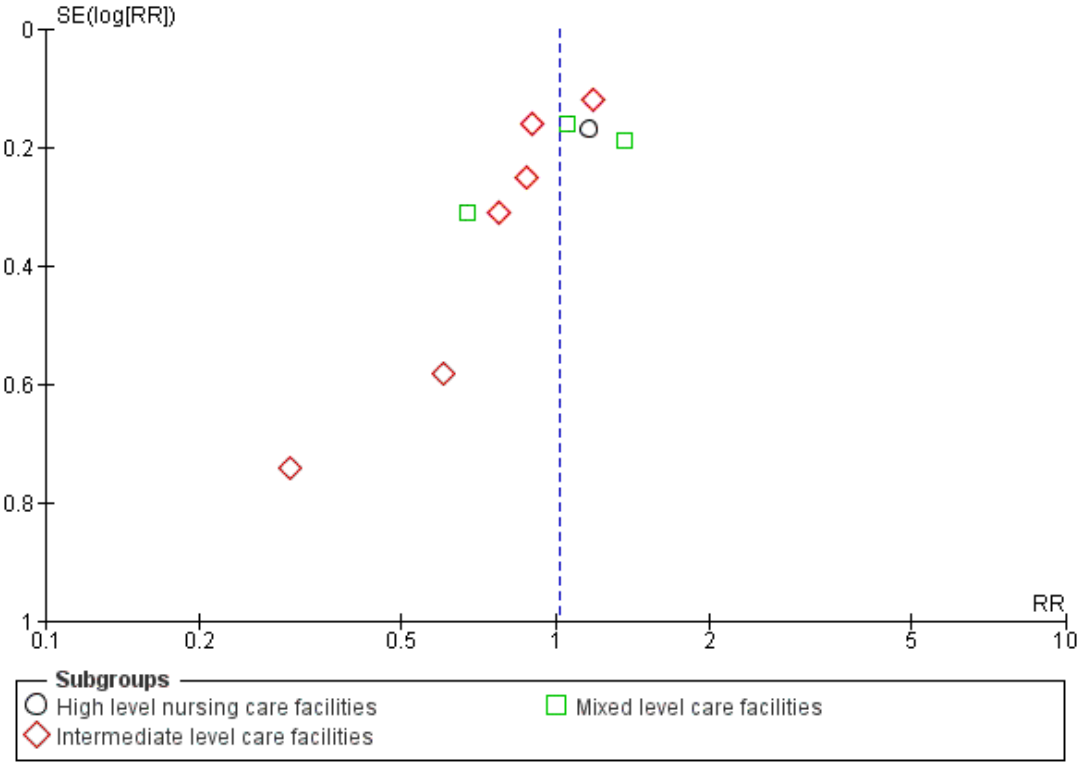


Figure 5. Funnel plot of comparison: 1 Care facilities: Exercise vs usual care (grouped by level of care), outcome: 1.2 Number of fallers. NB One additional trial with data not suitable for pooling reported no significant reduction in the risk of falling.



Comparisons of different exercise categories

Nine trials (584 participants) provided 12 comparisons of two different exercise programmes (Faber 2006; Fu 2015; Imaoka 2016; Kovacs 2012; Saravanakumar 2014; Shimada 2004; Serra-Rexach 2011; Sitja Rabert 2015; Tuunainen 2013). All trials were individually randomised. Seven trials (nine comparisons; 505 participants) had data suitable for pooling (Faber 2006; Fu 2015; Kovacs 2012; Saravanakumar 2014; Shimada 2004; Sitja Rabert 2015; Tuunainen 2013). Two trials provided data on the effectiveness of additional balance exercises (Shimada 2004; Tuunainen 2013). All other comparisons included only single trials; the quality of evidence was considered very low for these comparisons.

Rate of falls

Five trials (Faber 2006; Fu 2015; Saravanakumar 2014; Shimada 2004; Tuunainen 2013; 305 participants) with data suitable for analysis reported the effect of nine comparisons of different exercise programmes on the rate of falls (Analysis 4.1). For eight of these comparisons there was only a single trial with less than 200 participants; the quality of the evidence was considered very low so the relative effectiveness of these exercise programmes on reducing the rate of falls remains uncertain.

Pooled data from two trials (Shimada 2004; Tuunainen 2013) of additional balance exercises indicated a reduction in the rate of falls (Analysis 4.1.1: RaR 0.62, 95% CI 0.40 to 0.96; I^2 = 0%; 56 participants; 86 falls). We are uncertain of the effect of additional balance exercise on falls as the quality of the evidence has been assessed as very low (downgraded two levels due to serious risk of bias, and one level for imprecision).

Serra-Rexach 2011 (40 participants) compared training sessions of a combination of exercises in addition to usual physiotherapy and reported fewer falls in the intervention group (Analysis 4.2).

Risk of falling

Six trials (Faber 2006; Imaoka 2016; Kovacs 2012; Shimada 2004;

Sitja Rabert 2015; Tuunainen 2013; 327 participants) reported the effect of seven comparisons of different exercise categories on the risk of falling (Analysis 4.3). Six comparisons contained only a single trial and the quality of evidence for these comparisons was considered very low; the relative effectiveness of these exercise programmes on reducing the risk of falling remains uncertain.

Pooled data from two trials (Shimada 2004; Tuunainen 2013) of additional balance exercises did not show evidence of a strong effect on reducing the risk of falling Analysis 4.3.1 (RR 0.79, 95% CI 0.43 to 1.45; I^2 = 0%; 56 participants; 24 fallers). We are uncertain of the effect of additional balance exercise on falls as the quality of the evidence has been assessed as very low (downgraded two levels for risk of bias, and one level for imprecision).

In Imaoka 2016, there was no strong evidence for a reduction in the risk of falling in the post-intervention period with additional group exercise (RR 0.48, 95% CI 0.17 to 1.3).

Risk of fracture

Sitja Rabert 2015 (159 participants) compared exercise performed on a whole body vibration platform to the same land based exercises and reported one fracture in the intervention group and none in the control group (Analysis 4.4: RR 2.89, 95% CI 0.12 to 69.07; 1 fracture). We are uncertain whether or not whole body vibration reduces the risk of fracture.

Adverse events

Four trials (269 participants) comparing alternative exercise programmes reported on adverse events; no serious adverse events were reported. Saravanakumar 2014 (29 participants) reported an instance of a non-injurious fall during a yoga intervention. Sitja Rabert 2015 (159 participants) comparing exercise on a whole body vibration platform with land-based exercise reported that "statistical results showed no differences between groups (P= 0.430)" and that "ten percent of participants in the exercise group and 16.3% in the whole body vibration plus exercise group presented a possible or probable relation of causality with the intervention, but this difference was not statistically significant (P = 0.450)." The most commonly reported adverse events were pain (18%) and soreness (13%) but these data were not reported according to group allocation. Serra-Rexach 2011 (40 participants), testing additional physiotherapy, reported a case of transient lumbalgia. Lastly, Kovacs 2012 (41 participants), which compared a multimodel exercise programme based on Otago plus osteoporosis exercises with osteoporosis exercises, reported that there were no adverse events.

Medication (drug target) interventions

Medication review

Twelve studies (7366 participants) examined the effect of medication review interventions in care facilities on falls (Crotty 2004a; Crotty 2004b; Frankenthal 2014; Garcia Gollarte 2014; Juola 2015; Frankenthal 2014; Houghton 2014; Lapane 2011; Patterson 2010; Potter 2016; Streim 2012; Zermansky 2006). Seven trials (4536 participants) were individually randomised (Crotty 2004a; Frankenthal 2014; Frankenthal 2014; Lapane 2011; Potter 2016; Streim 2012; Zermansky 2006), and five trials (2830 participants) were cluster randomised (Crotty 2004b; Garcia Gollarte 2014; Juola 2015; Houghton 2014; Patterson 2010). Two studies (1054 participants) did not report falls data suitable for pooling (Garcia Gollarte 2014; Streim 2012). The primary aim of all medication review is generally to reduce psychoactive medications. Therefore, all trials were considered clinically similar except for one study of medication review for hyponatraemia (Peyro Saint Paul 2013). Further details of the interventions and comparisons are provided in Table 3. A summary of the evidence for general medication review for falls prevention in care facilities is provided in Summary of findings 2.

Rate of falls

Six trials (2409 participants) reporting data on the rate of falls in trials of general medication review were considered clinically appropriate to pool, despite considerable statistical heterogeneity. General medication review may make little or no difference to the rate of falls (Analysis 5.1.1: RaR 0.93, 95% CI 0.64 to 1.35, 6 trials, 2409 participants; I^2 = 93%; low-quality evidence). Subgroup analyses by level of care were not conducted as all trials were conducted in mixed settings.

Garcia Gollarte 2014 (716 participants) conducted a cluster-randomised trial of education of physicians on drug use in older people, plus medication review with feedback in 10% of patients. Data from this study were not pooled as falls during the six-month intervention period were not reported. Over the three months following the intervention, after adjustment for clustering, the rate of falls (RaR 0.74, 95% CI 0.49 to 1.13) did not provide strong evidence for an effect.

A post-hoc sensitivity analyses was conducted excluding Potter 2016 (93 participants), in which 3 participants in the intervention group had more than 30 falls. The heterogeneity in this analysis remained high (Analysis 5.4: I^2 = 87%) and there was no strong evidence of a reduction in the rate of falls.

One additional small trial examined medication review to avoid hyponatraemia (Peyro Saint Paul 2013; Analysis 5.1.2: nine participants), we are uncertain whether medication review reduces falls in adults with chronic moderate hyponatraemia (serum sodium level 123 mEq/L to 134 mEq/L).

Streim 2012 conducted a trial that included both randomised and a non-randomised patient-preference arm. The randomised

arms of the trial (36 participants), examined deprescribing of antidepressants. The authors reported that "the discontinuation and continuation groups exhibited similar non-significant increases in the odds of fall per week with an increase in odds of falls of 1.38 per week (95% CI 4.07 to 0.47); Z=0.59; p=0.55) in the discontinuation group and 1.50 per week (95% CI 0.55 to 4.07); Z= 0.80; p=0.43) in the continuation group. The similarity in odds ratios corresponds to discontinuation only reducing the odds ratio of falls relative to the continuation ratio by approximately 10% (ratio of ORs=0.92 (95% CI=(0.21, 4.01); Z=0.11; p=0.91)."

Risk of falling

Pooled data from six clinically similar trials (5139 participants) reporting falls risk data indicated that general medication review may make little or no difference to the risk of falling (Analysis 5.2.1: RR 0.93, 95% CI 0.80 to 1.09; 5139 participants: I^2 = 48%). The quality of the evidence was considered low (downgraded one level for risk of bias and one level for inconsistency).

In Garcia Gollarte 2014 (716 participants), after adjustment for clustering, the risk of falling (RR 0.86, 95% CI 0.59 to 1.26) did not provide strong evidence for an effect over the three months following the intervention.

We are uncertain of whether medication review reduces falls in adults with chronic moderate hyponatraemia (Analysis 5.2.2: RR 0.42, 95% CI 0.07 to 2.59: 1 trial; 9 participants).

Risk of fracture

Potter 2016 (93 participants) reported the effect of medication review on the risk of fracture (Analysis 5.3: RR 1.60, 95%CI 0.28 to 9.16; 5 fractures), we are uncertain of the effect of medication review on risk of fracture as the quality of the evidence has been assessed as very low.

Subgroup analysis by cognitive status

Juola 2015 provided data for subgroups according to cognitive status. After adjustment for clustering, the rate of falls was reduced for those with an Mini Mental State Examination (MMSE) greater than 15 (RaR 0.23, 95% CI 0.12 to 0.44; 49 participants) or an MMSE of 10-15 (RaR 0.27, 95%CI 0.17 to 0.44; 45 participants) but not for those with an MMSE <10 (RaR 1.27, 95% CI 0.95 to 1.69; 95 participants).

Adverse events

Two studies (102 participants) reported on adverse events; the remaining 10 studies did not clearly report on adverse events related to the intervention.

In a study of deprescribing (Potter 2016; 93 participants), serious vascular events occurred in three control participants and one intervention participant, and two intervention participants experienced significant adverse medicine withdrawal reactions (symptomatic rapid atrial fibrillation and agitation) (Analysis 5.5.1: RR 1.07, 95%CI 0.23 to 5.01; 1 trial).

Peyro Saint Paul 2013 (nine participants) reported one serious adverse event (a major gastrointestinal bleed) related to discontinuing a proton-pump inhibitor in the intervention arm.

We are uncertain of the effects of medication review on adverse events as the quality of the evidence has been assessed as very low (Summary of findings 2).

Vitamin D supplementation

Eight studies (9278 participants) examined vitamin D supplementation administered in some form (Bischoff 2003; Broe 2007; Chapuy 2002; Flicker 2005; Grieger 2009; Imaoka 2016; Kennedy 2015; Law 2006). Six trials (5561 participants) were individually randomised (Bischoff 2003; Broe 2007; Chapuy 2002; Flicker 2005; Grieger 2009; Imaoka 2016) and two trials (3717 participants) were cluster randomised (Kennedy 2015; Law 2006). Four trials (4512 participants) tested the effect of vitamin D supplementation on falls (Bischoff 2003; Broe 2007; Flicker 2005; Law 2006), one trial (583 participants) tested the effect of vitamin D and calcium supplementation (Chapuy 2002), two trials (166 participants) tested multivitamin supplementation that included vitamin D plus calcium (Grieger 2009; Imaoka 2016), and one trial (4017 participants) tested an educational intervention aimed at increasing prescription of adequate levels of vitamin D, calcium and osteoporosis medications (Kennedy 2015). Seven of the eight studies reported serum vitamin D levels at baseline (Bischoff 2003; Broe 2007; Chapuy 2002; Flicker 2005; Grieger 2009; Imaoka 2016; Law 2006). Vitamin D levels were low or very low in these studies enrolling residents of care facilities. Baseline vitamin D levels for one trial (Kennedy 2015) were not reported. A summary of the evidence for vitamin D supplementation for falls prevention in care facilities is provided in Summary of findings 3.

For the specific comparison of multivitamin supplementation including vitamin D and calcium versus placebo (Grieger 2009; Imaoka 2016), the quality of the evidence was considered very low.

Rate of falls

Pooled data from four trials (4512 participants) indicated that vitamin D supplementation probably reduces the rate of falls (Analysis 6.1.1: RaR 0.72, 95% CI 0.55 to 0.95; I^2 = 62%: moderate-quality evidence). The type of vitamin D administered is indicated in the footnotes.

We are uncertain whether multivitamin supplementation including vitamin D and calcium reduces the rate of falls as the quality of

the evidence is very low (Analysis 6.1.2: RaR 0.38, 95% CI 0.20 to 0.71; 91 participants; 1 study).
An education intervention aimed at increasing the prescription of vitamin D, calcium and osteoporosis medication (Kennedy 2015) may make little or no difference to the rate of falls (Analysis 6.1.3: RaR 1.03, 95% CI 0.85 to 1.25; 4017 participants; 1 study; low-quality evidence, downgraded two levels due to risk of bias).

Risk of falling

Pooled data from four trials (4512 participants) indicated that vitamin D supplementation probably makes little or no difference to the risk of falling (Analysis 6.2.1: RR 0.92, 95% CI 0.76 to 1.12; I^2 = 42%; moderate-quality evidence, downgraded one level for risk of bias).
Vitamin D plus calcium supplementation (Chapuy 2002), probably makes little or no difference to the risk of falling (Analysis 6.2.2: RR 1.03, 95% CI 0.90 to 1.18; 583 participants; 1 study; moderate-quality evidence downgraded one level for risk of bias).
We are uncertain whether multivitamin supplementation including vitamin D and calcium reduces the risk of falling (Analysis 6.2.3: RR 0.82, 95% CI 0.40 to 1.66; 91 participants; 1 study). Imaoka 2016 (75 participants), conducted a four-arm trial which found no strong evidence for an effect of daily nutritional supplementation including 900 IU vitamin D (including 400 IU vitamin D3 and 200mg calcium in a multivitamin supplement) in comparison with usual care over the six months following the three-month intervention period (RR 0.58, 95%CI 0.20 to 1.68, N = 34). Outcomes data were not pooled with other studies as they excluded the intervention period; falls are for six months post-intervention.
An education intervention aimed at increasing the prescription of vitamin D, calcium and osteoporosis medication (Kennedy 2015) may make little difference or no difference to the risk of falling (Analysis 6.2.4: RR 1.05, 95% CI 0.90 to 1.23; 4017 participants; 1 study; low-quality evidence, downgraded two levels for risk of bias).

Risk of fracture

Pooled data from three trials of vitamin D supplementation showed little effect on fall related fractures (Analysis 6.3.1: RR 1.09, 95% CI 0.58 to 2.03; I^2 = 63%; 4464 participants; 178 fractures: very low-quality evidence). Different trials reported different types of fractures; the type of fractures are shown in the footnotes to the analysis. We are uncertain whether vitamin D supplementation reduces the risk of fall related fractures as the evidence has been assessed as very low.
We are uncertain whether vitamin D plus calcium supplementation reduces the risk of fall related fractures (Analysis 6.3.2: RR 0.62, 95% CI 0.36 to 1.07; 583 participants; 48 hip fractures; very low-quality evidence, downgraded one level for risk of bias, one level for imprecision and one level as this review only includes a subset of the trials available reporting the effects of this intervention on fractures).
An education intervention aimed at increasing the prescription of vitamin D, calcium and osteoporosis medication (Kennedy 2015; 4017 participants) reported that 1.5% of falls in control participants and 1.6% of falls in intervention participants resulted in a fracture, the study was not powered to detect a difference in fall-related fractures, we are uncertain of the effects of this intervention on fractures (very low-quality evidence, downgraded two levels for risk of bias and two levels for imprecision).

Adverse events

Four trials (1365 participants) reported adverse-event data.
Two of four trials (747 participants) of vitamin D supplementation reported on adverse events (Bischoff 2003, Flicker 2005); no serious adverse events were reported. Bischoff 2003 reported two cases of increased constipation in the intervention arm and no cases of hypercalcaemia (Analysis 6.4.1: constipation RR 4.84, 95%CI 0.24 to 98.80; 122 participants). Flicker 2005 reported that there were no adverse events. We are uncertain of the effects of Vitamin D supplementation (up to 1000 IU daily) on adverse events as the quality of the evidence has been assessed as very low (Summary of findings 3).
One trial of vitamin D and calcium supplementation (800 IU of vitamin D3 + 1200 mg calcium carbonate daily) reported a similar rate of gastrointestinal disorders in each arm of the study and three cases of hypercalcaemia in the intervention arm, we are uncertain of the effects on adverse events (Chapuy 2002; Analysis 6.4.2; gastrointestinal adverse events RR 0.82, 95% CI 0.45 to 1.48; 583 participants; very low-quality, downgraded one level for risk of bias and two levels for imprecision).
Grieger 2009, which tested multivitamin supplementation including vitamin D and calcium, reported there were no serious adverse events; the three adverse events reported were in the control arm of the trial (rash/vertigo, behavioural issues, indigestion), we are uncertain of the effects on adverse events (Analysis 6.4.2: RR 0.13, 95% CI 0.01 to 2.41; 91 participants, 40 events; very low-quality evidence).

Environment/assistive technology

In a cross-over trial, Clifton 2009 (43 participants) tested a wireless position-monitoring device and found no strong evidence for a reduction in the rate of falls (Analysis 7.1: RaR 0.65, 95% CI 0.33 to 1.27; no adjustments for cross-over design made in the analysis). There were no serious adverse events. We are uncertain whether or not wireless position monitoring has an effect on the rate of falls in care facilities (very low-quality evidence).

Social environment

Seven cluster-randomised trials examined service change interventions in care facilities (13,127 participants in six trials Cox 2008; Chenoweth 2009; Meyer 2009; Van de Ven 2014; Van Gaal 2011a; Ward 2010, plus 982 facility beds in Colon-Emeric 2013). These included three trials of staff training interventions (Colon-Emeric 2013; and 7029 participants from Cox 2008 and Van Gaal 2011a) and four of a service model change (6098 participants; Chenoweth 2009; Meyer 2009; Van de Ven 2014; Ward 2010). These interventions target staff or caregivers and changes in the organisational system in which an intervention is delivered, rather than targeting patients directly. The rate of falls for these interventions were not pooled due to high clinical and statistical heterogeneity (test for subgroup differences: P = 0.0001, I^2 = 85.6%). Two studies (6516 participants) reported data on risk of fracture (Meyer 2009, Ward 2010). No studies reported on adverse events. Although there were only single trials for the comparisons within this category, the generally larger size of these trials meant that optimal information size criteria may be met and GRADE assessments were conducted by two review authors.

Staff training

Cox 2008 (5637 participants) studied a half day education programme about fall and fracture prevention for managers, nurses and health care assistants, given by specialist osteoporosis nurses. There was no strong evidence for a reduction in the rate of falls, we are uncertain of the effects as the quality of the evidence was assessed as very low (Analysis 8.1.1: RaR 1.19, 95% CI 0.92 to 1.53; very low-quality evidence, downgraded two levels for risk of bias and one level for imprecision). The intervention may make little or no difference to the rate of fracture (reported incidence rate ratio (IRR) for all fractures: IRR 0.94, 95% CI 0.71 to 1.26; for hip fractures: IRR 0.86, 95% CI 0.63 to 1.18; low-quality evidence downgraded two levels for risk of bias).

The intervention in Van Gaal 2011a (392 participants) consisted of education to implement a patient-safety programme directed at falls, urinary tract infection, and pressure ulcers based on available guidelines. There was no strong evidence for a reduction in rate of falls, we are uncertain of the effects on the rate of falls (Analysis 8.1.2: RaR 0.63, 95% CI 0.34 to 1.16; very low-quality evidence, downgraded two levels for risk of bias, one level for indirectness and one level for imprecision).

Colon-Emeric 2013 (number of resident participants not reported, 497 staff participants, 982 facility beds) conducted a pilot cluster-randomised trial testing a programme to improve staff connections, communication, and problem solving compared to usual care during implementation of a falls quality improvement programme. There was no strong evidence for an effect on the change in falls rate from baseline to post intervention periods between the two arms of the study, we are uncertain of the effects in reducing falls (RaR of change in falls rate 0.81, 95% CI 0.55 to 1.20; very low-quality evidence, downgraded one level for each of risk of bias, indirectness and imprecision).

Service model change

Meyer 2009 (1125 participants) found that use of a falls risk-assessment tool in comparison with nurses' judgement alone probably makes little or no difference to the rate of falls or risk of falling (Analysis 8.1.3: RaR 0.96, 95% CI 0.84 to 1.10; Analysis 8.2: RR 0.99, 95% CI 0.85 to 1.16; both outcomes moderate-quality evidence, downgraded one level for risk of bias). We are uncertain whether or not this intervention reduces the risk of fracture as the quality of the evidence was assessed as very low (Analysis 8.3.1: RR 0.96, 95% CI 0.57 to 1.63; 77 fractures in total; downgraded one level for risk of bias and two levels for imprecision).

Two studies examined dementia care mapping, but data from Chenoweth 2009 were not suitable for pooling. Chenoweth 2009 (289 participants) reported that "... at follow-up there were fewer falls with dementia-care mapping than in usual care (p=0·02) and more falls in person-centred care than in usual care (p=0·03)." Van de Ven 2014 (293 participants) delivered a four-month dementia care mapping intervention twice during the 12-month follow-up period after baseline. The rate of falls at study endpoint was greater in the intervention arm of the study (Analysis 8.1.4: RaR 1.84, 95% CI 1.40 to 2.42). We are uncertain of the effects of dementia care mapping on the rate of falls as the quality of the evidence has been assessed as very low (downgraded two levels for risk of bias, one level for inconsistency and one level for imprecision).

Ward 2010 (5391 participants) employed a practice nurse to encourage the adoption of best practice strategies and reported "0.13 fewer falls per 100 beds per month; 95% CI, −0.36 to 0.10; P = 0.259" for the intervention period. There was no difference in risk of hip fracture between intervention and control groups during the 17 months of intervention (Analysis 8.3.2; RR 0.95, 95% CI 0.63 to 1.44; 215 hip fractures). We are uncertain of the effects of this intervention on fractures as the quality of the evidence has been assessed as very low (downgraded two levels for risk of bias, and two levels for imprecision).

Psychological interventions

Two studies (163 participants) examined the impact of psychological interventions on falls (Huang 2016; Van het Reve 2014). Both trials were individually randomised, Huang 2016 is a three-arm trial for which falls excluded the intervention period; findings are also discussed under "Care facilities: multiple interventions". Neither trial reported data on the risk of fracture or adverse events. In Van het Reve 2014 (114 participants) a computer-based cognitive training programme focused on improving attention was combined with strength and balance training, and compared with strength and balance training alone. The intervention showed no strong evidence for an effect on falls rates (Analysis 9.1: RaR 1.22,

95% CI 0.78 to 1.92), risk of falling during the intervention period (Analysis 9.2.2; RR 1.35, 95% CI 0.23 to 7.88) or over 12 months post-intervention (RR 1.38, 95% CI 0.76 to 2.51; data not shown).

In a three-arm study, Huang 2016 tested the effects of a cognitive-behavioural intervention conducted by a trained facilitator in comparison with usual care in 49 participants. Over the three months following the intervention, there were 1.67 falls per person year in the usual care arm of the study (10 falls in seven fallers), but no falls in the cognitive-behavioural intervention arm. Data were not pooled as falls excluded the intervention period.

The quality of the evidence for both the rate and risk of falling was considered very low (downgraded one level for risk of bias, inconsistency and indirectness and two levels for imprecision), so we are uncertain of the effectiveness of psychological interventions in reducing falls.

Other single interventions

Three trials (564 participants) examined other single interventions of lavender olfactory stimulation (Sakamoto 2012), sunlight exposure (Sambrook 2012), and multisensory stimulation in a Snoezelen room (Klages 2011); two trials (169 participants) were individually randomised (Sakamoto 2012; Klages 2011) and one (Sambrook 2012; 395 participants) was cluster randomised. The quality of the evidence was considered very low for all of these single-trial comparisons.

For one year, Sakamoto 2012 (145 participants) tested the effect of lavender olfactory stimulation by applying lavender patches or placebo patches to clothing near the neck daily. This intervention did not show strong evidence for a reduction in the rate of falls (Analysis 10.1: RaR 0.57, 95% CI 0.32 to 1.01) or risk of falling (Analysis 10.2: RR 0.67, 95% CI 0.40 to 1.12). The authors reported that there were no adverse events. We are uncertain of the effectiveness of lavender olfactory stimulation as the quality of the evidence is very low.

In Sambrook 2012 (395 participants), a trial of increased sunlight exposure had low adherence to the sunlight intervention (Durvasula 2012). We are uncertain of the effects on falls as the quality of the evidence has been assessed as very low for all outcomes (downgraded one level for each of risk of bias, indirectness and imprecision; Analysis 10.1.2: RaR 1.05, 95% CI 0.71 to 1.56; Analysis 10.2.2: RR 1.09, 95% CI 0.88 to 1.36; Analysis 10.3: risk of fracture: RR 1.07, 95% CI 0.53 to 2.17, total 32 fractures). The authors reported no difference in the incidence rates of new skin cancers between arms of the trial and one fall on the way to a sunlight session. Adverse-event data for this three-arm trial are also reported below under Multiple interventions.

Klages 2011 (24 participants) compared the effect of multisensory stimulation in a Snoezelen room with control activities in people with dementia and reported, without providing data, that the "Group membership did not alter falls frequency". Adverse-event data were not reported. We are uncertain of the effectiveness of multisensory stimulation as the quality of the evidence is very low.

Care facilities: multiple interventions

In multiple interventions, the same combination of single categories of intervention was delivered to all participants in the group. Three trials (652 participants) examined multiple interventions in care facilities (Sambrook 2012; Schnelle 2003; Huang 2016). One trial (412 participants) was cluster randomised (Sambrook 2012) and two trials (240 participants) were individually randomised. The quality of the evidence was considered very low for the single trial comparisons of exercise plus management of urinary incontinence and fluid therapy with usual care (Schnelle 2003), and cognitive-behavioural therapy to address fear of falling with an exercise programme versus usual care (Huang 2016).

In Schnelle 2003 (190 participants), participants engaged in supervised exercises and were offered fluids and regular toileting. There was no strong evidence for an effect in reducing the rate of falls (Analysis 11.1.1: RaR 0.62, 95% CI 0.38 to 1.01), risk of falling (Analysis 11.2.1: RR 0.62, 95% CI 0.36 to 1.05) or risk of fracture (Analysis 11.3.1: RR 4.26, 95% CI 0.48 to 37.55; total five fractures). Adverse events were not reported. We are uncertain of the effectiveness of this intervention as the quality of the evidence is very low.

One intervention group in Sambrook 2012 (412 participants), which was based in Australia, tested the effect of increased sunlight exposure plus calcium supplementation, with low adherence to the sunlight intervention (Durvasula 2012). We are uncertain of the effects on falls as the quality of the evidence has been assessed as very low for all outcomes (downgraded one level for each of risk of bias, indirectness and imprecision; Analysis 11.1.2: RaR 1.03, 95% CI 0.85 to 1.25; Analysis 11.2.2: RR 0.96, 95% CI 0.77 to 1.19; Analysis 11.3.2: risk of fracture RR 0.78, 95% CI 0.36 to 1.67; total 31 fractures). The authors reported no significant difference in the incidence rates of new skin cancers between arms of the trial (18 new cancers total) and an increase in the adjusted all-cause mortality in the calcium-treated group compared with the UV alone group (HR 1.23 versus 0.76, P = 0.03; 40 deaths; adjusted for age, sex and season). There was a lack of evidence for a strong effect on increased death rates from myocardial infarction (age-adjusted HR 3.83, 95% CI 0.97 to 15.27, P = 0.06; sex-adjusted HR 4.17, 95% CI 0.69 to 25.16, P = 0.12; the authors reported that they did not record cardiovascular events prospectively). We are uncertain of the effects on adverse events as the quality of the evidence is very low (downgraded one level for each of risk of bias, indirectness and imprecision).

In a three-arm trial, Huang 2016 studied an intervention which combined cognitive-behavioural therapy to address fear of falling with an exercise programme in comparison with usual care in 50 participants. In the three months following the eight-week intervention the authors reported a reduction in falls in both the

combined intervention and the cognitive-behavioural intervention arm alone (reported Kruskal-Wallis P < 0.001). There were 1.67 falls per person year in the usual care arm of the study (10 falls in seven fallers), and no falls in the cognitive behavioural plus exercise intervention arm; data were not pooled as falls excluded the intervention period. Adverse events were not reported. We are uncertain of cognitive-behavioural therapy combined with an exercise programme as the quality of the evidence is very low.

Care facilities: multifactorial interventions

In multifactorial interventions, two or more categories of intervention are given, and these are linked to each individual's risk profile. An initial assessment is usually carried out by one or more health professionals and an intervention is then provided or recommendations given or referrals made for further action. A summary of the evidence for multifactorial interventions in comparison with usual care in care facilities is provided in Summary of findings 4. Thirteen trials (4226 participants) in care facilities studied multifactorial interventions (Beck 2016; Becker 2003; Dyer 2004; Jensen 2002; Kerse 2004; McMurdo 2000; Neyens 2009; Ray 1997; Rubenstein 1990; Salvà 2016; Shaw 2003; Walker 2015; Whitney 2017). Eleven trials were cluster-randomised trials (Beck 2016; Becker 2003; Dyer 2004; Jensen 2002; Kerse 2004; McMurdo 2000; Neyens 2009; Ray 1997; Salvà 2016; Walker 2015; Whitney 2017; 3470 participants), and two were individually randomised (Rubenstein 1990; Shaw 2003; 756 participants). Whitney 2017 was also a cross-over trial. None of these trials were sufficiently similar to allow analysis of subgroups of specific combinations of interventions. Two studies did not report data suitable for use in the quantitative analysis (Beck 2016; Ray 1997). Three studies (2160 participants) reported data on hip fractures (Becker 2003; Jensen 2002; Shaw 2003), and one reported total fractures (Salvà 2016). Three studies (312 participants) reported adverse-event data (Beck 2016; McMurdo 2000; Whitney 2017).

Rate of falls

Despite statistical heterogeneity between the trials for the rate of falls, trials were considered clinically similar enough for pooling to be meaningful. Pooled data from 10 trials (3439 participants) for rate of falls did not demonstrate strong evidence for a reduction in falls (Analysis 12.1: RaR random effects 0.88, 95% CI 0.66 to 1.18: I^2 = 84%). Beck 2016 (31 participants) reported falls outcomes in a cluster-randomised trial of an exercise programme plus nutritional support. There were zero falls in the intervention arm and two in the control arm over an 11-week period. Overall, we are uncertain of the effects of multifactorial interventions on the rate of falls in care facilities as the quality of evidence has been assessed as very low (Summary of findings 4).

Risk of falling

Pooled data from nine trials (3153 participants) for risk of falling (Analysis 12.2: RR random effects 0.92, 95% CI 0.81 to 1.05: I^2 = 42%) did not demonstrate strong evidence for a reduction in falls. Ray 1997 (482 participants) only recorded the number of people having two or more falls during follow-up (recurrent fallers) and reported a reduction in the proportion of recurrent fallers (difference 19%, 95% CI 2% to 36%: P = 0.03). Overall, multifactorial interventions in care facilities may make little or no difference to the risk of falling (low-quality evidence; Summary of findings 4).

Risk of fracture

Pooled results for five studies (2160 participants) reporting risk of fracture did not show strong evidence for an effect (Analysis 12.3: RR 0.79, 95% CI 0.30 to 2.07: I^2 = 44%; 76 fractures). Data from three of the five trials (1695 participants) were for hip fracture (Becker 2003; Jensen 2002; Salvà 2016) and two trials (465 participants) reported total fractures (Shaw 2003; Whitney 2017). Two trials (1255 participants) included hip protectors as an intervention (Becker 2003; Shaw 2003). We are uncertain of the effects of multifactorial interventions on the risk of fracture as the quality of evidence has been assessed as very low (Summary of findings 4).

Adverse events

Three studies (312 participants) reported adverse-event data. One trial reported an instance of a fall in the intervention arm (Whitney 2017), two studies reported that there were no adverse events (Beck 2016; McMurdo 2000). We are uncertain of the effects of multifactorial interventions on adverse events as the quality of evidence has been assessed as very low (Summary of findings 4).

Subgroup analyses exploring heterogeneity

To explore the heterogeneity in these results, we carried out post-hoc subgroup analysis by levels of care (high or intermediate or mixed levels of care). The test for subgroup differences showed a difference between subgroups for both the rate of falls (Analysis 13.1: P = 0.005, I^2 = 81%) and risk of falling (Analysis 13.2: P = 0.03, I^2 = 72%). Within care facilities providing either high or intermediate levels of care, statistical heterogeneity was not important and pooled data showed a reduction in both the rate of falls (Analysis 13.1.1: high-level care: RaR 0.59, 95% CI 0.44 to 0.79; I^2 = 8%, P = 0.30; Analysis 13.1.2: intermediate-level care: RaR 0.64, 95% CI 0.50 to 0.83; I^2 = 33%, P = 0.23), and the risk of falling (Analysis 13.2.1: high level care: RR 0.75, 95% CI 0.57 to 0.98; Analysis 13.2.2: intermediate level care: RR 0.75, 95% CI 0.60 to 0.94; I^2 = 0%, P = 0.44). However, heterogeneity

remained high in studies of mixed levels of care (Analysis 13.1.3: RaR 1.23, 95% CI 0.85 to 1.77; I^2 = 77%, P = 0.001; Analysis 13.2.3: RR 1.01, 95% CI 0.88 to 1.15; I^2 = 24%, P = 0.26).

We also carried out a subgroup analysis comparing trials recruiting people with cognitive impairment versus trials with participants with no cognitive impairment (based on inclusion/exclusion criteria) or a mixed sample. Two trials recruited residents with cognitive impairment only (Neyens 2009; Shaw 2003). In addition, two trials (Becker 2003; Jensen 2002) carried out pre-planned subgroup analyses by levels of cognition, which are reported in Rapp 2008 and Jensen 2003, respectively. Cognitive impairment was defined differently in all four studies (*see* footnotes to Analysis 14.1 and Analysis 14.2). There was no evidence of subgroup differences between those with higher or mixed levels of cognition and those with lower cognition for both rate of falls (Analysis 14.1: test for subgroup differences P = 0.97, I^2 = 0%) and risk of falling (Analysis 14.2: test for subgroup differences P = 0.41, I^2 = 0%).

Subgroup analysis based upon the individual components of the interventions was precluded by the study design.

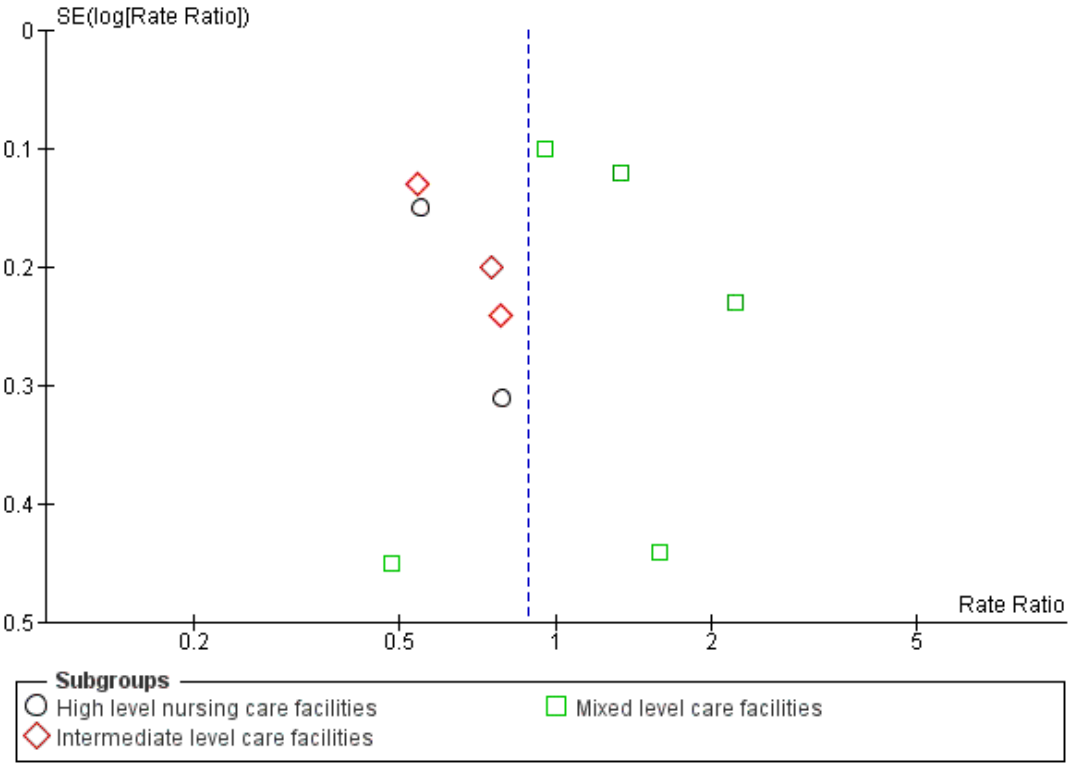
Sensitivity analysis

Considering statistical heterogeneity in the rate of falls, meta-analyses with a random-effects model was considered the most appropriate. However, there was only moderate heterogeneity in the risk of falling data, therefore trials were pooled using the fixed-effect model as a sensitivity analysis. Pooled data from 10 trials (3439 participants) using a fixed-effect model for rate of falls showed an RaR 0.87, 95% CI 0.79 to 0.97 (compare with Analysis 12.1: I^2 = 84%) and from nine trials (3153 participants) for risk of falling showed an RR 0.92, 95% CI 0.84 to 1.00 (compare with Analysis 12.2: I^2 = 42%).

Funnel plots testing for publication bias

A funnel plot of trials of multifactorial interventions in care facilities was conducted for the outcome of rate of falls (Figure 6). There was no obvious asymmetry on visual inspection.

Figure 6. Funnel plot of comparison: 11 Multifactorial interventions vs usual care grouped by level of care (care facilities), outcome: 11.1 Rate of falls.



Hospitals: single interventions

Exercise

Three individually-randomised trials (244 participants) tested the effect of additional physiotherapy in rehabilitation wards (Donald 2000; Jarvis 2007; Treacy 2015). One study tested additional strengthening exercises (Donald 2000), one additional balance training (Treacy 2015), and one additional physiotherapy (Jarvis 2007). A summary of the evidence for exercise for falls prevention in hospitals is provided in Summary of findings 5. No data on the risk of fractures were reported. One trial (161 participants) reported that there were no adverse events (Treacy 2015), two studies did not report adverse-event data.

Pooled data did not provide evidence for a reduction in rate of falls (Analysis 15.1: RaR 0.59, 95% CI 0.26 to 1.34; 215 participants, 2 trials; I^2 = 0%; very low-quality evidence). Pooled data from two trials (83 participants) showed a reduction in the risk of falling (Analysis 15.2: RR 0.36, 95% CI 0.14 to 0.93: I^2 = 0%: very low-quality evidence). We are uncertain whether additional exercise reduces the rate or risk of falling or has adverse events as the evidence has been assessed as very low.

Medication (drug target) interventions

Two trials (319 participants) examined medication target interventions, one examined medication review (Michalek 2014), and the other vitamin D supplementation (Burleigh 2007). These comparisons were from single trials only and the quality of evidence was considered very low.

Multiprofessional medication review

Michalek 2014 (114 participants) conducted a quasi cluster-randomised trial that examined the effect of review of suitability of medications for aged patients in comparison with usual care. After adjustment for clustering there was no strong evidence for an effect on the rate of falls (Analysis 16.1: RaR 0.14, 95% CI 0.00 to 6.63) or risk of falling (Analysis 16.2: RR 0.18, 95% CI 0.01 to 3.47). Adverse-event data were not reported. We are uncertain of the effectiveness of medication review in hospitals as the quality of the evidence is very low.

Vitamin D supplementation

Burleigh 2007 (205 participants) conducted an individually-randomised trial that investigated whether 800 IU of vitamin D plus 1200 mg of calcium supplements reduced falls compared with 1200 mg calcium supplements alone in participants with a median length of stay of 30 days. There was no strong evidence for an effect on risk of falling (Analysis 17.1: RR 0.82, 95% CI 0.59 to 1.14) or fractures (Analysis 17.2: RR 0.34, 95% CI 0.04 to 3.05; total four fractures). The rates of gastrointestinal complaints were similar between the arms of the trial (Analysis 17.3: RR 1.37, 95% CI 0.32 to 5.98). We are uncertain of the effectiveness of vitamin D supplementation in hospitals as the quality of the evidence is very low.

Environment/assistive technology interventions

Six trials (39,127 participants) examined environment or assistive technology interventions, two trials (11,153 participants) were of furnishing adaptations (Donald 2000; Haines 2010), and four (27,974 participants) were of communication aids (Mayo 1994; Shorr 2012; Tideiksaar 1993; Wolf 2013). Four trials (356 participants) were individually randomised (Donald 2000; Mayo 1994; Tideiksaar 1993; Wolf 2013), and two (38, 771 participants) were cluster randomised (Haines 2010; Shorr 2012). Donald 2000 was a 2 x 2 factorial design. The quality of the evidence was considered very low for the single trial comparisons of carpet in comparison with vinyl floors (Donald 2000) and identification bracelets for high-risk fallers (Mayo 1994).

Furnishing/adaptations

Donald 2000, in a factorial design with 54 participants, found that carpeted floors compared with existing vinyl floors in subacute hospital wards resulted in an increase in rate of falls (Analysis 18.1.1: RaR 14.73, 95% CI 1.88 to 115.35) and no strong evidence for an increase in the risk of falling (Analysis 18.2.1: RR 8.33, 95% CI 0.95 to 73.37). We are uncertain of the impact of carpeting on falls as the quality of the evidence is very low.

In a cluster-randomised trial, Haines 2010 (11,099 participants) examined an intervention which consisted of providing one low-low bed per 12 existing beds in acute and subacute wards. There was no strong evidence of an effect on the rate of falls; we are uncertain of the effectiveness of low-low beds as the quality of the evidence is considered very low (Analysis 18.1.2: RaR 1.39, 95% CI 0.22 to 8.78; very low-quality evidence downgraded two levels for risk of bias, one level for indirectness and two levels for imprecision).

Neither trial reported adverse event or fracture data.

Communication aids

Identification bracelet for high-risk fallers

Mayo 1994 (134 participants) studied the effect of wearing a blue identification bracelet on falls in high-risk patients in a subacute hospital setting. They found no reduction in rate of falls (Analysis 18.1.3: RaR 1.15, 95% CI 0.72 to 1.84) or risk of falling (Analysis 18.2.2: RR 1.34, 95% CI 0.76 to 2.36). In this study, there was no reduction in risk of falling in the subgroup with a Short Portable Mental Status Questionnaire (SPMSQ) score < 9 (low cognition) or the subgroup with SPMSQ score ≥ 9 (high cognition). Adverse events were not reported. We are uncertain of the effectiveness of

identification bracelets for reducing falls in hospitals as the quality of the evidence is very low.

Bed exit alarms

Three trials (28,717 participants) examined bed exit alarms in hospital (Shorr 2012; Tideiksaar 1993; Wolf 2013). One large trial (Shorr 2012) was cluster randomised. A summary of the evidence for bed exit alarms for falls prevention in hospitals is provided in Summary of findings 6. Shorr 2012 (27,672 participants) examined an educational intervention to support clinical judgement on the use of bed or chair exit alarms. Wolf 2013 (98 participants) enrolled patients with an increased risk of falling that required assistance with mobilisation during rest time. Pooled data from these two studies did not show a strong reduction in the rate of falls (Analysis 18.1.4: RaR 0.60, 95% CI 0.27 to 1.34: very low-quality evidence) or risk of falling (Analysis 18.2.3: RR 0.93, 95% CI 0.38 to 2.24: very low-quality evidence). We are uncertain whether bed exit alarms reduce the rate of falls or risk of falling as the quality of the evidence has been assessed as very low

Tideiksaar 1993 (70 participants) studied bed exit alarms for preventing falls in hospital. During the nine-month evaluation period, "There was no significant difference in the number of bed-falls between the two groups (p = 1.00)."

Two trials of bed alarms (27,742 participants) indicated that there were no adverse events (Shorr 2012; Tideiksaar 1993); we are uncertain of the effects of bed alarms on adverse events as the quality of the evidence has been assessed as very low (Summary of findings 6).

Social environment

Social environment interventions target staff members and changes in the organisational system, rather than targeting patients directly. Six trials (9074 participants) examined service model change interventions (Dykes 2010; Koh 2009; Mador 2004; Stenvall 2007; Van Gaal 2011b; Wald 2011). Three trials (8587 participants) were cluster randomised (Dykes 2010; Koh 2009; Van Gaal 2011b), and three (487 participants) were individually randomised (Mador 2004; Stenvall 2007; Wald 2011). Studies were not pooled as they were considered to examine clinically heterogenous interventions. One study reported data on risk of fracture (Stenvall 2007). None of the studies reported adverse-event data. We are uncertain of the effects of all social environment interventions in hospitals as the quality of the evidence was assessed as very low.

Service model change

Two studies examined implementation of guidelines in acute care settings in hospitals. Koh 2009 (1122 participants) compared multifaceted fall-prevention guideline implementation with routine dissemination. There was no strong evidence for an effect on the rate of falls (Analysis 19.1.1: RaR 1.82, 95% CI 0.23 to 14.55; very low-quality evidence, downgraded two levels for risk of bias, one level for indirectness and two levels for imprecision). Van Gaal 2011b (2201 participants) studied the implementation of three guidelines (falls, urinary tract infection, pressure ulcers) targeting nursing staff in comparison with usual care. There was no strong evidence for an effect on the rate of falls (Analysis 19.1.2: RaR 0.67, 95% CI 0.17 to 2.59; very low-quality evidence, downgraded two levels for risk of bias, and two levels for imprecision). We are uncertain of the effects of guideline implementation on falls as the quality of the evidence is considered very low.

Dykes 2010 (5264 participants) tested the effect of a computer-based fall-prevention tool kit in comparison with usual care. There was no strong evidence for an effect on the rate of falls (Analysis 19.1.3: RaR 0.55, 95% CI 0.02 to 16.29) or risk of falling (Analysis 19.2.1 RR 0.91, 95% CI 0.06 to 14.21). We are uncertain of the effectiveness of this intervention (very low-quality evidence, downgraded two levels for risk of bias, and two levels for imprecision).

Wald 2011 (217 participants) compared providing care in an acute ward for the elderly with care in general medical wards to usual care. There was no strong evidence for an effect on the rate of falls (Analysis 19.1.4: RaR 0.72, 95% CI 0.10 to 5.10).

Mador 2004 (71 participants) examined a new behavioural advisory service for people with confusion in comparison with usual care. There was no strong evidence for an effect on the risk of falling (Analysis 19.2.2: RR 2.44, 95% CI 0.85 to 7.02).

Stenvall 2007 (199 participants) compared post-operative care in a ward providing a comprehensive ortho-geriatric service with usual care in an orthopaedic ward following surgery for hip fracture. This intervention achieved a reduction in the rate of falls (Analysis 19.1.5: RaR 0.38, 95% CI 0.19 to 0.74) and the risk of falling (Analysis 19.2.3: RR 0.41, 95% CI 0.20 to 0.83) at discharge. There were four new fractures in the control group but none in the intervention group (Analysis 19.3.1: RR 0.11, 95% CI 0.01 to 1.52). These findings also applied to the subgroup analysis of patients with dementia (64 participants), i.e. the rate of falls and risk of falling was reduced (RaR 0.07, 95% CI 0.01 to 0.57; RR 0.12, 95% CI 0.02 to 0.85).

Knowledge interventions

Two trials (3028 participants) examined knowledge interventions in hospitals in individually-randomised trials. Neither trial reported data on the risk of fracture. Haines 2011 reported that there were no adverse events from interaction with the education materials; Ang 2011 did not report on adverse events.

Ang 2011 (1822 participants), testing an educational session by a trained research nurse targeting individual fall risk factors in patients at high risk of falling in an acute setting and achieved a reduction in risk of falling (Analysis 20.2: RR 0.29, 95% CI 0.11 to 0.74); however, we are uncertain of the effects of this intervention as the quality of the evidence has been assessed as

very low (downgraded two levels for risk of bias, one level for indirectness and one level for imprecision).

Haines 2011 (1206 participants) evaluated two forms of multimedia patient education compared with usual care in a mixture of acute and subacute wards. One intervention consisted of written and video-based materials plus one-on-one bedside follow-up from a physiotherapist (complete programme) and the other intervention group received educational materials only. Neither intervention showed strong evidence of a reduction in the rate of falls (Analysis 20.1.1 complete programme RaR 0.83, 95%CI 0.54 to 1.27; very low-quality evidence, downgraded one level for indirectness, one level for inconsistency and one level for imprecision; Analysis 20.1.2 educational materials only RaR 0.91, 95%CI 0.62 to 1.35; low-quality evidence, downgraded one level for indirectness and one level for imprecision) or risk of falling (Analysis 20.2.2 complete programme RR 0.74, 95%CI 0.48 to 1.14; very low-quality evidence, downgraded one level for indirectness, one level for inconsistency and one level for imprecision; Analysis 20.2.3 educational materials only RR 0.84, 95% CI 0.56 to 1.27; low-quality evidence, downgraded one level for indirectness and one level for imprecision). In a post-hoc subgroup analysis, in participants who were cognitively intact the authors reported that falls were less frequent in those receiving the complete programme, compared with those in the materials only group (adjusted hazard ratio (HR) for rate of falls 0.51, 95% CI 0.28 to 0.93; risk of falling 0.65, 95% CI 0.36 to 1.18; 626 participants) and the control group (adjusted HR for rate of falls 0.43, 95% CI 0.24 to 0.78; risk of falling 0.51, 95%CI 0.28 to 0.94; 590 participants) (test for subgroup differences $P < 0.05$). There was a higher risk of injurious falls in those with cognitive impairment with the complete programme (7.49 falls per 1000 patient days compared with 2.89 falls per 1000 patient days in the control group; 192 participants). We are uncertain of the effects of the complete educational programme with follow-up on falls (very low-quality evidence) but providing educational materials only may make little or no difference to the rate of falls or risk of falling (low-quality evidence).

Other single interventions

No included studies examined other single interventions in a hospital setting.

Hospitals: multiple interventions

No included studies examined multiple interventions in a hospital setting.

Hospitals: multifactorial interventions

Six trials (45,416 participants) tested the effect of multifactorial interventions in comparison with usual care in a hospital setting (Aizen 2015; Barker 2016; Cumming 2008; Haines 2004; Healey 2004; Hill 2015). Five trials (44,790 participants) were cluster randomised (Aizen 2015; Barker 2016; Cumming 2008; Healey 2004; Hill 2015), and one (626 participants) was individually randomised (Haines 2004). Two trials used a stepped-wedge design (Aizen 2015; Hill 2015). The categories of interventions for each trial are shown in Appendix 3 and further details are provided in the Characteristics of included studies. A summary of the evidence for multifactorial interventions for falls prevention in hospitals is provided in Summary of findings 7. Two studies (4625 participants) reported data on risk of fracture (Cumming 2008; Haines 2004). Four of six trials (39,763 participants) reported on adverse events (Aizen 2015; Barker 2016; Haines 2004; Hill 2015). We have shown whether the settings were acute or subacute in the footnotes of the analyses. Given most of these trials were large with important differences such as in the setting and in the format and delivery of their multifactorial intervention, we present some details of the individual trials first before reporting the pooled analyses.

Aizen 2015 (752 participants) conducted a two-stage (stepped-wedge) cluster randomised trial in five geriatric rehabilitation wards. The multifactorial intervention included medical, behavioural, cognitive and environmental modifications with additional orientation guidance and mobility restriction for moderate-risk patients and permanent personal supervision for high-risk patients. The usual care arm included any activities undertaken by the participants recommended or administered by their treating team. The authors reported that "No significant difference was found in fall rates during follow-up between intervention and control wards". The findings of this study were not pooled as some aspects of the study methodology and data collection could not be confirmed.

Barker 2016 (35,264 participants, 46,245 admissions) investigated a "6-PACK" intervention in comparison with usual care (which included standard falls prevention activities) with a cluster-randomised trial in 24 acute medical or surgical wards and found no change in rate of falls or risk of falling. There was no evidence of effect on the rate of injurious falls (RaR 0.96, 95% CI 0.72 to 1.27). Data were determined based on admissions, some patients were admitted more than once.

Cumming 2008 (3999 participants) examined an intervention in both acute and subacute wards in which a nurse and physiotherapist each worked for 25 hours per week for three months in all intervention wards. No trial interventions were delivered in the usual care arm. This trial also found no change in the rate of falls or risk of falling. The review authors consider both Barker 2016 and Cumming 2008 to be well-conducted trials. The interventions they studied would be regarded as sound falls prevention practice including use of falls risk-assessment tools and supervision for patients at risk but no effect on falls was observed.

The multidisciplinary intervention in Haines 2004 (626 participants) took place in three subacute wards. The programme in-

cluded a falls risk alert card with an information brochure, exercise, education programme, and hip protectors, in addition to usual care. In the control arm, patients received usual care but none of the interventions from the falls prevention programme; the study staff completed the risk assessment and generated recommendations but none of these recommendations were instituted. The authors reported that the difference in falls between the two groups was "most obvious after 45 days of observation", suggesting that this programme benefited people staying longer in hospital but it could also be explained by long staying frequent fallers in the control group.

Healey 2004 (1654 participants) examined a risk-factor reduction care plan for patients with a history of falls in a cluster-randomised trial in eight acute and subacute wards. Interventions included assessment and interventions targeted at eyesight, medications, blood pressure management, mobility, urine testing, bed rail use, bed height, footwear, ward positioning, environmental causes and call bells. In the usual care arm, the care plan was not introduced and no changes to practice or environment relevant to falls prevention were made during the study.

Hill 2015 conducted a stepped-wedge cluster-randomised controlled trial in eight hospital rehabilitation and geriatric wards (3121 participants, 3606 admissions), which tested the effect of an individualised multimedia education intervention (also tested in Haines 2011) provided to eligible patients with basic cognition, and staff, aiming to educate patients about falls prevention strategies and to motivate engagement in falls-prevention strategies (ProFaNE categories of social environment and knowledge). Usual care included patient's screening, assessment and implementation of individualised falls-prevention strategies, ongoing staff training and environmental strategies. There was a reduction in the rate of falls (Analysis 21.1: RaR 0.60, 95% CI 0.42 to 0.94). There was also a reduction in the rate of injurious falls (adjusted RaR 0.65, 95% CI 0.42 to 0.88; data analysed by number of admissions rather than participants).

In a pre-specified subgroup analysis, Hill 2015 reported that the rate of falls was reduced in people without significant cognitive impairment who received the educational intervention (MMSE > 23/30; adjusted RaR 0.53, 95%CI 0.36 to 0.77, P < 0.001; 1930 participants), but there was no strong evidence for an effect in the subgroup of patients who were cognitively impaired (who did not receive the patient intervention, but may have benefited from the staff training intervention component; adjusted RaR 0.65, 95% CI 0.40 to 1.05; 1676 participants).

Rate of falls

Pooled results from five trials (44,664 participants) of multifactorial interventions showed a borderline reduction in the rate of falls, with a reduction overall of 20%; the 95% confidence intervals indicated this estimate of effect may range as high as a reduction of 36% or result in an increase in falls rates of 1%; (Analysis 21.1: RaR random-effects 0.80, 95% CI 0.64 to 1.01; 5 trials: I^2 = 52%; low-quality evidence, downgraded one level for risk of bias and one level for imprecision; Summary of findings 7). These findings were further explored in a subgroup analysis by setting (see below).

Risk of falling

Pooled data from three trials (39,889 participants) of the five trials pooled for the rate of falls outcome were generally consistent with the effect estimate for the rate of falls with a reduction in the risk of falling that did not reach statistical significance (Analysis 21.2: RR random-effects 0.82, 95% CI 0.62 to 1.09; 3 trials: I^2 = 0%; very low-quality evidence; Summary of findings 7). Notably Hill 2015 reported a reduction in the risk of falling (adjusted odds ratio (OR) 0.55, 95% CI 0.38 to 0.81) in a subacute setting; however, these data were analysed by number of admissions, rather than participants, so these data were not pooled. The choice of model for the pooled analysis did not affect the estimate of effect as the statistical heterogeneity was 0%. We are uncertain of the effects of multifactorial interventions on risk of falling in hospitals (very low-quality evidence).

Risk of fracture

Two trials (4625 participants; Cumming 2008; Haines 2004) reported fracture data suitable for pooling. There was no strong evidence for a reduction in the number of people sustaining a fracture (Analysis 21.3: RR 0.76, 95% CI 0.14 to 4.10: I^2 = 0%; nine fractures; very low-quality evidence; Summary of findings 7).

In Barker 2016, there were very few fractures in an acute setting, with 11 (0.06%) people experiencing a fall-related fracture in the intervention arm and 13 (0.07%) in the control arm. In Hill 2015, there were six fractures in the control group (three hip fractures) and four in the intervention group (not hip) in a subacute setting; these data represent number of fractures and admissions rather than patients. The data from these two studies are not pooled; however, the results are consistent with the pooled estimate showing no strong effect on the risk of fracture.

We are uncertain whether multifactorial interventions reduce the risk of fracture as the quality of the evidence has been assessed as very low.

Adverse events

No adverse events were reported in the four trials (39,763 participants; Aizen 2015; Barker 2016; Haines 2004; Hill 2015) that reported this outcome. We are uncertain of the effects of multifactorial interventions on adverse events as the quality of the evidence has been assessed as very low (Summary of findings 7).

Subgroup analysis by type of care (acute, subacute or mixed settings)

A post-hoc subgroup analysis was conducted for multifactorial interventions conducted in hospitals for acute care settings, subacute settings or mixed (both subacute and acute) settings. The test for subgroup differences indicated a possible difference between the settings (types of care) for rate of falls (Analysis 22.1, P = 0.04). Pooled data indicate a reduction in the falls rate in trials conducted in the subacute setting (Analysis 22.1.3: RaR 0.67, 95% CI 0.54 to 0.83), but not in the acute (Analysis 22.1.1: RaR 1.04, 95% CI 0.79 to 1.37) or mixed settings (Analysis 22.1.2: RaR 0.88, 95% CI 0.61 to 1.27). There were no differences between subgroups for pooled data by setting for risk of falling (Analysis 22.2, test for subgroup differences P = 0.75) or risk of fracture (Analysis 22.3, test for subgroup differences P = 0.56). One additional study reporting data for the risk of falling and fracture that were not pooled was conducted in a subacute setting (Hill 2015).

Multifactorial interventions including targeted patient education may reduce the rate of falls in a subacute setting (low-quality evidence, downgraded one level for risk of bias and one level for inconsistency due to some uncertainty in the subgroup analysis).

Studies in participants with cognitive impairment

Eleven trials reported findings specifically for patients with dementia or cognitive impairment.

Care facilities

In care facilities, Juola 2015 (227 participants) included 93% of participants with a dementia diagnosis in a trial of nurse education on harmful medications. The intervention showed a reduction in the rate of falls in those with an MMSE score of 10 or greater, but no strong evidence of an effect in those with an MMSE of less than 10. In a trial of a multifactorial intervention (Whitney 2017; 191 participants), 97% of participants were cognitively impaired but the intervention did not show any strong evidence for an effect on the rate of falls or risk of falling. The effects of combination exercise, a multimodal exercise programme, a behaviour advisory service for people with confusion, dementia care mapping, and multisensory stimulation in a Snoezelen room have been examined in people with dementia in several studies (Chenoweth 2009; Klages 2011; Kovacs 2013; Mador 2004; Toulotte 2003; Van de Ven 2014). However, these interventions were tested in single small studies or the studies did not report data suitable for further analysis. Chenoweth 2009 and Buettner 2002 reported costs associated with interventions for participants with dementia in care facilities.

Hospitals

In hospitals, a knowledge-based intervention that did not show strong evidence for a reduction in the rate of falls overall showed a reduction in falls in those who were cognitively intact, but not in those with cognitive impairment in a post-hoc analysis (Haines 2011). When the intervention was applied as a multifactorial intervention, only delivered to those with basic cognition, a reduction in both the rate of falls and risk of falling was observed (Hill 2015). In an acute hospital setting, Stenvall 2007 found that a multifactorial intervention including comprehensive geriatric assessment and rehabilitation for people with femoral neck fractures reduced falls in a subgroup with dementia, however the number of participants was low and the evidence assessed as very low quality, so we are uncertain of the effectiveness of this intervention.

Economic evaluations

The 11 studies reporting economic outcomes (nine in care facilities and two in a hospital setting) are summarised in Appendix 10. Only one study (Haines 2013), reported an economic evaluation in terms of the cost to prevent falls.

In a subgroup of hospital inpatients who were cognitively intact, a falls patient education programme in a hospital setting had a cost of AUD 294 to prevent one fall and AUD 526 to prevent one faller (Haines 2013).

ADDITIONAL SUMMARY OF FINDINGS [*Explanation*]

General medication review compared with usual care for falls prevention in care facilities

Population and setting: older (≥ 65 years) residents of care facilities
Intervention: general medication review (NB: the primary aim of all medication review is to reduce psychoactive medications)
Comparison: usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Usual care	General medication review				
Rate of falls Length of follow-up: 6 to 12 months	Low-risk population[1]		**RaR 0.93** (0.64 to 1.35)	**2409** (6 studies)	++oo LOW[7]	The approaches taken in the six pooled studies were: • medication review meeting involving clinical pharmacist, pharmacy technician, care home staff and GP(s) • medication review with recommendations to chief physician based on STOPP/START criteria • nurse education on harmful medications in older people • monthly reviews of psychoactive

						medications • medication review and deprescribing • GP record + consultation with patient and care 2 additional RCTs found no strong evidence for an effect on the rate of falls (1 RCT of education of physicians on drug use in older people (716 participants, falls only reported following the intervention period); 1 trial of antidepressant deprescribing (36 participants randomised))
	1000 per 1000 py	**930** (640 to 1350)**per 1000 py**				
	High-risk population[2]					
	3500 per 1000 py	**3255** (2240 to 4725)**per 1000 py**				
Risk of falling Length of follow-up: 6 to 12 months	Low-risk population[3]		**RR 0.93** (0.80 to 1.09)	**5139** (6 studies)	++oo LOW[8]	The approaches taken in the six studies were: • pharmacist transition coordinator for patients discharged from hospital to nursing care facilities for the first time • a pharmacist-led outreach programme (audit + feedback + education of staff regarding medications and falls risk) • nurse education

						on harmful medications in older people • GRAM software for decision support for prescribing practices • GP and a geriatrician / pharmacologist independently identifying deprescribing targets using a list of potentially inappropriate medicines vs medication review without deprescribing • review of GP record + consultation with patient and carer 1 additional RCT of education of physicians on drug use in older people (716 participants) found no strong evidence for an effect on the risk of falling following the intervention period
	250 per 1000	**233** (200 to 273) **per 1000**				
	Moderate-risk population[4]					
	500 per 1000	**465** (400 to 545) **per 1000**				
	High-risk population[5]					
	700 per 1000	**651** (560 to 763) **per 1000**				
Risk of fracture Length of follow-up: 12 months	Average risk population[6]		**RR 1.60** (0.28 to 9.16)	**93** (1 trial)	+ooo VERY LOW[9]	Intervention was GP and a geriatrician/ pharmacologist independently identifying

						deprescribing targets using a list of potentially inappropriate medicines vs medication review without deprescribing
	42 per 1000	**67** (12 to 614) **per 1000**				
Adverse events Length of follow-up: 12 months	Average risk population[10]		**RR 1.07** (0.23 to 5.01)	**93** (1 trial)	+ooo VERY LOW[9]	Serious vascular events in both trial arms and significant withdrawal reactions in 2 intervention participants (Potter 2016).
	60 per 1000	**64** (14 to 301) **per 1000**				

Illustrative risks for the control group were derived from all or subgroups of trials in care facilities reporting the outcome. The exact basis for the **assumed risk for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Low risk was based on the mean control risk of the 17 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.07, rounded to 1.0 per person year; thus 1000 per 1000 person years.
[2] High risk was based on the mean control risk of the 18 (top third) trials with the highest rate of falls. The mean rate of falls = 3.69, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[3] Low risk was based on the mean control risk of the 20 trials with the lowest risk of falling. The mean risk of falling = 0.268, rounded to 0.25; thus 250 per 1000 people.
[4] Moderate risk was based on the mean control risk of the 20 trials reporting a moderate risk of falling, not described as high-risk populations. The mean risk of falling = 0.539, rounded to 0.5; thus 500 per 1000 people.
[5] High risk was based on the mean control risk of the 13 trials reporting a high risk of falling, including populations with a description as a high-risk population. The mean risk of falling = 0.680, rounded to 0.7; thus 700 per 1000 people.
[6] Risk based on the median control risk of fracture of the trials reporting this outcome. Median risk = 0.042; thus 42 per 1000.

[7] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and detection bias, high or unclear risk of method of ascertaining falls, and high risk of baseline imbalance) and one level due to inconsistency (unexplained heterogeneity, I^2 = 93%).

[8] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and detection bias, baseline imbalance, method of ascertaining falls and high or unclear risk of selection bias), and one level for inconsistency (I^2 = 48%, P > 0.05; inconsistency in point estimates between studies).

[9]The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and detection bias), one level for indirectness (sIngle trial conducted in rural Western Australia (Potter 2016) that may have limited applicability), two levels for imprecision (extremely wide confidence intervals that include the possibility of both important benefit and harm) and one level for publication bias (few studies reported this outcome).

[10] Determined from the control arm of Potter 2016.

Vitamin D supplementation compared with no vitamin D supplementation for falls prevention in care facilities

Population and setting: older (≥ 65 years) residents of care facilities[1]
Intervention: vitamin D supplementation (vitamin D or vitamin D + calcium)
Comparison: usual care (or calcium supplementation)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Vitamin D				
Rate of falls Length of follow-up: 3 to 24 months	Low-risk population[2]		**RaR 0.72** (0.55 to 0.95)	**4512** (4 studies)	+++o MODERATE[8]	Studies included two studies of vitamin D3 + calcium versus calcium, and 2 studies of vitamin D2 versus usual care or placebo
	1000 per 1000 py	**720** (550 to 950)**per 1000 py**				
	High-risk population[3]					
	3500 per 1000 py	**2520** (1925 to 3325)**per 1000 py**				
Risk of falling Length of follow-up: 3 to 24 months	Low-risk population[4]		**RR 0.92** (0.76 to 1.12)	**4512** (4 studies)	+++o MODERATE[9]	Studies included two studies of vitamin D3 + calcium versus calcium, and 2 studies of vitamin D2 versus usual care or placebo
	250 per 1000	**230** (190 to 280)**per 1000**				
	Moderate-risk population[5]					
	500 per 1000	**460** (380 to 515)**per 1000**				
	High-risk population[6]					
	700 per 1000	**644** (532 to 784)**per 1000**				

Risk of fracture Length of follow-up: 3 to 24 months	Average risk population[7]		**RR 1.09** (0.58 to 2.03)	**4464** (3 studies)	+ooo VERY LOW[10]	These studies represent only a subset of studies evaluating the effect of vitamin D on fractures. Included studies were two studies of vitamin D3 + calcium versus calcium, and 1 study of vitamin D2 versus usual care
	42 per 1000	**46** (24 to 85) **per 1000**				
Adverse events Length of follow-up: 3 to 24 months	ND[12]	ND[12]	**RR 4.84** (0.24 to 98.90)	**747** (2 studies)	+ooo VERY LOW[11]	No serious events reported. Studies tested supplementation with 800 IU oral cholecalciferol (vitamin D3) and 1000 IU oral ergocalciferol (vitamin D2) daily Data derived from just 2 cases of increased constipation in the intervention arm in 1 study (N = 122). No adverse events recorded in the other study (N = 625)

*Illustrative risks for the control group were derived from all or subgroups of trials in care facilities reporting the outcome. The exact basis for the **assumed risk** for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **ND**: not done; **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Studies confirmed the participants had low or very low serum vitamin D levels at baseline.

[2] Low risk was based on the mean control risk of the 17 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.07, rounded to 1.0 per person year; thus 1000 per 1000 person years.
[3] High risk was based on the mean control risk of the 18 (top third) trials with the highest rate of falls. The mean rate of falls = 3.69, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[4] Low risk was based on the mean control risk of the 20 trials with the lowest risk of falling. The mean risk of falling = 0.268, rounded to 0.25; thus 250 per 1000 people.
[5] Moderate risk was based on the mean control risk of the 20 trials reporting a moderate risk of falling, not described as high-risk populations. The mean risk of falling = 0.539, rounded to 0.5; thus 500 per 1000 people.
[6] High risk was based on the mean control risk of the 13 trials reporting a high risk of falling, including populations with a description as a high-risk population. The mean risk of falling = 0.680, rounded to 0.7; thus 700 per 1000 people.
[7] Risk based on the median control risk of fracture of the trials reporting this outcome. Median risk = 0.042; thus 42 per 1000.
[8] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and detection bias and method of ascertaining falls for one trial contributing 49%).
[9] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and detection bias and method of ascertaining falls for one trial contributing 56%).
[10] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and detection bias and method of ascertaining falls for one trial contributing 49%), and two levels for imprecision (small number of fractures, confidence intervals cross the range of strong effect and significant harm).
[11] The quality of the evidence was downgraded two levels for imprecision (low event rate, inadequate power to assess rare adverse events) and two levels for other reasons (concerns that adverse events were not recorded systematically and likely publication bias, few studies reported this outcome).
[12] Not done. Illustrative comparative risks not presented as considered uninformative due to paucity of data available.

Multifactorial interventions compared with usual care for falls prevention in care facilities

Population and setting: older (≥ 65 years) residents of care facilities
Intervention: multifactorial interventions (two or more categories of intervention given based on individual risk profile)
Comparison: usual care (without intervention)[1]

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Vitamin D				
Rate of falls Length of follow-up: 6 to 12 months	Low-risk population[2]		**RaR 0.88** (0.66 to 1.18)	**3439** (10 studies)	+ooo VERY LOW[8]	One additional study (31 participants) of exercise plus nutritional support reported zero falls in the intervention arm and two in the control arm
	1000 per 1000 py	**720** (550 to 950)**per 1000 py**				
	High-risk population[3]					
	3500 per 1000 py	**2520** (1925 to 3325)**per 1000 py**				
Risk of falling Length of follow-up: 6 to 12 months	Low-risk population[4]		**RR 0.92** (0.81 to 1.05)	**3153** (9 studies)	++oo LOW[9]	One additional study (482 participants) reported a reduction in the proportion of recurrent fallers (difference 19%, 95% CI 2% to 36%: P = 0.03)
	250 per 1000	**230** (190 to 280)**per 1000**				
	Moderate-risk population[5]					
	500 per 1000	**460** (380 to 515)**per 1000**				
	High-risk population[6]					
	700 per 1000	**644** (532 to 784)**per 1000**				

Risk of fracture Length of follow-up: 6 to 12 months	Average risk population[7]		**RR 0.79** (0.30 to 2.07)	**2160** (5 studies)	+ooo VERY LOW[10]	
	42 per 1000	**34** (13 to 87) **per 1000**				
Adverse events Length of follow-up: 11 weeks to 12 months	See comment	See comment	Not estimable.	312 (3 studies)	+ooo VERY LOW[11]	One trial reported a case of a fall in the intervention arm; two studies reported no adverse events

*Illustrative risks for the control group were derived from all or subgroups of trials in care facilities reporting the outcome. The exact basis for the **assumed risk** for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Nine of 11 trials described the control arm as usual care not receiving the intervention. In one trial contributing data to the risk of falling and fracture, the control arm received multidisciplinary assessment without the intervention in addition to usual care; in one trial contributing data to the rate of falls and risk of falling, the control included reminiscence therapy.
[2] Low risk was based on the mean control risk of the 17 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.07, rounded to 1.0 per person year; thus 1000 per 1000 person years.
[3] High risk was based on the mean control risk of the 18 (top third) trials with the highest rate of falls. The mean rate of falls = 3.69, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[4] Low risk was based on the mean control risk of the 20 trials with the lowest risk of falling. The mean risk of falling = 0.268, rounded to 0.25; thus 250 per 1000 people.
[5] Moderate risk was based on the mean control risk of the 20 trials reporting a moderate risk of falling, not described as high-risk populations. The mean risk of falling = 0.539, rounded to 0.5; thus 500 per 1000 people.
[6] High risk was based on the mean control risk of the 13 trials reporting a high risk of falling, including populations with a description as a high-risk population. The mean risk of falling = 0.680, rounded to 0.7; thus 700 per 1000 people.
[7] Risk based on the median control risk of fracture of the trials reporting this outcome. Median risk = 0.042; thus 42 per 1000.

[8] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and attrition bias and baseline imbalance), one level for serious inconsistency (high heterogeneity I^2 = 84%) and one level for imprecision (wide CIs despite large N).

[9] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and attrition bias and some uncertainty in selection bias) and one level for inconsistency (inconsistency in point estimates between studies).

[10] The quality of the evidence was downgraded one level for serious risk of bias (including high risk of performance and attrition bias and baseline imbalance), one level for inconsistency (moderate heterogeneity, I^2 = 60%, P = 0.04) and two levels for imprecision (extremely wide confidence intervals)

[11] The quality of the evidence was downgraded two levels for serious risk of bias (2 of 3 trials had a high risk of baseline imbalance or incomplete outcome data), two levels for imprecision (not powered for rare events) and two levels for other reasons (concerns that adverse events were not recorded systematically and few studies reported this outcome).

Additional exercise plus physiotherapy compared with usual physiotherapy for falls prevention in hospitals

Population and setting: older (≥ 65 years) patients in hospital settings
Intervention: additional exercise plus physiotherapy
Comparison: usual physiotherapy

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Usual physiotherapy	Additional Exercise				
Rate of falls Length of follow-up: in-patient stay (mean 29 days) or 2 weeks	Low-risk population[1]		**RaR 0.59** (0.26 to 1.34)	**215** (2 studies)	+ooo VERY LOW[7]	One study compared additional exercises versus conventional physiotherapy alone, and 1 study tested additional group standing balance circuit classes
	1300 per 1000 py	**767** (338 to 1742) **per 1000 py**				
	Moderate-risk population[2]					
	3500 per 1000 py	**2065** (910 to 4690) **per 1000 py**				
	High-risk population[3]					
	6000 per 1000 py	**3540** (1560 to 8040) **per 1000 py**				
Risk of falling Length of follow-up: in-patient stay (mean 29 days) or 8 weeks	Low-risk population[4]		**RR 0.36** (0.14 to 0.93)	**83** (2 studies)	+ooo VERY LOW[8]	One study compared additional exercises versus conventional physiotherapy alone, and 1 study tested additional daily physiotherapy sessions

	30 per 1000	**11** (4 to 28) **per 1000**				
	Moderate-risk population[5]					
	150 per 1000	**54** (21 to 140) **per 1000**				
	High-risk population[6]					
	340 per 1000	**122** (48 to 316) **per 1000**				
Risk of fracture	See comment	See comment	See comment			No data available
Adverse events Length of follow-up: 2 weeks	0 events	0 events	Not estimable	**161** (1 study)	+ooo VERY LOW[9]	One study reported no adverse events, two studies did not report this outcome

*Illustrative risks for the control group were derived from all or subgroups of trials in hospitals reporting the outcome. The exact basis for the **assumed risk** for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Low risk was based on the mean control risk of the 7 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.27, rounded to 1.3 per person year; thus 1300 per 1000 person years.
[2] Moderate risk was based on the mean control risk of the 7 (middle third) trials with a moderate rate of falls. The mean rate of falls = 3.23, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[3] High risk was based on the mean control risk of the 7 (top third) trials with the highest rate of falls. The mean rate of falls = 6.33, rounded to 6.0 per person year; thus 6000 per 1000 person years.
[4] Low risk was based on the mean control risk of 10 trials with the lowest risk of falling. The mean risk of falling = 0.034, rounded to 0.03; thus 30 per 1000 people.

[5] Moderate risk was based on the mean control risk of 7 (middle third) trials reporting the risk of falling. The mean risk of falling = 0.156, rounded to 0.15; thus 150 per 1000 people.

[6] High risk was based on the mean control risk of 6 (top third) trials reporting the risk of falling. The mean risk of falling = 0.340; thus 340 per 1000 people.

[7]The quality of the evidence was downgraded one level for risk of bias (including unclear risk of selection bias and method of ascertaining falls in one study) and two levels for very serious imprecision (the wide confidence intervals cross the range of estimates of harm and strong effect).

[8]The quality of the evidence was downgraded one level for risk of bias (including unclear risk of bias in both trials for selection bias and high risk of attrition bias for study contributing 69%), one level for indirectness (possibly limited applicability as both trials conducted in UK rehabilitation settings) and one level for imprecision (total N = 83, wide 95% confidence intervals).

[9]The quality of the evidence was downgraded one level for indirectness (single trial in Australian rehabilitation setting), two levels for imprecision (no events recorded, inadequate power to assess rare adverse events) and one level for other reasons (concerns that adverse events were not recorded systematically).

Bed alarms compared with usual care for falls prevention in hospitals

Population and setting: older (≥ 65 years) patients in hospital settings
Intervention: bed alarms
Comparison: usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	**Corresponding risk**				
	Usual care	**Bed alarms**				
Rate of falls Length of follow-up: in-patient stay (mean 19 days; not known)	Low-risk population[1]		**RaR 0.60** (0.27 to 1.34)	**28,649** (2 studies)	++oo VERY LOW[7]	One cluster-randomised study tested education and support on using bed/chair alarms; and one study tested sensor alarms fitted to patients' upper leg at rest time A third study (n = 70) reported no difference in the number of falls (data not suitable for pooling)
	1300 per 1000 py	**780** (351 to 1742) **per 1000 py**				
	Moderate-risk population[2]					
	3500 per 1000 py	**2100** (945 to 4690) **per 1000 py**				
	High-risk population[3]					
	6000 per 1000 py	**3600** (1620 to 8040) **per 1000 py**				
Risk of falling Length of follow-up: in-patient stay (mean 19 days; not known)	Low-risk population[4]		**RR 0.93** (0.38 to 2.24)	**28,649** (2 studies)	+ooo VERY LOW[8]	One cluster-randomised study tested education and support on using bed/chair alarms; and one study tested sensor alarms fitted to patients' upper leg at rest time

	30 per 1000	**28** (11 to 67) **per 1000**				
	Moderate-risk population[5]					
	150 per 1000	**140** (57 to 336) **per 1000**				
	High-risk population[6]					
	340 per 1000	**316** (129 to 762) **per 1000**				
Risk of fracture	See comment	See comment	See comment			No data available.
Adverse events Length of follow-up: in-patient stay (mean 19 days; not known)	0 events	0 events	Not estimable.	**27,742** (2 studies)	+ooo VERY LOW[9]	2 trials reported that there were no adverse events

*Illustrative risks for the control group were derived from all or subgroups of trials in hospitals reporting the outcome. The exact basis for the **assumed risk** for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Low risk was based on the mean control risk of the 7 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.27, rounded to 1.3 per person year; thus 1300 per 1000 person years.
[2] Moderate risk was based on the mean control risk of the 7 (middle third) trials with a moderate rate of falls. The mean rate of falls = 3.23, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[3] High risk was based on the mean control risk of the 7 (top third) trials with the highest rate of falls. The mean rate of falls = 6.33, rounded to 6.0 per person year; thus 6000 per 1000 person years.

[4] Low risk was based on the mean control risk of 10 trials with the lowest risk of falling. The mean risk of falling = 0.034, rounded to 0.03; thus 30 per 1000 people.

[5] Moderate risk was based on the mean control risk of 7 (middle third) trials reporting the risk of falling. The mean risk of falling = 0.156, rounded to 0.15; thus 150 per 1000 people.

[6] High risk was based on the mean control risk of 6 (top third) trials reporting the risk of falling. The mean risk of falling = 0.340; thus 340 per 1000 people.

[7]The quality of the evidence was downgraded one level for risk of bias (including high risk of selection bias and unclear risk of bias for balance in baseline characteristics in the larger trial, a cluster RCT, Shorr 2012; unclear or high risk of bias for all domains for trial with greatest weighting; risk of performance and detection bias due to lack of blinding although this is not feasible); one level for imprecision (despite the large sample size, the wide confidence intervals cross the range of strong effect and significant harm) and one level for indirectness (the larger trial, Shorr 2012, is of education and support on using bed alarms, rather than directly implementing bed alarms).

[8]The quality of the evidence was downgraded one level for risk of bias (including high risk of selection bias and unclear risk of bias for balance of baseline characteristics in the larger trial, Shorr 2012), one level for indirectness (the larger trial, Shorr 2012, is of education and support on using bed alarms, directly implementing bed alarms) and one level for imprecision, despite the large sample size, the wide confidence intervals cross the range of strong effect and significant harm).

[9]The quality of the evidence was downgraded one level for risk of bias (including high risk of selection bias and unclear risk of bias for balance of baseline characteristics, one level for indirectness (trial is of education and support on using bed alarms, directly implementing bed alarms) and one level for imprecision (no events recorded, low power to assess rare adverse events) and one level for other reasons (concerns that adverse events were not recorded systematically).

Multifactorial interventions compared with usual care for falls prevention in hospitals

Population and setting: older (≥ 65 years) patients in hospital settings
Intervention: multifactorial interventions (two or more categories of intervention given based on individual risk profile)
Comparison: usual care [1]

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Usual care	Multifactorial				
Rate of falls Length of follow-up: in-patient stay (median 4 days to mean 30 days)	Low-risk population[2]		**RaR 0.80** (0.64 to 1.01)	**44,664** (5 studies)	++oo LOW[9]	The 5 studies tested compared different multifactorial interventions versus usual care in acute, subacute or mixed care settings • 1 study (acute care) tested risk assessment and up to 6 interventions for high-risk patients, plus staff education • 1 study (acute and subacute care) tested risk assessment, staff and patient education, drug review, environmental modifications and exercise • 1 study (subacute care) tested risk

						assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) • 1 study (acute and subacute care) tested risk factor screening and targeted care plan in at-risk patients • 1 study (subacute care) tested a multimedia falls education with follow-up for patients plus staff education and feedback. See footnote[13] for comment on a post-hoc subgroup analysis by setting.
	1300 per 1000 py	**1040** (832 to 1313) **per 1000 py**				
	Moderate-risk population[3]					
	3500 per 1000 py	**2,800** (2240 to 3535) **per 1000 py**				
	High-risk population[4]					
	6000 per 1000 py	**4800** (3840 to 6060) **per 1000 py**				
Risk of falling Length of follow-up: inpatient stay (median 4 days to mean 30 days)	Low-risk population[5]		**RR 0.82** (0.62 to 1.09)	**39,889** (3 studies)	+ooo VERY LOW [10]	The 3 studies tested compared different multifactorial interventions versus usual care in acute, subacute or mixed care settings • 1 study (acute care) tested risk assessment and up to 6 interventions for high-risk patients, plus staff education • 1 study (acute and subacute care) tested

						risk assessment, staff and patient education, drug review, environmental modifications and exercise • 1 study (subacute care) tested risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) One additional study analysed fallers by the number of admissions, and found a reduction in the risk of falling (adjusted OR 0.55, 95% CI 0.38 to 0.81)
	30 per 1000	**25** (19 to 33) **per 1000**				
	Moderate-risk population[6]					
	150 per 1000	**123** (93 to 164) **per 1000**				
	High-risk population[7]					
	340 per 1000	**279** (211 to 371) **per 1000**				
Risk of fracture Length of follow-up: inpatient stay (mean in acute wards 8 days to mean 30 days)	Average risk population[8]		**RR 0.76** (0.14 to 4.10)	**4615** (2 studies)	+ooo VERY LOW[11]	The 2 studies pooled tested compared different multifactorial interventions versus usual care in subacute or mixed care settings • 1 study (acute and subacute care) tested risk assessment, staff and patient education, drug review, environmental modifications and exercise

						• 1 study (subacute care) tested risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) Two additional studies reported no difference in extremely low risk of fracture (1 study Intervention: 11/17698, 0.06% vs Control: 13/17566, 0.07%) or number of fractures (Intervention 4/1402 0.3% vs 6/1719, 0.3%)
	18 per 1000	**14** (3 to 74) **per 1000**				
Adverse events Length of follow-up: in-patient stay	0 events	0 events	Not estimable.	**39,763** (4 studies)	+ooo VERY LOW[12]	4 trials reported that there were no adverse events.

*Illustrative risks for the control group were derived from all or subgroups of trials in hospitals reporting the outcome. The exact basis for the **assumed risk** for each outcome is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **MultiF:** multifactorial; **OR:** Odds Ratio; **OT:** Occupational Therapist **py:** person years; **RaR:** Rate Ratio; **RR:** Risk Ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Usual care generally included some standard falls prevention activities. The degree to which this included components of the intervention was not always clear. Usual care falls prevention activities are likely to very over time and between settings.
[2] Low risk was based on the mean control risk of the 7 (bottom third) trials with the lowest rate of falls. The mean rate of falls = 1.27, rounded to 1.3 per person year; thus 1300 per 1000 person years.

[3] Moderate risk was based on the mean control risk of the 7 (middle third) trials with a moderate rate of falls. The mean rate of falls = 3.23, rounded to 3.5 per person year; thus 3500 per 1000 person years.
[4] High risk was based on the mean control risk of the 7 (top third) trials with the highest rate of falls. The mean rate of falls = 6.33, rounded to 6.0 per person year; thus 6000 per 1000 person years.
[5] Low risk was based on the mean control risk of 10 trials with the lowest risk of falling. The mean risk of falling = 0.034, rounded to 0.03; thus 30 per 1000 people.
[6] Moderate risk was based on the mean control risk of 7 (middle third) trials reporting the risk of falling. The mean risk of falling = 0.156, rounded to 0.15; thus 150 per 1000 people.
[7] High risk was based on the mean control risk of 6 (top third) trials reporting the risk of falling. The mean risk of falling = 0.340; thus 340 per 1000 people.
[8] Risk based on the median risk of fracture in the control arm of trials reporting this outcome. Median risk = 0.018; thus 18 per 1000 people.
[9] The quality of the evidence was downgraded one level for risk of bias (including high risk of selection bias, performance and detection bias) and one level for imprecision (confidence intervals overlap no effect but fail to exclude important benefit)
[10] The quality of the evidence was downgraded one level for risk of bias (including high risk of selection bias, performance and detection bias), one level for imprecision (confidence intervals overlap no effect but fail to exclude important benefit) and one level for other bias (one study not included in pooled estimate creating uncertainty in overall point estimate).
[11] The quality of the evidence was downgraded one level for risk of bias (including high risk of selection bias, performance and detection bias) and two levels for imprecision (small number of fractures, the extremely wide confidence intervals include both possible benefit and possible harm).
[12] The quality of the evidence was downgraded one level for risk of bias (including high risk of selection and performance bias and baseline imbalance), one level for imprecision (no events recorded, low power to assess rare adverse events) and one level for other reasons (concerns that adverse events were not recorded systematically).
[13] A post-hoc subgroup analysis by setting found a reduction in the rate of falls in 2 trials conducted in a subacute setting (RaR 0.67, 95% CI 0.54 to 0.83; 2 trials; 3747 participants; test for subgroup differences P = 0.04). These trials included including targeted patient education as a component of the multifactorial intervention.

DISCUSSION

Summary of main results

This review now includes 95 trials (138,164 participants) of which 71 trials (40,374 participants; mean age 84 years; 75% women) were in care facilities and 24 trials (97,790 participants; mean age 78 years; 52% women) were in hospitals. Despite the addition of 35 trials (77,869 participants) to the previous review, many of the results from the pooled analyses remain inconsistent and inconclusive. Although 24 trials reported data on fractures suitable for use in the analyses, all fracture data were very low-quality evidence and thus we are uncertain of the effects of any intervention on risk of fracture. Twenty-nine trials clearly reported data on adverse events, although in several it was to report an absence of adverse events. There were very few serious adverse events and minor complications, where reported, were usually similar in the intervention and control groups. Overall, we are uncertain of the effects on adverse events as the quality of the evidence has been assessed as very low.

Care facilities

Exercise

Twenty-five trials in care facilities investigated exercise as a single intervention. Despite the large number of trials, many were small (< 100 participants). Only two trials reported the effects of exercise on risk of fracture and nine on adverse events.

Seventeen trials compared an exercise intervention with usual care. A summary of the evidence for exercise in comparison with usual care in care facilities is provided in Summary of findings for the main comparison. Funnel plots of the pooled trials (10 trials each for rate of falls and risk of falling; plus positive findings in an additional four trials reporting rate of falls that could not be pooled) indicated potential publication bias for this comparison.

In the 10 trials of exercise compared with usual care that were pooled reporting rate of falls, there was considerable heterogeneity in the results, which was only partially explained by a subgroup analysis grouping trials according to level of nursing care provided. We are uncertain whether exercise had an effect on the rate of falls in care facilities as the quality of the evidence has been assessed as very low. Subgroup analyses by type of exercise did not explain the heterogeneity.

There was less statistical heterogeneity in the data on risk of falling for trials of exercise compared with usual care. Pooled data indicated exercise may make little or no difference to the risk of falling (low-quality evidence).

There was limited evidence for exercise types other than gait, balance and functional training or trials testing a combination of exercise categories in comparison with usual care. Whilst three trials tested Tai Chi programmes (which have been demonstrated to be effective at reducing the risk of falling in a community setting), data were not suitable for pooling.

We are uncertain of the impact of exercise on the risk of fracture or adverse events (very low-quality evidence).

Nine trials provided 12 comparisons of two different exercise programmes. Comparisons of different types of exercise were all considered of very low quality so we are uncertain of the relative effectiveness of different types of exercise.

While no clear effect on reduction in falls from exercise was identified within the current review, either overall or by subgroups according to level of care or type of exercise, there was a high degree of heterogeneity between the studies. The range of different types of exercise, populations and settings investigated plus the small size of many trials has resulted in only limited evidence being available for any particular combination of these factors. Importantly, the limited evidence does not represent convincing evidence of a lack of effect and the possibility of some types, intensity or duration of exercise being effective for specific populations remains.

Medication (drug target)

Medication review

Twelve studies examined medication review in care facilities. One study reported on the risk of fracture. Two studies reported instances of adverse events.

A summary of the evidence for general medication review in care facilities is provided in Summary of findings 2. Pooled results from five trials of general medication review indicated that this intervention may make little or no difference to the rate of falls or risk of falling (low-quality evidence). We are uncertain of the effect of general medication review on risk of fracture or adverse events as the quality of the evidence has been assessed as very low.

Vitamin D supplementation

Eight trials examined vitamin D interventions in care facilities. Five trials examined the effect of vitamin D supplementation, two trials investigated the effect of daily multivitamin supplementation which included vitamin D and calcium and one tested an education intervention aimed at increasing prescription of adequate levels of vitamin D, calcium and osteoporosis medications. Only three trials reported data on the risk of fracture and five on adverse events.

A summary of the evidence for vitamin D supplementation in care facilities is provided in Summary of findings 3. Vitamin D supplementation probably reduces the rate of falls (moderate-quality evidence) but vitamin D supplementation (with or without calcium) probably makes little or no difference to the risk of falling (moderate-quality evidence). The 28% reduction in falls rate observed (RaR 0.72, 95% CI 0.55 to 0.95) is substantial. Average serum vitamin D levels at baseline were reported to be low or very low in seven of eight studies (including the five studies of vitamin D with or without calcium supplementation), indicating that these results are applicable to residents of care facilities with low vitamin D levels. Based on other studies, the reduction in the rate of falls

may be related to improvement in muscle function (De Spiegeleer 2018).

We are uncertain of the effect of vitamin D supplementation (up to 1000 IU daily) on the risk of fall-related fractures or adverse events as the quality of the evidence has been assessed as very low. These studies represent only a subset of the studies evaluating the effect of vitamin D on fractures.

We are uncertain whether multivitamin supplementation including vitamin D and calcium reduces the rate or risk of falling based on two studies as the quality of the evidence is very low.

One study of an education intervention aimed at increasing the prescription of vitamin D, calcium and osteoporosis medication may make little or no difference to the rate of falls or risk of falling (low-quality evidence).

Environment/assistive technology

There were no large trials of this type in care facilities. We are uncertain of the effect on rate of falls of wireless position monitoring in care facilities (very low-quality evidence).

Social environment

Seven trials in care facilities targeted staff training or implemented service model changes. Two studies reported data on the risk of fracture and no studies reported adverse-event data. None of the interventions showed strong evidence for a reduction in falls. These interventions included staff education on fall and fracture prevention, a project nurse facilitating best-practice falls injury prevention strategies, guideline implementation (falls, urinary tract infection, and pressure ulcers), dementia care mapping, a risk-assessment tool versus nurses' judgement and a programme to improve staff connections, communication, and problem-solving. Results were inconsistent in two trials of dementia care mapping. Use of a falls risk-assessment tool in comparison with nurses' judgement alone probably makes little or no difference to the rate of falls or risk of falling (moderate-quality evidence). We are uncertain of the effect on falls of a half-day education programme about fall and fracture prevention for staff (very low-quality evidence). We are uncertain of the impact of the other social environment interventions on falls.

Knowledge/education

There were no trials of knowledge interventions in care facilities.

Psychological interventions

Two studies in care facilities evaluated the effect of psychological interventions on falls. Neither trial reported data on the risk of fracture and adverse-event data were not reported.

One trial examined a cognitive-behavioural intervention with a focus on falls-risk reduction, the other examined a computer-based cognitive training programme focused on improving attention combined with strength and balance training, compared with strength and balance training alone. We are uncertain of the effects of psychological interventions on rate of falls or risk of falling as the quality of the evidence is very low.

Other single interventions

Three trials (564 participants) examined other single interventions. We are uncertain whether lavender olfactory stimulation, multisensory stimulation in a Snoezelen room or sunlight exposure reduces falls as the quality of the evidence has been assessed as very low.

Multiple interventions

An intervention for incontinent residents in high-level nursing care facilities that included exercise, offering regular fluids and toileting, showed no strong evidence for an effect and we are uncertain of the effectiveness as the quality of the evidence is very low (Schnelle 2003).

Increased sunlight exposure plus calcium supplementation had low adherence to sunlight exposure; we are uncertain of the effects on falls or adverse events as the quality of the evidence is very low (Sambrook 2012). There was no difference in the incidence rates of new skin cancers, but an increase in the adjusted all-cause mortality in the calcium-treated group compared with the UV alone group (hazard ratio (HR) 1.23 versus 0.76, P = 0.03). Despite documented concerns about increased risk of cardiovascular events, in particular myocardial infarction, with calcium supplementation (Bolland 2010), there was a lack of evidence for a strong effect on increased death rates from myocardial infarction, so the biological reason for the observed increase in all-cause mortality is uncertain. We are uncertain of the effects on adverse events as the quality of evidence is very low.

Multifactorial interventions

In multifactorial interventions, two or more categories of intervention are given, and these are linked to each individual's risk profile. An initial assessment is usually carried out by one or more health professionals and an intervention is then provided or recommendations given or referrals made for further action. All trials compared a multifactorial intervention with 'usual care', which in many cases included some falls-prevention activities. These standard care practices may have changed over time; however, the degree to which the comparator arm does or does not include components of the intervention activities is not clear enough to base any additional analysis on. A summary of the evidence for multifactorial interventions in comparison with usual care in care facilities is provided in Summary of findings 4.

This review included 13 multifactorial trials in care facilities. Five studies reported data on risk of fractures. One study reported an

instance of a fall as an adverse event, two studies reported that there were no adverse events, and the remaining studies did not report on adverse events. The interpretation of pooled data from multifactorial interventions is problematic because of variation in components between trials, and variation of combinations of components delivered to individuals in the trials.

Pooled results did not show strong evidence for a reduction in the risk of falling or risk of fracture; however, there was considerable statistical heterogeneity. Multifactorial interventions may make little or no difference to the risk of falling in care facilities (low-quality evidence). We are uncertain of the effects of multifactorial interventions in care facilities on the rate of falls or risk of fractures as the quality of evidence has been assessed as very low. A post-hoc subgroup analysis based on high, intermediate or mixed levels of nursing care showed a statistical difference between subgroups, with a reduction in falls in high- and intermediate-level care facilities, but not in studies or facilities with a mixed level of care. As there is no clear external evidence that could explain these subgroup results, and the finding is not completely consistent across studies, the finding is not considered credible (Guyatt 2011a), and no conclusion based on these subgroups is made. Subgroup analysis by level of cognition did not explain the heterogeneity.

Hospitals

Exercise

Three trials in hospitals (244 participants) investigated exercise as a single intervention. Two of these were small, including less than 60 participants. Only one trial reported on adverse events.

The three trials tested the effect of additional physiotherapy in rehabilitation wards (Summary of findings 5); however, we are uncertain of the effect of this intervention on rate of falls or whether it reduces risk of falling as the quality of the evidence has been assessed as very low. There were no data available on fractures and the one study reporting on adverse events found none.

Medication (drug target)

Medication review

In hospitals, we are uncertain of the effects of medication review on either rate of falls or risk of falling; this was tested in only one trial (very low-quality evidence).

Vitamin D supplementation

One trial in an acute geriatric unit found no strong evidence of an effect of vitamin D supplementation on risk of falling, despite the low levels of vitamin D at baseline. The median length of stay was only 30 days. We are uncertain of the effects of vitamin D in hospitals on rate of falls or risk of falling, rate of fracture or adverse events as the quality of the evidence has been assessed as very low.

Environment/assistive technology

Six trials in hospitals investigated environment/assistive technology interventions.

Pooled data from two trials (28,649 participants) were available on the use of bed alarms in hospitals (Summary of findings 6). The larger trial, which was a cluster-randomised trial with 28,551 participants, of bed/chair alarms was an education intervention to support judgement on their use. We are uncertain of the effects of bed alarms on the rate of falls, risk of falling or adverse events as the quality of the evidence has been assessed as very low.

We are uncertain whether carpet flooring, tested in one small trial, increases the rate of falls and risk of falling compared with vinyl flooring (very low-quality evidence). We are uncertain of the effects on rate of falls or risk of falling of using identification bracelets for patients at high risk.

A large trial of the use of one low-low bed per 12 existing beds in hospitals had no effect on rate of falls. However, large confidence intervals indicate a lack of precision in the estimate and we are uncertain of the effect of providing low-low beds on the rate of falls (very low-quality evidence).

Social environment

Six trials in hospitals targeted staff training or implemented service model changes. One trial in a hospital setting reported data on the risk of fracture. No studies reported adverse-event data. Trials tested a comprehensive post-operative ortho-geriatric service in a geriatric ward for patients with proximal femoral fracture surgery compared with usual care in an orthopaedic ward, guideline implementation, fall-prevention toolkit software, a new acute care service for elderly patients, and a new behavioural advisory service for people with confusion. We are uncertain of the effects of these interventions on falls as the quality of the evidence has been assessed as very low.

Knowledge/education

Two trials examined knowledge interventions in hospitals. Neither trial reported data on the risk of fracture and one study reported that there were no adverse events.

We are uncertain of the effects of an educational session based on identified risk factors and usual fall-prevention care in acute medical wards as the quality of the evidence was assessed as very low.

In a mixture of acute and subacute wards, a trial providing patients with educational materials alone and educational materials with professional follow-up did not show strong evidence for a reduction in the rate of falls (Haines 2011). Providing patients with educational materials alone may make little or no difference to the rate of falls or risk of falling (low-quality evidence).

In a post-hoc subgroup analysis, educational materials with professional follow-up showed a reduction in falls in participants with no cognitive impairment in comparison with usual care. There is moderate credibility for this post-hoc subgroup analysis (Guyatt 2011a); however, we are uncertain of the effectiveness of this intervention in reducing the rate of falls as the quality of the evidence has been assessed as very low. Due to the contrast between the effectiveness of providing this intervention as a single intervention and its effectiveness when provided as a multifactorial intervention targeted at cognitively intact participants (Hill 2015; which further supports the credibility of the result found in the subgroup analysis within Haines 2011), no conclusion on the effectiveness of this intervention when delivered as a single intervention is made as this is likely to result in difficulty in interpretation.

Psychological interventions

There were no trials of psychological interventions in hospitals.

Other single interventions

There were no trials of other single interventions in hospitals.

Multiple interventions

There were no trials of multiple interventions in hospitals.

Multifactorial interventions

In multifactorial interventions, two or more categories of intervention are given, and these are linked to each individual's risk profile. An initial assessment is usually carried out by one or more health professionals and an intervention is then provided or recommendations given or referrals made for further action. All trials included a comparison with 'usual care' that in many cases included some falls prevention activities. These standard care practices may have changed over time; however, the degree to which the comparator arm does or does not include components of the intervention activities was not clear enough to explore this.

This review included six multifactorial trials in hospitals. Five trials provided data suitable for pooling for the rate of falls, three for the risk of falling. Two studies reported data on risk of fractures. Four studies reported adverse-event data, there were no adverse events. The evidence for multifactorial interventions in hospitals is summarised in Summary of findings 7. Pooled results showed a borderline reduction in the rate of falls with a point estimate of a reduction of 20%; the 95% confidence intervals indicated this estimate of effect may range as high as a reduction of 36% or result in an increase in falls rates of 1% (Analysis 21.1: RaR random effects 0.80, 95% CI 0.64 to 1.01; 5 trials: I^2 = 52%); however, there was moderate heterogeneity. The interpretation of pooled data from multifactorial interventions is problematic because of variation in components between trials, and variation of combinations of components delivered to individuals in the trials. A subgroup analysis based on the setting demonstrated a likely significant difference between subgroups. Pooled data from two trials in a subacute setting showed that multifactorial interventions, both included targeted patient education, may reduce the rate of falls (RaR 0.67, 95%CI 0.54 to 0.83; low-quality evidence).

Pooled results on the risk of falling included only three of the five trials that were pooled for the rates of falls, but the overall effect estimate was generally consistent with the rate of falls, giving a point estimate of a 18% reduction in the risk of falling, with wider 95% confidence intervals indicating this may range between a 38% reduction and a 9% increase (Analysis 21.2: RR random-effects 0.82, 95% CI 0.62 to 1.09; 3 trials: I^2 = 0%). This did not achieve statistical significance, but one of the additional trials that was not pooled also reported a reduction in the risk of falling based on admissions in a subacute setting (Hill 2015; 3121 participants). No difference between subgroups by setting was observed. We are uncertain of the effects on risk of falling as the quality of the evidence was assessed as very low.

We are uncertain of the effect of multifactorial interventions on the risk of fracture or adverse events as the quality of the evidence has been assessed as very low.

Subgroup analyses by level of care partly explained the heterogeneity, but due to variations in study design there is some uncertainty if findings are due to the setting or other factors, including the specific combination of interventions provided. Multifactorial interventions that include targeted patient education may reduce the rate of falls in a subacute setting (low-quality evidence).

A cost-effectiveness analysis from one trial of multifactorial interventions is to be published (Hill 2014 protocol for Hill 2015).

Studies in participants with cognitive impairment

There is limited evidence for interventions to reduce falls in people with cognitive impairment where these people are a clearly defined group. Although only 11 trials reported findings specifically for patients with dementia or cognitive impairment, many participants in care facilities trials, including those testing interventions that probably or may reduce falls (e.g. vitamin D supplementation), had cognitive impairment.

Economic evaluations

A cost-effectiveness analysis of a patient education programme reduced falls in a subgroup of hospital patients who were cognitively intact (Haines 2011). In this subgroup the intervention, which consisted of written and video-based materials plus one-on-one bedside follow-up from a trained health professional, cost AUD 294 to prevent one fall and AUD 526 to prevent one person falling (2008 dollars; reported in Haines 2013).

No conclusions can be drawn from the other 10 trials reporting economic outcomes.

Overall completeness and applicability of evidence

Although we have included 95 trials in this review, these have tested a very wide variety of interventions, sometimes with different comparators rather than control or usual care, in various types of facility. Approximately three quarters of included trials were conducted in care facilities, however many of these were small.

In this review, we have reported results from care facilities and hospitals separately to improve applicability of the interventions to each setting. Careful consideration of the context of effective interventions is required. As Becker 2010 points out, the type of care provided in care facilities differs between countries and healthcare systems. Also, consideration needs to be taken of cultural and organisational contexts when generalising the results from this review. Unfortunately, the level of care and case mix in each facility in this review was often not clearly defined. In addition there is striking variability in type, targeting, intensity and duration of the falls prevention programmes that were studied. Reports of trials in hospitals are also unlikely to adequately describe the complex interaction that is likely to occur between the intervention and the usual falls-prevention practices occurring within hospitals.

Twenty-five trials of exercise in care facilities were included, 17 of which tested exercise with usual care. However, many of these were small and whilst there were a number of trials examining balance, gait or functional training exercise programmes, there were few trials on flexibility, strength/resistance training and 3D exercise (including Tai Chi). There were several comparisons of different exercise programs; however, there was generally only one small trial for each comparison so the data were too few to be informative.

The quality of available evidence for vitamin D supplementation was reasonable (moderate-quality evidence). However, there were few studies of vitamin D supplementation taken in the form of a multivitamin. Trials of environmental/assistive technologies and social environment (e.g. staff training, service model changes) generally studied clinically different interventions, precluding pooling of trial results. Whilst there was a very large trial of bed alarms conducted in hospitals, this trial was of education, training and support for their use and there were no trials of bed alarms in care facilities. Medication review is generally aimed at reducing psychoactive medications. There were a number of trials of medication review in care facilities considered clinically similar enough to justify pooling. However, there was a large degree of inconsistency in the trial findings.

The interpretation of the multifactorial interventions is complex because of the variation in components, duration and intensity of the intervention, and how the interventions were implemented. The study design does not allow evaluation of individual components of the interventions in either care facilities or hospitals.

Only one trial specifically assessed the benefit of using a validated falls risk-assessment tool in comparison with clinical judgement in a care facility (Meyer 2009) and none did in hospital, although this approach is widely used in both settings. Some multifactorial trials (e.g. Barker 2016) used validated falls risk-assessment tools to determine the application of appropriate interventions, but the effects of the falls risk-assessment tool cannot be separated from that of the interventions. This lack of evidence calls into question the wide use of these tools internationally and further trials examining the effectiveness of the tools are warranted.

Few trials have incorporated interventions relating to the circumstances of falls, e.g. assistance with toileting, rather than targeting individual risk factors, as in the continuous quality improvement model used to develop a falls-prevention programme in Lohse 2012.

The comparator in many trials is 'usual care'. Frequently, what falls-prevention activities are included as a component of usual care is not clearly reported. This hinders interpretation of how 'usual' care may change over time and any potentially useful subgroup analyses based on this.

In terms of outcomes, 30 of the included trials did not report usable data for calculating rate of falls and 36 trials for risk of falling (*see* Appendix 7). Many studies reporting data suitable for pooling reported data for one but not both of these outcomes. This may explain some of the inconsistency between the findings. Even fewer studies reported the impact of the interventions on fractures or adverse events. Within those studies that did report on adverse events, it was often unclear if these data were recorded systematically. Studies that reported data on fractures reported outcomes for different types of fractures (e.g. hip fractures only versus total fractures). Other studies not eligible for inclusion in this review may provide additional evidence for the impact of the interventions on fractures. In particular, whilst a larger proportion of included studies reported data on the risk of fracture following vitamin D supplementation, it is important to consider that these trials represent only a subset of the studies evaluating the effect of vitamin D on fractures available. In addition, some trials of interventions that may increase falls during the intervention period (exercise, medication review) only reported falls during the post-intervention period. Other studies report only a subset of falls (e.g. bedside falls in Sahota 2014), and therefore do not meet the inclusion criteria for this review. Many cluster-randomised trials did not adjust for clustering, therefore this was performed post-hoc by the review authors (as indicated by a "c" in Appendix 7, for details see Unit of analysis issues).

Vitamin D supplementation in care facilities reduced the rate of falls but not the risk of falling. This discrepancy might be explained by differential effects on multiple fallers (i.e. those falling more than once over the study period). However, too few of these trials reported data on multiple fallers to enable meaningful analysis of this outcome.

Only Haines 2011 included a cost-effectiveness evaluation of their hospital patient education programme in terms of falls prevented

to inform the value for money for the intervention tested. An economic evaluation of the intervention tested in Hill 2015 is still to be published.

Many of the interventions studied would be difficult to sustain in usual clinical practice due to competing factors in the clinical environment. In aged-care settings, vitamin D supplementation is relatively cheap, and once it commences as part of a person's regular medication regimen it can be continued indefinitely. In hospital settings, educating staff and patients regarding falls prevention would be regarded as good clinical practice and is sustainable in the long term provided the necessary resources are available.

There is scope for realigning clinical practice with less emphasis on use of scales to assess falls risk (because there is no convincing research evidence of their effectiveness) and encouraging clinical staff to focus on factors that may be more effective, for example educating patients and families about falls and how to avoid them.

Quality of the evidence

This review containing 95 trials (138,164 participants) does not provide robust evidence regarding effective interventions for reducing falls in the settings considered. We assessed the quality of the evidence using the GRADE approach which considers the risk of bias, inconsistency, indirectness, imprecision and other biases (including publication bias) for the evidence for each outcome of the main comparisons. The GRADE assessments are reported in Summary of findings for the main comparison to Summary of findings 7 and the findings are cross-referenced in the relevant results sections. The GRADE quality of evidence for many outcomes was low or very low. This largely reflects the risk of bias in the individual studies and also the significant heterogeneity and imprecision in many of the pooled study estimates.

Despite the addition of 35 trials in this update, this has generally not improved the robustness of the results compared with the previous version of this review (Cameron 2012). Although there are now a number of trials conducted for some interventions types (e.g. exercise, medication review and vitamin D supplementation in care facilities and multifactorial interventions in hospitals), the overall quality of the evidence was low to very low for all outcomes and comparisons except for rate and risk of falling for vitamin D supplementation, and use of a falls risk-assessment tool, all in care facilities. There was also evidence indicating potential publication bias in trials of exercise conducted in care facilities.

Studies in this review varied widely in their risk of bias (*see* Table 4). The majority of included studies all contained some risk of bias. The included studies illustrated the wider problems of variation in the methods of ascertaining, recording, analysing, and reporting falls described in Hauer 2006. Many trials have used a single approach for ascertaining the number of falls, the limitations of this have been demonstrated in a study of falls data derived from a large hospital based randomised controlled trial (Hill 2010). For some aspects of study design, minimisation of bias is difficult. For example, it is not possible to blind participants and treatment providers for exercise, bed alarms and other types of interventions. Falls were generally recorded by nursing or care home staff who were frequently not blinded to the intervention. In addition, not all studies met the contemporary standards of the extended CONSORT statement (Schulz 2010), including the extensions for cluster-randomised trials (Campbell 2004), non-pharmacological trials (Boutron 2008), and pragmatic randomised trials (Zwarenstein 2008), so reporting was unclear in many instances, particularly for allocation concealment or selective outcome reporting when no protocol could be identified.

There is a potential for differences between individually- and cluster-randomised trials. This review included a large proportion of cluster-randomised trials (44%). Within this review, in general trials were more likely to be cluster randomised or not depending on the intervention being investigated and the setting. Thus, whilst five of six trials of multifactorial interventions in hospitals (enrolling 99% of participants), and 85% of those conducted in care facilities (82% of participants) were cluster randomised, in contrast for trials of exercise in care facilities, 88% of trials with 65% of participants were individually randomised. Similarly, for trials of vitamin D supplementation in care facilities, 75% of trials (with 60% of participants) were individually randomised. Although it has been reported that contamination, or 'herd effects' in individually-randomised trials conducted in facilities may result in decreasing the estimate of effect (Hahn 2005), this is considered unlikely to have had a major impact on the estimates of effect or conclusions for this review. The reasons for this according to the major categories of intervention are described below.

For trials of exercise in care facilities, the estimates of effect of the three cluster-randomised trials that contributed to pooling (Kerse 2008; Rosendahl 2008; Yokoi 2015), did not appear to differ to the range of estimates for the individually-randomised trials. For vitamin D in care facilities, as the single cluster-randomised trial contributing to the pooled result (Law 2006) had a smaller estimate of effect compared to the individually-randomised trials, this indicates that contamination of the control group was unlikely to have played a role in the estimate of effect, which increases the confidence in the effect estimate. For medication review in care facilities, there was a more even balance of individually- and cluster-randomised trials; 58% of trials (62% of participants) were individually randomised. The estimates of effect from the trials were inconsistent within both the cluster- and individually-randomised trials, thus the high inconsistency of findings between trials for this intervention cannot be explained by the type of randomisation used. Two cluster-randomised trials contributed only 18% of the participants for the evidence for multifactorial interventions in care facilities, the estimates of effect in these two trials were similar to that for the pooled overall effect estimates. All trials of additional exercise in care facilities were individually randomised. In trials of bed exit alarms in hospitals, only two trials contributed to pooled data; 96% of participants were enrolled in one trial that

was cluster randomised, thus consideration of the findings of trials that were individually in comparison with cluster randomised is uninformative. Similarly, comparisons of individually- and cluster-randomised trials within multifactorial interventions in hospitals are not feasible given 99% of participants were enrolled in cluster-randomised trials.

There was significant unexplained heterogeneity in the findings for the rate of falls for several comparisons (exercise, medication review and multifactorial in care facilities), which limited the confidence in the results (*see* Summary of findings for the main comparison, Summary of findings 2 and Summary of findings 4), and was reflected in the generally low quality of evidence. The heterogeneity may be due to variations in intervention components, duration, intensity and settings as well as variations in the populations.

The evidence for some ProFaNE categories of interventions contained a degree of indirectness, where the intervention was a recommendation for, or education on, use of the intervention, rather than implementing the intervention for all participants (e.g. Kennedy 2015 for vitamin D, Shorr 2012 for bed alarms). In addition, where evidence was from a single trial or setting, it was likely to be considered to have a degree of limited applicability, or indirectness to other settings, (e.g. Sambrook 2012 which examined sunlight exposure in Australia).

There was also imprecision in some estimates, where the number and size of trials was small (*see* Summary of findings 5) or in particular for the risk of fracture where few trials reported this outcome and events were infrequent (e.g. vitamin D Summary of findings 3).

There was some evidence for likely publication bias for trials in exercise, where the included studies appeared to include a disproportionate number of small studies with positive findings (*see* Figure 4, Figure 5) .

Potential biases in the review process

We attempted to minimise publication bias in the review by searching multiple databases, and drew on the handsearch results published in the Cochrane Library in the Cochrane Central Register of Controlled Trials (CENTRAL). We also contacted authors of studies identified in trials registers that were completed, but for which full reports had not been identified, studies where only conference abstracts were identified, and many studies where it was unclear whether or not they met the inclusion criteria. We placed no foreign language restrictions in our search strategy; two studies were published in languages other than English (Peyro Saint Paul 2013; Salvà 2016), correspondence with authors provided information on study methods and results. However, despite these efforts, evidence of likely publication bias in trials of exercise conducted in care facilities remained.

Although the majority of screening of search citations for potentially eligible studies in this update was performed by only one author, we suggest this was not a source of bias given that the screening was over-inclusive with the onus being given to obtaining full-text reports for all potentially eligible studies. We observe also that where screening was undertaken by two review authors, the progression to full-text review was reduced.

Five newly published studies that were identified in the top-up search in August 2017 await classification (Dever 2016; Hewitt 2014; Raymond 2017; Van der Linden 2017; Wylie 2017). This was a pragmatic decision taken in view of the delay that would have resulted from their likely inclusion and after consideration of the potential impact of these trials on review findings. We concluded that our decision to postpone the inclusion of these five trials was not an important source of bias.

Whilst we strictly applied *a priori* inclusion and exclusion criteria to the selection of studies for this review, which should minimise bias, this does result in the inclusion of a subset of the available evidence and this applies in particular to risk of fracture outcome. All included studies were required to present data on the overall rate of falls or risk of falling, those reporting only a subset of falls (e.g. injurious falls, bedside falls) were excluded. We also excluded 22 trials reporting falls as adverse effects, although in some instances the intervention might plausibly have reduced falls. For a more comprehensive systematic review of the effect of vitamin D supplementation on fractures, see Avenell 2014.

For single-trial comparisons, we took a different approach to GRADE assessment where a single rater checked whether the trial findings for each outcome met pre-specified criteria for downgrading the evidence. The criteria were established before this alternative assessment took place. For 26 single-trial comparisons these criteria were met. For 18 comparisons in 16 trials these criteria did not apply, generally because of a large trial size, and GRADE assessment was conducted in duplicate. For these assessments, in two trials (three outcomes), the quality of the evidence was considered moderate (Chapuy 2002; Meyer 2009), and in three trials (five outcomes) the quality of the evidence was considered low (Cox 2008; Haines 2011; Kennedy 2015); for all other comparisons and outcomes the quality of the evidence was considered very low.

There are potential biases within the data included in the review in terms of non-normal distribution of falls rates in the included studies (as seen in Potter 2016), missing data including the loss of clusters within some trials, selective outcome reporting (*see* Table 4), decisions regarding pooling of studies where there is high heterogeneity and selection of models used for meta-analyses where there is heterogeneity for one falls outcome, but not another (e.g. high heterogeneity for rate of falls but not risk of falling). The potential biases due to these factors are captured by the GRADE assessments of the overall quality of evidence (Summary of findings for the main comparison to Summary of findings 7). There are also potential biases in decisions to conduct post-hoc subgroup and sensitivity analyses (e.g. Analysis 5.4; *see* Subgroup analysis and investigation of heterogeneity and Sensitivity analysis). This has been taken into account in conducting GRADE assessments (e.g.

confidence in the credibility of subgroup analysis is considered in the inconsistency rating for the subgroup analysis by setting for multifactorial interventions in hospitals), making cautious interpretations of the findings (e.g. considering findings based on subgroup analysis by setting for multifactorial interventions in care facilities of low credibility) and transparently reporting these analyses under Differences between protocol and review.

We explored the possibility of publication bias by constructing funnel plots of trials of exercise in care facilities and multifactorial interventions in care facilities (Figure 4, Figure 5, Figure 6). There was some asymmetry in the falls outcomes for trials of exercise in care facilities indicating potential publication bias.

Using the generic inverse variance method in this review enabled us to pool results as reported by trial authors with our own calculated from raw data, and results adjusted for clustering.

The ProFaNE falls prevention taxonomy enabled us to pool similar interventions in the analyses using a systematic approach. However, classification of some interventions according to this taxonomy was unclear and required judgement in some cases. We consulted with the ProFaNE authors when necessary.

Agreements and disagreements with other studies or reviews

We searched for other systematic reviews of falls prevention initiatives in care facilities and hospitals published since 2012 within our search described in Appendix 1. We compared our review results with the Cochrane Review '*Interventions for preventing falls in older people living in the community*' (Gillespie 2012), and identified six other systematic reviews incorporating meta-analyses (Chan 2015; Le Blanc 2015; Sherrington 2017; Silva 2013; Stubbs 2015; Vlaeyen 2015).

Comparison with trials in community-living older people

In contrast to the findings in this review for residents of care facilities and hospital inpatients, the evidence is clear that falls can be prevented using exercise in older people living in the community (Gillespie 2012). The effectiveness of group-based and home-based exercise programmes and Tai Chi in particular is well established in the community setting. There is the potential for falls to be reduced in care facilities using the same multiple-component exercise programmes, but despite 25 trials in this review testing exercise programmes in care facilities, the results were inconsistent. Only three trials examined exercises in hospitals; the quality of the evidence was considered very low.

Vitamin D supplementation may reduce falls in community-living people with lower vitamin D levels (Gillespie 2012). This is consistent with the finding in this review that vitamin D is effective in reducing falls in care facilities as most residents have low vitamin D levels (Pilz 2012).

The effects of multifactorial approaches are inconsistent between trials and settings. In the community setting, multifactorial interventions, including falls-risk assessment, reduced the rate of falls but not the risk of falling (Gillespie 2012). Similarly, multifactorial interventions overall may make little or no difference to the risk of falling in care facilities. However, findings on the rate of falls were inconsistent. In hospitals, multifactorial interventions (that include targeted patient education) may reduce the rate of falls in a subacute hospital setting.

There is some evidence that falls prevention strategies in the community can be cost saving (Gillespie 2012), but there were no economic evaluations conducted within the care facilities and only one in hospital trials (Haines 2011) to provide information on value for money for effective interventions.

Supplementary review

Nyman 2011 conducted a supplementary review of the 41 trials included in Cameron 2010 with specific reference to people's recruitment, retention in the trial, and adherence to intervention components. Adherence was high for individually-targeted and group-based exercise (72% to 89%) and for medication interventions (68% to 88%). The authors reported that adherence was related to treatment effectiveness in three studies testing medication and multifactorial interventions in care facilities. They estimated that by 12 months, on average, only a third of care-facility residents are likely to be adhering to falls prevention interventions. The current review was not able to comment on adherence or retention. Nyman 2011 provides an important perspective giving context to interpretation of the research.

Exercise

Chan 2015 conducted a systematic review of exercise interventions for older adults with cognitive impairment, only three of seven trials in a pooled analysis enrolled participants living in a care setting. Two of these studies were included in this review (Toulotte 2003 and Rosendahl 2008), but Chan 2015 included unpublished subgroup data for Rosendahl 2008, and Rolland 2007 and was excluded from this review as falls were monitored as adverse events. Sherrington 2017 conducted a systematic review and meta-analysis of exercise interventions to prevent falls in older adults. This review included 14 RCTs (15 comparisons) of exercise interventions in care settings and found no significant effect on the rate of falls. These authors observed possible asymmetry in the funnel plot, which was not statistically significant on Egger's test. Three of the trials included in Sherrington 2017 were excluded from this review (DeSure 2013; Resnick 2002; Rolland 2007; *see* Characteristics of excluded studies). Two of the trials included in the pooled estimate in Sherrington 2017 were considered as multiple interventions under the ProFaNE classification system in this review (Huang 2016, ; *see* Appendix 3). Data reported for one study were considered not suitable for pooling in this review (Toulotte 2003). All other trials were included.

Silva 2013 included 12 studies of exercise in care facilities. This review pooled studies of exercise as a single intervention with studies of exercise as a component of a multifactorial intervention. The authors found a significant reduction in the risk of falling (RR 0.71, 95% CI 0.64 to 0.92, I^2 = 72%). There was no significant effect on the risk of fracture (RR 0.57, 95% CI 0.21 to 1.57). All of the included trials were included in our review.

Lee 2017 included 21 studies of exercise in care facilities, 15 with exercise as a single intervention, six with exercise combined with one or more interventions. Data were pooled from studies comparing exercise with other interventions, usual care or placebo. In the current review, comparisons of alternate exercise programs were not pooled with trials of exercise in comparison with usual care (for details see Table 2). Three of the trials included in Lee 2017 were excluded from this review (DeSure 2013; Lord 2003b; Wolf 2003); two of these were considered to be conducted in a community setting. Data from one trial were not pooled in our review as there were zero falls in the intervention arm (Cadore 2014); this study has a weighting of 0.4% in the meta-analysis in Lee 2017. Pooled data of trials of exercise as a single intervention in Lee 2017 found no difference in the rate of falls or risk of falling, consistent with the findings of our review.

The current review found inconsistent effects for exercise in care facilities and is broadly consistent with Silva 2013 and Sherrington 2017 although pooling combinations differed. Our review contrasts with Chan 2015 as Chan 2015 pooled trials across both community and care facility settings and much of the impact observed in their meta-analysis may have been from trials conducted in the community.

Vitamin D supplementation

A systematic review conducted for the US Preventative Services Task Force (Le Blanc 2015), examining trials conducted in both institutionalised or community settings, found that vitamin D significantly reduced the number of falls per person but did not significantly reduce the risk of falling, consistent with the findings in care facilities in this review. The authors reported that sensitivity analysis based on institutionalised status "resulted in similar estimates". The two included studies conducted in institutionalised settings are included in this Cochrane Review. The authors concluded that "Treatment of vitamin D deficiency in asymptomatic persons might reduce mortality risk in institutionalised elderly persons and risk for falls but not fractures."

Bolland 2014 pooled outcomes from six randomised trials conducted in care facilities or hospitals and found no significant reduction in falls with vitamin D supplementation with or without calcium supplementation (RR 0.96, 95% CI 0.88 to 1.05). The authors concluded that supplementation with vitamin D does not reduce risk of falling by a 'clinically relevant' threshold of 15% or more and that future trials are unlikely to alter this conclusion. One study included as institutional in the Bolland 2014 review was excluded from this review as 51% of participants were residing in the community (Graafmans 1996); all other studies were included in this review. This Cochrane Review has analysed studies conducted in care facilities or hospitals separately and found that whilst vitamin D supplementation did not reduce the risk of falling, it did reduce the rate of falls in care facilities. Our analysis included data on the rate of falls in care facilities from the same four studies pooled for the risk of falling and whilst there was heterogeneity for the pooled rate of falls outcome (I^2 = 62%), it was lower than observed in Bolland 2014 when pooling studies in either setting (I^2 = 92%).

Other recent systematic reviews

Vlaeyen 2015 included 13 randomised controlled trials of fall-prevention programmes conducted in nursing homes. The authors found no significant effect of the interventions overall on the number of falls (10 studies) or risk of falling (six studies). They reported that multifactorial interventions significantly reduced the number of falls (four studies) and the number of recurrent fallers (four studies), but not the risk of falling (four studies). They reported that staff training and education had a significant harmful effect on the number of falls (two studies). All trials were included in our review.

Stubbs 2015 conducted an umbrella review of meta-analyses in care facilities and hospitals and concluded that there was consistent evidence that multifactorial interventions reduce falls in care facilities and hospitals and reported that there was consistent evidence that exercise and vitamin D reduces falls in care facilities, based on the inclusion of nine individual meta-analyses including Cameron 2012, Bolland 2014 and Sherrington 2011 (Sherrington 2017 is discussed above). Other meta-analyses included in Stubbs 2015 and published since 2012 were Choi 2012, Guo 2014 and Santesso 2014. Choi 2012 pooled three studies conducted in care settings, all of which were included in this review: a vitamin D trial (Broe 2007), a multifactorial trial (Neyens 2009), and Rapp 2008, which is included as a subgroup analysis of Becker 2003 in our review. Guo 2014 conducted an 'exploratory meta-analysis' examining fall-prevention interventions for those with or without cognitive impairment in institutionalised and non-institutionalised settings. Eight trials included in Guo 2014 were not considered for our review as they had been assessed as being conducted in the community setting: all eight trials were considered in Gillespie 2012, seven of which were included (Conroy 2010, Davison 2005, Haines 2009, Hendriks 2008, Latham 2003, Lightbody 2002, Lord 2005) and one of which was excluded because falls were reported as adverse events (Vogler 2009). Santesso 2014 conducted a meta-analysis of hip protectors; as we consider hip protectors are intended to reduce fractures rather than falls, this intervention is not included in our review.

AUTHORS' CONCLUSIONS

Implications for practice

We found evidence of effectiveness for some fall-prevention interventions in care facilities and hospitals, although for many the quality of the evidence was considered low or very low. For all interventions, we are uncertain of their effects on fractures and on adverse events as the quality of the evidence for both outcomes was assessed as very low. For each setting, the summary is structured by the main categories of interventions evaluated in at least one setting in the review: exercise, medication (medication review; vitamin D supplementation); psychological interventions, environment/assistive technology, social environment, interventions to increase knowledge, other interventions, multiple interventions and multifactorial interventions. There was a lack of evidence on surgery, management of urinary incontinence, or fluid or nutrition therapy in both settings.

Care facilities

- Exercise
  - We are uncertain of the effect of exercise on the rate of falls as the quality of the evidence was assessed as very low. Exercise may make little or no difference to the risk of falling (low-quality evidence; Summary of findings for the main comparison).
- Medication
  - General medication review may make little or no difference to the rate of falls or risk of falling (low-quality evidence); Summary of findings 2.
  - The prescription of vitamin D in care facilities probably reduces rate of falls (moderate-quality evidence), but prescription of vitamin D (with or without calcium) probably makes little or no difference to the risk of falling (moderate-quality evidence); Summary of findings 3.
  - An education intervention aimed at increasing the prescription of vitamin D, calcium and osteoporosis medication may make little or no difference to the rate of falls or risk of falling (low-quality evidence).
- Environment/assistive technology
  - There is a general lack of evidence on these interventions in care facilities.
  - We are uncertain of the effect on rate of falls of wireless position monitoring in care facilities (very low-quality evidence).
- Social environment
  - Use of a falls risk-assessment tool in comparison with nurses' judgement alone probably makes little or no difference to the rate of falls or risk of falling (moderate-quality evidence).
  - We are uncertain of the effects on falls of a half-day education programme about fall and fracture prevention for staff given by specialist osteoporosis nurses in care facilities (very low-quality evidence).
  - We are uncertain of the effects on falls of other interventions targeting staff and the organisation of care on falls, including guideline implementation and dementia care mapping (very low-quality evidence).
- Knowledge/education
  - There is a lack of evidence on these interventions in care facilities.
- Psychological interventions
  - We are uncertain of the effects on falls of a cognitive-behavioural intervention with a focus on falls risk reduction (very low-quality evidence).
  - We are uncertain of the effects on falls of a computer-based cognitive training programme focused on improving attention (very low-quality evidence).
- Other single interventions
  - We are uncertain whether lavender olfactory stimulation, multisensory stimulation in a Snoezelen room or sunlight exposure reduces falls (very low-quality evidence).
- Multiple interventions
  - We are uncertain about the effect on falls of a multiple intervention for incontinent residents that included exercise, offering regular fluids and toileting (very low-quality evidence).
  - We are uncertain about the effect on falls of a multiple intervention comprising increased sunlight exposure plus calcium supplementation (very low-quality evidence).
- Multifactorial
  - We are uncertain of the effects of multifactorial interventions on the rate of falls (very low-quality evidence). Multifactorial interventions may make little or no difference to the risk of falling (low-quality evidence); Summary of findings 4.

Hospitals

- Exercise.
  - We are uncertain whether providing additional physiotherapy in subacute wards has an effect on the rate of falls or whether it reduces the risk of falling (very low-quality evidence); Summary of findings 5.
- Medication
  - We are uncertain of the effect of medication review on either rate of falls or risk of falling (very low-quality evidence).
  - We are uncertain of the effect of vitamin D supplementation on either rate of falls or risk of falling (very low-quality evidence).
- Environment/assistive technology
  - We are uncertain of the effect of bed sensor alarms on the rate of falls or risk of falling (very low-quality evidence); Summary of findings 6.
  - We are uncertain whether carpet flooring, tested in one small trial, increases the rate of falls and risk of falling compared with vinyl flooring (very low-quality evidence).

- We are uncertain of the effects on rate of falls or risk of falling of using identification bracelets for patients at high risk of falling (very low-quality evidence).
- We are uncertain of the effect of providing low-low beds on the rate of falls (very low-quality evidence).

- Social environment
  - We are uncertain of the effects of interventions targeting staff and the organisation of care (including guideline implementation) on rate of falls or risk of falling (very low-quality evidence).
- Knowledge or education
  - We are uncertain of the effects on falls of an educational session based on identified risk factors and usual fall-prevention care in acute medical wards (very low-quality evidence).
  - Providing patients with educational materials alone may make little or no difference to the rate of falls or risk of falling (low-quality evidence).

- Psychological interventions
  - There is a lack of evidence on these interventions in hospitals.
- Other single interventions
  - There is a lack of evidence on whether or not falls risk-assessment tools and associated interventions reduce falls.
- Multiple interventions
  - There is a lack of evidence on these interventions in hospitals.
- Multifactorial intervention
  - Multifactorial interventions may reduce the rate of falls, although subgroup analysis suggest this may apply mostly to a subacute setting (low-quality evidence). We are uncertain of the effects of multifactorial interventions on the risk of falling (very low-quality evidence); Summary of findings 7.

Implications for research

Further research, primarily randomised controlled trials, is warranted to help inform decisions in this key area. We suggest the following guide to help discussions on future priorities.

- Further research into supervised exercise programmes in both settings. There is a particular need for larger trials in care facilities and trials that clearly describe the care needs of the participants.

- Further research to strengthen the evidence for multifactorial interventions in both settings. Of note is that there are some substantial individual trials that have shown an important effect in reducing the rate of falls. A key feature of these multifactorial interventions is the individualised nature of the interventions delivered. This implies that further research with emphasis on an individualised, standardised approach to delivery of interventions with consistent description and application within further trials is warranted, including as a clear description of existing falls prevention practices in the control arm of any trials and the interaction of the intervention arm of the trial with usual care. A mixed methods approach may be necessary to achieve this.

- Further trials of patient-directed interventions, especially in care facilities; for example, with a psychological and educational focus.

- Trials with interventions incorporating approaches based on the circumstances of falls in addition to individual risk factors, e.g. regular assisted toileting in both care facilities and hospitals (Lohse 2012; Schnelle 2003).

- Further trials testing the routine use of validated falls risk-assessment tools.

- Further research is required testing interventions targeting staff, and changes to the organisational system in which an intervention is delivered or the introduction of new healthcare models.

- In care facilities, additional trials on medication review, vitamin D plus calcium supplementation, environmental/assistive technologies and social environment interventions are required. There should be an emphasis on large trials.

- In hospitals, more trials of additional exercise, social environment and knowledge interventions are needed.

- Further research focusing on participants with dementia.

Other aspects, including research methods, that need to be adopted in all future studies are as follows.

- Classification of the components of the fall-prevention intervention using the taxonomy developed by the Prevention of Falls Network Europe (ProFaNE) (Lamb 2007; Lamb 2011). This will produce consistency between trials allowing for more effective pooling of data.

- Consideration is needed of the nature of 'usual care' and its potential interaction with the intervention group.

- For multifactorial trials, clear descriptions are needed of the components and the proportion of the participants receiving the different interventions.

- Falls data should be collated by a researcher blind to group allocation.

- Fall events should be reported by group as total number of falls, fallers, and people sustaining a fall-related fracture or brain injury; rate of falls (falls per person year or per 1000 patient days); multiple fallers and number in each analysis.

- Results should be analysed using appropriate, pre-specified methodology (e.g. negative binomial regression, survival

analysis) (Robertson 2005). Group comparisons should be expressed as incidence rate ratios and risk ratios with 95% confidence intervals.

- Authors of trials not excluding people with cognitive impairment should plan to report the results by level of cognitive impairment to indicate whether degree of impairment is an effect modifier.

- Design and reporting of trials should meet the contemporary standards of the extended CONSORT statement including those relating to randomised sequence generation and allocation concealment prior to randomisation (Schulz 2010). Pragmatic trials and those testing non-pharmacological interventions should incorporate the requirements defined in Zwarenstein 2008 and Boutron 2008.

- Clear description of usual care in the control arms of trials and discussion of the interaction of the intervention with this is needed.

- Design and reporting of cluster randomised trials should follow contemporary guidance (Campbell 2004) including the reporting of intra-class correlation coefficients.

- Where factorial designs are employed, data for each treatment cell should be reported to allow interpretation of possible interactions between different intervention components (McAlister 2003).

- There is a clear need for further research clearly reporting on the cognitive status of the included participants and including those with cognitive impairment.

- Economic evaluations should be conducted alongside randomised controlled trials to establish the cost-effectiveness of each intervention being tested. This involves measuring health-related quality of life as an outcome, defining the perspective and timeframe for costs, collecting data on healthcare use, costing healthcare resources, calculating cost-effectiveness ratios (if the intervention is effective in reducing falls), and evaluating uncertainty. Guidelines for carrying out and reporting economic evaluations in falls prevention trials have been published (Davis 2011).

ACKNOWLEDGEMENTS

The authors would like to acknowledge the considerable contributions of Leslie Gillespie and Clare Robertson to earlier versions of the review, and Clare Robertson for support on data management and statistical calculations. The authors would like to thank Lindsey Elstub and Joanne Elliott for their support at the editorial base.

We thank the following for their useful and constructive comments on this version of the review: Dr Joanne Elliott, Dr Helen Handoll, Prof Finbarr Martin, and Prof Cameron Swift.

We thank the following for their useful and constructive comments on earlier versions of the protocol and/or review: Assoc Prof Jacqueline Close, Dr Simon Gates, Dr Helen Handoll, Prof Peter Herbison, Prof Finbarr Martin, Prof Cathie Sherrington, and Dr Janet Wale. We are grateful to Prof Sarah Lamb, Prof Clemens Becker and Dr Klaus Pfeiffer for their assistance with use of the ProFaNE taxonomy, and to Prof Peter Herbison for his advice on statistical issues in previous versions of the review. We are also grateful to Prof William Gillespie and Dr Mohit Arora for assistance in assessing the risk of bias for some previously included studies and Dr Mohit Arora for screening trial registry records in the top-up search. We thank Geraldine Wallbank of the George Institute for Global Health, Sydney for her assistance in completing Appendix 7 for previously included studies.

This project was supported by the National Institute for Health Research (NIHR) via Cochrane Infrastructure funding to the Cochrane Bone, Joint and Muscle Trauma Group. The views and opinions expressed therein are those of the review authors and do not necessarily reflect those of the Systematic Reviews Programme, the NIHR, the National Health Service (NHS) or the Department of Health. Funding for immediate open access was received from the NIHR Cochrane review gold open access scheme.

REFERENCES

References to studies included in this review

Aizen 2015 *{published data only}*
Aizen E, Lutsyk G, Wainer L, Carmeli S. Effectiveness of individualized fall prevention program in geriatric rehabilitation hospital setting: a cluster randomized trial. *Aging Clinical and Experimental Research* 2015;**27**(5):681–8.

Ang 2011 *{published data only}*
Ang E, Mordiffi SZ, Wong HB. Evaluating the use of a targeted multiple intervention strategy in reducing patient falls in an acute care hospital: a randomized controlled trial. *Journal of Advanced Nursing* 2011;**67**(9):1984–92.

Barker 2016 *{published and unpublished data}*
Ayton DR, Barker AL, Morello RT, Brand CA, Talevski J, Landgren FS, et al. Barriers and enablers to the implementation of the 6-PACK falls prevention program: A pre-implementation study in hospitals participating in a cluster randomised controlled trial. PLOS One [Electronic

Resource] 2017; Vol. 12, issue 2:e0171932.
Barker A. RE: Queries re: your 6-pack trial. [personal communication] Email to: S. Dyer 24 October 2016.
Barker A, Brand C, Haines T, Hill K, Brauer S, Jolley D, et al. The 6-PACK programme to decrease fall-related injuries in acute hospitals: protocol for a cluster randomised controlled trial. *Injury Prevention* 2011;**17**(4):e5. DOI: 10.1136/injuryprev-2011-040074
Barker AL, Morello RT, Ayton DR, Hill KD, Brand CA, Livingston PM, et al. Acceptability of the 6-PACK falls prevention program: A pre-implementation study in hospitals participating in a cluster randomized controlled trial. PLOS One 2017; Vol. 12, issue 2) (no pagination.
Barker AL, Morello RT, Ayton DR, Hill KD, Landgren FS, Brand CA. Development of an implementation plan for the 6-PACK falls prevention programme as part of a randomised controlled trial: protocol for a series of preimplementation studies. Injury Prevention 2016; Vol. 22, issue 6:446–52.
* Barker AL, Morello RT, Wolfe R, Brand CA, Haines TP, Hill KD, et al. 6-PACK programme to decrease fall injuries in acute hospitals: cluster randomised controlled trial. *BMJ* 2016;**352**:h6781. DOI: 10.1136/bmj.h6781
Morello R, Barker A, Zavarsek S, Watts JJ, Haines T, Hil K, et al. The 6-PACK programme to decrease falls and fall-related injuries in acute hospitals: protocol for an economic evaluation alongside a cluster randomised controlled trial. *Injury Prevention* 2012;**18**(2):e2.
Morello RT, Barker AL, Haines T, Zavarsek S, Watts JJ, Hill K, et al. In-hospital falls and fall-related injuries: A protocol for a cost of fall study. *Injury Prevention* 2013;**19**(5):363.

Beck 2016 ***{published and unpublished data}***
Beck AM. Re: NCT01873456: Trial of multidisciplinary nutrition in nursing home and home-care. Email to: S Dyer 4 July 2016.
* Beck AM, Christensen AG, Hansen BS, Damsbo-Svendsen S, Kreinfeldt Skovgaard Moller T. Multidisciplinary nutritional support for undernutrition in nursing home and home-care: A cluster randomized controlled trial. *Nutrition* 2016;**32**(2):199–205. DOI: 10.1016/j.nut.2015.08.009
Beck AM, Gogsig CA, Stenbaek HB, Damsbo-Svendsen S, Kreinfeldt SM, Boll HE, et al. Study protocol: Cost-effectiveness of multidisciplinary nutritional support for undernutrition in older adults in nursing home and home-care: Cluster randomized controlled trial. *Nutrition Journal* 2014;**13**(1):86.
Beck Docent AM, Christensen AG, Hansen BS, Damsbo-Svendsen S, Moller TK. Author response re. "Rehabilitation nutrition for undernourished participants in nursing home and home care: Cluster randomized controlled study". Nutrition 2016; Vol. 32, issue 4:504. PUBMED: 26803332]
Nishioka S, Wakabayashi H, Maeda K. Rehabilitation nutrition for undernourished participants in nursing home and home care: Cluster randomized controlled study. Nutrition (Burbank, Los Angeles County, Calif.) 2016; Vol. 32, issue 4:503. PUBMED: 26732831]

Becker 2003 ***{published data only}***
* Becker C, Kron M, Lindemann U, Sturm E, Eichner B, Walter-Jung B, et al. Effectiveness of a multifaceted intervention on falls in nursing home residents. *Journal of the American Geriatrics Society* 2003;**51**(3):306–13.
Becker C, Lindemann U, Nikolaus T. Multifactorial intervention on falls and fractures in nursing homes (abstract). *Age and Ageing* 2000;**29**(Suppl 2):18.
Becker C, Walter-Jung B, Nikolaus T. The other side of hip protectors [letter]. *Age and Ageing* 2000;**29**(2):186.
Becker C, Walter-Jung B, Scapan K, Kron M, Nikolaus T. Effectiveness of multi-factorial intervention for reducing falls with proximal femoral fractures in homes for the aged and nursing homes. Goals and study design of a population-based study [Effektivitat einer multifaktoriellen Intervention zur Reduktion von Sturzen mit proximalen Femurfrakturen in Alten– und Pflegeheimen. Ziele und Studiendesign einer populationsbasierten Untersuchung]. *Zeitschrift fur Gerontologie und Geriatrie* 1997;**30**(4):293–7.
Rapp K, Lamb SE, Buchele G, Lall R, Lindemann U, Becker C. Prevention of falls in nursing homes: subgroup analyses of a randomized fall prevention trial. *Journal of the American Geriatrics Society* 2008;**56**(6):1092–7.

Bischoff 2003 ***{published data only}***
Bischoff HA, Hannes BS, Dick W, Akos R, Knecht M, Salis C, et al. Effects of vitamin D supplementation on falls: a randomized controlled trial. *Journal of Bone and Mineral Research* 2003;**18**(2):343–51.

Broe 2007 ***{published data only}***
Broe KE, Chen TC, Weinberg J, Bischoff-Ferrari HA, Holick MF, Kiel DP. A higher dose of vitamin D reduces the risk of falls in nursing home residents: A randomized, multiple-dose study. *Journal of the American Geriatrics Society* 2007;**55**(2):234–9.

Buckinx 2014 ***{published and unpublished data}***
Beaudart C, Buckinx F, Demonceau M, Maquet D, Crielaard JM, Reginster JY, et al. Evaluation of the impact of a 6-month training by whole body vibration on the risk of falls among nursing home residents. *Osteoporosis International.* 2013;**24**(1 Suppl):S243. Conference Abstract P435]
Beaudart C, Buckinx F, Maquet D, Crielaard JM, Reginster JY, Bruyere O. What are the clinical characteristics of patients improving their gait and body balance with whole body vibration? Results of a 3-month randomized controlled trial. 9th Congress of the European Union Geriatric Medicine Society, EUGMS13; 2013; Venice, Italy. 2013.
Beaudart C, Maquet D, Mannarino M, Buckinx F, Demonceau M, Crielaard JM, et al. Effects of 3 months of short sessions of controlled whole body vibrations on the risk of falls among nursing home residents. *BMC Geriatrics* 2013;**13**:42.
Buckinx F. Re: Trial NCT 01759680: 6 month whole body vibration training in nursing home residents [personal communication]. Email to: S Dyer. 14 July 2016.
Buckinx F, Beaudart C, Demonceau M, Maquet D,

Crielaard JM, Reginster JY, et al. Impact of a 6-month training by whole body vibration on functional and motor abilities among nursing home residents observed over a 12-months period. 9th Congress of the European Union Geriatric Medicine Society, EUGMS13; 2013; Venice, Italy. 2013:S55.

* Buckinx F, Beaudart C, Maquet D, Demonceau M, Crielaard JM, Reginster JY, et al. Evaluation of the impact of 6-month training by whole body vibration on the risk of falls among nursing home residents, observed over a 12-month period: a single blind, randomized controlled trial. *Aging-Clinical & Experimental Research* 2014;**26**(4):369–76.

Buettner 2002 ***{published data only}***

Buettner LL. *Efficacy of prescribed therapeutic recreation protocols on falls and injuries in nursing home residents with dementia (Research monograph)*. Fort Myers (FL): Florida Gulf Coast University, 2001. ISBN–13: 978–1889435190]

* Buettner LL. Focus on caregiving. Falls prevention in dementia populations. *Provider* 2002;**28**(2):41–3.

Burleigh 2007 ***{published data only}***

* Burleigh E, McColl J, Potter J. Does vitamin D stop inpatients falling? A randomised controlled trial. *Age and Ageing* 2007;**36**(5):507–13.

Burleigh E, Potter J, McColl J. Does vitamin D stop hospital inpatients falling? - a randomised controlled trial [abstract]. *Age and Ageing* 2006;**35**(Suppl 3):i40.

Burleigh E, Potter J, McColl J. Does vitamin D stop hospital inpatients falling? A randomized controlled trial [abstract]. *Internal Medicine Journal* 2006;**36**(Suppl 5):A165.

ISRCTN18282824. Does vitamin D stop inpatients falling? - a randomised control trial. controlled-trials.com/ISRCTN18282824 (first received 25 August 2005).

Cadore 2014 ***{published and unpublished data}***

* Cadore EL, Casas-Herrero A, Zambom-Ferraresi F, Idoate F, Millor N, Gomez M, et al. Multicomponent exercises including muscle power training enhance muscle mass, power output, and functional outcomes in institutionalized frail nonagenarians. *Age* 2014;**36**(2):773–85.

Izquierdo M. Re: Your trial of multicomponent exercises in institutionalised frail elderly [personal communication]. Email to: S Dyer. 5 October 2016.

Chapuy 2002 ***{published data only}***

Chapuy MC, Pamphile R, Paris E, Kempf C, Schlichting M, Arnaud S, et al. Combined calcium and vitamin D3 supplementation in elderly women: Confirmation of reversal of secondary hyperparathyroidism and hip fracture risk: The Decalyos II study. *Osteoporosis International* 2002;**13**(3):257–64.

Chenoweth 2009 ***{published data only}***

ACTRN12608000084381. Dementia Care Mapping in residential aged care. www.anzctr.org.au/trial_view.aspx?ID=82599 (first received 14 February 2008).

* Chenoweth L, King MT, Jeon YH, Brodaty H, Stein-Parbury J, Norman R, et al. Caring for Aged Dementia Care Resident Study (CADRES) of person-centred care, dementia-care mapping, and usual care in dementia: a cluster-randomised trial. [Erratum appears in Lancet Neurology 09;8(5):419]. *Lancet Neurology* 2009;**8**(4):317–25.

Norman R, Haas M, Chenoweth L, Jeon Y-H, King M, Brodaty H, et al. *Dementia care mapping and patient-centred care in Australian residential homes: an economic evaluation of the CARE Study*. Sydney: Centre for Health Economics Research and Evaluation, 2008.

Choi 2005 ***{published data only}***

Choi JH, Moon JS, Song R. Effects of Sun-style Tai Chi exercise on physical fitness and fall prevention in fall-prone older adults. *Journal of Advanced Nursing* 2005;**51**(2):150–7.

Clifton 2009 ***{unpublished data only}***

* Clifton GD, Shonkwiler JS, Kelly KE. Report of a randomized, controlled trial to assess reduction in falls and related injuries using the FallSaver™ position monitor. Unpublished report 2009.

NCT00249743. Clinical evaluation of a wireless monitoring device to reduce falls in the elderly and others at high risk of falling. clinicaltrials.gov/ct2/show/NCT00249743 (first received 03 November 2005).

Colon-Emeric 2013 ***{published data only}***

Colon-Emeric C. RE: CONNECT for better fall prevention in nursing homes: Results from a pilot intervention study. Email to: C Panagoda. 12 July 2016.

Colon-Emeric C, Pinheiro SM, Simpson K, Porter K, Corazzini K, Anderson RA. Improving uptake of a falls educational program by focusing on staff interactions. *Journal of the American Geriatrics Society Conference 2012 May 3-5; Seattle (WA)* 2012;**60**(Suppl s4):S157. C74]

Colon-Emeric CS, McConnell E, Pinheiro S, Corazzini K, Porter K, Anderson R. CONNECT for fall prevention: A randomized controlled pilot study. *Journal of the American Geriatrics Society Conference 2013 May 3-5; Grapevine (TX)* 2013;**61**(Suppl s1):S1. P2]

* Colon-Emeric CS, McConnell E, Pinheiro SO, Corazzini K, Porter K, Earp KM, et al. CONNECT for better fall prevention in nursing homes: results from a pilot intervention study. *Journal of the American Geriatrics Society* 2013;**61**(12):2150–9.

Colon-Emeric CS, Pinheiro SO, Anderson RA, Porter K, McConnell E, Corazzini K, et al. Connecting the learners: improving uptake of a nursing home educational program by focusing on staff interactions. *Gerontologist* 2014;**54**(3):446–59.

NCT00836433. CONNECT for better falls prevention in VA Community Living Centers. clinicaltrials.gov/show/NCT00836433 (first received 02 February 2009).

Cox 2008 ***{published data only}***

Cox H, Puffer S, Morton V, Cooper C, Hodson J, Masud T, et al. Educating nursing home staff on fracture prevention: a cluster randomised trial. *Age and Ageing* 2008;**37**(2):167–72.

Crotty 2004a ***{published data only}***

Crotty M, Rowett D, Spurling L, Giles LC, Phillips PA. Does the addition of a pharmacist transition coordinator

improve evidence-based medication management and health outcomes in older adults moving from the hospital to a long-term care facility? Results of a randomized, controlled trial. *American Journal of Geriatric Pharmacotherapy* 2004;**2**(4):257–64.

Crotty 2004b *{published data only}*
Crotty M, Whitehead C, Rowett D, Halbert J, Weller D, Finucane P, et al. An outreach intervention to implement evidence based practice in residential care: A randomized controlled trial [ISRCTN67855475]. *BMC Health Services Research* 2004;**4**(1):6.

Cumming 2008 *{published data only}*
Cumming RG, Sherington C, Lord SR, Simpson JM, Vogler C, Cameron ID, et al. Cluster randomised trial of a targeted multifactorial intervention to prevent falls among older people in hospital. *BMJ* 2008;**336**(7647):758–60.

da Silva Borges 2014 *{published data only}*
da Silva Borges E, de Souza Vale RG, Cader SA, Leal S, Miguel F, Pernambuco CS, et al. Postural balance and falls in elderly nursing home residents enrolled in a ballroom dancing program. *Archives of Gerontology and Geriatrics* 2014;**59**(2):312–6.

Donald 2000 *{published data only}*
Donald IP, Pitt K, Armstrong E, Shuttleworth H. Preventing falls on an elderly care rehabilitation ward. *Clinical Rehabilitation* 2000;**14**(2):178–85.

Dyer 2004 *{published data only}*
Dyer CA, Taylor GJ, Halpin M, Dyer CA, Robertson DR, Harrington R. Falls prevention in residential homes: a randomised controlled trial (abstract). *Age and Ageing* 2003;**32**(Suppl 1):16.
* Dyer CA, Taylor GJ, Reed M, Dyer CA, Robertson DR, Harrington R. Falls prevention in residential care homes: a randomised controlled trial. *Age and Ageing* 2004;**33**(6):596–602.
N0037081503. Preventing falls in residential homes: a multi-agency pilot study. www.nihr.ac.uk/Profiles/NRR.aspx?Publication_ID=N0037081503 (accessed 04 March 2012).

Dykes 2010 *{published data only}*
Dykes PC, Carroll DL, Hurley A, Lipsitz S, Benoit A, Chang F, et al. Fall prevention in acute care hospitals: A randomized trial. *JAMA - Journal of the American Medical Association* 2010;**304**(17):1912–8.

Faber 2006 *{published data only}*
Faber MJ, Bosscher RJ, Chin A, Paw MJ, van Wieringen PC. Effects of exercise programs on falls and mobility in frail and pre-frail older adults: A multicenter randomized controlled trial. *Archives of Physical Medicine & Rehabilitation* 2006;**87**(7):885–96.

Flicker 2005 *{published data only}*
Flicker L, MacInnis R, Stein M, Scherer S, Mead K, Nowson C, et al. Erratum: Should older people in residential care receive vitamin D to prevent falls? Results of a randomized trial (Journal of the American Geriatrics Society (2005) 53 (1881-1888)). *Journal of the American Geriatrics Society* 2012;**60**(8):1599.
Flicker L, MacInnis R, Stein M, Scherer S, Mead K, Nowson C, et al. Should all older people in residential care be supplemented with vitamin D to prevent falls? Results of a randomised trial [abstract]. 14th National conference on falls and postural instability 2003 Sept 8; London.
Flicker L, MacInnis R, Stein M, Scherer S, Mead K, Nowson C, et al. Should all older people in residential care receive vitamin D to prevent falls? Results of a randomised trial [abstract]. *Journal of Bone and Mineral Research* 2004; **19(Suppl 1)**:S99.
* Flicker L, MacInnis R, Stein M, Scherer S, Mead K, Nowson C, et al. Should older people in residential care receive Vitamin D to prevent falls? Results of a randomized trial. *Journal of the American Geriatrics Society* 2005;**53**(11):1881–8.
Flicker L, MacInnis R, Stein M, Scherer S, Mead K, Nowson C, et al. Vitamin D to prevent falls in older people in residential care. *Asia Pacific Journal of Clinical Nutrition* 2005;**14 Suppl**:S18.

Frankenthal 2014 *{published data only}*
Frankenthal D, Israeli A, Caraco Y, Kalendaryev E, Zandman-Goddard G, Lerman Y. Long-term outcomes of medication intervention using the screening tool of older persons potentially inappropriate prescriptions screening tool to alert doctors to right treatment criteria. Journal of the American Geriatrics Society 2017; Vol. 65, issue 2: e33–8.
Frankenthal D, Kalendaryev E, Lerman Y. Intervention with the STOPP/START criteria among elderly residents of a chronic geriatric facility: A randomized clinical trial. 10th International Congress of the European Union Geriatric Medicine Society - Geriatric Medicine Crossing Borders, EUGMS14; 2014; Rotterdam, Netherlands. 2014:S69.
* Frankenthal D, Lerman Y, Kalendaryev E, Lerman Y. Intervention with the screening tool of older persons potentially inappropriate prescriptions/screening tool to alert doctors to right treatment criteria in elderly residents of a chronic geriatric facility: a randomized clinical trial. *Journal of the American Geriatrics Society* 2014;**62**(9):1658–65.
Frankenthal D, Lerman Y, Lerman Y, Kalendaryev E. Response to Lavan and colleagues. Journal of the American Geriatrics Society 2015; Vol. 63, issue 5:1044–5. PUBMED: 25989578]
Lavan AH, O'Mahony D, Gallagher P. Comments on "intervention with the screening tool of older persons potentially inappropriate prescriptions/screening tool to alert doctors to right treatment criteria in elderly residents of a chronic geriatric facility: a randomized clinical trial". Journal of the American Geriatrics Society 2015; Vol. 63, issue 5:1043–4. PUBMED: 25989577]
NCT01602744. Wolfson Medical Center. The use of STOPP/START criteria for medication intervention among elderly population living in a geriatric hospital. ClinicalTrials.gov/show/NCT01602744 (first received 1

May 2012).

Fu 2015 *{published data only}*

* Fu AS, Gao KL, Tung AK, Tsang WW, Kwan MM. Effectiveness of exergaming training in reducing risk and incidence of falls in frail older adults with a history of falls. *Archives of Physical Medicine & Rehabilitation* 2015;**96**(12): 2096–102.

Tsang WW, Fong SS, Tung KK, Fu AS. Is virtual reality exercise effective in reducing falls among older adults with a history of falls?. *Physiotherapy (United Kingdom). World Confederation for Physical Therapy Congress* 2015;**101**:Suppl 1. Conference Abstract 201552]

Garcia Gollarte 2014 *{published data only}*

Garcia-Gollarte F, Baleriola-Julvez J, Ferrero-Lopez I, Cuenllas-Diaz A, Cruz-Jentoft AJ. An educational intervention on drug use in nursing homes improves health outcomes resource utilization and reduces inappropriate drug prescription. *Journal of the American Medical Directors Association* 2014;**15**(12):885–91.

Grieger 2009 *{published data only}*

Grieger JA, Nowson CA, Jarman HF, Malon R, Ackland LM. Multivitamin supplementation improves nutritional status and bone quality in aged care residents. *European Journal of Clinical Nutrition* 2009;**63**(4):558–65.

Haines 2004 *{published data only}*

* Haines TP, Bennell KL, Osborne RH, Hill KD. Effectiveness of targeted falls prevention programme in subacute hospital setting: randomised controlled trial. *BMJ* 2004;**328**(7441):676–9.

Haines TP, Hill KD, Bennell KL, Osborne RH. Additional exercise for older subacute hospital inpatients to prevent falls: benefits and barriers to implementation and evaluation. *Clinical Rehabilitation* 2007;**21**(8):742–53.

Haines TP, Hill KD, Bennell KL, Osborne RH. Patient education to prevent falls in subacute care. *Clinical Rehabilitation* 2006;**20**(11):970–9.

Haines 2010 *{published data only}*

ACTRN12609000243213. Cluster randomized trial to evaluate the effectiveness of low-low beds for the prevention of in-hospital falls. www.anzctr.org.au/trial_view.aspx?ID=83489 (first received 12 May 2009).

* Haines TP, Bell RA, Varghese PN. Pragmatic, cluster randomized trial of a policy to introduce low-low beds to hospital wards for the prevention of falls and fall injuries. *Journal of the American Geriatrics Society* 2010;**58**(3): 435–41.

Haines 2011 *{published data only}*

Haines TP, Hill AM, Hill KD, Brauer SG, Hoffmann T, Etherton-Beer C, et al. Cost effectiveness of patient education for the prevention of falls in hospital: economic evaluation from a randomized controlled trial. *BMC Medicine* 2013;**11**:135.

* Haines TP, Hill AM, Hill KD, McPhail S, Oliver D, Brauer S, et al. Patient education to prevent falls among older hospital inpatients: a randomized controlled trial. *Archives of Internal Medicine* 2011;**171**(6):516–24.

Hill AM, Hill K, Brauer S, Oliver D, Hoffmann T, Beer C, et al. Evaluation of the effect of patient education on rates of falls in older hospital patients: Description of a randomised controlled trial. *BMC Geriatrics* 2009;**9**:14.

Hill AM, Hoffmann T, Beer C, McPhail S, Hill KD, Oliver D, et al. Falls after discharge from hospital: is there a gap between older peoples' knowledge about falls prevention strategies and the research evidence?. *Gerontologist* 2011;**51** (5):653–62.

Hill AM, Hoffmann T, Haines TP. Circumstances of falls and falls-related injuries in a cohort of older patients following hospital discharge. *Clinical Interventions In Aging* 2013;**8**:765–74.

Hill AM, Hoffmann T, McPhail S, Beer C, Hill KD, Oliver D, et al. Evaluation of the sustained effect of inpatient falls prevention education and predictors of falls after hospital discharge--follow-up to a randomized controlled trial. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences* 2011;**66**(9):1001–12.

Healey 2004 *{published data only}*

Healey F, Monro A, Cockram A, Adams V, Heseltine D. Using targeted risk factor reduction to prevent falls in older in-patients: a randomised controlled trial. *Age and Ageing* 2004;**33**(4):390–5.

Hill 2015 *{published and unpublished data}*

ACTRN12612000877886. Improving older patients' safety in Western Australian hospitals - reducing falls in rehabilitation units. www.anzctr.org.au/ACTRN12612000877886.aspx (first received 16 August 2012).

Hill AM, McPhail S, Waldron N, Etherton-Beer C, Flicker L, Ingram K, et al. Reducing falls in hospital through education to change patient and staff behaviour: A stepped wedge cluster randomised controlled effectiveness trial. *Physiotherapy (United Kingdom)* 2015;**101**(Suppl 1):e984.

Hill AM, McPhail SM, Waldron N, Etherton-Beer C, Ingram K, Flicker L, et al. Fall rates in hospital rehabilitation units after individualised patient and staff education programmes: a pragmatic, stepped-wedge, cluster-randomised controlled trial. *Lancet* 2015;**385** (9987):2592–9.

* Hill, AM, Waldron N, Etherton-Beer C, McPhail SM, Ingram K, Flicker L, et al. A stepped-wedge cluster randomised controlled trial for evaluating rates of falls among inpatients in aged care rehabilitation units receiving tailored multimedia education in addition to usual care: a trial protocol. *BMJ Open* 2014;**4**(1):e004195.

Houghton 2014 *{published and unpublished data}*

Desborough J. Re: Multi-professional clinical medication reviews in care homes for the elderly: study protocol for a randomised controlled trial with cost effectiveness analysis [personal communication]. Email to: S. Dyer. 26 July 2016.

Desborough J, Houghton J, Wood J, Wright D, Holland R, Sach T, et al. Multi-professional clinical medication

reviews in care homes for the elderly: study protocol for a randomised controlled trial with cost effectiveness analysis. *Trials* 2011;**12**:218.
Desborough J, Somally D. Multi-professional medication reviews in care homes for older people: What were the interventions in the CAREMED study?. *International Journal of Pharmacy Practice* 2014;**22**(Suppl S2):10. Abstract 0008]
* Houghton J, Kirthisinga V, Desborough JA. Multi-professional medication reviews in care homes for older people: Results from the CARE MED randomised controlled trial. *Age and Ageing* 2014;**43**(suppl 2):ii19–20. DOI: 10.1093/ageing/afu131.5
ISRCTN90761620. Multi-professional clinical medication reviews in care homes for the elderly. controlled-trials.com/ISRCTN90761620 (first received 21 October 2010).
Sach TH, Desborough J, Houghton J, Holland R, CAREMED study team. Applying micro-costing methods to estimate the costs of pharmacy interventions: an illustration using multi-professional clinical medication reviews in care homes for older people. *International Journal of Pharmacy Practice* 2015;**23**:237–47. DOI: 10.1111/ijpp.12162

Huang 2016 ***{published data only}***
Huang TT, Chung ML, Chen FR, Chin YF, Wang BH. Evaluation of a combined cognitive-behavioural and exercise intervention to manage fear of falling among elderly residents in nursing homes. *Aging & Mental Health* 2016; **20**(1):2–12. DOI: 10.1080/13607863.2015.1020411

Imaoka 2016 ***{published and unpublished data}***
Imaoka M. RE: "Intervention for falls: Reduced exercise and vitamin D supplementation among the institutionalized frail elderly" [personal communication]. Email to: S Dyer. 12 September 2016.
Imaoka M, Higuchi Y, Todo E, Hirasima K, Kitagawa T, Ueda T, et al. Intervention for falls: Reduced exercise and vitamin D supplementation among the institutionalized frail elderly. *Physiotherapy (United Kingdom)* 2015;**101** (Suppl 1):e641. Conference Abstract 201552]
* Imaoka M, Higuchi Y, Todu E, Kitagwa T, Ueda T. Low-frequency exercise and vitamin D supplementation reduce falls among institutionalized frail elderly. *International Journal of Gerontology* 2016;**10**(4):202–6.

Irez 2011 ***{published data only}***
Irez GB, Ozdemir RA, Evin R, Irez SG, Korkusuz F. Integrating pilates exercise into an exercise program for 65+ year-old women to reduce falls. *Journal of Sports Science & Medicine* 2011;**10**(1):105–11.

Jarvis 2007 ***{published data only}***
Clague N, Kerr KM, Mockett SP. A pilot randomised trial to assess the effects of inpatient physiotherapy on falls in the elderly (poster 49). Proceedings of the Chartered Society of Physiotherapy Annual Congress and Exhibition; 2003 Oct 17-19; Birmingham (UK). London: Chartered Society of Physiotherapy, 2003:81.
* Jarvis N, Kerr K, Mockett S. Pilot study to explore the feasibility of a randomised controlled trial to determine the dose effect of physiotherapy on patients admitted to hospital following a fall. *Practical Evidence* 2007;**2**(2):4–12.

Jensen 2002 ***{published data only}***
* Jensen J, Lundin-Olsson L, Nyberg L, Gustafson Y. Falls and injury prevention in older people living in residential care facilities. A cluster randomized trial. *Annals of Internal Medicine* 2002;**136**(10):733–41.
Jensen J, Nyberg L, Gustafson Y, Lundin-Olsson L. Fall and injury prevention in residential care-effects in residents with higher and lower levels of cognition. *Journal of the American Geriatrics Society* 2003;**51**(5):627–35.
Jensen J, Nyberg L, Rosendahl E, Gustafson Y, Lundin-Olsson L. Effects of a fall prevention program including exercise on mobility and falls in frail older people living in residential care facilities. *Aging-Clinical & Experimental Research* 2004;**16**(4):283–92.

Juola 2015 ***{published data only}***
Juola AL, Bjorkman MP, Kautiainen H, Pylkkanen S, Finne-Soveri UH, Soini H, et al. Nursing staff education to reduce potentially harmful medication use among older people in assisted living facilities: Effects of randomized controlled trial on cognition and falls. 10th International Congress of the European Union Geriatric Medicine Society - Geriatric Medicine Crossing Borders, EUGMS14; Rotterdam, Netherlands. 2014. S51]
* Juola AL, Bjorkman MP, Pylkkanen S, Finne-Soveri H, Soini H, Kautiainen H, et al. Nurse education to reduce harmful medication use in assisted living facilities: effects of a randomized controlled trial on falls and cognition. *Drugs and Aging* 2015;**32**(11):947–55.
Pitkälä KH, Juola AL, Kautiainen H, Soini H, Finne-Soveri UH, Bell JS, et al. Education to reduce potentially harmful medication use among residents of assisted living facilities: a randomized controlled trial. *Journal of the American Medical Directors Association* 2014;**15**(12):892–8. DOI: 10.1016/j.jamda.2014.04.002

Kennedy 2015 ***{published data only}***
Ioannidis G, Papaioannou A, Kennedy C, Giangregorio L, Pickard L, Johnson J, et al. Vitamin D and calcium supplementation in women and men living in long term care (LTC) homes: The vitamin D osteoporosis study (VIDOS). *Journal of Bone and Mineral Research* 2010;**25**: S1. Abstract SU0416]
Kennedy C, Papaioannou A, Ioannidis G, Giangregorio L, Pickard L, Johnson J, et al. The vitamin D in osteoporosis study (VIDOS): A novel knowledge translation initiative in Canadian long-term care homes. *Journal of Bone and Mineral Research* 2010;**25**:S1. Abstract S344]
Kennedy CC, Ioannidis G, Giangregorio LM, Adachi JD, Thabane L, Morin SN, et al. An interdisciplinary knowledge translation intervention in long-term care: Study protocol for the vitamin D and osteoporosis study (ViDOS) pilot cluster randomized controlled trial. *Implementation Science* 2012;**7**(1):48.
* Kennedy CC, Ioannidis G, Thabane L, Adachi JD, Marr S, Giangregorio LM, et al. Successful knowledge translation intervention in long-term care: final results from

the vitamin D and osteoporosis study (ViDOS) pilot cluster randomized controlled trial. *Trials [Electronic Resource]* 2015;**16**:214.

Papaioannou A, Kennedy C, Ioannidis G, Giangregorio L, Thabane L, Morin S, et al. A successful knowledge translation intervention in long-term care: Results from the vitamin D and osteoporosis (ViDOS) cluster randomized trial. *Journal of Bone and Mineral Research* 2013;**28**:Suppl 1. Abstract SU0394]

Papaioannou A, Marr S, Ioannidis G, Kennedy C, Giangregorio L, Pickard L, et al. Bisphosphonate use in women and men who are at high risk for new fractures and living in long-term care homes: The vitamin D osteoporosis study (ViDOS). *Journal of Bone and Mineral Research* 2010; **25**:S1. Abstract SA0385]

Kerse 2004 ***{published data only}***

Kerse N, Butler M, Robinson E, Todd M. Fall prevention in residential care: a cluster, randomized, controlled trial. *Journal of the American Geriatrics Society* 2004;**52**(4): 524–31.

Kerse 2008 ***{published data only}***

ACTRN12605000667617. Promoting Independence in residential care. www.anzctr.org.au/trial_view.aspx?id=735 (first received 20 October 2005).

* Kerse N, Peri K, Robinson E, Wilkinson T, von Randow M, Kiata L, et al. Does a functional activity programme improve function, quality of life, and falls for residents in long term care? Cluster randomised controlled trial. *BMJ* 2008;**337**(7675):a1445.

Peri K, Kerse N, Kiata L, Wilkinson T, Robinson E, Parsons J, et al. Promoting independence in residential care: successful recruitment for a randomized controlled trial. *Journal of the American Medical Directors Association* 2008;**9** (4):251–6.

Klages 2011 ***{published data only}***

Klages K, Zecevic A, Orange JB, Hobson S. Potential of Snoezelen room multisensory stimulation to improve balance in individuals with dementia: a feasibility randomized controlled trial. *Clinical Rehabilitation* 2011; **25**(7):607–16.

Koh 2009 ***{published and unpublished data}***

Koh S. personal communication June 8 2012.

* Koh SL, Hafizah N, Lee JY, Loo YL, Muthu R. Impact of a fall prevention programme in acute hospital settings in Singapore. *Singapore Medical Journal* 2009;**50**(4):425–32.

Kovacs 2012 ***{published data only}***

Kovacs E, Toth K, Denes L, Valasek T, Hazafi K, Molnar G, et al. Effects of exercise programs on balance in older women with age-related visual problems: a pilot study. *Archives of Gerontology & Geriatrics* 2012;**55**(2):446–52.

Kovacs 2013 ***{published data only}***

Kovacs E, Sztruhar JI, Karoczi CK, Korpos A, Gondos T. Effects of a multimodal exercise program on balance, functional mobility and fall risk in older adults with cognitive impairment: a randomized controlled single-blind study. *European Journal of Physical & Rehabilitation Medicine* 2013;**49**(5):639–48.

Lapane 2011 ***{published data only}***

Lapane KL, Hughes CM, Daiello LA, Cameron KA, Feinberg J. Effect of a pharmacist-led multicomponent intervention focusing on the medication monitoring phase to prevent potential adverse drug events in nursing homes. *Journal of the American Geriatrics Society* 2011;**59**(7): 1238–45.

Law 2006 ***{published data only}***

ISRCTN47348080. A trial of vitamin D in preventing hip fracture. controlled-trials.com/ISRCTN47348080 (first received 19 December 2002).

Law M, Withers H, Morris J. Vitamin D supplementation and the prevention of fractures and falls [reply]. *Age and Ageing* 2007;**36**(2):233. DOI: 10.1093/ageing/afl176

* Law M, Withers H, Morris J, Anderson F. Vitamin D supplementation and the prevention of fractures and falls: results of a randomised trial in elderly people in residential accommodation. *Age and Ageing* 2006;**35**(5):482–6.

Mador 2004 ***{published data only}***

Mador JE, Giles L, Whitehead C, Crotty M. A randomized controlled trial of a behavior advisory service for hospitalized older patients with confusion. *International Journal of Geriatric Psychiatry* 2004;**19**(9):858–63.

Mayo 1994 ***{published data only}***

Mayo NE, Gloutney L, Levy AR. A randomized trial of identification bracelets to prevent falls among patients in a rehabilitation hospital. *Archives of Physical Medicine & Rehabilitation* 1994;**75**(12):1302–8.

McMurdo 2000 ***{published data only}***

* McMurdo ME, Millar AM, Daly F. A randomized controlled trial of fall prevention strategies in old peoples' homes. *Gerontology* 2000;**46**(2):83–7.

Millar AM, McMurdo MET. A trial of falls prevention [abstract]. *Age and Ageing* 1999;**28**(Suppl 1):15.

N0405062690. Preventing falls in residents of old peoples homes. www.nihr.ac.uk/Profiles/NRR.aspx?Publication_ID=N0405062690 (accessed 13 March 2012).

Meyer 2009 ***{published data only}***

ISRCTN37794278. Predicting the risk of falling - efficacy of a risk assessment tool compared to nurses' judgement: a cluster-randomised controlled trial. www.controlled-trials.com/ISRCTN37794278 (first received 11 August 2005).

Meyer G, Kopke S, Bender R, Muhlhauser I. Predicting the risk of falling--efficacy of a risk assessment tool compared to nurses' judgement: a cluster-randomised controlled trial [ISRCTN37794278]. *BMC Geriatrics* 2005;**5**:14.

* Meyer G, Kopke S, Haastert B, Muhlhauser I. Comparison of a fall risk assessment tool with nurses' judgement alone: a cluster-randomised controlled trial. *Age and Ageing* 2009; **38**(4):417–23.

Michalek 2014 ***{published data only}***

* Michalek C, Wehling M, Schlitzer J, Frohnhofen H. Effects of "Fit fOR The Aged" (FORTA) on pharmacotherapy and

clinical endpoints--a pilot randomized controlled study. *European Journal of Clinical Pharmacology* 2014;**70**(10): 1261–7.
Wehling M, Kuhn-Thiel A, Throm C, Burkhardt H, Frohnhofen H, Pazan F, et al. Clinical validation of the FORTA (Fit fOR The Aged) List in a prospective randomized controlled clinical study. 11th International Congress of the European Union Geriatric Medicine Society, EUGMS15; 2015 September 16-18; Oslo Norway. 2015:S146. Abstract P–426; Available from: www.europeangeriaticmedicine.com/]

Mulrow 1994 ***{published data only}***
Mulrow CD, Gerety MB, Kanten DN. Effects of physical therapy on functional status of nursing home residents. *Journal of the American Geriatrics Society* 1993;**41**(3):326–8.
* Mulrow CD, Meghan BG, Kanten D, Cornell JE, DeNino LA, Chioda L, et al. A randomized trial of physical rehabilitation for very frail nursing home residents. *JAMA* 1994;**271**(7):519–24.

Neyens 2009 ***{published data only}***
* Neyens JC, Dijcks BP, Twisk J, Schols JM, Van Haastregt JC, Van Den Heuvel WJ, et al. A multifactorial intervention for the prevention of falls in psychogeriatric nursing home patients, a randomised controlled trial (RCT). *Age and Ageing* 2009;**38**(2):194–9.
Neyens JC, Schols JM, Dijcks BP, Van Haastregt JC, Van den Heuvel WJ, Crebolder HF, et al. Development and implementation of a multifactorial intervention for psychogeriatric nursing home patients targeted on the prevention of falls and fall-related injuries [Ontwikkeling en implementatie van een multifactoriele interventie gericht op de preventie van valincidenten en de gevolgen daarvan bij psychogeriatrische verpleeghuispatienten]. *Tijdschrift voor Verpleeghuisgeneeskunde* 2002;**26**(Special Issue):24–8.

Nowalk 2001 ***{published data only}***
Nowalk MP, Prendergast JM, Bayles CM, D'Amico FJ, Colvin GC. A randomized trial of exercise programs among older individuals living in two long-term care facilities: the FallsFREE program. *Journal of the American Geriatrics Society* 2001;**49**(7):859–65.

Patterson 2010 ***{published data only}***
ISRCTN18113807. An evaluation of an adapted United States model of pharmaceutical care to improve psychoactive prescribing for nursing home residents in Northern Ireland. controlled-trials.com/ISRCTN18113807 (accessed 07 March 2012).
Patterson SM, Hughes CM, Cardwell C, Lapane KL, Murray AM, Crealey GE. A cluster randomized controlled trial of an adapted U.S. model of pharmaceutical care for nursing home residents in Northern Ireland (Fleetwood Northern Ireland study): a cost-effectiveness analysis. *Journal of the American Geriatrics Society* 2011;**59**(4): 586–93.
* Patterson SM, Hughes CM, Crealey G, Cardwell C, Lapane KL. An evaluation of an adapted U.S. model of pharmaceutical care to improve psychoactive prescribing for nursing home residents in Northern Ireland (Fleetwood Northern Ireland study). *Journal of the American Geriatrics Society* 2010;**58**(1):44–53.
Patterson SM, Hughes CM, Lapane KL. Assessment of a United States pharmaceutical care model for nursing homes in the United Kingdom. *Pharmacy World and Science* 2007; **29**(5):517–25.

Peyro Saint Paul 2013 ***{published and unpublished data}***
* Peyro Saint Paul L, Martin J, Gaillard C, Mosquet B, Coquerel A, de la Gastine B. Moderate potentially drug-induced hyponatremia in older adults: benefit in drug reduction. [L'hyponatrémie modérée potentiellement médicamenteuse du sujet âgé : bénéfice de la réduction des medicaments]. *Therapie* 2013;**68**(6):341–6.
Peyro Saint Paul L, Martin J, Gaillard C, Mosquet B, Coquerel A, de la Gastine B, et al. Moderate, potentially drug-induced hyponatremia in older adults: is there a benefit in drug reduction?. *Journal of the American Geriatrics Society* 2012;**60**(10):1991–3. DOI: 10.1111/j.1532-5415.2012.04186.x
Peyro Saint Paul L, Martin J, Mosquet B, Gaillard C, De La Gastine B. Benefit of pharmacological intervention on drug-induced mild hyponatremia in elderly: A prospective randomised trial. *Fundamental and Clinical Pharmacology* 2012;**26**(Suppl s1):74. Abstract P229]

Potter 2016 ***{published and unpublished data}***
ACTRN12611000370909. Deprescribing in frail older people: a randomised controlled trial [A randomised controlled trial in frail older people living in residential aged care facilities in Western Australia designed to test the effect of deprescribing on medication burden at one year]. www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=336757&isReview=true (first received 4 April 2011).
Potter K. Re: Falls data from deprescribing study [personal communication]. Email to: S Dyer. 4 March 2017. ACTRN12611000370909]
* Potter K, Flicker L, Page A, Etherton-Beer C. Deprescribing in frail older people: a randomised controlled trial. *PLOS One* 2016;**11**(3):e0149984. DOI: 10.1371/journal.pone.0149984
Potter KB, Flicker L, Beer C. Deprescribing in frail older people: Protocol for a randomised controlled trial. *Basic and Clinical Pharmacology and Toxicology* 2011;**109**(Suppl s1):1–176. Abstract P335]

Ray 1997 ***{published data only}***
Ray WA, Taylor JA, Meador KG, Thapa PB, Brown AK, Kajihara HK, et al. A randomized trial of a consultation service to reduce falls in nursing homes. *JAMA* 1997;**278** (7):557–62.

Rosendahl 2008 ***{published data only}***
ISRCTN31631302. The Frail Older People-Activity and Nutrition [FOPANU] study in Umeå: a cluster-randomised controlled trial. controlled-trials.com/ISRCTN31631302 (accessed 09 March 2012).
Littbrand H, Carlsson M, Lundin Olsson L, Lindelöf N, Håglin L, Gustafson Y, et al. Effect of a high-intensity functional exercise program on functional balance: preplanned subgroup analyses of a randomized controlled

trial in residential care facilities. *Journal of the American Geriatrics Society* 2011;**59**(7):1274–82.
Littbrand H, Lundin Olsson L, Gustafson Y, Rosendahl E. The effect of a high-intensity functional exercise program on activities of daily living: a randomized controlled trial in residential care facilities. *Journal of the American Geriatrics Society* 2009;**57**(10):1741–9.
Littbrand H, Rosendahl E, Lindelof N, Lundin-Olsson L, Gustafson Y, Nyberg L. A high-intensity functional weight-bearing exercise program for older people dependent in activities of daily living and living in residential care facilities: evaluation of the applicability with focus on cognitive function. *Physical Therapy* 2006;**86**(4):489–98.
Rosendahl E. *Fall prediction and a high-intensity functional exercise programme to improve physical functions and to prevent falls among older people living in residential care facilities [thesis]*. Umeå, Sweden: Umeå University, 2009.
* Rosendahl E, Gustafson Y, Nordin E, Lundin-Olsson L, Nyberg L. A randomized controlled trial of fall prevention by a high-intensity functional exercise program for older people living in residential care facilities. *Aging Clinical and Experimental Research* 2008;**20**(1):67–75.
Rosendahl E, Lindelof N, Littbrand H, Yifter Lindgren E, Nordin E, Lundin Olssen L, et al. High-intensity functional exercise program for older people dependent in ADL: an RCT evaluating the effects on physical functions and falls [abstract]. *Physiotherapy* 2007;**93**(Suppl 1):S377.
Rosendahl E, Lindelof N, Littbrand H, Yifter-Lindgren E, Lundin-Olsson L, Haglin L, et al. High-intensity functional exercise program and protein-enriched energy supplement for older persons dependent in activities of daily living: a randomised controlled trial. *Australian Journal of Physiotherapy* 2006;**52**(2):105–13.

Rubenstein 1990 ***{published data only}***
Rubenstein LZ, Robbins AS, Josephson KR, Schulman BL, Osterweil D. The value of assessing falls in an elderly population. A randomized clinical trial. *Annals of Internal Medicine* 1990;**113**(4):308–16.

Sakamoto 2006 ***{published data only}***
Sakamoto K, Nakamura T, Hagino H, Endo N, Mori S, Muto Y, et al. Effects of unipedal standing balance exercise on the prevention of falls and hip fracture among clinically defined high-risk elderly individuals: A randomized controlled trial. *Journal of Orthopaedic Science* 2006;**11**(5): 467–72.

Sakamoto 2012 ***{published data only}***
Ebihara S, Shannon F, Sakamoto Y, Ebihara T, Kohzuki M. Response Letter to Lakin and Doe...Lakin JR, Doe EW. Something just doesn't smell quite right. J Am Geriatr Soc. 2013;61:313. Journal of the American Geriatrics Society 2013; Vol. 61, issue 2:313–4.
JPRN-UMIN000004222. Prevention of fall by lavender oil olfactory stimulation. apps.who.int/trialsearch/trial.aspx?trialid=JPRN-UMIN000004222 (accessed 12 June 2012).
Lakin JR, Doe EW. Something just doesn't smell quite right...Sakamoto Y, Ebihara S, Ebihara T, et al. Fall prevention using olfactory stimulation with lavender odor in elderly nursing home residents: a randomized controlled trial. J Am Geriatr Soc. 2012;60:1005-1011... [corrected][published erratum appears in J AM GERIATR SOC 2013 Apr; 61(4): 673]. Journal of the American Geriatrics Society 2013; Vol. 61, issue 2:313.
Sakamoto Y, Ebihara S, Ebihara T, Tomita N, Toba K, Freeman S, et al. Erratum. Fall prevention using olfactory stimulation with lavender odor in elderly nursing home residents: a randomized controlled trial. Journal of the American Geriatrics Society 2012; Vol. 60, issue 11:2193.
* Sakamoto Y, Ebihara S, Ebihara T, Tomita N, Toba K, Freeman S, et al. Fall prevention using olfactory stimulation with lavender odor in elderly nursing home residents: a randomized controlled trial. *Journal of the American Geriatrics Society* 2012;**60**(6):1005–11.

Salvà 2016 ***{published and unpublished data}***
Figuls, M.R. Randomized clinical trial of a fall-prevention strategy for institutionalized elderly based on the Mini Falls Assessment Instrument [personal communication]. Email to: S Dyer. 11 October 2016. NCT00888953]
* Salvà A, Rojano X, Coll-Planas L, Domenèch S, Roqué I Figuls M. [Randomized clinical trial of a fall-prevention strategy for institutionalized elderly based on the Mini Falls Assessment Instrument]. [Spanish]. *Revista Espanola de Geriatria y Gerontologia* 2016;**51**(1):18–24.

Sambrook 2012 ***{published data only}***
Durvasula S, Gies P, Mason RS, Chen JS, Henderson S, Seibel MJ, et al. Vitamin D response of older people in residential aged care to sunlight-derived ultraviolet radiation. *Archives of Osteoporosis* 2014;**9**(1):1–7.
Durvasula S, Kok C, Sambrook PN, Cumming RG, Lord SR, March LM, et al. Sunlight and health: attitudes of older people living in intermediate care facilities in southern Australia. *Archives of Gerontology & Geriatrics* 2010;**51**(3): e94–9.
Durvasula S, Sambrook PN, Cameron ID. Factors influencing adherence with therapeutic sunlight exposure in older people in intermediate care facilities. *Archives of Gerontology & Geriatrics* 2012;**54**(2):e234–41.
March LM, Seibell MJ, Simpson JM, Sambrook P, Cameron ID, Durvasula S, et al. A randomised controlled trial of increased sunlight exposure to reduce vitamin D deficiency and falls risk in the elderly [abstract]. *Journal of Bone and Mineral Research* 2009;**24**(Suppl 1):S73.
NCT00322166. The FREEDOM study: a randomised controlled trial of sunlight and calcium in older people. clinicaltrials.gov/ct2/show/NCT00322166 (first received 4 May 2006).
* Sambrook PN, Cameron ID, Chen JS, Cumming RG, Durvasula S, Herrmann M, et al. Does increased sunlight exposure work as a strategy to improve vitamin D status in the elderly: a cluster randomised controlled trial. *Osteoporosis International* 2012;**23**(2):615–24.
Wilson N, Hilmer S, March L, Cameron I, Lord S, Mason R, et al. Physical functioning measures and risk of falling in older people living in residential aged care facilities.

Therapeutic Advances in Musculoskeletal Disease 2011;**3**(1): 9–15.
Wilson NM, Hilmer SN, March LM, Cameron ID, Lord SR, Seibel MJ, et al. Associations between drug burden index and falls in older people in residential aged care. *Journal of the American Geriatrics Society.* 2011;**59**(5): 875–80.
Wilson NM, Hilmer SN, March LM, Cameron ID, Lord SR, Seibel MJ, et al. Associations between drug burden index and physical function in older people in residential aged care facilities. *Age and Ageing* 2010;**39**(4):503–7.

Saravanakumar 2014 ***{published data only}***
ACTRN12612000103864, University of Newcastle. Preventing falls in older people in residential care settings: improving balance through Tai Chi and Yoga - a randomised controlled study. www.anzctr.org.au/ACTRN12612000103864.aspx (first received 8 January 2012).
* Saravanakumar P, Higgins IJ, van der Riet PJ, Marquez J, Sibbritt D. The influence of tai chi and yoga on balance and falls in a residential care setting: A randomised controlled trial. *Contemporary Nurse* 2014;**48**(1):76–87.

Schnelle 2003 ***{published data only}***
Bates-Jensen BM, Alessi CA, Al-Samarrai NR, Schnelle JF. The effects of an exercise and incontinence intervention on skin health outcomes in nursing home residents. *Journal of the American Geriatrics Society* 2003;**51**(3):348–55.
Schnelle JF, Alessi CA, Simmons SF. Translating clinical records into practice. A randomized controlled trial of exercise and incontinence care with nursing home residents. *Journal of the American Geriatrics Society* 2002;**50**(9): 1476–83.
* Schnelle JF, Kanika K, Alessi C, Osterweil D, Beck JG, Al-Samarrai N, et al. Does an exercise and incontinence intervention save healthcare costs in a nursing home population?. *Journal of the American Geriatrics Society* 2003; **51**(2):161–8.

Schoenfelder 2000 ***{published data only}***
* Schoenfelder DP. A fall prevention program for elderly individuals. Exercise in long-term care settings. *Journal of Gerontological Nursing* 2000;**26**(3):43–51.
Schoenfelder DP, Rubenstein LM. An exercise program to improve fall-related outcomes in elderly nursing home residents. *Applied Nursing Research* 2004;**17**(1):21–31.

Serra-Rexach 2011 ***{published data only}***
NCT00848978. Strength training in nonagenarians (STRONG). clinicaltrials.gov/ct2/show/NCT00848978 (accessed 04 June 2012).
Serra Rexach JA, Ruiz JR, Bustamante-Ara N, Villaran MH, Gil PG, Sanz Ibanez MJ, et al. Health enhancing strength training in nonagenarians (STRONG): rationale, design and methods. *BMC Public Health* 2009;**9**:152.
* Serra-Rexach JA, Bustamante-Ara N, Hierro Villaran M, Gonzalez Gil P, Sanz Ibanez M, Blanco Sanz N, et al. Short-term, light- to moderate-intensity exercise training improves leg muscle strength in the oldest old: A randomized controlled trial. *Journal of the American Geriatrics Society* 2011;**59**(4):594–602.

Shaw 2003 ***{published data only}***
Chapman KL, Dawson P, Shaw FE, Kenny RA. Physiotherapy intervention for cognitively impaired elderly fallers attending casualty [abstract]. *Age and Ageing* 1997;**26** (Suppl 1):13.
Dawson P, Chapman KL, Shaw FE, Kenny RA. Measuring the outcome of physiotherapy in cognitively impaired elderly patients who fall. *Physiotherapy* 1997;**83**(7):352.
Shaw FE. *Risk modification of falls in older patients with cognitive impaired and dementia attending a casualty department [thesis]*. Newcastle upon Tyne (UK): Univ. of Newcastle upon Tyne, 2001.
* Shaw FE, Bond J, Richardson DA, Dawson P, Steen IN, McKeith IG, et al. Multifactorial intervention after a fall in older people with cognitive impairment and dementia presenting to the accident and emergency department: randomised controlled trial. *BMJ* 2003;**326**(7380):73–5.
Shaw FE, Richardson DA, Dawson P, Steen IN, McKeith IG, Bond J, et al. Can multidisciplinary intervention prevent falls in patients with cognitive impairment and dementia attending a casualty department (abstract). *Age and Ageing* 2000;**29**(Suppl 1):47.

Shimada 2004 ***{published data only}***
Shimada H, Obuchi S, Furuna T, Suzuki T. New intervention program for preventing falls among frail elderly people: The effects of perturbed walking exercise using a bilateral separated treadmill. *American Journal of Physical Medicine and Rehabilitation* 2004;**83**(7):493–9.

Shorr 2012 ***{published and unpublished data}***
Shorr RI. RE: Information on number of patients. Email to: S. Dyer. 28 November 2016.
Shorr RI, Chandler AM, Kessler LA, Miller ST, Waters TM, Daniels MJ, et al. Trial of proximity alarms to prevent patient falls in hospitals. *Journal of the American Geriatrics Society* 2010;**58**(Suppl S1):S103. Abstract B80]
* Shorr RI, Chandler AM, Mion LC, Waters TM, Liu M, Daniels MJ, et al. Effects of an intervention to increase bed alarm use to prevent falls in hospitalized patients: a cluster randomized trial. *Annals of Internal Medicine* 2012;**157** (10):692–9.

Sihvonen 2004 ***{published data only}***
* Sihvonen S, Sipila S, Taskinen S, Era P. Fall incidence in frail older women after individualized visual feedback-based balance training. *Gerontology* 2004;**50**(6):411–6.
Sihvonen SE, Sipila S, Era PA. Changes in postural balance in frail elderly women during a 4-week visual feedback training: a randomized controlled trial. *Gerontology* 2004; **50**(2):87–95.

Sitja Rabert 2015 ***{published and unpublished data}***
NCT01375790. Whole-body vibration training in older people (GERIAPLAT). clinicaltrials.gov/ct2/show/NCT01375790 (first received 16 June 2011).
Sitja-Rabert M, Martinez-Zapata MJ, Fort-Vanmeerhaeghe A, Rey-Abella F, Romero-Rodriguez D, Bonfill X. Whole body vibration for older persons: an open randomized,

multicentre, parallel, clinical trial. *BMC Geriatrics* 2011; **11**:89.

* Sitja-Rabert M, Martinez-Zapata MJ, Vanmeerhaeghe AF, Abella FR, Romero-Rodriguez D, Bonfill X. Effects of a whole body vibration (WBV) exercise intervention for institutionalized older people: a randomized, multicentre, parallel, clinical trial. *Journal of the American Medical Directors Association* 2015;**16**(2):125–31.

Stenvall 2007 ***{published data only}***

Berggren M, Stenvall M, Olofsson B, Gustafson Y. Evaluation of a fall-prevention program in older people after femoral neck fracture: A one-year follow-up. *Osteoporosis International* 2008;**19**(6):801–9.

Gustafson Y. Outcomes of hip fractures: Rehabilitation programmes: Comprehensive Geriatric Assessment and Rehabilitation-a prerequisite for successful treatment of people who have suffered a hip-fracture. *European Geriatric Medicine* 2012;**3**(Suppl S1):S19.

Lundström M, Olofsson B, Stenvall M, Karlsson S, Nyberg L, Englund U, et al. Postoperative delirium in old patients with femoral neck fracture: A randomized intervention study. *Aging Clinical and Experimental Research.* 2007;**19** (3):178–86.

Stenvall M, Berggren M, Lundstrom M, Gustafson Y, Olofsson B. A multidisciplinary intervention program improved the outcome after hip fracture for people with dementia--subgroup analyses of a randomized controlled trial. *Archives of Gerontology & Geriatrics* 2012;**54**(3): e284–e289.

* Stenvall M, Olofsson B, Lundstrom M, Englund U, Borssen B, Svensson O, et al. A multidisciplinary, multifactorial intervention program reduces postoperative falls and injuries after femoral neck fracture. *Osteoporosis International* 2007;**18**(2):167–75.

Stenvall M, Olofsson B, Lundstrom M, Svensson O, Nyberg L, Gustafson Y. Inpatient falls and injuries in older patients treated for femoral neck fracture. *Archives of Gerontology and Geriatrics* 2006;**43**(3):389–99.

Stenvall M, Olofsson B, Nyberg L, Lundstrom M, Gustafson Y. Improved performance in activities of daily living and mobility after a multidisciplinary postoperative rehabilitation in older people with femoral neck fracture: a randomized controlled trial with 1-year follow-up. *Journal of Rehabilitation Medicine* 2007;**39**(3):232–8.

Streim 2012 ***{published data only}***

NCT00076622. Medication treatment for depression in nursing home residents. https://clinicaltrials.gov/ct2/show/NCT00076622 (first received 28 January 2004).

Streim 2016. NCT00076622 - Trial of antidepressant discontinuation in older residents of long term care facilities [personal communication]. Email to: S Dyer. 3 November 2016.

* Streim JE, Di FilippoS, Ten HaveT, Mavandadi S, Weintraub D, Oslin D. Antidepressant discontinuation associated with cognitive decline in older adult residents of long-term care facilities. *American Journal of Geriatric Psychiatry* 2012;**20**(Suppl 1):S148–9.

Tideiksaar 1993 ***{published data only}***

Tideiksaar R, Feiner CF, Maby J. Falls prevention: the efficacy of a bed alarm system in an acute-care setting. *Mount Sinai Journal of Medicine* 1993;**60**(6):522–7.

Toulotte 2003 ***{published data only}***

* Toulotte C, Fabre C, Dangremont B, Lensel G, Thevenon A. Effects of physical training on the physical capacity of frail, demented patients with a history of falling: a randomised controlled trial. *Age and Ageing* 2003;**32**(1): 67–73.

Toulotte C, Fabre C, Dangremont B, Thevenon A. Prevention of falls by physical training in frail demented and faller elderly subjects [Prevention des chutes par l'entrainement chez des seniors dements et chuteurs]. *Revue de Geriatrie* 2003;**28**(3):221–6.

Treacy 2015 ***{published and unpublished data}***

Treacy 2016. Your trial ACTRN12611000412932 of balance circuit classes in hospital [personal communication]. Email to: S Dyer. 14 October 2016. ACTRN12611000412932]

Treacy D, Schurr K, Lloyd B, Sherrington C. Additional standing balance circuit classes during inpatient rehabilitation improved balance outcomes: An assessor blinded randomised controlled trial. *Physiotherapy (United Kingdom)* 2015;**101**(Suppl 1):e1533–e1534.

* Treacy D, Schurr K, Lloyd B, Sherrington C. Additional standing balance circuit classes during inpatient rehabilitation improved balance outcomes: An assessor-blinded randomised controlled trial. *Age and Ageing* 2015; **44**(4):580–6.

Treacy D, Schurr K, Sherrington C. Balance circuit classes to improve balance among rehabilitation inpatients: a protocol for a randomised controlled trial. *BMC Geriatrics* 2013;**13**:75.

Tuunainen 2013 ***{published and unpublished data}***

Tuunainen E, Jantti P, Pyykko I, Rasku J, Moisio-Vilenius P, Makinen E, et al. Intervention to prevent falls in elderly adults living in a residential home. *Journal of the American Geriatrics Society* 2013;**61**(8):1426–7.

* Tuunainen E, Rasku J, Jantti P, Moisio-Vilenius P, Makinen E, Toppila E, et al. Postural stability and quality of life after guided and self-training among older adults residing in an institutional setting. *Clinical Interventions In Aging* 2013;**8**:1237–46.

Van de Ven 2014 ***{published and unpublished data}***

Van de Ven G. RE: NTR 2314: Dementia-care mapping in nursing homes: a cluster-randomised controlled trial [personal communication]. Email to: S Dyer. 7 March 2017. NTR 2314]

van de Ven G, Draskovic I, Adang EM, Donders R, Zuidema SU, Koopmans RT, et al. Effects of dementia-care mapping on residents and staff of care homes: a pragmatic cluster-randomised controlled trial. *PLOS One* 2013;**8**(7): e67325. PUBMED: 23844003]

van de Ven G, Draskovic I, Adang EM, Donders RA, Post A, Zuidema SU, et al. Improving person-centred care in nursing homes through dementia-care mapping: design of a

cluster-randomised controlled trial. *BMC Geriatrics* 2012; **12**:1. DOI: 10.1186/1471-2318-12-1

* van de Ven G, Draskovic I, van Herpen E, Koopmans RT, Donders R, Zuidema SU, et al. The economics of dementia-care mapping in nursing homes: a cluster-randomised controlled trial. *PLOS One* 2014;**9**(1):e86662.

Van Gaal 2011a ***{published data only}***

NCT00365430. SAFE or SORRY? Patient safety study of the prevention of adverse patient outcomes. clinicaltrials.gov/ct/show/NCT00365430 (accessed 11 March 2012).

van Gaal BG, Schoonhoven L, Hulscher ME, Mintjes JA, Borm GF, Koopmans RT, et al. The design of the SAFE or SORRY? study: a cluster randomised trial on the development and testing of an evidence based inpatient safety program for the prevention of adverse events. *BMC Health Services Research* 2009;**9**:58.

* van Gaal BG, Schoonhoven L, Mintjes JA, Borm GF, Hulscher ME, Defloor T, et al. Fewer adverse events as a result of the SAFE or SORRY? programme in hospitals and nursing homes. part i: primary outcome of a cluster randomised trial. *International Journal of Nursing Studies* 2011;**48**(9):1040–8.

van Gaal BG, Schoonhoven L, Mintjes JA, Borm GF, Koopmans RT, van Achterberg T. The SAFE or SORRY? programme. part II: effect on preventive care. *International Journal of Nursing Studies* 2011;**48**(9):1049–57.

van Gaal BG, Schoonhoven L, Vloet LC, Mintjes JA, Borm GF, Koopmans RT, et al. The effect of the SAFE or SORRY? programme on patient safety knowledge of nurses in hospitals and nursing homes: a cluster randomised trial. *International Journal of Nursing Studies* 2010;**47**(9): 1117–25.

Van Gaal 2011b ***{published data only}***

NCT00365430. SAFE or SORRY? Patient safety study of the prevention of adverse patient outcomes. clinicaltrials.gov/ct/show/NCT00365430 (accessed 11 March 2012).

van Gaal BG, Schoonhoven L, Hulscher ME, Mintjes JA, Borm GF, Koopmans RT, et al. The design of the SAFE or SORRY? study: a cluster randomised trial on the development and testing of an evidence based inpatient safety program for the prevention of adverse events. *BMC Health Services Research* 2009;**9**:58.

* van Gaal BG, Schoonhoven L, Mintjes JA, Borm GF, Hulscher ME, Defloor T, et al. Fewer adverse events as a result of the SAFE or SORRY? programme in hospitals and nursing homes. part i: primary outcome of a cluster randomised trial. *International Journal of Nursing Studies* 2011;**48**(9):1040–8.

van Gaal BG, Schoonhoven L, Mintjes JA, Borm GF, Koopmans RT, van Achterberg T. The SAFE or SORRY? programme. part II: effect on preventive care. *International Journal of Nursing Studies* 2011;**48**(9):1049–57.

van Gaal BG, Schoonhoven L, Vloet LC, Mintjes JA, Borm GF, Koopmans RT, et al. The effect of the SAFE or SORRY? programme on patient safety knowledge of nurses in hospitals and nursing homes: a cluster randomised trial. *International Journal of Nursing Studies* 2010;**47**(9): 1117–25.

Van het Reve 2014 ***{published and unpublished data}***

Van het Reve E. AW: Strength-balance supplemented with computerized cognitive training to improve dual task gait and divided attention in older adults: a multicenter randomized-controlled trial [personal communication]. Email to: C Panagoda 11 September 2016.

* Van het Reve E, de Bruin ED. Strength-balance supplemented with computerized cognitive training to improve dual task gait and divided attention in older adults: a multicenter randomized-controlled trial. *BMC Geriatrics* 2014;**14**:134.

Wald 2011 ***{published data only}***

Wald HL, Glasheen JJ, Guerrasio J, Youngwerth JM, Cumbler EU. Evaluation of a hospitalist-run acute care for the elderly service. *Journal of Hospital Medicine (Online)* 2011;**6**(6):313–21.

Walker 2015 ***{published data only}***

Logan PA, Walker G, Gladman JR, Robertson K, Ward M, Armstrong S, et al. A feasibility study of a cluster randomised controlled trial to evaluate a falls prevention intervention in care homes for older people. *European Geriatric Medicine* 2014;**5**(Suppl 1):s240. Abstract P497]

* Walker GM, Armstrong S, Gordon AL, Gladman J, Robertson K, Ward M, et al. The Falls In Care Home study: A feasibility randomized controlled trial of the use of a risk assessment and decision support tool to prevent falls in care homes. *Clinical Rehabilitation* 2016;**30**(10):972-83. Epub 2015 Sep 18. DOI: 10.1177/0269215515604672

Ward 2010 ***{published data only}***

Ward JA, Harden M, Gibson RE, Byles JE. A cluster randomised controlled trial to prevent injury due to falls in a residential aged care population. *Medical Journal of Australia* 2010;**192**(6):319–22.

Whitney 2017 ***{published and unpublished data}***

ISRCTN00695885, King's College, Hospital. PROF-COG prevention of falls in older people with cognitive impairment. isrctn.com/ISRCTN00695885 (first received 26 March 2013).

Whitney J. RE: PROF-COG trial ISRCTN00695885 [personal communication]. Email to: S Dyer 10 October 2016.

* Whitney J, Jackson SH, Martin FC. Feasibility and efficacy of a multi-factorial intervention to prevent falls in older adults with cognitive impairment living in residential care (ProF-Cog). A feasibility and pilot cluster randomised controlled trial. *BMC Geriatrics* 2017;**17**(1): 115. PUBMED: 28558714]

Wolf 2013 ***{published and unpublished data}***

Wolf K-H. Re: Trial "Development and pilot study of a bed-exit alarm based on a body-worn accelerometer". Email to: S Dyer 9 November 2016.

* Wolf, K-H, Hetzer K, zu Schwabedissen HM, Wiese B, Marschollek M. Development and pilot study of a bed-exit

alarm based on a body-worn accelerometer. *Zeitschrift fur Gerontologie und Geriatrie* 2013;**46**(8):727–33.

Yokoi 2015 ***{published and unpublished data}***

Yokoi H. Re: "short stick exercises for fall prevention among older adults: a cluster randomised trial" [personal communication]. Email to: C Panagoda 19 July 2016.

* Yokoi K, Yoshimasu K, Takemura S, Fukumoto J, Kurasawa S, Miyashita K. Short stick exercises for fall prevention among older adults: a cluster randomized trial. *Disability & Rehabilitation* 2015;**37**(14):1268–76.

Zermansky 2006 ***{published data only}***

ISRCTN45416155. Can a review of the medication of elderly nursing and residential home patients improve the quality of prescribing and residents' outcomes?. controlled-trials.com/ISRCTN45416155 (accessed 04 April 2012).

* Zermansky AG, Alldred DP, Petty DR, Raynor DK, Freemantle NE, Eastaugh J, et al. Clinical medication review by a pharmacist of elderly people living in care homes-randomised controlled trial. *Age and Ageing* 2006; **35**(6):586–91.

References to studies excluded from this review

Barreca 2004 ***{published data only}***

Barreca S, Sigouin CS, Lambert C, Ansley BA. Effects of extra training on the ability of stroke survivors to perform an independent sit-to-stand: A randomized controlled trial. *Journal of Geriatric Physical Therapy* 2004;**27**(2):59–68.

Bernhardt 2008 ***{published data only}***

Bernhardt J, Dewey H, Thrift A, Collier J, Donnan G. A very early rehabilitation trial for stroke (AVERT): phase II safety and feasibility. *Stroke* 2008;**39**(2):390–6.

Bosner 2012 ***{published data only}***

Bosner S, Keller H, Wohner A, Wohner C, Sonnichsen A, Baum E, et al. Prevention of falls by outdoor-walking in elderly persons at risk ("power") - A pilot study. *European Geriatric Medicine* 2012;**3**(1):28–32.

Bouwen 2008 ***{published data only}***

Bouwen A, De Lepeleire J, Buntinx F. Rate of accidental falls in institutionalised older people with and without cognitive impairment halved as a result of a staff-orientated intervention. *Age and Ageing* 2008;**37**:306–10.

Capezuti 1998 ***{published data only}***

Capezuti E. *The relationship between physical restraint removal and fall-related incidents and injuries among nursing home residents [thesis]*. Pittsburgh (PA): Univ. of Pennsylvania, 1995.

Capezuti E, Evans L, Strumpf N, Maislin G. Physical restraint use and falls in nursing home residents. *Journal of the American Geriatrics Society* 1996;**44**(6):627–33.

Capezuti E, Maislin G, Strumpf N, Evans LK. Side rail use and bed-related fall outcomes among nursing home residents. *Journal of the American Geriatrics Society* 2002;**50** (1):90–6.

Capezuti E, Strumpf N, Evans L, Maislin G. Outcomes of nighttime physical restraint removal for severely impaired nursing home residents. *American Journal of Alzheimer's Disease* 1999;**14**(3):157–64.

* Capezuti E, Strumpf NE, Evans LK, Grisso JA, Maislin G. The relationship between physical restraint removal and falls and injuries among nursing home residents. *Journals of Gerontology Series A-Biological Sciences and Medical Sciences* 1998;**53**(1):M47–52.

Evans LK, Strumpf NE, Allen-Taylor SL, Capezuti E, Maislin G, Jacobsen B. A clinical trial to reduce restraints in nursing homes. *Journal of the American Geriatrics Society* 1997;**45**:675–81.

Siegler EL, Capezuti E, Maislin G, Baumgarten M, Evans L, Strumpf N. Effects of a restraint reduction intervention and OBRA '87 regulations on psychoactive drug use in nursing homes. *Journal of the American Geriatrics Society* 1997;**45**: 791–6.

Crotty 2002 ***{published data only}***

Crotty M, Whitehead CH, Gray S, Finucane PM. Early discharge and home rehabilitation after hip fracture achieves functional improvements: a randomized controlled trial. *Clinical Rehabilitation* 2002;**16**(4):406–13.

Cucca 2017 ***{published data only}***

Cucca A, Migdadi H, Son AY, Gallo EC, Fisher SJ, Agarwal S, et al. Feasibility and safety of combining rtms with physical therapy: Preliminary data in Parkinson's disease. Brain Stimulation 2017; Vol. 10, issue 4:e63–4. 1876–4754]

Cummings 2015 ***{published data only}***

Cummings J, Lyketsos CG, Tariot P, Peskind E, Nguyen U, Knowles N, et al. Dextromethorphan/quinidine (AVP-923) efficacy and safety for treatment of agitation in persons with Alzheimer's disease: Results from a phase 2 study (NCT01584440). *American Journal of Geriatric Psychiatry* 2015;**23**(Supplement 3):S165. Conference Abstract 201512]

Dattalo 2015 ***{published data only}***

Dattalo M, Ford J, Kedrowski K, Mahoney J. Bringing healthy aging to scale: Will quality improvement coaching facilitate the dissemination of evidence-based health promotion programs to rural communities?. *Journal of the American Geriatrics Society* 2015;**63**(Suppl 1):S134.

Davison 2005 ***{published data only}***

Davison J, Bond J, Dawson P, Steen IN, Kenny RA. Patients with recurrent falls attending Accident & Emergency benefit from multifactorial intervention - a randomised controlled trial. *Age and Ageing* 2005;**34**(2):162–8.

de Morton 2007 ***{published data only}***

de Morton NA, Keating JL, Berlowitz DJ, Jackson B, Lim WK. Additional exercise does not change hospital or patient outcomes in older medical patients: a controlled clinical trial. *Australian Journal of Physiotherapy* 2007;**53**(2): 105–11.

de Souto 2016 ***{published data only}***

de Souto Barreto P, Denormandie P, Lepage B, Armaingaud D, Rapp T, Chauvin P, et al. Effects of a long-term exercise programme on functional ability in people with dementia

living in nursing homes: Research protocol of the LEDEN study, a cluster randomised controlled trial. Contemporary Clinical Trials 2016; Vol. 47:289–95. 1559–2030]

de Souto 2017 *{published data only}*
de Souto Barreto P, Cesari M, Denormandie P, Armaingaud D, Vellas B, Rolland Y. Exercise or social intervention for nursing home residents with dementia: a pilot randomized, controlled trial. Journal of the American Geriatrics Society. 2017; Vol. 65, issue 9:E123–9. 0002–8614: 1532–5415]

DeSure 2013 *{published data only}*
DeSure AR, Peterson K, Gianan FV, Pang L. An exercise program to prevent falls in institutionalized elderly with cognitive deficits: a crossover pilot study. *Hawaii Journal of Medicine & Public Health : A Journal of Asia Pacific Medicine & Public Health* 2013;**72**(11):391–5.

Donat 2007 *{published data only}*
Donat H, Ozcan A. Comparison of the effectiveness of two programmes on older adults at risk of falling: unsupervised home exercise and supervised group exercise. *Clinical Rehabilitation* 2007;**21**(3):273–83.

Drahota 2013 *{published data only}*
Drahota A, Gal D, Windsor J, Dixon S, Udell J, Ward D, et al. Pilot cluster randomised controlled trial of flooring to reduce injuries from falls in elderly care units: study protocol. *Injury Prevention* 2011;**17**(6):e7.
* Drahota AK, Ward D, Udell JE, Soilemezi D, Ogollah R, Higgins B, et al. Pilot cluster randomised controlled trial of flooring to reduce injuries from falls in wards for older people. *Age & Ageing* 2013;**42**(5):633–40.
NCT00817869. The HIP-HOP flooring study: helping injury prevention in hospitalised older people. clinicaltrials.gov/ct2/show/NCT00817869 (first received 6 January 2009).

Fiatarone 1994 *{published data only}*
Fiatarone MA, O'Neill EF, Doyle N, Clements KM, Roberts SB, Kehayias JJ, et al. The Boston FICSIT study: the effects of resistance training and nutritional supplementation on physical frailty in the oldest old. *Journal of the American Geriatrics Society* 1993;**41**(3):333–7.
* Fiatarone MA, O'Neill EF, Ryan ND, Clements KM, Solares GR, Nelson ME, et al. Exercise training and nutritional supplementation for physical frailty in very elderly people. *New England Journal of Medicine* 1994;**330**(25):1769–75.

Forster 2017 *{published data only}*
Forster A, Airlie J, Birch K, Cicero R, Cundill B, Ellwood A, et al. Research Exploring Physical Activity in Care Homes (REACH): study protocol for a randomised controlled trial. Trials 2017; Vol. 18, issue 1:182. 1745–6215]

Fossey 2006 *{published data only}*
Fossey J, Ballard C, Juszczak E, James I, Alder N, Jacoby R, et al. Effect of enhanced psychosocial care on antipsychotic use in nursing home residents with severe dementia: cluster randomised trial. *BMJ* 2006;**332**(7554):756–61.

Galik 2014 *{published data only}*
Galik E, Resnick B, Hammersla M, Brightwater J. Optimizing function and physical activity among nursing home residents with dementia: testing the impact of function-focused care. *Gerontologist* 2014;**54**(6):930–43.

Galik 2015 *{published data only}*
Galik E, Resnick B, Lerner N, Hammersla M, Gruber-Baldini AL. Function focused care for assisted living residents with dementia. Gerontologist 2015; Vol. 55: S13–26.

Gianturco 2013 *{published data only}*
Gianturco V, Troisi G, Ripani M, Marigliano V. Impact of exoskeleton Human Body Posturizer on the fall risk in the elderly: A randomized controlled trial. *European Geriatric Medicine* 2013;**4**(Suppl 1):S46.

Ginde 2017 *{published data only}*
Ginde AA, Blatchford P, Breese K, Zarrabi L, Linnebur SA, Wallace JI, et al. High-dose monthly vitamin D for prevention of acute respiratory infection in older long-term care residents: a randomized clinical trial. Journal of the American Geriatrics Society 2017; Vol. 65, issue 3: 496–503.

Graafmans 1996 *{published data only}*
Graafmans WC, Ooms ME, Hofstee HM, Bezemer PD, Bouter LM, Lips P. Falls in the elderly: a prospective study of risk factors and risk profiles. *American Journal of Epidemiology* 1996;**143**(11):1129–36. PUBMED: 8633602]

Grant 2005 *{published data only}*
Grant AM, Avenell A, Campbell MK, McDonald AM, MacLennan GS, McPherson GC, et al. Oral vitamin D3 and calcium for secondary prevention of low-trauma fractures in elderly people (Randomised Evaluation of Calcium Or vitamin D, RECORD): a randomised placebo-controlled trial. *Lancet* 2005;**365**(9471):1621–8.

Greenspan 2013 *{published data only}*
Greenspan S, Ferchak M, Lee C, Nace D, Perera S, Resnick NM. Are the iom vitamin d guidelines sufficient for long term care residents?. *Journal of the American Geriatrics Society* 2013;**61**(Suppl s1):S99.

Greenspan 2015 *{published data only}*
Greenspan SL, Perera S, Ferchak MA, Nace DA, Resnick NM. Efficacy and safety of single-dose zoledronic acid for osteoporosis in frail elderly women: a randomized clinical trial. *JAMA Internal Medicine* 2015;**175**(6):913–21.

Gruber-Baldini 2011 *{published data only}*
Gruber-Baldini AL, Resnick B, Hebel JR, Galik E, Zimmerman S. Adverse events associated with the Res-Care Intervention. *Journal of the American Medical Directors Association* 2011;**12**(8):584–9.

Gu 2006 *{published data only}*
Gu MO. personal communication May 22 2012.
* Gu MO, Jeon MY, Eun Y. The development & effect of a tailored falls prevention exercise for older adults [Korean]. *Taehan Kanho Hakhoe Chi* 2006;**36**(2):341–52.

Hardin 2013 *{published and unpublished data}*

Hardin SR. RE: FW: "Inpatient Fall Prevention: Use of In-room Webcams" [personal communication]. Email to: S Dyer. 28 June 2016.

* Hardin SR, Dienemann J, Rudisill P, Mills KK. Inpatient fall prevention: use of in-room Webcams. *Journal of Patient Safety* 2013;**9**(1):29–35.

Harwood 2004 *{published data only}*

Harwood RH, Sahota O, Gaynor K, Masud T, Hosking DJ. A randomised, controlled comparison of different calcium and vitamin D supplementation regimens in elderly women after hip fracture: The Nottingham Neck of Femur (NONOF) Study. *Age and Ageing* 2004;**33**(1):45–51.

Hauer 2001 *{published data only}*

Hauer K, Pfisterer M, Schuler M, Bartsch P, Oster P. Two years later: a prospective long-term follow-up of a training intervention in geriatric patients with a history of severe falls. *Archives of Physical Medicine and Rehabilitation* 2003; **84**(10):1426–32.

* Hauer K, Rost B, Rutschle K, Opitz H, Specht N, Bartsch P, et al. Exercise training for rehabilitation and secondary prevention of falls in geriatric patients with a history of injurious falls. *Journal of the American Geriatrics Society* 2001;**49**(1):10–20.

Heiberg 2017 *{published data only}*

Heiberg KE, Bruun-Olsen V, Bergland A. The effects of habitual functional training on physical functioning in patients after hip fracture: the protocol of the HIPFRAC study. BMC Geriatrics 2017; Vol. 17, issue 1:23. 1471–2318]

Herrmann 2016 *{published data only}*

* Herrmann N, O'Regan J, Ruthirakuhan M, Kiss A, Eryavec G, Williams E, et al. A randomized placebo-controlled discontinuation study of cholinesterase inhibitors in institutionalized patients with moderate to severe Alzheimer disease. *Journal of the American Medical Directors Association* 2016;**17**(2):142–7. DOI: 10.1016/j.jamda.2015.08.019

Lanctot KL, Ruthirakuhan M, O'Regan JN, Eryavec G, Williams E, Black SE, et al. Cholinesterase inhibitor discontinuation in institutionalized persons with moderate to severe AD: Results of a double-blind, placebo controlled trial. *American Journal of Geriatric Psychiatry* 2015;**23**(3 Suppl):S172–3.

Hopman-Rock 1999 *{published data only}*

Hopman-Rock M, Staats PG, Tak EC, Droes R. The effects of a psychomotor activation programme for use in groups of cognitively impaired people in homes for the elderly. *International Journal of Geriatric Psychiatry* 1999; **14**:633–42.

Huang 2005 *{published data only}*

Huang TT, Liang SH. A randomized clinical trial of the effectiveness of a discharge planning intervention in hospitalized elders with hip fracture due to falling. *Journal of Clinical Nursing* 2005;**14**(10):1193–201.

Il'nitskii 2014 *{published data only}*

Il'nitskii AN, Proshchayev KI, Schwartsman GI, Bahmutova IuV, Pozdnyakova NM, Krivetskiy VV, et al. The use of piribedil for the prevention of falls in elderly patients with metabolic syndrome. [Russian]. *Klinicheskaia Meditsina* 2014;**92**(5):46–50.

Ilfeld 2010 *{published data only}*

Ilfeld BM, Loland VJ, Donovan JF, Le LT, Mariano ER. A multicenter, randomized, triple-masked, placebo-controlled trial of the effect of ambulatory continuous femoral nerve blocks on discharge-readiness following total knee arthroplasty in patients on general orthopaedic wards. *Regional Anesthesia and Pain MedicineConference:* 2010;**35th Annual Regional Anesthesia Meeting and Workshops, ASRA10 Toronto, ON Canada**(var.pagings): Conference–October.

Jarret 2015 *{published data only}*

Jarret G, Orpanna A, Helbostad J. Can a three weeks program in a rehabilitation center improve balance in elderly people? A randomized clinical controlled trial. *Physiotherapy (United Kingdom)* 2015;**101**(Suppl 1):eS671–2.

Jeon 2015 *{published data only}*

Jeon YH, Simpson JM, Li Z, Cunich MM, Thomas TH, Chenoweth L, et al. Cluster randomized controlled trial of an aged care specific leadership and management program to improve work environment, staff turnover, and care quality. *Journal of the American Medical Directors Association* 2015;**16**(7):629–e28.

JPRN-UMIN0000167 *{published data only}*

JPRN-UMIN000016716. Effect of fall prevention education programs for patients in convalescent phase after stroke: a randomized controlled trial. upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000019395 (first received 5 May 2015).

Kato 2006 *{published data only}*

Kato M, Izumi K, Hiramatsu T, Shogenji M. Development of an exercise program for fall prevention for elderly persons in a long-term care facility. *Japan Journal of Nursing Science* 2006;**3**(2):107–17.

Katz 2004 *{published data only}*

Katz IR, Jeste DV, Mintzer JE, Clyde C, Napolitano J, Brecher M. Comparison of risperidone and placebo for psychosis and behavioral disturbances associated with dementia: a randomized, double-blind trial. Risperidone Study Group. *Journal of Clinical Psychiatry* 1999;**60**(2): 107–15.

* Katz IR, Rupnow M, Kozma C, Schneider L. Risperidone and falls in ambulatory nursing home residents with dementia and psychosis or agitation: secondary analysis of a double-blind, placebo-controlled trial. *American Journal of Geriatric Psychiatry* 2004;**12**(5):499–508.

Katz 2005 *{published data only}*

Katz IR. Atypical antipsychotics and falls in the elderly [abstract]. 158th Annual Meeting of the American Psychiatric Association; 2005 May 21-26; Atlanta (GA).

Kenny 2001 *{published data only}*

Kenny RA, Richardson DA, Steen N, Bexton RS, Shaw FE, Bond J. Carotid sinus syndrome: a modifiable risk factor for nonaccidental falls in older adults (SAFE PACE). *Journal of the American College of Cardiology* 2001;**38**(5):1491–6.

Koczy 2011 *{published data only}*

Branitzki S, Koczy P. ReduFix - a study of reducing physical restraint: preventing risk of injury [ReduFix – Eine Studie zur Reduktion von körpernaher Fixierung: Heimbewohner vor Schaden bewahren]. *Pflege Zeitschrift* 2005;**58**(5): 310–3.

* Koczy P, Becker C, Rapp K, Klie T, Beische D, Buchele G, et al. Effectiveness of a multifactorial intervention to reduce physical restraints in nursing home residents. *Journal of the American Geriatrics Society* 2011;**59**(2):333–9.

Koczy P, Klie T, Kron M, Bredthauer D, Rissmann U, Branitzki S, et al. Effectiveness of a multifactorial intervention to reduce physical restraints in nursing home residents with dementia [Effektivität einer multifaktoriellen Intervention zur Reduktion von körpernaher Fixierung bei demenzerkrankten Heimbewohnern: Ziele und Studiendesign einer prospektiven clusterrandomisierten Interventionsstudie]. *Zeitschrift fur Gerontologie und Geriatrie* 2005;**38**(1):33–9.

Kopke 2012 *{published data only}*

Haut A, Kopke S, Gerlach A, Muhlhauser I, Haastert B, Meyer G. Evaluation of an evidence-based guidance on the reduction of physical restraints in nursing homes: a cluster-randomised controlled trial [ISRCTN34974819]. *BMC Geriatrics* 2009;**9**:42.

ISRCTN34974819. Evaluation of an evidence-based guidance on the reduction of physical restraints in nursing homes EBAGRAP. controlled-trials.com/ISRCTN34974819 (first received 29 April 2009).

* Kopke S, Muhlhauser I, Gerlach A, Haut A, Haastert B, Mohler R, et al. Effect of a guideline-based multicomponent intervention on use of physical restraints in nursing homes: a randomized controlled trial. *JAMA* 2012;**307**(20): 2177–84.

Kwok 2006 *{published data only}*

Kwok T, Mok F, Chien WT, Tam E. Does access to bed-chair pressure sensors reduce physical restraint use in the rehabilitative care setting?. *Journal of Clinical Nursing* 2006; **15**(5):581–7.

Lackner 2008 *{published data only}*

Lackner TE, Wyman JF, McCarthy TC, Monigold M, Davey C. Randomized, placebo-controlled trial of the cognitive effect, safety, and tolerability of oral extended-release oxybutynin in cognitively impaired nursing home residents with urge urinary incontinence. *Journal of the American Geriatrics Society* 2008;**56**:862–70.

Li 2017 *{published data only}*

Li CM, Chang CI, Yu WR, Yang W, Hsu CC, Chen CY. Enhancing elderly health examination effectiveness by adding physical function evaluations and interventions. Archives of Gerontology and Geriatrics 2017; Vol. 70: 38–43.

Lord 2003b *{published data only}*

Lord SR, Castell S, Corcoran J, Dayhew J, Matters B, Shan A, et al. The effect of group exercise on physical functioning and falls in frail older people living in retirement villages: a randomized, controlled trial. *Journal of the American Geriatrics Society* 2003;**51**(12):1685–92.

Mailhot 2012 *{published data only}*

Mailhot T, Cossette S, Van T. Nursing intervention to support self-efficacy of family caregivers to improve delirium management in cardiac surgery patients: A randomized study protocol. *Canadian Journal of CardiologyConference* 2012;**65th Annual Meeting of the Canadian Cardiovascular Society Toronto, ON Canada. Conference Start**(var.pagings):121027–October.

Mailhot 2014 *{published and unpublished data}*

Mailhot T. RE: MENTOR˙D trial [personal communication]. Email to: S Dyer. 20 April 2016.

* Mailhot T, Cossette S, Denault AY, Lamarche Y, Cote MC, Carbonneau MH, et al. Nursing intervention involving family caregiver to improve the management of post-cardiac surgery delirium: Results from a randomized pilot study. *Canadian Journal of Cardiology.* 2014;**67th Annual Meeting of the Canadian Cardiovascular Society Vancouver, BC Canada.**:141025.

Mailhot T, Cossette S, Van, Tassel J. Nursing intervention to support self-efficacy of family caregivers to improve delirium management in cardiac surgery patients: A randomized study protocol. *Canadian Journal of Cardiology.* 2012;**65th Annual Meeting of the Canadian Cardiovascular Society Toronto, ON Canada.**:121027–October.

Mak 2016 *{published data only}*

Mak JC, Klein LA, Finnegan T, Mason RS, Cameron ID. An initial loading-dose vitamin D versus placebo after hip fracture surgery: baseline characteristics of a randomized controlled trial (REVITAHIP). *BMC Geriatrics* 2014;**14**: 101. DOI: 10.1186/1471-2318-14-101

Mak JC, Mason R, Klein L, Cameron ID. Improving mobility and reducing disability in older people through early high-dose vitamin d replacement following hip fracture: a protocol for a randomized controlled trial and economic evaluation. *Geriatric Orthopaedic Surgery & Rehabilitation* 2011;**2**(3):94–9.

* Mak JC, Mason RS, Klein L, Cameron ID. An initial loading-dose vitamin D versus placebo after hip fracture surgery: randomized trial. *BMC Musculoskeletal Disorders* 2016;**17**:336. DOI: 10.1186/s12891-016-1174-9

Mak JCS, Cameron ID, Mason RS, Klein L, Soong M, Ohn K. Improving mobility and reducing disability in older people through early high-dose vitamin D replacement following hip fracture (the revitahip trial): Preliminary results. *Osteoporosis International* 2011;**22**(4 Suppl):S587.

Mansfield 2015 *{published data only}*

Mansfield A, Wong JS, Bryce J, Brunton K, Inness EL, Knorr S, et al. Use of accelerometer-based feedback of walking activity for appraising progress with walking-related goals in inpatient stroke rehabilitation: A randomized

controlled trial. *Neurorehabilitation and Neural Repair* 2015;**29**(9):847–57.

McRae 1996 ***{published data only}***
MacRae PG, Asplund LA, Schnelle JF, Ouslander JG, Abrahamse A, Morris C. A walking program for nursing home residents: effects on walk endurance, physical activity, mobility, and quality of life. *Journal of the American Geriatrics Society* 1996;**44**(2):175–80.

Mudge 2008 ***{published data only}***
Mudge AM, Giebel AJ, Cutler AJ. Exercising body and mind: an integrated approach to functional independence in hospitalized older people. *Journal of the American Geriatrics Society* 2008;**56**(4):630–5.

NCT00973297 ***{published data only}***
NCT00973297. A targeted falls prevention program in rehabilitation in-patients. clinicaltrials.gov/show/NCT00973297 (first received 25 August 2009).

NCT01054287 ***{published and unpublished data}***
NCT01054287. Falls prevention in acute care hospital PRECEPT. clinicaltrials.gov/show/NCT01054287 (first received 21 January 2010).
Waeber G. RE:NCT01054287 (2010). Falls Prevention in Acute Care Hospital [personal communication]. Email to: S Dyer. 1 August 2016.

NCT01523600 ***{published data only}***
NCT01523600. Whole body vibration training among older people using sheltered housing. ClinicalTrials.gov/show/NCT01523600 (first received 27 January 2012).

NCT01618786 ***{published data only}***
NCT01618786. Flooring for injury prevention trial (FLIP). clinicaltrials.gov/show/NCT01618786 (first received 11 June 2012).

NCT02686515 ***{published data only}***
Chang Gung University, Mackay Memorial Hospital. Comparative effectiveness research of dual-task and single-task balance training in people with stroke [NCT02686515]. ClinicalTrials.gov/show/NCT02686515 (first received 16 February 2016). NCT02686515]

Nyaruhirira 2013 ***{published data only}***
Nyaruhirira I, Di Romana S, Peretz A, Baillon R, Rozenberg S. Management of hip fracture patients in a network of public hospitals in Brussels II: Protocol of a randomized controlled trial, added value of coordinator. *Osteoporosis International.* 2013;**24**(1 Suppl):S370. Abstract P729]

Ouslander 2005 ***{published data only}***
Ouslander JG, Griffiths P, McConnell E, Riolo L, Schnelle J. Functional Incidental Training: applicability and feasibility in the Veterans Affairs nursing home patient population. *Journal of the American Medical Directors Association* 2005;**6**(2):121–7.
* Ouslander JG, Griffiths PC, McConnell E, Riolo L, Kutner M, Schnelle J. Functional incidental training: a randomized, controlled, crossover trial in Veterans Affairs nursing homes. *Journal of the American Geriatrics Society* 2005;**53**(7):1091–100.

Parasurum 2011 ***{published data only}***
Parasurum R, Chua PS, Kannusamy P. A randomised controlled study examining the impact of a staffing model and nursing care delivery system on patient, nurse and organisational outcomes. *Annals of the Academy of Medicine Singapore.Conference: Singapore Health and Biomedical Congress, SHBC* 2011;**40**(11 Suppl):S8. Abstract SG–NA–03]

Pedreira 2014 ***{published data only}***
Pedreira ÉMD. Use of virtual reality games for the treatment of balance and reducing the occurrence of falls in patients after stroke. ClinicalTrials.gov/show/NCT02475083 (first received 12 June 2015).

Peng 2014 ***{published data only}***
Peng L, Ren L, Qin P, Chen J, Feng P, Lin H, et al. Continuous femoral nerve block versus intravenous patient controlled analgesia for knee mobility and long-term pain in patients receiving total knee replacement: A randomized controlled trial. *Evidence-based Complementary and Alternative Medicine* 2014;**2014**:Article ID 569107. DOI: 10.1155/2014/569107

Peri 2008 ***{published data only}***
Peri K, Kerse N, Robinson E, Parsons M, Parsons J, Latham N. Does functionally based activity make a difference to health status and mobility? A randomised controlled trial in residential care facilities (The Promoting Independent Living Study; PILS). *Age and Ageing* 2008;**37**(1):57–63.

Rantz 2001 ***{published data only}***
Rantz MJ, Popejoy L, Petroski GF, Madsen RW, Mehr DR, Zwygart-Stauffacher M, et al. Randomized clinical trial of a quality improvement intervention in nursing homes. *Gerontologist* 2001;**41**(4):525–38.

Ray 2005 ***{published data only}***
Ray WA, Taylor JA, Brown AK, Gideon P, Hall K, Arbogast P, et al. Prevention of fall-related injuries in long-term care: a randomized controlled trial of staff education. *Archives of Internal Medicine* 2005;**165**(19):2293–8.

Reinhardt 2014 ***{published data only}***
Reinhardt KR, Duggal S, Umunna BP, Reinhardt GA, Nam D, Alexiades M, et al. Intraarticular analgesia versus epidural plus femoral nerve block after TKA: a randomized, double-blind trial. *Clinical Orthopaedics & Related Research* 2014;**472**(5):1400–8.

Resnick 2002 ***{published data only}***
Resnick B. Testing the effect of the WALC intervention on exercise adherence in older adults. *Journal of Gerontological Nursing* 2002;**28**(6):40–9.

Resnick 2012 ***{published data only}***
Resnick B, Galik E, Gruber-Baldini A, Zimmerman S. Testing the effect of function-focused care in assisted living. *Journal of the American Geriatrics Society* 2011;**59**(12):2233–40.
* Resnick B, Galik E, Gruber-Baldini AL, Zimmerman S. Falls and fall-related injuries associated with function-focused care. *Clinical Nursing Research* 2012;**21**(1):43–63.

Richter 2015 *{published data only}*
Richter C, Berg A, Fleischer S, Kopke S, Balzer K, Fick EM, et al. Effect of person-centred care on antipsychotic drug use in nursing homes (EPCentCare): study protocol for a cluster-randomised controlled trial. Implementation Science 2015; Vol. 10:82. 1748–5908]

Rolland 2007 *{published data only}*
Rolland Y, Pillard F, Klapouszczak A, Reynish E, Thomas D, Andrieu S, et al. Exercise program for nursing home residents with Alzheimer's disease: a 1-year randomized, controlled trial. *Journal of the American Geriatrics Society* 2007;**55**(2):158–65.

Sackley 2009 *{published data only}*
ISRCTN79859980. A randomised trial of an occupational therapy and physiotherapy intervention to enhance mobility and activity in a nursing or residential home setting after stroke. controlled-trials.com/ISRCTN79859980 (accessed 11 March 2012).
* Sackley CM, van den Berg ME, Lett K, Patel S, Hollands K, Wright CC, et al. Effects of a physiotherapy and occupational therapy intervention on mobility and activity in care home residents: a cluster randomised controlled trial. *BMJ* 2009;**339**:b3123.

Sahota 2014 *{published data only}*
ISRCTN44972300. REducing Falls in IN-patient Elderly: a randomised controlled trial. www.controlled-trials.com/ISRCTN44972300 (first received 17 August 2009).
* Sahota O, Drummond A, Kendrick D, Grainge MJ, Vass C, Sach T, et al. REFINE (REducing Falls in In-patieNt Elderly) using bed and bedside chair pressure sensors linked to radio-pagers in acute hospital care: a randomised controlled trial. *Age & Ageing* 2014;**43**(2):247–53.
Vass CD, Sahota O. REFINE - reducing falls in inpatient elderly using bed and bedside pressure sensors linked to radio-pagers in acute hospital care: a randomized controlled trial. *Clinical Rehabilitation* 2014;**28**(4):403.
Vass CD, Sahota O, Drummond A, Kendrick D, Gladman J, Sach T, et al. REFINE (Reducing Falls in In-patient Elderly)--a randomised controlled trial. *Trials [Electronic Resource]* 2009;**10**:Article number 83.
Vass CD, Sahota O, Drummond A, Kendrick D, Grainge M, Gladman J, et al. Refine-reducing falls in in-patient elderly using bed and chair pressure sensors in acute hospital care: A randomised controlled trial. *Age and Ageing* 2012; **41**:Suppl 1.

Said 2012 *{published data only}*
Said CM, Morris ME, Woodward M, Churilov L, Bernhardt J. Enhancing physical activity in older adults receiving hospital based rehabilitation: a phase II feasibility study. *BMC Geriatrics* 2012;**12**:26.

Said 2015 *{published data only}*
Said CM, Morris ME, McGinley JL, Szoeke C, Workman B, Liew D, et al. Evaluating the effects of increasing physical activity to optimize rehabilitation outcomes in hospitalized older adults (MOVE Trial): study protocol for a randomized controlled trial. *Trials* 2015;**16**:13. PUBMED: 25588907]

Sato 2000 *{published data only}*
Sato Y, Asoh T, Kaji M, Oizumi K. Beneficial effect of intermittent cyclical etidronate therapy in hemiplegic patients following an acute stroke. *Journal of Bone and Mineral Research* 2000;**15**(12):2487–94.

Sato 2005a *{published data only}*
Sato Y, Iwamoto J, Kanoko T, Satoh K. Low-dose vitamin D prevents muscular atrophy and reduces falls and hip fractures in women after stroke: a randomized controlled trial. *Cerebrovascular Diseases* 2005;**20**:187–92.

Sato 2005b *{published data only}*
* Sato Y, Honda Y, Iwamoto J, Kanoko T, Satoh K. Effect of folate and mecobalamin on hip fractures in patients with stroke: a randomized controlled trial. *JAMA* 2005;**293**(9): 1082–8.
Sato Y, Honda Y, Iwamoto J, Kanoko T, Satoh K. Inaccurate description of collaborating hospitals in a study of the effect of folate and mecobalamin on hip fractures after stroke. *JAMA* 2006;**296**(4):396.

Sato 2011 *{published data only}*
Sato Y, Iwamoto J, Honda Y. An open-label trial comparing alendronate and alphacalcidol in reducing falls and hip fractures in disabled stroke patients. *Journal of Stroke and Cerebrovascular Diseases* 2011;**20**(1):41–6.

Schneider 2006 *{published data only}*
* Schneider LS, Tariot PN, Dagerman KS, Davis SM, Hsiao JK, Ismail MS, et al. Effectiveness of atypical antipsychotic drugs in patients with Alzheimer's disease. *New England Journal of Medicine* 2006;**355**(15):1525–38.
Tariot PN. Efficacy and tolerability of atypical antipsychotics in agitation and psychosis: Research results [abstract]. 158th Annual Meeting of the American Psychiatric Association; 2005 May 21-26; Atlanta (GA) 31A.
Tariot PN, Schneider L, Katz IR, Mintzer JE, Street J, Copenhaver M, et al. Quetiapine treatment of psychosis associated with dementia: A double-blind, randomized, placebo-controlled clinical trial. *American Journal of Geriatric Psychiatry* 2006;**14**(9):767–76.
Zhong KX, Tariot PN, Mintzer J, Minkwitz MC, Devine NA. Quetiapine to treat agitation in dementia: a randomized, double-blind, placebo-controlled study. *Current Alzheimer Research* 2007;**4**(1):81–93.

Schwendimann 2006 *{published data only}*
Schwendimann R. personal communication 22 April 2005.
Schwendimann R, Buhler H, De Geest S, Milisen K. Falls and consequent injuries in hospitalized patients: effects of an interdisciplinary falls prevention program. *BMC Health Services Research* 2006;**6**:69.
* Schwendimann R, Milisen K, Buhler H, De Geest S. Fall prevention in a Swiss acute care hospital setting Reducing multiple falls. *Journal of Gerontological Nursing* 2006;**32**(3): 13–22.

Sherrington 2016a *{published data only}*
Sherrington C, Fairhall N, Kirkham C, Clemson L, Howard K, Vogler C, et al. Exercise and fall prevention self-management to reduce mobility-related disability and

falls after fall-related lower limb fracture in older people: protocol for the RESTORE (Recovery Exercises and STepping On afteR fracturE) randomised controlled trial. *BMC Geriatrics* 2016;**16**(1):34.

Shimada 2003 ***{published data only}***
Shimada H, Uchiyama Y, Kakurai S. Specific effects of balance and gait exercises on physical function among the frail elderly. *Clinical Rehabilitation* 2003;**17**(5):472–9.

Shimada 2009 ***{published data only}***
Shimada H, Tiedemann A, Lord S, Suzuki T. The effect of enhanced supervision on fall rates in residential aged care. *American Journal of Physical Medicine and Rehabilitation* 2009;**88**(10):823–8.

Siddiqi 2016 ***{published data only}***
Heaven, A, Cheater F, Clegg A, Collinson M, Farrin A, Forster A, et al. Pilot trial of Stop Delirium! (PiTStop)--a complex intervention to prevent delirium in care homes for older people: study protocol for a cluster randomised controlled trial. *Trials [Electronic Resource]* 2014;**15**:47.
* Siddiqi N, Cheater F, Collinson M, Farrin A, Forster A, George D, et al. The PiTSTOP study: a feasibility cluster randomized trial of delirium prevention in care homes for older people. *Age and Ageing* 2016;**45**(5):652–61. DOI: 10.1093/ageing/afw091

Sjoberg 2013 ***{published data only}***
Sjoberg C. SV: Effects of medication reviews performed by a physician on treatment with fracture-preventing and fall-risk-increasing drugs in older adults with hip fracture-a randomized controlled study [personal communication]. Email to: S Dyer. 19 September 2016.
* Sjoberg C, Wallerstedt SM. Effects of medication reviews performed by a physician on treatment with fracture-preventing and fall-risk-increasing drugs in older adults with hip fracture-a randomized controlled study. *Journal of the American Geriatrics Society* 2013;**61**(9):1464–72.

Smith 2017 ***{published data only}***
Smith T, Clark A, Dodd E, Khoo ME, Heneker S, Cross J, et al. Feasibility study suggests no impact from protected engagement time on adverse events in mental health wards for older adults. International Journal of Mental Health Nursing 2017 Jul 6 Epub ahead of print]. DOI: 10.1111/inm.12362

Sola 2014 ***{published data only}***
Sola SM, Lopez del Amo JL, Valero O. The effect of 24 weeks of moderate-to-high intensity strength training on the elderly [Spanish]. *Revista Espanola de Geriatria y Gerontologia* 2014;**49**(3):115–20.

Southard 2006 ***{published data only}***
Southard V. A randomized control trial of the application of efficacy training to balance assessment. *Physical and Occupational Therapy in Geriatrics* 2006;**25**(2):51–66.

Steadman 2003 ***{published data only}***
Steadman J, Donaldson N, Kalra L. A randomized controlled trial of an enhanced balance training program to improve mobility and reduce falls in elderly patients. *Journal of the American Geriatrics Society* 2003;**51**(6):847–52.

Tanikawa 2014 ***{published data only}***
Tanikawa H, Sato T, Nagafuchi M, Takeda K, Oshida J, Okuma K. Comparison of local infiltration of analgesia and sciatic nerve block in addition to femoral nerve block for total knee arthroplasty. *Journal of Arthroplasty* 2014;**29**(12):2462–7.

Tariot 2004 ***{published data only}***
* Tariot PN, Farlow MR, Grossberg GT, Graham SM, McDonald S, Gergel I. Memantine treatment in patients with moderate to severe Alzheimer disease already receiving donepezil: a randomized controlled trial. *JAMA* 2004;**291**(3):317.
Van Dyck CH, Tariot PN, Meyers B, Malca Resnick E. A 24-week randomized, controlled trial of memantine in patients with moderate-to-severe Alzheimer disease. *Alzheimer Disease and Associated Disorders* 2007;**21**(2):136–43.

Tariot 2005 ***{published data only}***
Profenno LA, Jakimovich L, Holt CJ, Porsteinsson A, Tariot PN. A randomized, double-blind, placebo-controlled pilot trial of safety and tolerability of two doses of divalproex sodium in outpatients with probable Alzheimer's disease. *Current Alzheimer Research* 2005;**2**(5):553–8.
* Tariot PN, Raman R, Jakimovich L, Schneider L, Porsteinsson A, Thomas R, et al. Divalproex sodium in nursing home residents with possible or probable Alzheimer disease complicated by agitation: A randomized, controlled trial. *American Journal of Geriatric Psychiatry* 2005;**13**(11):942–9.

Teresi 2013 ***{published data only}***
Teresi JA, Ramirez M, Remler D, Ellis J, Boratgis G, Silver S, et al. Comparative effectiveness of implementing evidence-based education and best practices in nursing homes: Effects on falls, quality-of-life and societal costs. *International Journal of Nursing Studies.* 2013;**50**(4):448–63.

Underwood 2011 ***{published data only}***
Ellard DR, Taylor SJ, Parsons S, Thorogood M. The OPERA trial: a protocol for the process evaluation of a randomised trial of an exercise intervention for older people in residential and nursing accommodation. *Trials [Electronic Resource]* 2011;**12**:Article number 28.
ISRCTN43769277. Older people's exercise intervention in residential and nursing accommodation. www.controlled-trials.com/ISRCTN43769277 (accessed 11 March 2012).
* Underwood M, Eldridge S, Lamb S, Potter R, Sheehan B, Slowther AM, et al. The OPERA trial: protocol for a randomised trial of an exercise intervention for older people in residential and nursing accommodation. *Trials [Electronic Resource]* 2011;**12**:Article number 27.

van Ooijen 2013 ***{published and unpublished data}***
van Ooijen KMV. RE: C-mill gait-adaptability training following hip fracture study [personal communication]. Email to: S Dyer. 6 May 2016.
* van Ooijen MW, Roerdink M, Trekop M, Visschedijk J, Janssen TW, Beek PJ. Functional gait rehabilitation in elderly people following a fall-related hip fracture using

a treadmill with visual context: design of a randomized controlled trial. *BMC Geriatrics* 2013;**13**:34.

Vassallo 2004 *{published data only}*

Vassallo M, Vignaraja R, Sharma JC, Briggs RS, Allen SC. Can intervention prevent falls and injury in geriatric wards? Hospital injury prevention (HIP) study [abstract]. *Age and Ageing* 2001;**30**(Suppl 2):15.

* Vassallo M, Vignaraja R, Sharma JC, Hallam H, Binns K, Briggs R, et al. The effect of changing practice on fall prevention in a rehabilitative hospital: the Hospital Injury Prevention Study. *Journal of the American Geriatrics Society* 2004;**52**(3):335–9.

Visvanathan 2015 *{published data only}*

ACTRN12617000981325. Effectiveness of an Ambient Intelligence Geriatric Management system (AmbIGeM) to prevent falls in older people in hospitals: the AmbIGeM stepped wedge pragmatic trial. http://www.anzctr.org.au/ACTRN12617000981325.aspx (first received 7 July 2017).

Von Koch 2001 *{published data only}*

Thorsen AM, Widen Holmqvist L, de Pedro-Cuesta J, von Koch L. A randomized controlled trial of early supported discharge and continued rehabilitation at home after stroke: five-year follow-up of patient outcome. *Stroke* 2005;**36**(2): 297–303.

Thorsen AM, Widen Holmqvist L, von Koch L. Early supported discharge and continued rehabilitation at home after stroke: 5-year follow-up of resource use. *Journal of Stroke and Cerebrovascular Diseases* 2006;**15**(4):139–43.

Widen Holmqvist L, von Koch L, Kostulas V, Holm M, Widsell G, Tegler H, et al. A randomized controlled trial of rehabilitation at home after stroke in southwest Stockholm. *Stroke* 1998;**29**:591–7.

* von Koch L, Pedro-Cuesta J, Kostulas V, Almazan J, Widen Holmqvist L. Randomized controlled trial of rehabilitation at home after stroke: one-year follow-up of patient outcome, resource use and cost. *Cerebrovascular Disease* 2001;**12**(2):131–8.

von Koch L, Widen Holmqvist L, Kostulas V, Almazan J, de Pedro-Cuesta J. A randomized controlled trial of rehabilitation at home after stroke in Southwest Stockholm: outcome at six months. *Scandinavian Journal of Rehabilitation Medicine* 2000;**32**(2):80–6.

Wolf 2003 *{published data only}*

Sattin RW, Easley KA, Wolf SL, Chen Y, Kutner MH. Reduction in fear of falling through intense tai chi exercise training in older, transitionally frail adults. *Journal of the American Geriatrics Society* 2005;**53**(7):1168–78.

Wolf SL, O'Grady M. The influence of intense Tai Chi training on physical performance and hemodynamic outcomes in transitionally frail, older adults. *Journals of Gerontology Series A-Biological Sciences and Medical Sciences* 2006;**61**(2):184–9.

* Wolf SL, Sattin RW, Kutner M, O'Grady M, Greenspan AI, Gregor RJ. Intense Tai Chi exercise training and falls occurrences in older, transitionally frail adults: A randomized controlled trial. *Journal of the American Geriatrics Society* 2003;**51**:1693–1701.

Wolf SL, Sattin RW, O'Grady M, Freret N, Ricci L, Greenspan AI, et al. A study design to investigate the effect of intense Tai Chi in reducing falls among older adults transitioning to frailty. *Controlled Clinical Trials* 2001;**22** (6):689–704.

Zhong 2007 *{published data only}*

Zhong KX, Tariot PN, Mintzer J, Minkwitz MC, Devine NA. Quetiapine to treat agitation in dementia: a randomized, double blind, placebo-controlled study. *Current Alzheimers Research* 2007;**4**(1):81–93.

References to studies awaiting assessment

Dever 2016 *{published data only}*

Dever Fitzgerald T, Hadjistavropoulos T, Williams J, Lix L, Zahir S, Alfano D, et al. The impact of fall risk assessment on nurse fears, patient falls, and functional ability in long-term care. Disability and Rehabilitation 2016; Vol. 38, issue 11:1041–52.

Frohnhofen 2013 *{published data only (unpublished sought but not used)}*

Frohnhofen H, Schlitzer J, Wehling M. Pharmacotherapy and frequency of inhospital falls among geriatric patients. *Osteoporosis International.Conference: European Congress on Osteoporosis and Osteoarthritis, ESCEO13-IOF Rome Italy.Conference Start:* 2013;**24**(1 Suppl):P590.

Hewitt 2014 *{published data only}*

* Hewitt J, Goodall S, Clemson L, Henwood T, Refshauge K. Progressive resistance and balance training for falls Prevention in long-term residential aged care: A cluster randomized trial of the SUNBEAM program. *Journal of the American Medical Directors Association* 2018;**19**(4):361–9. PUBMED: 29402651]

Hewitt J, Refshauge KM, Goodall S, Henwood T, Clemson L. Does progressive resistance and balance exercise reduce falls in residential aged care? Randomized controlled trial protocol for the SUNBEAM program. *Clinical Interventions In Aging* 2014;**9**:369–76.

MacRitchie 2001 *{published data only}*

MacRitchie RF. *Reducing the incidence of falls among elderly nursing home residents: An evaluation of an ameliorative pilot program [thesis]*. Dublin, USA: Southern Connecticut State Univ, 2001. CENTRAL: CN–00691333]

Raymond 2017 *{published data only}*

Raymond MJ, Jeffs KJ, Winter A, Soh SE, Hunter P, Holland AE. The effects of a high-intensity functional exercise group on clinical outcomes in hospitalised older adults: an assessor-blinded, randomised-controlled trial. Age and Ageing 2017; Vol. 46, issue 2:208–13.

Tallon 2013 *{published data only}*

Tallon G, Ramdani S, Jaussent A, Decker L, Bernard PL, Blain H. Effect of whole-body-vibration training in institutionalized older adults. *European Geriatric Medicine* 2013;**4**(Suppl 1):S55.

Van der Linden 2017 *{published data only}*
Van Der Linden L. Reduction of inappropriate prescribing in older persons using the RASP list: A cluster-randomised controlled trial. European Journal of Hospital Pharmacy 2014; Vol. 21:A212.
* Van der Linden L, Decoutere L, Walgraeve K, Milisen K, Flamaing J, Spriet I, et al. combined use of the rationalization of home medication by an adjusted STOPP in older patients (RASP) list and a pharmacist-led medication review in very old inpatients: impact on quality of prescribing and clinical outcome. Drugs and Aging 2017; Vol. 34, issue 2:123–33.

Wylie 2017 *{published data only}*
NCT02178527. Podiatry Intervention to Reduce Falls in Elderly Care Trial (PIRFECT). ClinicalTrials.gov/show/NCT02178527 (first received 17 June 2014). N. H.S.Tayside; University of Stirling; La Trobe University; University of Toronto; Uniersity of Dundee]
* Wylie G, Menz HB, McFarlane S, Ogston S, Sullivan F, Williams B, et al. Podiatry intervention versus usual care to prevent falls in care homes: pilot randomised controlled trial (the PIRFECT study). BMC Geriatrics 2017; Vol. 17, issue 1:143.

References to ongoing studies

ACTRN12613000228785 *{published data only}*
ACTRN12613000228785. Preventing falls and fractures in low-level aged-care residents by increasing dairy food intake by two serves per day. www.anzctr.org.au/ACTRN12613000228785.aspx (first received 20 February 2013).

ACTRN12615000817549 *{published data only}*
ACTRN12615000817549. Establishing the effectiveness, cost-effectiveness and student experience of simulation training for the prevention of falls amongst hospitalised inpatients. www.anzctr.org.au/ACTRN12615000817549.aspx (first received 27 July 2015).
* Williams C, Bowles KA, Kiegaldie D, Maloney S, Nestel D, Kaplonyi J, et al. Establishing the effectiveness, cost-effectiveness and student experience of a Simulation-based education Training program On the Prevention of Falls (STOP-Falls) among hospitalised inpatients: A protocol for a randomised controlled trial. BMJ Open 2016; Vol. 6, issue 6:e010192. DOI: 10.1136/bmjopen-2015-010192.

ACTRN12617000314325 *{published data only}*
ACTRN12617000314325. Does abbreviating patient falls risk screening in documentation impact on falls in hospital inpatients: A stepped wedge cluster randomised control trial. https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12617000314325 (first received 21 February 2017).

Dal Bello-Haas 2012 *{published data only}*
Dal Bello-Haas VP, Thorpe LU, Lix LM, Scudds R, Hadjistavropoulos T. The effects of a long-term care walking program on balance, falls and well-being. *BMC Geriatrics* 2012;**12**:76.

Hassett 2016 *{published data only}*
Hassett L, Van Den Berg M, Lindley RI, Crotty M, McCluskey A, Van Der Ploeg HP, et al. Effect of affordable technology on physical activity levels and mobility outcomes in rehabilitation: A protocol for the Activity and MObility UsiNg Technology (AMOUNT) rehabilitation trial. BMJ Open 2016; Vol. 6, issue 6:e012074.

ISRCTN34353836 *{published data only}*
ISRCTN34353836. Finch: Falls in care homes study. http://www.isrctn.com/ISRCTN34353836 (first received 22 March 2016).

ISRCTN42003273 *{published data only}*
Loffler C, Drewelow E, Paschka SD, Frankenstein M, Eger J, Jatsch L, et al. Optimizing polypharmacy among elderly hospital patients with chronic diseases--study protocol of the cluster randomized controlled POLITE-RCT trial. *Implementation Science* 2014;**9**:151.

JPRN-UMIN000000555 *{published data only}*
JPRN-UMIN000000555. The effects of whole body vibration for the prevention of falls in elderly. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000000555 (first received 25 December 2006).

JPRN-UMIN000008361 *{published data only}*
JPRN-UMIN000008361. Multicenter, randomized, double-blind, placebo controlled, parallel group trial to evaluate the effect of Vitamin D supplementation for fall prevention. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000008361 (first received 6 July 2012).

McCullagh 2016 *{published data only}*
McCullagh R, O'Connell E, O'Meara S, Perry I, Fitzgerald A, O'Connor K, et al. A study protocol of a randomised controlled trial to measure the effects of an augmented prescribed exercise programme (APEP) for frail older medical patients in the acute setting. BMC Geriatrics 2016; Vol. 16:79.

Mestres 2017 *{published data only}*
Mestres Gonzalvo C, de Wit HA, van Oijen BP, Hurkens KP, Janknegt R, Schols JM, et al. Supporting clinical rules engine in the adjustment of medication (SCREAM): protocol of a multicentre, prospective, randomised study. BMC Geriatrics 2017; Vol. 17, issue 1:35.

Mudge 2017 *{published data only}*
Mudge AM, Banks MD, Barnett AG, Blackberry I, Graves N, Green T, et al. CHERISH (collaboration for hospitalised elders reducing the impact of stays in hospital): protocol for a multi-site improvement program to reduce geriatric syndromes in older inpatients. BMC Geriatrics 2017; Vol. 17, issue 1:11.

NCT00636675 *{published data only}*
Anderson RA, Corazzini K, Porter K, Daily K, McDaniel RR, Colón-Emeric C. CONNECT for quality: protocol of a cluster randomized controlled trial to improve fall

prevention in nursing homes. *Implementation Science* 2012; **7**(1):11. DOI: 10.1186/1748-5908-7-11
NCT00636675. CONNECT for Quality: A Study to Reduce Falls in Nursing Homes (CONNECT). clinicaltrials.gov/ct2/show/NCT00636675 (first received 14 March 2008).

NCT01483456 ***{published and unpublished data}***
* Grangier C, Mouchoux C, Le Pogam M, Toulouze B, Colin C, Krolak-Salmon P. A multidisciplinary program for preventing falls "identify, prevent and get up": Impact on falls in elderly inpatients. European Geriatric Medicine. 7th Congress of the European Union Geriatric Society (EUGMS); 2011 Sept 28-30; Malaga, Spain. 2011. Conference abstract 201200]
Mouchoux C. NCT01483456 "Impact of multidisciplinary program on falls in elderly inpatients (IPR)" [personal communication]. Email to: S Dyer. 29 June 2016.
NCT01483456. Impact of multidisciplinary program on falls in elderly inpatients (IPR). clinicaltrials.gov/show/NCT01483456 (first received 21 July 2007).

NCT01551121 ***{published and unpublished data}***
NCT01551121. Assessment of an automated telesurveillance system on the incidence of serious falls in nursing homes (TELEHPAD). clinicaltrials.gov/show/NCT01551121 (first received 12 March 2012).

NCT01561872 ***{published data only}***
NCT01561872. Assessment of an automated telesurveillance system on serious falls prevention in an elderly suffering from dementia specialized care unit: the URCC (GET-BETTER). clinicaltrials.gov/show/NCT01561872 (first received 21 March 2012).
* Saulnier I, Lachal F, Tchalla A, Trimouillas J, Gourdeau-Nauche F, Bernard-Bourzeix L, et al. Assessment of an automated tele vigilance system on serious falls prevention in a dementia specialized care unit: The URCC. *Journal of Nutrition, Health and Aging* 2012;**16**(9):865. Abstract P80]

NCT01735682 ***{published data only}***
NCT01735682. Whole body vibration exercise training for institutionalized elderly. ClinicalTrials.gov/show/NCT01735682 (first received 20 November 2012). The Hong Kong Polytechnic University, Shatin Hospital Hong Kong]

NCT01876095 ***{published data only}***
NCT01876095. Discontinuing inappropriate medication in nursing home residents. ClinicalTrials.gov/show/NCT01876095 (first received 10 June 2013).
* Wouters H, Quik EH, Boersma F, Nygard P, Bosman J, Bottger WM, et al. Discontinuing inappropriate medication in nursing home residents (DIM-NHR Study): protocol of a cluster randomised controlled trial. *BMJ Open* 2014;**4**(10):e006082.

NCT02295462 ***{published data only}***
NCT02295462. Effect of person-centred-care on antipsychotic drug use in nursing homes: a cluster-randomised trial. ClinicalTrials.gov/show/NCT02295462 (first received 12 November 2014).

NCT02570945 ***{published data only}***
NCT02570945. Trial of a pharmacist-physician intervention model to reduce high-risk drug use by hospitalised elderly patients. ClinicalTrials.gov/show/NCT02570945 (first received 25 September 2015).

NCT02604056 ***{published data only}***
NCT02604056. Pragmatic cluster trial for nursing home antipsychotic prescribing. ClinicalTrials.gov/show/NCT02604056 (first received 5 November 2015).

NCT02702037 ***{published data only}***
NCT02702037. Older person's exercise and nutrition study. ClinicalTrials.gov/show/NCT02702037 (first received 26 February 2016).

NCT02714257 ***{published data only}***
NCT02714257. Seniors avoiding falls through exercise study. https://ClinicalTrials.gov/show/NCT02714257 (first received 22 February 2016).

NCT02714582 ***{published data only}***
NCT02714582. Feasibility, appropriateness, meaningfulness and effectiveness of bedside shift reporting. ClinicalTrials.gov/show/NCT02714582 (first received 29 February 2016).

NCT02757131 ***{published data only}***
NCT02757131. Dedicated ambulator-assisted physical activity to improve hospital outcome measures in elderly patients. https://clinicaltrials.gov/show/NCT02757131 (first received 29 April 2016).

NCT02969343 ***{published data only}***
NCT02969343. Patient safety learning laboratory: making acute care more patient-centered. https://clinicaltrials.gov/show/NCT02969343 (first received 21 November 2016).

NCT03014570 ***{published data only}***
NCT03014570. Testing implementation of EIT-4-BPSD. https://clinicaltrials.gov/show/NCT03014570 (first received 9 January 2017).

NCT03019211 ***{published data only}***
NCT03019211. Feasibility aquatic physical exercise to reduce falls in institutionalized elderly. https://clinicaltrials.gov/show/NCT03019211 (first received 12 January 2017).

NCT03192384 ***{published data only}***
NCT03192384. A service intervention to reduce falls in hospital. https://clinicaltrials.gov/show/NCT03192384 (first received 20 June 2017).

NTR5015 ***{published data only}***
NTR5015. Prevention of falling. www.trialregister.nl/trialreg/admin/rctview.asp?TC=5015 (first received 22 December 2014).

Scheffers-Barnhoorn 2017 ***{published data only}***
NTR5695. A randomised controlled trial to improve outcomes of hip fracture patients with fear of falling in geriatric rehabilitation. http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=5695 (first received 7 March 2016).
* Scheffers-Barnhoorn MN, van Haastregt JC, Schols JM, Kempen GI, van Balen R, Visschedijk JH, et al. A multi-

component cognitive behavioural intervention for the treatment of fear of falling after hip fracture (FIT-HIP): protocol of a randomised controlled trial. BMC Geriatrics 2017; Vol. 17, issue 1:71.

Additional references

Avenell 2014
Avenell A, Mak JC, O'Connell D. Vitamin D and vitamin D analogues for preventing fractures in post-menopausal women and older men. *Cochrane Database of Systematic Reviews* 2014, Issue 4. DOI: 10.1002/14651858.CD000227.pub4

Becker 2010
Becker C, Rapp K. Fall prevention in nursing homes. [Review]. *Clinics In Geriatric Medicine* 2010;**26**(4): 693–704.

Becker 2012
Rapp K, Becker C, Cameron ID, Konig HH, Buchele G. Epidemiology of falls in residential aged care: analysis of more than 70,000 falls from residents of Bavarian nursing homes. *Journal of the American Medical Directors Association* 2012;**13**(2):187.e1–6.

Bolland 2010
Bolland MJ, Avenell A, Baron JA, Grey A, MacLennan GS, Gamble GD, et al. Effect of calcium supplements on risk of myocardial infarction and cardiovascular events: meta-analysis. *BMJ* 2010;**341**:c3691. [PUBMED: 20671013]

Bolland 2014
Bolland MJ, Grey A, Gamble GD, Reid IR. Vitamin D supplementation and falls: a trial sequential meta-analysis. *Lancet. Diabetes and Endocrinology* 2014;**2**(7):573–80. [PUBMED: 24768505]

Bolland 2016
Bolland MJ, Avenell A Gamble GD, Grey A. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. *Neurology* 2016;**87**(23): 2391–402.

Boutron 2008
Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine* 2008;**148**(4):295–309. MEDLINE: 8283207

Campbell 2004
Campbell MK, Elbourne DR, Altman DG, CONSORT group. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;**328**(7441):702–8. MEDLINE: 15031246

Chan 2015
Chan WC, Yeung JW, Wong CS, Lam LC, Chung KF, Luk JK, et al. Efficacy of physical exercise in preventing falls in older adults with cognitive impairment: a systematic review and meta-analysis. *Journal of the American Medical Directors Association* 2015;**16**(2):149–54.

Choi 2012
Choi M, Hector M. Effectiveness of intervention programs in preventing falls: a systematic review of recent 10 years and meta-analysis. *Journal of the American Medical Directors Association* 2012;**13**(2):188.e13–21. [PUBMED: 21680249]

Conroy 2010
Conroy S, Kendrick D, Harwood R, Gladman J, Coupland C, Sach T, et al. A multicentre randomised controlled trial of day hospital-based falls prevention programme for a screened population of community-dwelling older people at high risk of falls. *Age and Ageing* 2010;**39**(6):704–10.

Covidence [Computer program]
Veritas Health Innovation. Covidence systematic review software. Melbourne, Australia: Veritas Health Innovation, Accessed 3 March 2016.

Davis 2011
Davis JC, Robertson MC, Comans T, Scuffham PA. Guidelines for conducting and reporting economic evaluation of fall prevention strategies. *Osteoporosis International* 2011;**22**(9):2449–59.

De Spiegeleer 2018
De Spiegeleer A, Beckwée D, Bautmans I, Petrovic M. Pharmacological interventions to improve muscle mass, muscle strength and physical performance in older people: an umbrella review of systematic reviews and meta-analyses. *Drugs & Aging* 2018;**35**(8):719–34.

Deandrea 2013
Deandrea S, Bravi F, Turati F, Lucenteforte E, La Vecchia C, Negri E. Risk factors for falls in older people in nursing homes and hospitals. A systematic review and meta-analysis. *Archives of Gerontology and Geriatrics* 2013;**56**(3):407–15. [PUBMED: 23294998]

Durvasula 2012
Durvasula S, Sambrook PN, Cameron ID. Factors influencing adherence with therapeutic sunlight exposure in older people in intermediate care facilities. *Archives of Gerontology and Geriatrics* 2012;**54**(2):e234–41.

Excel [Computer program]
Microsoft. Excel X for Mac. Version 8. Microsoft, 2001.

Gillespie 2003
Gillespie LD, Gillespie WJ, Robertson MC, Lamb SE, Cumming RG, Rowe BH. Interventions for preventing falls in elderly people. *Cochrane Database of Systematic Reviews* 2003, Issue 4. DOI: 10.1002/14651858.CD000340.pub2

Gillespie 2012
Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, Lamb SE. Interventions for preventing falls in older people living in the community. *Cochrane Database of Systematic Reviews* 2012, Issue 9. DOI: 10.1002/14651858.CD007146.pub3

Guo 2014
Guo JL, Tsai YY, Liao JY, Tu HM, Huang CM. Interventions to reduce the number of falls among older adults with/without cognitive impairment: an exploratory meta-

analysis. *International Journal of Geriatric Psychiatry* 2014; **29**(7):661–9. [PUBMED: 24318959]

Guyatt 2011

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *Journal of Clinical Epidemiology* 2011;**64**(12):1283–93. [PUBMED: 21839614]

Guyatt 2011a

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *Journal of Clinical Epidemiology* 2011;**64**(12):1294–302. [PUBMED: 21803546]

Hahn 2005

Hahn S, Puffer S, Torgerson DJ, Watson J. Methodological bias in cluster randomised trials. *BMC Medical Research Methodology* 2005;**5**:10. [PUBMED: 15743523]

Haines 2009

Haines TP, Russell T, Brauer SG, Erwin S, Lane P, Urry S, et al. Effectiveness of a video-based exercise programme to reduce falls and improve health-related quality of life among older adults discharged from hospital: a pilot randomized controlled trial. *Clinical Rehabilitation* 2009;**23**(11): 973–85.

Haines 2013

Haines TP, Hill AM, Hill KD, Brauer SG, Hoffmann T, Etherton-Beer C, et al. Cost effectiveness of patient education for the prevention of falls in hospital: economic evaluation from a randomized controlled trial. *BMC Medicine* 2013;**11**:135.

Hannan 2010

Hannan MT, Gagnon MM, Aneja J, Jones RN, Cupples LA, Lipsitz LA, et al. Optimizing the tracking of falls in studies of older participants: comparison of quarterly telephone recall with monthly falls calendars in the MOBILIZE Boston Study. *American Journal of Epidemiology* 2010;**171** (9):1031–6. [PUBMED: 20360242]

Hauer 2006

Hauer K, Lamb SE, Jorstad EC, Todd C, Becker C, ProFaNE-Group. Systematic review of definitions and methods of measuring falls in randomised controlled fall prevention trials. *Age and Ageing* 2006;**35**(1):5–10.

Hendriks 2008

Hendriks MR, Evers SM, Bleijlevens MH, Van Haastregt JC, Crebolder HF, Van Eijk JT. Cost-effectiveness of a multidisciplinary fall prevention program in community-dwelling elderly people: A randomized controlled trial (ISRCTN 64716113). *International Journal of Technology Assessment in Health Care* 2008;**24**(2):193–202.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327** (7414):557–60.

Higgins 2011a

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011b

Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8.5: The Cochrane Collaboration's tool for assessing risk of bias. In: Higgins JPT, Green S (editors). Cochrane Handbook of Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011c

Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16.3.4: Approximate analyses of cluster-randomized trials for meta-analysis: effective sample sizes. In: Higgins JPT, Green S (editors). Cochrane Handbook of Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011d

Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9.6.3.1 Is the effect different in different subgroups? In: Higgins JPT, Green S (editors). Cochrane Handbook of Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Hill 2010

Hill AM, Hoffmann T, Hill K, Oliver D, Beer C, McPhail S, et al. Measuring falls events in acute hospitals-a comparison of three reporting methods to identify missing data in the hospital reporting system. *Journal of the American Geriatrics Society* 2010;**58**(7):1347–52. [PUBMED: 20487077]

Hill 2014

Hill, AM, Waldron N, Etherton-Beer C, McPhail SM, Ingram K, Flicker L, et al. A stepped-wedge cluster randomised controlled trial for evaluating rates of falls among inpatients in aged care rehabilitation units receiving tailored multimedia education in addition to usual care: a trial protocol. *BMJ Open* 2014;**4**(1):e004195.

Ibrahim 2015

Ibrahim JE, Murphy BJ, Bugeja L, Ranson D. Nature and extent of external-cause deaths of nursing home residents in Victoria, Australia. *Journal of the American Geriatrics Society* 2015;**63**(5):954–62. [PUBMED: 25940003]

Jensen 2003

Jensen J, Nyberg L, Gustafson Y, Lundin-Olsson L. Fall and injury prevention in residential care-effects in residents with higher and lower levels of cognition. *Journal of the American Geriatrics Society* 2003;**51**(5):627–35.

Kropelin 2013

Kropelin TF, Neyens JC, Halfens RJ, Kempen GI, Hamers JP. Fall determinants in older long-term care residents with dementia: a systematic review. *International Psychogeriatrics* 2013;**25**(4):549–63. [PUBMED: 23253253]

Lamb 2007

Lamb SE, Hauer K, Becker C. Manual for the fall prevention classification system. www.profane.eu.org/documents/Falls_Taxonomy.pdf (accessed 18 July 2012).

Lamb 2011
Lamb SE, Becker C, Gillespie LD, Smith JL, Finnegan S, Potter R, et al. Reporting of complex interventions in clinical trials: development of a taxonomy to classify and describe fall-prevention interventions. *Trials [Electronic Resource]* 2011;**12**:125.

Latham 2003
Latham NK, Anderson CS, Lee A, Bennett DA, Moseley A, Cameron ID. A randomized, controlled trial of quadriceps resistance exercise and vitamin D in frail older people: The Frailty Interventions Trial in Elderly Subjects (FITNESS). *Journal of the American Geriatrics Society* 2003;**51**:291–9.

Le Blanc 2015
Le Blanc ES, Zakher B, Daeges M, Pappas M, Chou R. Screening for vitamin D deficiency: a systematic review for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 2015;**162**(2):109–22.

Lee 2017
Lee SH, Kim HS. Exercise interventions for preventing falls among older people in care facilities: a meta-analysis. *Worldviews on Evidence-Based Nursing* 2017;**14**(1):74–80. [PUBMED: 27984675]

Lefebvre 2011
Lefebvre C, Manheimer E, Glanville J. Chapter 6.4.11.1 The Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Lightbody 2002
Lightbody E, Watkins C, Leathley M, Sharma A, Lye M. Evaluation of a nurse-led falls prevention programme versus usual care: a randomized controlled trial. *Age and Ageing* 2002;**31**(3):203–10.

Lohse 2012
Lohse GR, Leopold SS, Theiler S, Sayre C, Cizik A, Lee MJ. Systems-based safety intervention: Reducing falls with injury and total falls on an orthopaedic ward. *Journal of Bone and Joint Surgery. American Volume* 2012;**94**(13): 1212–22.

Lord 2003
Lord SR, March LM, Cameron ID, Cumming RG, Schwarz J, Zochling J, et al. Differing risk factors for falls in nursing home and intermediate-care residents who can and cannot stand unaided. *Journal of the American Geriatrics Society* 2003;**51**(11):1645–50.

Lord 2005
Lord SR, Tiedemann A, Chapman K, Munro B, Murray SM, Gerontology M, et al. The effect of an individualized fall prevention program on fall risk and falls in older people: a randomized, controlled trial. *Journal of the American Geriatrics Society* 2005;**53**(8):1296–304.

McAlister 2003
McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. *JAMA* 2003;**289**(19):2545–53.

Muir 2012
Muir SW, Gopaul K, Montero Odasso MM. The role of cognitive impairment in fall risk among older adults: a systematic review and meta-analysis. *Age and Ageing* 2012; **41**(3):299–308. [PUBMED: 22374645]

Murray 2007
Murray GR, Cameron ID, Cumming RG. The consequences of falls in acute and subacute hospitals in Australia that cause proximal femoral fractures. *Journal of the American Geriatrics Society* 2007;**55**:577–82.

NLM 2012
National Library of Medicine. Medical Subject Headings: MeSH Browser (2012 MeSH). www.nlm.nih.gov/mesh/2012/mesh_browser/MBrowser.html (accessed 21 August 2012).

Norman 2008
Norman R, Haas M, Chenoweth L, Jeon Y-H, King M, Brodaty H, et al. *Dementia Care Mapping and Patient-centred Care in Australian Residential Homes: an Economic Evaluation of the CARE Study*. Sydney: Centre for Health Economics Research and Evaluation, 2008.

Nurmi 2002
Nurmi I, Luthje P. Incidence and costs of falls and falls injuries among elderly in institutional care. *Scandinavian Journal of Primary Health Care* 2002;**20**(2):118–22.

Nyberg 1997
Nyberg L, Gustafson Y, Janson A, Sandman PO, Eriksson S. Incidence of falls in three different types of geriatric care. A Swedish prospective study. *Scandinavian Journal of Social Medicine* 1997;**25**(1):8–13.

Nyman 2011
Nyman SR, Victor CR. Older people's recruitment, sustained participation, and adherence to falls prevention interventions in institutional settings: a supplement to the Cochrane systematic review. [Review]. *Age and Ageing* 2011;**40**(4):430–6.

OECD 2011
OECD, Eurostat, WHO. Part 1, Chapter 6: ICHA-HP Classification of Health Care Providers. In: A System of Health Accounts. 2011 edition. OECD Publishing. 2011:133-4. Available from http://dx.doi.org/10.1787/9789264116016-en (accessed 22 August 2012).

Oliver 2004
Oliver D, Daly F, Martin FC, McMurdo ME. Risk factors and risk assessment tools for falls in hospital in-patients: a systematic review. *Age and Ageing* 2004;**33**(2):122–30.

Pilz 2012
Pilz S, Dobnig H, Tomaschitz A, Kienreich K, Meinitzer A, Friedl C, et al. Low 25-hydroxyvitamin D is associated with increased mortality in female nursing home residents. *Journal of Clinical Endocrinology and Metabolism* 2012;**97** (4):E653-7.

Rapp 2008
Rapp K, Lamb SE, Buchele G, Lall R, Lindemann U, Becker C. Prevention of falls in nursing homes: subgroup analyses of a randomized fall prevention trial. *Journal of the American Geriatrics Society* 2008;**56**(6):1092–7.

Retraction Watch
Retraction Watch. JAMA journals pull 3 papers by same authors for misconduct. retractionwatch.com/2016/06/03/jama-journals-pull-3-papers-by-same-authors-for-misconduct/ (accessed 1 July 2017).

Review Manager [Computer program]
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

Robertson 2005
Robertson MC, Campbell AJ, Herbison P. Statistical analysis of efficacy in falls prevention trials. *Journals of Gerontology Series A-Biological Sciences and Medical Sciences* 2005;**60**(4):530–4.

Rubenstein 2006
Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. *Age and Ageing* 2006; **35 Suppl 2**:ii37–41. [PUBMED: 16926202]

Sach 2015
Sach TH, Desborough J, Houghton J, Holland R, CAREMED study team. Applying micro-costing methods to estimate the costs of pharmacy interventions: an illustration using multi-professional clinical medication reviews in care homes for older people. *International Journal of Pharmacy Practice* 2015;**23**:237–47. DOI: 10.1111/ijpp.12162

Santesso 2014
Santesso N, Carrasco-Labra A, Brignardello-Petersen R. Hip protectors for preventing hip fractures in older people. *Cochrane Database of Systematic Reviews* 2014, Issue 3. DOI: 10.1002/14651858.CD001255.pub5; PUBMED: 24687239

Schulz 2010
Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology* 2010;**63**(8):834–40. [PUBMED: 20346629]

Schünemann 2011
Schünemann HJ, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Chapter 11: Presenting results and 'Summary of findings' tables. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Sherrington 2011
Sherrington C, Tiedemann A, Fairhall N, Close JC, Lord SR. Exercise to prevent falls in older adults: an updated meta-analysis and best practice recommendations. *New South Wales Public Health Bulletin* 2011;**22**(3-4):78–83. [PUBMED: 21632004]

Sherrington 2017
Sherrington C, Michaleff ZA, Fairhall N, Paul SS, Tiedemann A, Whitney J, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. *British Journal of Sports Medicine* 2017;**51**(24):1750–8. [PUBMED: 27707740]

Silva 2013
Silva RB, Eslick GD, Duque G. Exercise for falls and fracture prevention in long term care facilities: a systematic review and meta-analysis. *Journal of the American Medical Directors Association* 2013;**14**(9):685–9.e2. [PUBMED: 23860265]

Stata [Computer program]
StataCorp LP. Stata Statistical Software. Version 8. StataCorp LP, 2003.

Stenvall 2012
Stenvall M, Berggren M, Lundstrom M, Gustafson Y, Olofsson B. A multidisciplinary intervention program improved the outcome after hip fracture for people with dementia--subgroup analyses of a randomized controlled trial. *Archives of Gerontology and Geriatrics* 2012;**54**(3):e284–9.

Stubbs 2015
Stubbs B, Denkinger MD, Brefka S, Dallmeier D. What works to prevent falls in older adults dwelling in long term care facilities and hospitals? An umbrella review of meta-analyses of randomised controlled trials. *Maturitas* 2015;**81**(3):335–42. [PUBMED: 25935294]

Sutton 1994
Sutton JC, Standen PJ, Wallace WA. Patient accidents in hospital: incidence, documentation and significance. *British Journal of Clinical Practice* 1994;**48**(2):63–6. [PUBMED: 8024991]

Verheyden 2013
Verheyden GS, Weerdesteyn V, Pickering RM, Kunkel D, Lennon S, Geurts AC, et al. Interventions for preventing falls in people after stroke. *Cochrane Database of Systematic Reviews* 2013, Issue 5. DOI: 10.1002/14651858.CD008728.pub2; CD008728

Vieira 2011
Vieira ER, Freund-Heritage R, da Costa BR. Risk factors for geriatric patient falls in rehabilitation hospital settings: a systematic review. *Clinical Rehabilitation* 2011;**25**(9):788–99.

Vlaeyen 2015
Vlaeyen E, Coussement J, Leysens G, Van der Elst E, Delbaere K, Cambier D, et al. Characteristics and effectiveness of fall prevention programs in nursing homes: a systematic review and meta-analysis of randomized controlled trials. *Journal of the American Geriatrics Society* 2015;**63**(2):211–21. [PUBMED: 25641225]

Vogler 2009
Vogler CM, Sherrington C, Ogle SJ, Lord SR. Reducing risk of falling in older people discharged from hospital: a randomized controlled trial comparing seated exercises, weight-bearing exercises, and social visits. *Archives of Physical Medicine & Rehabilitation* 2009;**90**(8):1317–24.

Zwarenstein 2008
Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;**337**:a2390.

References to other published versions of this review

Cameron 2005
Cameron I, Murray GR, Gillespie LD, Cumming RG, Robertson MC, Hill KD, et al. Interventions for preventing falls in older people in residential care facilities and hospitals. *Cochrane Database of Systematic Reviews* 2005, Issue 3. DOI: 10.1002/14651858.CD005465

Cameron 2010
Cameron ID, Murray GR, Gillespie LD, Robertson MC, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in nursing care facilities and hospitals. *Cochrane Database of Systematic Reviews* 2010, Issue 1. DOI: 10.1002/14651858.CD005465.pub2

Cameron 2012
Cameron ID, Gillespie LD, Robertson MC, Murray GR, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in care facilities and hospitals. *Cochrane Database of Systematic Reviews* 2012, Issue 12. DOI: 10.1002/14651858.CD005465.pub3

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Aizen 2015

Methods	Stepped-wedge, cluster-randomised controlled trial.
Participants	**Setting:** hospital, subacute, 5 geriatric rehabilitation wards, Israel N = 508 participants; 5 clusters Sample: 52% women Age (years): mean 83.2 **Baseline characteristics** Individualised fall prevention programme • *N*: 200 • *Age - mean (SD)* : 84.6 (5.6) • *Female - N (%)*: 92 (46.0) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 308 • *Age - mean (SD)* : 84.1 (7.7) • *Female - N (%)*: 173 (56.1) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Inclusion criteria:** Over 65 years; admitted to rehabilitation ward **Exclusion criteria:** Restricted to bed; refused to participate **Pretreatment differences:** Phase 1: Longer stay in the control group patients (P < 0. 001); higher percentage of females in the control group (P = 0.03)
Interventions	• Individualised fall prevention programme. Falls risk assessment and management: including medical interventions, environmental modifications, equipment modifications, cognitive and behavioural treatment, family guidance. Mobility restrictions and optimising location on weekly assessment. Environmental modifications unclear. • Usual care. Any activities undertaken by the participants recommended or administered by their treating team
Outcomes	• Rate of falls • Rate ratio • Adverse events
Duration of the study	Period of inpatient admission

Notes	Outcomes of phase one used only. Outcomes data for phase one and two only reported separately, attempts to contact authors unsuccessful. Excluded from pooling as group allocation of clusters unclear

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Insufficient information for judgement.
Allocation concealment (selection bias)	High risk	Allocation not concealed as consent only required for those receiving the intervention
Blinding of participants and personnel (performance bias) All outcomes	High risk	Blinding not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Senior nursing staff in control wards were aware of the study because the researchers were collecting study data. Researchers were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Attrition balanced across groups and missing outcomes not great enough to have a clinically relevant impact on observed effect size
Selective reporting (reporting bias)	Unclear risk	Protocol not available
Method of ascertaining falls	Low risk	Quote: "from notes in medical records themselves, and by asking a senior nurse each day about any falls on the ward in the past 24 h." Quote: "Information on falls was collected by the researchers from incident reports filed in patients' medical records,"
Baseline imbalance	High risk	Longer length of stay in control group at baseline suggests greater dependency in this group at baseline and not adjusted for in analysis
Other bias	Unclear risk	Quote: "some falls prevention activities were already occurring in control (and intervention) wards before the start of our study. These activities continued during the study period, making it more difficult

Aizen 2015 (*Continued*)

		to show any effect of our interventions." Impact of other falls intervention approaches unclear. Stepped-wedge trial but only data from phase 1 used as falls data not reported for both phases in combination

Ang 2011

Methods	RCT (individually randomised)
Participants	**Setting:** acute care hospital, Singapore N = 1822 participants Sample: newly admitted patients from 8 medical wards (50% women) Age (years): mean (SD) intervention group 70.3 (14.2), control group 69.7 (14.7) **Inclusion criteria:** aged ≥ 21; Hendrich II Fall Risk Model score ≥ 5 **Exclusion criteria:** admitted before start of study; fallen prior to falls risk assessment
Interventions	• Education + usual care: participants received one educational session (no more than 30 minutes) based on identified risk factors. Designed to increase awareness of risk of falling during hospitalisation and teach risk-reduction strategies. Relatives of confused participants received the educational session • Control: usual care and including usual fall-prevention interventions
Outcomes	• Number of people falling
Duration of the study	8 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Allocation of the participants to control or intervention groups was determined using block randomisation with the aid of a computer program and stratified by ward to ensure an even mix in the ward. "
Allocation concealment (selection bias)	Low risk	Quote: "Sealed, opaque, serially numbered envelopes were produced from the randomizations sequence separately for each stratum."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded

Ang 2011 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: "The research investigator scanned the electronic hospital occurrence report (eHOR) daily during weekday for entries of fall incidences reported by the nurses from the wards and ascertained if the entries were on participants involved in the study." Nursing staff recording falls described as blind to group allocation. Not clear if the research investigator was blind to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: all data analysed according to ITT.
Selective reporting (reporting bias)	High risk	Judgement comment: methods mention incidence of falls but only data on risk of falling reported
Method of ascertaining falls	High risk	Judgement comment: falls not clearly defined.
Baseline imbalance	Low risk	No important differences at baseline.
Other bias	Unclear risk	Unclear impact of standard falls prevention activities.

Barker 2016

Methods	Cluster-randomised controlled trial
Participants	**Setting:** 24 acute medical and surgical wards from 6 hospitals, Australia N = 31,411 unique participants, including 3853 admitted to both intervention and control wards at different times; 24 clusters Sample: 48.5% women Age (years): median 67 (interquartile range 51-79) **Baseline characteristics:** **6-PACK programme** • *N*: 22,670 admissions; 17,698 participants • *Age Median (IQR)*: 68 (51-80) • *Female N (%)*: 11,476 (50.6) • *Medical status defined? - Y/N* : Y (3+ comorbidities 21.2%) • *Falls risk defined?- Y/N*: N • *Dependency defined? - Y/N*: No • *Cognitive status defined? - Y/N*: Y **Usual care** • *N*: 23,575 admissions; 17,566 participants • *Age Median (IQR)*: 67 (51-79) • *Female N (%)*: 11,424 (48.5) • *Medical status defined? - Y/N* : Y (3+ comorbidities 25.3%) • *Falls risk defined?- Y/N*: N • *Dependency defined? - Y/N*: N • *Cognitive status defined? - Y/N*: Y **Inclusion criteria:** Wards: where fall-related injuries have been identified as a problem,

	acute medical or surgical wards; average patient length of stay <10 days; wards to have one or less low-low beds to each six standard beds on medical wards and one or less low-low beds to each, 29 standard beds on surgical wards; a fall risk assessment and/or prevention strategy checklist is not already included in the daily patient care plan documentation. Wards that have a fall risk assessment and/or prevention strategy checklist included on admission documentation but do not have a policy that this must be updated daily will not be excluded from participating in the study **Exclusion criteria:** No patient level exclusion criteria. **Pretreatment differences:** Nil
Interventions	• 6-PACK programme comprising a 9 item falls risk assessment tool and delivery of one or more of six interventions to high risk patients: 1) Placement of a 'falls alert' sign above the patient's bed. 2) Supervision of patients while in the bathroom. 3) Use of a low-low bed. 4) Ensuring that the patient's walking aid is within reach at all times. 5) Establishment of a toileting regimen. 6) Use of a bed/chair alarm when the patient is positioned in the bed/chair. Staff education integral to implementation.Nurses were asked to update the fall risk tool for each of their patients each shift and to apply a falls alert sign and one or more of the remaining 6-PACK interventions to patients classified as being at high risk • Usual care. Any standard hospital practice provided by wards as part of existing hospital policy relating to fall prevention, which may have included some components of the 6-PACK programme and other interventions such as non-slip socks, constant patient observers, and falls alert wrist bands.
Outcomes	• Rate of falls • Number of fallers (number of unique fallers provided by author correspondence) • Number of injurious falls • Fracture falls (number of unique patients with fractures provided by author correspondence) • Multiple falls • Adverse events
Duration of the study	12 months intervention period plus 3 month pre-randomisation baseline period
Notes	ACTRN12611000332921 "The use of all 6-PACK programme components (fall risk tool and six interventions) was threefold higher on intervention wards than on control wards (incidence rate ratio 3.05, 95% confidence interval 2.14 to 4.34; P<0.001)."

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "used the RALLOC command in Stata to develop the randomisation schedule, using a random sequence in blocks of two generated by the study statistician." Judgement comment: random sequence allocation done.

Allocation concealment (selection bias)	High risk	Quote: "Concealment of allocation was ensured, as the schedule was accessible only by the study statistician, who was not involved in ward recruitment or data collection." Judgement comment: although allocation sequence initially concealed, subjects were enrolled after cluster randomisation, and sequence would have been known at this time
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "It will not be feasible to blind ward nurses or patients to the intervention." Judgement comment: not done.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Blinding of the assessors collecting the fall and falls prevention practice data was also not possible. Assessors blinded to group allocation did the secondary coding of characteristics of falls and injuries, and the primary assessor completed the coding. A statistician blinded to group allocation (RW) did the data analysis." Judgement comment: not done.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: no loss to follow-up
Selective reporting (reporting bias)	Low risk	Judgement comment: all falls outcomes reported as per trial registry record
Method of ascertaining falls	Low risk	Quote: "by daily auditing of patient medical records and verbal report of the nurse unit managers. These data will then be triangulated with hospital incident reporting and administrative patient episode datasets. Concurrent to this will be hospital-wide education and reminders of the fall definition and incident reporting best practice, facilitated by use of an existing training package. 23 Patient" Judgement comment: multiple methods of concurrent recording of falls data used
Baseline imbalance	Low risk	Quote: "Characteristics of admitted patients and length of stay were similar for intervention and control groups and across baseline and randomised controlled trial periods" Judgement comment: no imbalance across groups.

Barker 2016 (*Continued*)

Other bias	Unclear risk	Unclear impact of any ongoing falls prevention activities.

Beck 2016

Methods	RCT (cluster randomised), nursing home subgroup data.
Participants	**Setting:** 3 residential care homes, high-level care, Denmark N = 31 participants; 3 clusters. Sample: 65% women Age (years): mean 88 **Baseline Characteristics** **Multidisciplinary nutritional support** • *N*: 9 nursing home • *Age Range*: 88.1 (9.6) • *Sex (% female)*: 6 (67) • *Medical status defined?*: N • *Falls risk defined?*: N • *Dependency defined? (ADL problem, No. social services for home help, nursing)*: Y **Control** • *N*: 22 nursing home • *Age Range*: 87.8 (7.0) • *Sex (% female)*: 14 (64) • *Medical status defined?*: N • *Falls risk defined?*: N • *Dependency defined? (ADL problem, No. social services for home help, nursing)*: Y **Inclusion criteria:** 65+ years, at nursing home or receiving home care (assistance with meals) with 2 points according to Eating Validation Scheme (EVS) completed by nursing staff caregivers (would benefit from intervention) able to completed planned tests **Exclusion criteria:** not able or willing to give informed consent **Pretreatment differences:** living in a nursing home: intervention 16%, control 55% (P < 0.001); 30-seconds chair-stand modified, mean (SD) 4.9 (3.3) intervention, 2.5 (2.7) control (P = 0.004); cognitive problem 56% intervention versus 78% control (P = 0.03)
Interventions	• Multidisciplinary nutritional support. Nutrition co-ordinator involvement, multidisciplinary project group meetings, plan of action in the municipality care register system, Exercise, nutritional support, support for dysphagia and eating problems as indicated by EVS screening. 30 to 45 minutes moderate-intensity exercise sessions including strength and balance training twice a week, oral training supplements after exercise, weekly assessment of weight, individual dietetics treatment plan and regular reviews by dietician, multidisciplinary meeting weekly to evaluate and adjust individual treatment plans, OT involvement if indicated. *Health professional involvement*: Nutrition co-ordinator, physiotherapist twice weekly, dietician performs initial interview, then regular consultations and phone or group follow-up, occupational therapist to consults with patients who suffer from eating dependency or chewing and swallowing problems and initiate interventions if indicated. • Control. Nutrition co-ordinator involvement plus standard interventions from physiotherapist, registered dietician and occupational therapist requested through the

	municipality's normal assessment and referral system will be maintained.
Outcomes	• Number of falls • Adverse events
Duration of the study	11 weeks
Notes	A trial of nutritional support using a structured and multidisciplinary approach, focusing on nutritional risk factors, in undernourished older adults in both home care and nursing home settings, with results reported separately

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Judgement comment: sequence generation by drawing a lot for an opaque envelope
Allocation concealment (selection bias)	High risk	Judgement comment: randomisation by researcher not involved in the study (2016 p200). Author correspondence quote: "participants were invited by means of the staff who did not know about the result of the group allocation". and "we did not include new admissions". However: "Due to the limited knowledge about the benefit of nutritional support among home-care clients, the aim was to randomly assign 2 of the 3 home-care clusters to the intervention group", this is likely to enable the randomisation sequence to be predicted, concealment not possible for the final cluster
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "The researchers for this study included the research assistants (AGC, BSH, SD-S, and TKSM) and the primary investigator (AB), who were not blinded for the intervention." Judgement comment: blinding not done.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The researchers for this study included the research assistants (AGC, BSH, SD-S, and TKSM) and the primary investigator (AB), who were not blinded for the intervention. Before starting the analysis the primary investigator (AB) was reblinded for participants' group assignment. " Judgement comment: not done, falls data

Beck 2016 (*Continued*)

		collected by unblinded research nurse. Although primary investigator "reblinded" before analysis no details were reported on the method for this and it is considered likely to include a risk of residual unblinding
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: author correspondence clarified data and indicated one withdrawal, no other loss to follow-up
Selective reporting (reporting bias)	Low risk	Judgement comment: trial protocol available and falls outcomes consistently reported
Method of ascertaining falls	Low risk	Quote: "The information was gathered by means of data from the RAI-NH version 2.0 and RAI-HC version 2.0 assessments and the municipality care register system. For each participant, the same trained nurse collected" Judgement comment: concurrent falls data collection with clear definition
Baseline imbalance	High risk	Judgement comment: baseline imbalance in nursing home subgroup for cognition, no adjustment performed
Other bias	Low risk	None detected.

Becker 2003

Methods	RCT (cluster randomised by facility).
Participants	**Setting:** 6 long-term care facilities (high-level nursing care), Germany N = 981 participants; 6 clusters. Sample: 79% women Age (years): mean (SD) intervention group 83.5 (7.5), control group 84.3 (6.9) **Inclusion criteria:** resident of facility. Inclusion criteria for exercise programme: able to stand while holding a chair, able to lift one foot **Exclusion criteria:** none stated
Interventions	• Fall prevention programme for staff and residents. Residents chose to participate in any combination of interventions for any length of time. Those choosing to participate in fall registration only also received environmental modification and modification of nursing care ○ Staff training on risk factors and preventive measures (60 minutes), audit and monthly feedback re falls and injuries ○ Check list of 76 environmental hazards (lighting, chair and bed height, floor

	surfaces, etc). Feedback to staff and administrators ○ Resident education: all received written information, offered personal consultation by study nurse or exercise instructor ○ Group exercise programme (progressive balance and resistance training) 75 minutes, 2 x per week ○ Hip protectors • Control: usual care, no specific program activities.
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (hip fractures) • Adverse events
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Cluster randomisation of 6 facilities using sealed envelopes selected by an independent person. Insufficient information to permit judgement
Allocation concealment (selection bias)	Unclear risk	Allocation in sealed envelopes, but individuals admitted after group allocation by a person who may have been unblinded and may have had knowledge of participant characteristics
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	All residents included in analysis.
Selective reporting (reporting bias)	Low risk	All expected falls outcomes completely and thoroughly reported. Adjustment for clustering conducted
Method of ascertaining falls	Low risk	Fall definition provided and concurrent recording of falls

Becker 2003 (*Continued*)

Baseline imbalance	High risk	Greater proportion of intervention group were taking 4 or more medications
Other bias	Low risk	None identified.

Bischoff 2003

Methods	RCT (individually randomised)
Participants	**Setting:** 2 hospitals with long-stay geriatric care units, Basel, Switzerland N = 122 participants Sample: 100% women Age (years): mean (SD) intervention group 85.4 (5.9), control group 84.9 (7.7) **Inclusion criteria:** female; aged ≥ 60; able to walk 3 metres **Exclusion criteria:** primary hyperparathyroidism; hypercalcaemia; hypercalcuria; renal insufficiency; fracture or stroke in last 3 months
Interventions	1. 800 IU oral cholecalciferol (vitamin D3) plus 1200 mg calcium daily for 12 weeks 2. Control: 1200 mg calcium daily for 12 weeks
Outcomes	• 1. Rate of falls • 2. Number of people falling • 3. Number sustaining a fracture (hip fractures) • 4. Adverse events
Duration of the study	12 weeks
Notes	50% of participants had a baseline serum vitamin level < 30 nmol/L

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The randomization was performed by an independent statistician."
Allocation concealment (selection bias)	Low risk	Participants randomised in groups of four by an independent statistician
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: double-blind
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Patients, nurses, and all investigators were blinded to the treatment assignment throughout the study

Bischoff 2003 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: high loss to follow-up (31% in vit D and 25% in control group); however, analysed as ITT with rate ratio accounting for days of follow-up and balanced between groups
Selective reporting (reporting bias)	Low risk	Judgement comment: no study protocol identified, but data on falls, fallers, multiple falls as adjusted and adjusted outcomes reported
Method of ascertaining falls	Low risk	Quote: "Falls were recorded by the nurses on the in-patient units who had received training in the use of the fall protocol (date, time, circumstances, injuries). Falls were defined as "unintentionally coming to rest on the ground, floor, or other lower level." Coming to rest against furniture or a wall was not counted as a fall. (24) Nurses completed the fall protocol if they observed or received a report of a fall."
Baseline imbalance	Low risk	Judgement comment: characteristics and number of falls balanced at baseline
Other bias	Low risk	Judgement comment: none identified. Small groups randomised however given trial is double-blinded randomisation unlikely to be predictable

Broe 2007

Methods	RCT (individually randomised)
Participants	**Setting:** 1 long-term care facility (high-level care), USA N = 48 participants included in review (total of 124 in the study) Sample: 73% women Age (years): mean 89 (SD 6) **Inclusion criteria:** life expectancy > 6 months; able to swallow medications; resident for > 3 months **Exclusion criteria:** taking glucocorticoids; anti-seizure medications; pharmacological doses of vitamin D; calcium metabolism disorders; severe mobility restriction; fracture within previous 6 months
Interventions	• 200 IU of vitamin D2 daily for 5 months (not included in review) • 400 IU of vitamin D2 daily for 5 months (not included in review) • 600 IU of vitamin D2 daily for 5 months (not included in review) • 800 IU of vitamin D2 daily for 5 months • Control: placebo daily for 5 months
Outcomes	• Rate of falls • Number of people falling

Broe 2007 (*Continued*)

Duration of the study	5 months
Notes	Mean baseline serum vitamin D level for 800 IU group and control group combined was 53 nmol/L

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "... computer-generated randomisation list."
Allocation concealment (selection bias)	Low risk	Pharmacy conducted randomisation and supplied medication in blister packs with name and patient identification number only
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: double-blind
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Nursing staff completing incident forms blinded to treatment status because blister packs and tablets identical in appearance. Also, quote: "a programmer, not involved with this study and not aware of participant study group assignments, created the falls dataset linking the participant identification number with falls reported during the study period"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: low loss to follow-up and ITT analysis performed
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified, however all expected outcomes (falls, rate of falls and fallers) reported
Method of ascertaining falls	Low risk	Judgement comment: falls concurrently recorded and clearly defined
Baseline imbalance	Unclear risk	Judgement comment: few differences at baseline; however baseline cognition, medical comorbidities and function not reported
Other bias	Low risk	None identified.

Buckinx 2014

Methods	RCT. Individually randomised.
Participants	**Setting:** 2 residential care facilities, intermediate-level care, Belgium N = 62 participants Sample: 76% women Age (years): mean 83.2 (SD 7.9) **Baseline characteristics** **Whole body vibration** • *N*: 31 • *Age - mean (SD)*: 82.2 (9.02) • *Female N (%)*: 20 (64.5) • *Medical status defined? - Y/N*: Y • *Falls risk defined (Y/N)?*: Y • *Dependency defined (Y/N)?*: Y • *Cognitive status defined? Y/N*: Y **Control** • *N*: 31 • *Age - mean (SD)*: 84.2 (6.83) • *Female N (%)*: 27 (87.1) • *Medical status defined? - Y/N*: Y • *Falls risk defined (Y/N)?*: Y • *Dependency defined (Y/N)?*: Y • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** residents from two nursing homes; able to remain standing; able to move with or without technical assistance **Exclusion criteria:** weight greater than 150 kg; electronic implants; knee or hip prostheses; epilepsy; bleeding disorders; inflammatory abdominal disorders; high risk of thromboembolism; malignancy; unconsolidated fracture; refusal of doctor or family **Pretreatment differences:** gender (more women in control group) P = 0.04; lower body mass in control group P < 0.01; lower MMSE in control group, P = 0.04
Interventions	• Whole body vibration. Exercise programme on a sinusoidal vibration platform (Vibrosphere), standing without shoes with knees flexed, cushion placed under vibrosphere. 3 x weekly, 5 series of 15 seconds of vibrations at 30 Hz, 2mm amplitude, alternate with 30 seconds rest, total vibration time 1 minute 15 seconds, minimum 1 day between sessions. Supervised by one of 4 people, 2 physiotherapists and 2 authors. • Usual care. No change to lifestyle during study, no involvement in any new type of physical activity
Outcomes	• Number of falls • Number of people falling
Duration of the study	6 months intervention, follow-up to 12 months.
Notes	Compliance: 91.9% of exercise sessions performed.

Risk of bias

Bias	Authors' judgement	Support for judgement

Buckinx 2014 (*Continued*)

Random sequence generation (selection bias)	Low risk	Quote: "We performed the randomisation by blocks of four with a computer-generated randomisation procedure." Judgement comment: computer-generated randomisation.
Allocation concealment (selection bias)	Unclear risk	Quote: "An identification number and a randomisation number were created for each participant." Judgement comment: method of concealment of allocation sequence from those enrolling participants was unclear
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not done. Blinding not possible.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: not done. Nurses recorded falls, they were not blinded. Blinded assessment unlikely to include falls outcomes
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: numbers and reasons balanced between groups
Selective reporting (reporting bias)	Low risk	Judgement comment: number of falls not defined as outcome in trial registry. Trials registry indicates 3 months outcomes. Reporting of falls data appears complete although not predefined
Method of ascertaining falls	Low risk	Quote: "Nurses completed the fall record with the date, time, and circumstances of the falls." Judgement comment: likely that falls were recorded at time of event
Baseline imbalance	High risk	Baseline differences in weight, gender, MMSE may impact on falls rates
Other bias	Low risk	Judgement comment: none identified

Buettner 2002

Methods	RCT (individually randomised)
Participants	**Setting:** 3 nursing care facilities, USA (1 high-level nursing facility, 1 skilled nursing facility, 1 intermediate-level facility) N = 27 participants Sample: 44% women Age (years): mean 83.3 (range 60 to 98) **Inclusion criteria:** ≥ 2 falls in past 2 months between 7.00 am to 9 am; MMSE score

	< 23; aged > 60; walking independently, or with 1 assistant or assistive device **Exclusion criteria:** not resident for ≥ 60 days; a healing fracture; attending physiotherapy
Interventions	• Supervised group exercises: walking group daily at 6.30 am; exercise to improve function (balance, strength, and flexibility) 3 x per week in mid afternoon; sensory air mat therapy (movement, relaxation) 2 x per week in evenings. Intervention overseen by Certified Therapeutic Recreational Specialist with assistance of staff members. The interventions were scheduled at the time of day when most falls occur and in the locations where the falls occur • Control: usual care
Outcomes	• Number of falls
Duration of the study	2 months
Notes	Published data incomplete. Further data provided by authors could not be analysed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No description of method of sequence generation
Allocation concealment (selection bias)	Unclear risk	No description of allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff collecting falls data do not appear to have been blinded to allocation status
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: insufficient detail on which patients are included in data analysis for judgement
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified. Number of falls analysed as t-test of absolute numbers without rate, considering pre-test falls
Method of ascertaining falls	High risk	Judgement comment: falls not clearly defined
Baseline imbalance	Unclear risk	Judgement comment: Baseline characteristics not reported by allocation group
Other bias	Low risk	Judgement comment: None detected

Burleigh 2007

Methods	RCT (individually randomised)
Participants	**Setting:** general assessment and rehabilitation wards in an acute geriatric unit, Glasgow, Scotland N = 205 participants Sample: 59% women, median serum vitamin D (25 OHD) = 22.00 nmol/L, IQR 15.00 to 30.50 at baseline. Age (years): mean (SD) intervention 82.3 (7.6), control 83.7 (7.6) **Inclusion criteria:** admitted to a ward in the acute geriatric unit; aged ≥ 65 **Exclusion criteria:** hypercalcaemia; urolithiasis; renal dialysis; terminal illness; bed bound; reduced Glasgow Coma Score; already prescribed vitamin D and calcium; 'nil by mouth' on admission
Interventions	• 800 IU oral cholecalciferol (vitamin D3) plus 1200 mg calcium daily until separation from the facility • Control: 1200 mg calcium daily until discharge or death
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures) • Adverse events
Duration of the study	Aproximately 9 months. Median length of stay 30 days
Notes	Baseline serum vitamin D (25 OHD) = median 22.00 nmol/L, IQR 15.00 to 30.50

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "... randomised using a random numbers table"
Allocation concealment (selection bias)	Low risk	Quote: "Randomisation was known only to the statistician and pharmacist who subsequently issued an appropriate uniquely numbered drug blister pack to each patient's ward."
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: double-blind
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Staff completing falls data may have been aware of treatment status as there was no placebo in place of vitamin D. Insufficient information to permit judgement
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis and losses balanced between groups

Burleigh 2007 (*Continued*)

Selective reporting (reporting bias)	Low risk	Judgement comment: falls and fallers outcomes reported as per trial registration
Method of ascertaining falls	Low risk	Judgement comment: falls defined and recorded concurrently
Baseline imbalance	High risk	Judgement comment: 2x2 indicates significant difference in proportion with Zimmer frame between groups (P = 0.02)
Other bias	Low risk	Judgement comment: falls and fallers outcomes reported as per trial registration

Cadore 2014

Methods	RCT (individually randomised)
Participants	**Setting:** residential care facility, mixed-level care, Spain N = 24 participants Sample: 70% women Age (years): mean 91.9 (SD 4.1) **Baseline Characteristics** **Multicomponent exercises** • *N*: 11 • *Age Range - mean (SD) (overall 91.9 +/- 4.1 years)*: 93.4 ± 3.2 • *Female (17/24 overall) n (%)*: 8/11 • *Medical status defined? (admission diagnosis & co-morbidities)*: N • *Falls risk defined?*: Y, Dual task walking • *Dependency defined?*: Y • *Mean no falls pre-training*: 0.77+/-0.44 • *Cognitive status defined?*: Y **Control** • *N*: 13 • *Age Range - mean (SD) (overall 91.9 +/- 4.1 years)*: 90.1 ± 1.1 • *Female (17/24 overall) n (%)*: 9/13 (69) • *Medical status defined? (admission diagnosis & co-morbidities)*: N • *Falls risk defined?*: Y, Dual task walking • *Dependency defined?*: Y • *Mean no falls pre-training*: 0.93n+/-0.3 • *Cognitive status defined?*: Y **Inclusion criteria:** nursing home residents from Pamplona, Spain; 85 years or older; frail (as per Fried's criteria): 3 or more of slowness, weakness, weight loss, exhaustion, and low physical activity **Exclusion criteria:** the absence of frailty or pre-frailty syndrome; dementia; disability (defined as a Barthel Index (BI) lower than 60 and inability to walk independently without help of another person); recent cardiac arrest; unstable coronary syndrome; active cardiac failure; cardiac block; any unstable medical condition **Pretreatment differences:** baseline demographic data not reported

Interventions	• Multicomponent exercises. Muscle power training (8-10 repetitions, 40% to 60 % of the one-repetition maximum) combined with balance and gait retraining, including warm up and cool down periods. Twice weekly, 40 minute duration, at least 2 consecutive days between sessions • Control. Mobility exercises: small active and passive movements applied as a series of stretches in a rhythmic fashion to the individual joints. Such exercises are routinely encouraged in most Spanish nursing homes. 30 minutes per day at least 4 days per week
Outcomes	• Rate of falls
Duration of the study	12 weeks
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The randomisation sequence was generated by http://www.randomization.com and concealed until interventions were assigned."
Allocation concealment (selection bias)	Low risk	Quote: "and concealed until interventions were assigned." Judgement comment: author correspondence. Quote: "The group allocation was concealed. A researcher with no previous contact with subjects as well as not involved with assessment and training made the allocation of subjects."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding of participants not possible due to active involvement in intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: blinding mentioned is not for falls outcomes. Residents who were not blinded recorded falls
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss to follow-up low and balanced between groups
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol or trials registry records identified
Method of ascertaining falls	High risk	Quote: "Data on the incidence of falls were assessed retrospectively using questionnaires to residents." Judgement comment: based on recall of participants.

Cadore 2014 (*Continued*)

Baseline imbalance	Unclear risk	Judgement comment: baseline demographic data not reported
Other bias	Low risk	Judgement comment: none identified

Chapuy 2002

Methods	RCT (individually randomised)
Participants	**Setting:** 55 intermediate nursing care facilities, France N = 610 participants Sample: 100% women Age (years): mean 85.2 (SD 7.1) **Inclusion criteria:** ambulatory; life expectancy > 2 years **Exclusion criteria:** malabsorption; serum calcium > 2.63 mmol/L; chronic renal failure (serum creatinine >150 μmol/L), taking bone metabolism altering medications within the past year, e.g. corticosteroids, anticonvulsants or high doses of thyroxine; fluoride salts (43 months), bisphosphonates, calcitonin (41 month), calcium (4500 mg/day) and vitamin D (4100 IU/day) during the last 12 months
Interventions	• 800 IU of vitamin D3 + 1200 mg calcium carbonate fixed combination daily • 800 IU of vitamin D3 + 1200 mg calcium carbonate separately daily • Control: placebo
Outcomes	• Number of people falling • Number sustaining a fracture (hip fracture) • Adverse events
Duration of the study	24 months
Notes	Described as "apartment houses for elderly people" in Chapuy 2002 but provision of drugs supervised by nursing staff "to ensure compliance". Mean baseline serum vitamin D level 22 nmol/L

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'
Allocation concealment (selection bias)	Unclear risk	Although described as multicentre, randomised, double-masked, placebo-controlled, the method of concealment prior to allocation is not described in sufficient detail to allow a definite judgement

Chapuy 2002 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: double-blind
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Blinding of treatment status to outcome assessors not mentioned. Participants were asked if they had an adverse event (including falls) in last 3 months. Not clear if the person asking would have known allocation status
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: loss to follow-up over 2-year period unclear
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified.
Method of ascertaining falls	High risk	Falls events poorly defined.
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none detected

Chenoweth 2009

Methods	RCT (cluster randomised by unit).
Participants	**Setting:** 15 residential dementia care units (high-level nursing care), Sydney, Australia N = 289 residents; 15 clusters Sample: people with dementia (78% women) Age (years): mean (SD) person-centred care group 83 (7.6), dementia-care mapping group 84 (6.4), usual care group 83 (7.6) **Inclusion criteria** (facilities): task-focused (not person-centred) care systems. Inclusion criteria (residents): dementia and low cognitive function; aged >60; high dependency needs; persistent need-driven dementia compromised behaviours **Exclusion criteria** (residents): serious co-morbidities complicating or masking dementia; palliative care; unremitting pain; distressing physical symptoms; respite placement
Interventions	• Person-centred care: one researcher trained 2 care staff per site in allocated method of care (*see* 'Notes'), worked with trained staff to implement care plans, provided two site visits to give ongoing support for staff, then regular telephone contact for 4 months • Dementia care mapping: two researchers trained 2 care staff per site in allocated method of care (*see* 'Notes'), carried out "mapping" with trained staff, developed care plans with trained staff, trained staff helped colleagues implement plans, regular telephone contact from researchers for 4 months • Usual care: non person-centred care that is task-focused and concerned mostly with physical care needs
Outcomes	• Number of people falling

Duration of the study	8 months
Notes	Person-centred care emphasised social interactions at affective level based on life histories; aimed to preserve personal identity and foster meaningful relationships Dementia-care mapping: "mapping" consisted of observation of each participant for 6 hours per day for 2 days to identify factors related to well-being

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Allocation was done by the study statistician (MTK), who was unaware of the identity of sites, using an SAS20 program."
Allocation concealment (selection bias)	Low risk	Eligible residents were selected by facility managers or directors before randomisation of sites
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Treatment allocation was masked to assessors." Three separate research assistants collected outcome data from each cluster of five facilities. Staff of facilities instructed not to inform assessors of interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: >20% loss from person-centred care and usual care arms
Selective reporting (reporting bias)	Unclear risk	Judgement comment: falls stated as outcome - Incidents, quote: "and subsequent admissions to hospital were discerned from official records of incidents including residents' falls, fractures, lacerations, bruises, medication errors, and behavioural incidents" (p320, column 1, para 2). However, falls not stated as outcome in initial trial registry record (added retrospectively)
Method of ascertaining falls	High risk	Judgement comment: falls poorly defined and multiple sites enrolled
Baseline imbalance	Low risk	Judgement comment: differences at baseline adjusted for in analysis
Other bias	Low risk	Judgement comment: none identified.

Choi 2005

Methods	RCT (cluster randomised).
Participants	**Setting:** 2 residential care facilities (intermediate-level care), Korea N = 68 participants; 2 clusters. Sample: 75% women Age (years): mean 77.9 (range 61 to 91) **Inclusion criteria:** ambulatory; age > 60; at least one fall risk factor (impaired gait, impaired balance; a fall in the last year; postural hypotension; four or more medications affecting balance) **Exclusion criteria:** severe dementia; physical illness that may prevent completion of 12-week course of exercise; involvement in any other exercise
Interventions	• Supervised Tai Chi: 35-minute group sessions with certified Tai Chi leader, 3 x per week for 12 weeks • Usual routine activities
Outcomes	• Number of people falling
Duration of the study	3 months
Notes	Cluster randomised, described as quasi-experimental design with a non-equivalent control group

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "... two facilities with similar characteristics were selected and randomly assigned to either the experimental or control group by coin tossing."
Allocation concealment (selection bias)	High risk	After first toss the allocation of the second facility would be known. No description of whether individual participant recruitment was undertaken after group allocation by a person who was unblinded and may have had knowledge of participant characteristics
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss similar between groups.
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified.

Choi 2005 (*Continued*)

Method of ascertaining falls	High risk	Judgement comment: falls defined but only recorded weekly.
Baseline imbalance	High risk	Judgement comment: significant difference between groups in muscle strength and balance measures - addressed for balance and strength scores by using difference scores - but no adjustment apparent for falls data
Other bias	Low risk	Judgement comment: assignment predicable as cluster randomised and only 2 facilities, however this accounted for under allocation concealment. No other sources of bias identified

Clifton 2009

Methods	RCT (individually randomised)
Participants	**Setting:** 1 veterans skilled nursing facility (high-level nursing care), Washingon state, USA N = 43 participants Sample: 5% women Age (years): mean 82.2 (SD 7.1) **Inclusion criteria:** expected length of stay > 120 days; high risk of falling (Morse Scale score ≥ 50); unable to ambulate or transfer without assistance **Exclusion criteria:** history of adverse reaction to medical adhesives; mechanobullous disease; skin breakdown on the legs > 10 cm; skin eruption on the legs
Interventions	• FallSaver system: wireless position-monitoring patch fixed to the thigh. Transmitted signal to receiver/alarm unit when angle of declination reached about 45 degrees from horizontal, indicating the individual was moving into a weight-bearing position • No FallSaver use
Outcomes	• Rate of falls
Duration of the study	Cross-over after 60 days for second 60-day period
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation sequence generated using a web-based programme

Clifton 2009 (*Continued*)

Allocation concealment (selection bias)	Low risk	Quote: "Allocation of sequence, performed by the study coordinator, was masked until informed consent was obtained from each respective subject."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Caregivers recorded falls. Not blind to FallSaver use
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: higher loss to follow-up in intervention arm due to discontinuing intervention
Selective reporting (reporting bias)	Low risk	Judgement comment: falls outcomes reported as per trial registration
Method of ascertaining falls	Low risk	Judgement comment: falls defined and recorded concurrently.
Baseline imbalance	Unclear risk	Judgement comment: characteristics not reported by group allocation
Other bias	High risk	Judgement comment: author employed by company making FallSaver devices

Colon-Emeric 2013

Methods	RCT (cluster randomised), pilot study
Participants	**Setting:** 8 residential care facilities, 4 veterans affair, 4 community, USA N = Not Reported (NR). 8 clusters, 982 facility beds. Sample: NR Age (years): NR **Baseline Characteristics** **CONNECT & FALLS** • *N*: NR. 4 facilities, average bed size 131.3. 243 staff participants. • *Age - mean (SD)*: NR • *Female (%)*: NR • *Medical status defined? - Y/N* : NR • *Falls risk defined? - Y/N* : NR • *Dependency defined? - Y/N*: NR • *Cognitive status defined? - Y/N*: NR **FALLS only** • *N*: NR. 4 facilities, average bed size 114.3. 254 staff participants. • *Age - mean (SD)*: NR

	• *Female (%)*: NR • *Medical status defined? - Y/N* : NR • *Falls risk defined? - Y/N* : NR • *Dependency defined? - Y/N*: NR • *Cognitive status defined? - Y/N*: NR **Inclusion criteria:** residents: aged 50 years or over; experienced one or more falls during the study period, and remained in the NH at least 72 hours after the fall. Staff: all NH employees aged 18 and older who had direct resident contact were eligible for participation. Emloyees from nursing, rehabilitation, social work, dietary services, environmental services, activities, medical services and administration **Exclusion criteria:** Staff: temporary agency staff and staff working only as needed **Pretreatment differences:** more patients who fell had visual impairment in intervention nursing homes, more Caucasian staff in intervention nursing homes
Interventions	• CONNECT followed by FALLS: CONNECT is an intervention which is a process to implement quality improvement programs, aiming to improve nursing home (NH) staff connections, communication, and problem solving. Uses storytelling, relationship mapping, mentoring, self-monitoring, and feedback to help staff identify communication gaps and practice interaction strategies. CONNECT for 12 weeks consisting of 2 in-class sessions plus mentoring for 2 weeks after each session; then FALLS for 12 weeks. • FALLS only. Falls quality improvement programme which includes group training, modules, teleconferences, academic detailing, and audit and feedback on multifactorial falls prevention (addressing orthostatic hypotension, sensory impairment, footwear, gait and assistive devices, toileting needs, environmental problems, fall-related medications, and vitamin D). One half-day training session followed by 11 weekly teleconferences. Case-based self study modules. Academic detailing sessions for small groups of staff conducted twice at each nursing unit.
Outcomes	• Rate of falls • Ratio of change in rate of falls • Adverse events
Duration of the study	24 weeks intervention (12 weeks CONNECT/control plus 12 weeks FALLS), 6 months post-intervention follow-up
Notes	NCT00836433. Baseline data and N for all residents not known, confirmed by author correspondence

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: sequence by random number generator.
Allocation concealment (selection bias)	Unclear risk	Judgement comment: person assigning treatment groups was blinded to nursing home identity, but unclear if individual participant

Colon-Emeric 2013 (*Continued*)

		recruitment (staff) was completed prior to assignment of the cluster
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: unable to blind personnel.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: as staff would have recorded falls and staff were the subject of the intervention, it is unlikely that blinding would have been possible for those recording falls data
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: there were no missing data for fall rates
Selective reporting (reporting bias)	Low risk	Judgement comment: data on falls are reported as per trial record and includes the main expected falls outcomes
Method of ascertaining falls	Low risk	Quote: "Falls were ascertained from facility fall logs, incident reports, and the Minimum Data Set; occupied bed days were calculated from daily census data that each facility provided." Judgement comment: falls were clearly defined and likely to be recorded concurrently in facility fall logs used as the data source
Baseline imbalance	Low risk	Judgement comment: the analysis accounted for clustering and potential confounders
Other bias	Low risk	Judgement comment: none identified.

Cox 2008

Methods	RCT (cluster randomised by Primary Care Organisation (PCO) each containing nursing care facilities)
Participants	**Setting:** 209 care homes (high and intermediate level care), England and Wales N = 5637 participants. 29 clusters Sample: 77% women Age (years): not stated **Inclusion criteria** (facilities): if local ethics and research governance procedures were swift enough to enable enrolment **Exclusion criteria** (facilities): if demographic information was not provided
Interventions	• Half day training sessions for managers, nurses and health care assistants in each PCO. Training delivered by specialist osteoporosis nurses and included information on falls and falls prevention

Cox 2008 (*Continued*)

	• Control group received training 12 months later
Outcomes	• Number of people falling • Number sustaining a fracture (all fractures, hip fractures)
Duration of the study	12 months
Notes	5 of 29 clusters lost to follow-up in intervention group compared with 16 of 29 clusters in control group

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The PCOs were stratified into two groups, larger PCOs and smaller PCOs based on the median number of care homes. Within each stratum, a single block of allocations was undertaken using a computer package to ensure equivalent numbers of PCOs in each group."
Allocation concealment (selection bias)	Low risk	Quote: "All PCO demographic data were forwarded to the Department of Health Science at the University of York for randomisation and allocation." "The allocation was undertaken by an independent researcher."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: no statement re blinding. Facilities and staff (including manager reporting outcome data) knew of allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: 16% loss to follow-up for control group
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol, but all expected outcomes reported (number of falls, risk of falls and fractures) and as described in methods
Method of ascertaining falls	High risk	Judgement comment: no fall definition reported. Fall and fracture data collected via questionnaire to each facility manager -

Cox 2008 (*Continued*)

		likely variability
Baseline imbalance	Unclear risk	Judgement comment: baseline data on cognition, comorbidities, function not reported
Other bias	Low risk	Judgement comment: none identified.

Crotty 2004a

Methods	RCT (individually randomised)
Participants	**Setting:** patients awaiting transfer from a hospital to a long-term care facility, Australia N = 110 participants Sample: 61% women Age (years): mean 82.7 (SD 6.4) **Inclusion criteria:** acute and subacute hospital patients being transferred to nursing care facility; life expectancy greater than a month **Exclusion criteria:** none stated
Interventions	• Pharmacist transition coordinator for patients transferring from hospital to a care facility for the first time: medication management transfer summaries from hospitals, timely coordinated medication reviews by accredited community pharmacists, and case conferences with physicians and pharmacists • Control: usual hospital discharge process
Outcomes	• Number of people falling
Duration of the study	12 months. Participants followed up for 8 weeks post discharge
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The study biostatistician provided a computer-generated allocation sequence that used block randomization and was stratified by hospital."
Allocation concealment (selection bias)	Low risk	Quote: "Randomization was coordinated by a centralized hospital pharmacy service."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible

Crotty 2004a (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not clear whether staff recording falls were aware of existence of transfer summaries and case conferences
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: reasons for loss to follow-up similar between groups
Selective reporting (reporting bias)	Low risk	Judgement comment: falls were a secondary outcome measure.
Method of ascertaining falls	High risk	Judgement comment: no clear definition or staff training described
Baseline imbalance	Low risk	Judgement comment: no significant difference between groups at baseline
Other bias	Low risk	None identified.

Crotty 2004b

Methods	RCT (cluster randomised) Cluster randomisation of regions such that each metropolitan health area allocated to intervention or control. Facility in an intervention region selected at random and matched to a facility in a control region. Matching facilities not randomised
Participants	**Setting:** 20 residential care facilities (10 high- and 10 low-level care), Adelaide, Australia N = 715 participants. 20 clusters. Sample: 84% women Age (years): mean 84.1 (SD 7.8) **Inclusion:** none stated **Exclusion criteria:** none stated
Interventions	• Pharmacist outreach intervention: intervention physicians received two 30 minutes academic detailing visits from pharmacist based on evidence-based guidelines, audit of prescribing practice (psychotropic and/or antihypertensive medication, use of aspirin or warfarin) and number of falls in previous 12 months. One nurse per facility received four 2-hour education sessions (change management, management of the behavioural symptoms of dementia, medication management and falls prevention techniques). Pharmacist educated each facility on reducing use of psychotropic drugs • Usual care
Outcomes	• Number of people falling
Duration of the study	7 months
Notes	
Risk of bias	

Crotty 2004b (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "All randomisation was conducted using a computer-generated random allocation program by a person external to the project."
Allocation concealment (selection bias)	High risk	Cluster randomisation of regions. Facility in an intervention region selected at random and matched to a facility in a control region
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: outcome was assessed blind to group allocation but intervention facilities would have been aware of intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses to follow-up even between groups
Selective reporting (reporting bias)	Unclear risk	Judgement comment: injurious falls included as outcome measure but not reported
Method of ascertaining falls	High risk	Judgement comment: no clear definition of falls, no staff training
Baseline imbalance	Low risk	Judgement comment: adjusted for baseline differences.
Other bias	Low risk	None identified.

Cumming 2008

Methods	RCT (cluster randomised) Cluster randomisation of 12 matched pairs of wards
Participants	**Setting:** 24 acute and subacute wards in 12 hospitals, Sydney, Australia N = 24 wards, 3999 patients. 24 clusters. Sample: 59% women Age (years): mean 79.0 (SD 12.8) **Inclusion criteria:** all admitted patients **Exclusion criteria:** none stated

Interventions	• Targeted multifactorial intervention: a nurse and physiotherapist each worked for 25 hours per week for 3 months in all intervention wards. Provided risk assessment of falls, staff and patient education sessions, drug review, arranged walking aids, eyewear, modification of bedside and ward environments, increased supervision, liaison with staff about confusion and foot problems, an exercise programme, and sock alarms for selected patients (maximum of 2 per ward) who staff considered unsafe to walk unsupported • Usual care. No trial interventions. NB. Continuation of existing pre-trial falls prevention activities in control and intervention wards during the study
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures)
Duration of the study	3 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Randomisation of each matched pair of wards was usually done during the week before the study started for that pair of wards. Randomisation involved sealed, opaque envelopes and was supervised by a study investigator ... unaware of ward characteristics."
Allocation concealment (selection bias)	Low risk	Quote: "We included all patients in study wards during each three month study period." "Randomisation of each matched pair of wards was usually done during the week before the study started for that pair of wards. Randomisation involved sealed, opaque envelopes and was supervised by a study investigator ... unaware of ward characteristics."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the wards who recorded falls were likely to be aware of their ward's allocation status

Cumming 2008 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis.
Selective reporting (reporting bias)	Low risk	Judgement comment: outcomes reported as per trial registration
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and collected concurrently
Baseline imbalance	Low risk	Judgement comment: groups well-balanced at baseline.
Other bias	Unclear risk	Quote: "Another limitation is that some falls prevention activities were already occurring in control (and intervention) wards before the start of our study. These activities would have continued during the study period, making it more difficult to show any effect of our interventions." Judgement comment: some other falls prevention activities ongoing - impact of this unclear

da Silva Borges 2014

Methods	RCT (individually randomised)
Participants	**Setting:** residential care facilities, intermediate-level care, Brazil N = 59 Sample: NR Age (years): 68 **Baseline Characteristics** **Ballroom dancing programme** • *N* : 30 • *Age : mean (SD)* : 68 (8.33) • *Female (%)*: NR • *Medical status defined? -Y/N* : N • *Falls risk defined? -Y/N*: N • *Dependency defined? - Y/N*: Y (all functionally autonomous) • *Cognitive status defined? - Y/N*: Y **Control** • *N* : 29 • *Age : mean (SD)* : 67 (7.70) • *Female (%)*: NR • *Medical status defined? -Y/N* : N • *Falls risk defined? -Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y

	Inclusion criteria: resident of long-stay institution in Rio de Janeiro state, Brazil, functionally autonomous in ADL, had not engaged in any regular physical activity for at least three months **Exclusion criteria:** any condition that could prevent a participant from undergoing tests or interventions (such as cardiopathy, hypertension, uncontrolled asthmatic bronchitis, osteoarthritis, recent fracture, tendinitis, neurological problems and severe obesity, as well as the use of a prosthesis or medication that could cause attention disorders); cognitive impairment, especially memory function **Pretreatment differences:** unclear, baseline characteristics not reported
Interventions	• Ballroom dancing programme. Ballroom dancing with 10 minute warm-up with flexibility exercises and low-intensity dance movements, then higher-intensity rhythms for 30 minutes, then 10 minutes relaxation to music. 3 x 50-minute sessions weekly on alternate days. • Control. Normal daily activities. Advised not to engage in any regular physical therapy until after study period
Outcomes	• Analysis of falls outcome
Duration of the study	12 weeks
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "randomly allocated by simple draw" Judgement comment: unclear how the draw was conducted and whether or not this would result in a truly random sequence
Allocation concealment (selection bias)	Unclear risk	Judgement comment: allocation concealment not described.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Judgement comment: it is unclear who reported the falls data
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: to exclude patients due to falls may have a significant impact on falls data if these patients were multiple fallers or at high risk. Group allocation is not reported

Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol available.Falls data not published due to error in article
Method of ascertaining falls	High risk	Judgement comment: falls data were not recorded concurrently, it is unclear what type of medical records were accessed to confirm falls, this may not include records of non-injurious falls
Baseline imbalance	Unclear risk	Judgement comment: baseline characteristics of participants not reported
Other bias	Low risk	Judgement comment: none detected

Donald 2000

Methods	RCT (2 x 2 factorial design)
Participants	**Setting:** 1 elderly care rehabilitation (subacute) ward, Gloucester, UK N = 54 Sample: individuals admitted to one elderly care rehabilitation ward over an 8-month period (81% women) Age (years): mean 83 **Inclusion criteria:** patients admitted for rehabilitation **Exclusion criteria:** none stated
Interventions	• Assigned to ward area with vinyl floor covering and conventional physiotherapy (functional based physiotherapy, once or twice daily) • As above (1) plus seated leg strengthening exercises (hip flexors and ankle dorsiflexors • Assigned to ward area with carpet and conventional physiotherapy • As above (3) plus seated leg strengthening exercises (hip flexors and ankle dorsiflexors)
Outcomes	• Rate of falls • Number of people falling
Duration of the study	9 months. Follow-up of individual patients was duration of admission (mean length of stay 29 days)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Sequence generation not described. Quote: "Using randomized envelopes for each risk group, patients were as-

Donald 2000 (*Continued*)

		signed a floor group (carpet or vinyl) and a physiotherapy group (conventional physiotherapy or additional exercise)."
Allocation concealment (selection bias)	Unclear risk	Randomised achieved by randomising envelopes. Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcome assessors do not appear to have been blinded to treatment status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: high loss to follow-up but ITT analysis for falls outcomes
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified, but falls and fallers data reported completely
Method of ascertaining falls	Unclear risk	Judgement comment: falls clearly defined, but insufficient information on frequency of recording of falls data for judgement
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline.
Other bias	Low risk	Judgement comment: none identified.

Dyer 2004

Methods	RCT (cluster randomised)
Participants	**Setting:** 20 residential care homes (intermediate-level care), UK N = 196 participants. 20 clusters. Sample: 78% women Age (years): mean (SD) intervention group 87.4 (6.9), control group 87.2 (6.9) **Inclusion criteria** (facilities)**:** ≥ 5 residents; not specializing in mental illness; without nursing services. Inclusion criteria (residents): aged ≥ 60 **Exclusion criteria:** temporary residents or terminal illness
Interventions	• Multifactorial, multidisciplinary intervention: baseline assessments by physiotherapist, nurse and OT and interventions based on these. ○ Exercise: supervised gait, balance, co-ordination and functional + strength/resistance + flexibility + general physical exercises. 3 x 40-minute sessions per week for 3 months. Progressive exercises individually tailored and delivered by exercise assistants supported by physiotherapists. Carried out in groups or individually if residents unable to participate in groups because of frailty or cognitive impairment

	○ Staff education ○ Medical review: baseline assessments screened by geriatrician. Recommendations re medication review, orthostatic hypotension, and osteoporosis prevention sent to participant's GP for GP to implement ○ Environmental modification: OT assistant visited facilities to assess and report on falls hazards, with facilities being alerted of major hazards ○ Optician and podiatry referrals based on baseline assessment • Usual care, no intervention.
Outcomes	• Rate of falls • Number of people falling
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Allocation sequence used computer-generated random number tables
Allocation concealment (selection bias)	Low risk	Quote: "The allocation sequence was performed and kept secure by a researcher independent of the study, and blinded to baseline assessment results."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses balanced between groups.
Selective reporting (reporting bias)	Low risk	No protocol identified but all expected falls data comprehensively reported as falls, fallers, multiple fallers and fractures reported
Method of ascertaining falls	Low risk	Data collected concurrently and clear definition.
Baseline imbalance	Unclear risk	Differences in cognition and medications at baseline, unclear if adjusted for in analysis
Other bias	Low risk	None identified.

Dykes 2010

Methods	RCT (cluster randomised) randomised 2 units matched on fall rates and patient days within each of 4 hospitals
Participants	**Setting:** 8 acute medical units, Boston, Massachusetts, USA N = 5264 patients aged ≥ 65. 8 clusters. Sample used in this review: patients aged ≥ 65 (% women not available) Age (years): mean 78.8 (SD 8.4) in patients aged ≥ 65 **Inclusion criteria** (units): fall rates higher than institution's mean rate for previous year; had a match within the institution (unit with similar fall rate and length of stay). Inclusion criteria (patients): all patients admitted to randomised units during study **Exclusion criteria** (units): involved in other performance improvement efforts relating to fall prevention
Interventions	• Falls Prevention Tool Kit (FPTK) software with strategies to improve unit-level buy-in: Morse Falls Scale completed using FPTK; software automatically-generated evidence-based/feasible interventions, tailored by nurse based on knowledge of patient; software automatically printed bed poster for patients at risk (updated with change in status); software generates tailored handout to educate patient/family (updated with change in status); tailored fall prevention plan automatically generated by software for documentation • Control: usual care in relation to fall prevention: Morse Falls Scale (MFS) completed using existing paper or electronic forms; "high risk of falls" signs above beds for patients with MFS > 45 points; educate patient/families with booklets or other handouts as needed; document plan manually in paper or electronic record Both groups used Morse Falls Scale to assess risk of falls on admission, daily and with change in status
Outcomes	• Rate of falls • Number of people falling
Duration of the study	6 months
Notes	Data for participants aged < 65 and ≥ 65 reported separately in Dykes 2010. Only data for participants aged ≥ 65 included in this review

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote "Matched units were randomised" Insufficient information to permit judgement
Allocation concealment (selection bias)	High risk	At each hospital pairs of wards were allocated to intervention and control, then patients admitted to these wards were recruited

Dykes 2010 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: Not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "the intervention was not blinded and falls were reported by unit-based caregivers who implemented fall prevention interventions."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: all patients included in ITT analysis.
Selective reporting (reporting bias)	Low risk	Judgement comment: falls outcomes reported consistent with trial registration
Method of ascertaining falls	Low risk	Judgement comment: falls recorded concurrently and would be defined in hospital system
Baseline imbalance	Low risk	Judgement comment: no significant differences at baseline, potential confounders adjusted for
Other bias	Low risk	Judgement comment: none identified.

Faber 2006

Methods	RCT (individually randomised) Facilities randomised to one of two interventions, then residents individually randomised to intervention or control group within facilities
Participants	**Setting:** 15 long-term care residences (combined high- and intermediate-level care within each), the Netherlands N = 238 Sample: 79% women Age (years): mean 84.9 (range 63 to 98) **Inclusion criteria:** resident of facility **Exclusion criteria:** unable to walk 6 metres unaided; poor cognition as judged by staff; GP contraindication
Interventions	• Functional Walking (FW) (7 residences): 10 exercises (gait, balance, and coordination + strength/resistance), 1 session per wk for 4 weeks then 2 sessions per week for 16 weeks; 90 minutes per session. Exercises individually tailored and delivered by an instructor • In Balance (IB) (8 residences): 3D exercises (based on Tai Chi). 1 session per week for 4 weeks followed by 2 sessions per week for 16 weeks. 90-minute sessions. Exercises individually tailored and delivered by an instructor Usual care (same 15 residences as above)

Faber 2006 (*Continued*)

Outcomes	• Rate of falls • Number of people falling	
Duration of the study	12 months	
Notes	Only data for combined control groups reported in Faber 2006	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	15 centres cluster randomised to one of two exercise regimens using "sealed envelopes". Individuals then randomised into intervention and control within each participating centre using computer generated random numbers
Allocation concealment (selection bias)	Unclear risk	Unclear whether initial randomisation to clusters used envelopes which were sequentially numbered, opaque and sealed. Insufficient information to permit judgement in relation to randomisation of individuals after cluster allocation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: withdrawals balanced across interventions
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified but all expected outcomes - falls and fallers thoroughly reported
Method of ascertaining falls	Low risk	Judgement comment: falls defined and recorded concurrently.
Baseline imbalance	Low risk	Judgement comment: no differences at baseline.
Other bias	Low risk	Judgement comment: none detected

Flicker 2005

Methods	RCT (individually randomised)
Participants	**Setting:** 60 assisted living facilities and 89 nursing homes (intermediate- and high-level nursing care facilities), urban and rural Australia N = 693 Sample: 95% women Age (years): mean 83.4 **Inclusion criteria:** serum 25-hydroxyvitamin D between 25 nmol/L and 90 nmol/L **Exclusion criteria:** use of medications affecting bone and mineral metabolism; thyrotoxicosis within 3 years; primary hyperparathyroidism treated within 3 years; multiple myeloma; Paget's disease of bone, history of malabsorption, intercurrent active malignancy, other disorders affecting bone and mineral metabolism
Interventions	• 10,000 IU oral ergocalciferol (vitamin D2) weekly (or 1000 IU oral ergocalciferol daily) plus 600 mg calcium carbonate daily • Placebo + 600 mg calcium carbonate daily
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures) • Adverse events
Duration of the study	24 months
Notes	58% of participants had a serum vitamin D between 25 nmol/L and 40 nmol/L at baseline

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Subjects were randomized via computer-generated lists," "Within each institution … in blocks of eight."
Allocation concealment (selection bias)	Low risk	Quote: "Subjects were randomized to receive sequentially numbered bottles containing vitamin D supplementation or placebo." Individual not involved in contact with subjects or facilities performed randomisation
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: double-blind
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Residential staff recording falls events blinded to whether participants were receiving vitamin D or placebo

Flicker 2005 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analyses performed.
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified but falls reported extensively as number of falls, fallers, fracture and ITT, raw and adjusted and additional analyses
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified

Frankenthal 2014

Methods	RCT (individually randomised)
Participants	**Setting:** 1 residential care facility, mixed level of care, Israel N = 359 residents Sample: 67% female, 46.8% 84 or over Age (years): mean 82.7 (SD 8.7) **Baseline Characteristics** **Medication intervention (STOP/START)** • *N*: 183 • *Age - mean (SD)* : Overall 82.7 (8.7) • *Female - N (%)*: 129 (70.5%) • *Medical status defined? - Y/N*: Y • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y **Control** • *N*: 176 • *Age - mean (SD)* : Overall 82.7 (8.7) • *Female - N (%)*: 110 (62.5%) • *Medical status defined? - Y/N*: Y • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y **Inclusion criteria:** all residents aged 65 and older in a chronic care geriatric facility in Israel, prescribed at least one daily medicine **Exclusion criteria:** terminally ill residents, those whose stay in the facility was shorter than 3 months **Pretreatment differences:** no significant differences

Interventions	• Medication review by pharmacist with Screening Tool of Older Persons potentially inappropriate Prescriptions/Screening Tool to Alert doctors to Right Treatment (STOPP/START). Pharmacist made recommendations to chief physician who decided whether to implement changes. Review at study opening, 6 and 12 months later. • Control. No interventional recommendations made by pharmacist to chief physician.
Outcomes	• Number of falls
Duration of the study	12 months
Notes	24 month follow-up data reported as retrospective cohort data for those alive at 24 months. These data not considered eligible for inclusion in the review

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: simple list generation. Fixed stratified randomisation - level of independence and cognition levels
Allocation concealment (selection bias)	Low risk	Judgement comment: physician who were not involved in the study did randomisation. Use of sealed envelopes. Study pharmacist (main person determining intervention recommendations) not involved in allocation, but aware of group allocation after randomisation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: pharmacist was aware of group allocation when making recommendations and implementing intervention group recommendations. Was also aware of control group medication use as well, as recommendations were made but not implemented for this group
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Nurses who were unaware of participants' group assignments assessed the outcome measures in the study population. The chief nurses routinely report falls, hospitalizations, and FIM in residents' records."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: similar loss to follow-up between groups
Selective reporting (reporting bias)	Low risk	Judgement comment: protocol registered and outcome measures are reported as per protocol

Frankenthal 2014 (*Continued*)

Method of ascertaining falls	Low risk	Judgement comment: clear definition, concurrent reporting by nurses
Baseline imbalance	Low risk	Judgement comment: no significant difference on main reported baseline measures
Other bias	Low risk	Judgement comment: none detected

Fu 2015

Methods	RCT (individually randomised)
Participants	**Setting:** 1 residential care facility, China N = 60 Sample: 65% women Age (years): 82 **Baseline Characteristics** **Wii Exercise** • *N*: 30 • *Age - mean (SD)* : 82.3 (4.3) • *Female - N (%)*: 20 (67) • *Medical status defined? - Y/N*: N • *Falls risk defined? - Y/N*: Y (PPA) • *Dependency defined? - Y/N*: Y (FAC) • *Cognitive status defined? - Y/N*: N **Conventional exercise** • *N*: 30 • *Age - mean (SD)* : 82.4 (3.8) • *Female - N (%)*: 19 (63) • *Medical status defined? - Y/N*: N • *Falls risk defined? - Y/N*: Y • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: N **Inclusion criteria:** 65 years and older, living in a nursing home, Functional Ambulation Category (FAC) grade 2 or 3, alert, medically stable and able to follow instructions, history of falls in the previous year **Exclusion criteria:** visual problems that might affect their training, unable to follow instructions, history of seizure, stroke, parkinsonism, or uncontrolled cardiovascular disease **Pretreatment differences:** no important differences between groups on a wide range of potential confounders
Interventions	• Exercise using a Wii Fit balance board to perform three balance training games: Soccer Heading, Table Tilt, and Balance Bubble. Tasks became progressively more difficult with improvements in performance. 1-hour sessions, 3 sessions a week • Usual care. Conventional exercise: balance exercise regimen consisting of: lower limb strengthening; tandem standing, tandem walking, sideways and turnaround walking exercises in parallel bars; stepping exercise; sitting to standing exercise; and

Fu 2015 (*Continued*)

	half-squats (Otago balance programme). 1-hour sessions, 3 sessions a week.	
Outcomes	• Number of falls	
Duration of the study	6 weeks	
Notes		

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Participants were randomly assigned to the conventional or Wii Fit balance training group by using a random number produced by the computerized method of minimization"
Allocation concealment (selection bias)	Unclear risk	Judgement comment: allocation not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not possible given nature of intervention
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Nurses at the nursing home who documented falls were unaware of participants' group allocation."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: less than 10% missing from each group.
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol available
Method of ascertaining falls	Low risk	Judgement comment: falls were recorded by the nursing staff according to a clear definition and reported to the investigator for each participants monthly over the 12-month period after randomisation
Baseline imbalance	Low risk	Quote: "There was no statistically significant difference in age, sex, height, weight, body mass index, FAC distribution, or number of falls in the previous year between the 2 groups."
Other bias	Low risk	Judgement comment: none identified

Garcia Gollarte 2014

Methods	RCT (cluster randomised)
Participants	**Setting:** residential care facilities, mixed-level care, 60 physicians, Spain N = 1018 residents. 59 physicians, 37 nursing home clusters. Sample: 73% women. Age (years): 84.4 (SD 12.7) **Baseline Characteristics** **Educational intervention** • *N*: 516 • *Age - mean (SD)*: 84.24 (14.6) • *Female (%)*: 382 (74.0) • *Medical status defined? - Y/N* : No • *Falls risk defined? - Y/N*: No • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y **Control** • *N*: 502 • *Age - mean (SD)*: 84.5 (10.4) • *Female (%)*: 362 (72.1) • *Medical status defined? - Y/N* : N • *Falls risk defined? - Y/N*: No • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y **Inclusion criteria:** facilities: owned by the same private company in Spain; Physicians: at included nursing homes Residents: older than 65 years; living in nursing home for at least 3 months; expected to stay for 6 months or longer; clinically stable (no changes in prescription in the last 2 months); accepted that their clinical data were used for the study **Exclusion criteria:** residents: receiving palliative care; usually cared by other primary care providers outside the nursing home **Pretreatment differences:** significant difference in Barthel index at baseline P = 0.003, indicated made no difference to results but methods of adjustment not reported
Interventions	• Educational intervention. Structured educational intervention directed to nursing home physicians in reducing inappropriate prescription and improving health outcomes and resource utilisation. 10 hours educational programme, on demand support by phone for 6 months. • Control. No intervention or information about an educational intervention
Outcomes	• Number of falls (post-intervention) • Number of fallers (post-intervention)
Duration of the study	12 months total, 6 months intervention period. Baseline recorded following 3 months pre-intervention. Endpoint at 12 months, for 3 months post-intervention
Notes	
Risk of bias	

Garcia Gollarte 2014 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomization was done using random number tables and" Judgement comment: random number tables.
Allocation concealment (selection bias)	Unclear risk	Judgement comment: details of timing of individual participant recruitment/person recruiting not reported (i.e. whether completed before cluster randomisation or not)
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: physicians were blinded to purpose of trial. Unclear if participants were blinded but unlikely to be aware of educational interventions of physicians
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Judgement comment: do not know who did outcome assessment or how
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: loss of one nursing home cluster after randomisation
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol available. Falls only reported for pre and post-intervention periods
Method of ascertaining falls	High risk	Quote: "We did not use a daily systematic registry of falls and delirium, therefore, some episodes may have gone unnoticed, as is suggested by our lower rates of both syndromes compared with similar studies."
Baseline imbalance	Unclear risk	Judgement comment: significant difference in Barthel index at baseline. Results indicate that adjusting for this imbalance made no difference in results, however no details of how adjustment was performed are provided
Other bias	Low risk	Judgement comment: none identified

Grieger 2009

Methods	RCT (individually randomised)
Participants	**Setting:** 1 aged care facility (high and intermediate-level care), Victoria, Australia N = 115 Sample: 65% women in analysis Age (years): not stated **Inclusion criteria:** able to consume food orally **Exclusion criteria:** residents in the dementia, rehabilitation and palliative care wards

Interventions	• One multivitamin tablet (Heron Women's Multivitamin) daily for 6 months. Tablets included 400 IU vitamin D3 and 360 mg calcium carbonate. • Control: one placebo tablet daily for 6 months
Outcomes	• Rate of falls • Number of people falling • Adverse events Other outcomes not included in this review
Duration of the study	6 months
Notes	Mean baseline serum vitamin D level 36 nmol/L

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Random number generator used in Excel
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement comment: double-blind
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Double-blind administration of tablets but no mention of maintaining blinding of researchers when falls were extracted from medical histories at the end of the 6-month trial
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: large loss from groups as randomised. 25% loss as randomised from placebo group, 16% from intervention group
Selective reporting (reporting bias)	High risk	Judgement comment: excluded multiple faller from number of falls data as outlier
Method of ascertaining falls	High risk	Judgement comment: falls not clearly defined
Baseline imbalance	Unclear risk	Judgement comment: baseline age, cognition, medical comorbidities not reported
Other bias	Low risk	Judgement comment: none identified

Haines 2004

Methods	RCT (individually randomised)
Participants	**Setting:** one hospital (three subacute wards), specialising in rehabilitation and care of elderly patient, sMelbourne, Australia, N = 626 Sample: 67% women Age (years): mean 80 (SD 9) **Inclusion criteria:** all patients admitted to three subacute wards **Exclusion criteria:** none stated
Interventions	• Targeted falls risk prevention programme based on identified falls risk (Peter James Centre Falls Risk Assessment Tool) in additional to usual care. Potential interventions were: ○ supervised exercise programme: 45-minute sessions 3 x per week from commencement of intervention until discharge. Exercises comprised gait, balance and coordination + strengthening/resistance + 3D (Tai Chi). Exercises were individually tailored. Exercises were delivered by physiotherapist ○ falls risk alert card ○ up to four educational sessions from OT at bedside to individual participants of up to 30-minute duration ○ hip protectors • Usual care. Received usual care but none of the interventions from the falls prevention programme. Staff completed risk assessment and generated recommendations these recommendations were not instituted.
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures) • Adverse events
Duration of the study	10 months recruitment. Follow-up time was until participants were discharged from hospital
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "We randomly allocated participants by using a random number table held at the centre by one investigator (TPH) who revealed allocation on receipt of written consent."
Allocation concealment (selection bias)	Unclear risk	See above. Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: unblinding likely

Haines 2004 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Staff recorded falls on incident report forms likely to be aware of individual's allocation status. Survey of staff indicated they were relatively unaware of participant group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis
Selective reporting (reporting bias)	Low risk	Judgement comment: all outcome measures reported
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: groups similar at baseline
Other bias	Low risk	Judgement comment: none identified.

Haines 2010

Methods	RCT (cluster randomisation of pairs of hospital wards matched on rate of falls in preceding 6 months)
Participants	**Setting:** 18 publicly funded hospital wards (acute and subacute), Queensland, Australia N = 11,099 patients. 18 clusters. Sample: patients admitted to study wards after October 2007 when beds provided to intervention wards (% women not stated) Age (years): not stated **Inclusion criteria:** no previous access to or provision of low-low beds **Exclusion criteria:** none described
Interventions	• Low-low beds: provision of one low-low bed for every 12 beds on a hospital ward. Lowered bed height 28.5 cm from the ground, highest bed height 64 cm. Written guidance on their use and for prioritising patients at greatest risk of falls • Control: usual care Staff on intervention and control wards received falls incident reporting training video
Outcomes	• Rate of falls
Duration of the study	6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement

Haines 2010 (*Continued*)

Random sequence generation (selection bias)	Low risk	Quote: "...18 wards were then matched into pairs ... and ordered alphabetically within pairs. A research assistant in a separate location and blinded to this ordering flipped a coin to determine whether the first or second listed ward in the pair was to be allocated to the intervention group."
Allocation concealment (selection bias)	Unclear risk	See above, but patients could have been allocated to a specific ward with the knowledge that it was an intervention or control ward
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Falls recorded by ward staff using routine computer-based incident reporting scheme. Would not be blind to allocation. No mention of blinding in relation to the person extracting data from centrally held database
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis performed.
Selective reporting (reporting bias)	Low risk	Quote: "(ANZCTR registration number: 12609000243213)." Judgement comment: all outcome measures reported
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	High risk	Judgement comment: patient level characteristics at baseline not reported. Intervention wards included 2 stand-alone acute medicine wards, no standalone acute medicine wards in control arm
Other bias	Low risk	Judgement comment: none identified.

Haines 2011

Methods	RCT (individually randomised)
Participants	**Setting:** 6 acute and subacute wards in 2 hospitals, Brisbane and Perth, Australia N = 1206 Sample: patients admitted to acute (orthopedic and acute-respiratory medicine) and subacute (geriatric assessment and rehabilitation) wards of one hospital, and to the acute (medical-surgical) and subacute (restorative-stroke rehabilitation) wards of a second hospital (53% women) Age (years): mean (SD) intervention group (complete programme) 75.3 (11.0), intervention group (materials only programme) 74.7 (11.7), control group 75.3 (10.1) **Inclusion criteria:** aged > 60; expected to stay at least 3 days (acute wards only) **Exclusion criteria:** medically too unwell; previously participated in the trial
Interventions	• Complete programme: multimedia patient education programme involving written and video-based materials combined with physiotherapist follow-up • Materials only programme: multimedia patient education materials without physiotherapist follow-up • Control: usual care
Outcomes	• Rate of falls • Number of people falling • Adverse events
Duration of the study	22 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "a computer-generated random allocation sequence"
Allocation concealment (selection bias)	Low risk	Quote: "opaque, consecutively numbered envelopes"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: participants not blinded, blinded assessment but treatment providers not blinded
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "research assistants ... completed weekly falls reviews ... were blind to group allocation"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis, no loss.
Selective reporting (reporting bias)	Low risk	Judgement comment: falls reported as per publication, To check ACTRN12608000015347

Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: baseline characteristics similar.
Other bias	Low risk	Judgement comment: none identified

Healey 2004

Methods	RCT (cluster randomised by ward in matched pairs)
Participants	**Setting:** 8 elderly care wards (acute and subacute) in 1 hospital, York, UK N = 1654 participants, 32,528 bed days during intervention. 8 clusters. Sample: approximately 60% women Age (years): mean 81.3 (range 63 to 102) **Inclusion criteria:** all patients admitted to target wards **Exclusion criteria:** none specified
Interventions	• Targeted risk factor reduction care plan for patients with a history of falls or a near fall during admission. Based on assessment (and subsequent referral/action) relating to: eyesight (referral to ophthalmologist); medications check for sedatives, anti-depressants, diuretics, polypharmacy, etc (medical review of benefit vs harm); lying and standing blood pressure (advice to participant and referral to medical staff); ward urine test (mid-stream urine if positive for nitrites, blood or protein); difficulty with mobility (referral to physiotherapist); review of bed rail use; footwear safety (advice on replacement); bed height (kept at lowest height); position in ward (placing high risk patients near nurses' station); environmental causes (act to correct); nurse call bell (explained and in reach) • Usual care. Managers on control wards were made aware of the study, and the need not to introduce the care plan in their area. Control wards made no other changes to practice or environment relevant to falls prevention during the study. Whilst nurses instigated the process, remedial interventions were multi-disciplinary, including mobility assessment by physiotherapists and medication review by medical staff.
Outcomes	• Rate of falls
Duration of the study	6 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Method of randomisation not described. Quote: "The study wards were divided into matched pairs. In each pair, one ward was

Healey 2004 (*Continued*)

		randomly allocated to control or intervention by lottery ..."
Allocation concealment (selection bias)	Unclear risk	Individual study wards aware of their allocation from beginning of study. It is unclear whether knowledge of group status could have influenced admission of new patients during the study
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the wards who recorded falls were likely to be aware of their ward's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis. all occupied bed days and falls analysed, unlikely to be loss in hospital
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified.
Method of ascertaining falls	High risk	Judgement comment: no definition of falls used. Used accident and incidence reporting forms
Baseline imbalance	High risk	Judgement comment: imbalance in length of stay and dementia diagnoses
Other bias	Low risk	Judgement comment: none detected.

Hill 2015

Methods	RCT (cluster-randomised, stepped-wedge)
Participants	**Setting:** 24 wards in 8 rehabilitation or geriatric evaluation and management units in Australian hospitals, Western Australia N = 3606 admissions; 3121 unique patients. 24 clusters. Sample: 62% women Age (years): 82 **Baseline Characteristics** **Individualized fall education programme** • *N* : 1623 admissions, 1402 unique patients • *Age - mean (SD)* : 81.4 (9.3) • *Female - N (%)*: 999 (62%) • *Medical status defined? - Y/N* : Y

	• *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive impairment defined? Y/N*: Y **Usual care** • *N* : 1983 admissions, 1719 unique patients • *Age - mean (SD)* : 82.1 (8.3) • *Female - N (%)*: 1211 (61%) • *Medical status defined? - Y/N* : Y • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive impairment defined? Y/N*: Y **Inclusion criteria:** for individuals on units to receive intervention: aged more than 60 years, projected length of stay of at least 3 days, basic cognitive functioning (MMSE > 23/30 and AMTS > 7/10), when the treating clinical team judged that the patient had a high enough level of cognition to benefit from the education **Exclusion criteria:** for individuals on units not to receive intervention: diagnosis of delirium, patients with moderate or severe cognitive impairment (MMSE of less than 24/30 or AMTS of less than 8/10), permanently unable to mobilise and remain bed-bound or are receiving palliative care **Pretreatment differences:** significant difference in comorbidities at baseline (more co-morbidities in intervention period), but confounding adjusted for in analysis
Interventions	• Individualised fall education programme. Safe Recovery programme for patients and staff. For patients, an individually-tailored multimedia falls prevention education package (DVD and workbook) with further face to face follow-up education (including workbook completion and goal setting) with a health professional was provided. Aimed to alert patients to their personal risk of falls, raise their knowledge about falls epidemiology and falls prevention, and to motivate them to engage in falls-prevention strategies. Patients were eligible to receive the individualised education if they were aged more than 60 years, had a projected length of stay of at least 3 days, had basic cognitive functioning, and when the treating clinical team judged that the patient had a high enough level of cognition to benefit from the education. Basic cognition was defined as having a Mini-Mental State Examination (MMSE) score of more than 23/30 or an Abbreviated Mental Test Score (AMTS) of more than 7/10.Staff education on the programme and feedback about patients' goals and perceived barriers, plus unit managers receive feedback on perceived barriers. Patient education sessions ranged between 15 and 35 minutes with 1-4 sessions per patient. Staff training in the week of the start of the intervention on their unit and feedback to staff weekly, 56% of patients in the intervention arm were eligible to receive the intervention based on their cognitive status. • Usual care. Usual care includes patient's screening, assessment and implementation of individualised falls prevention strategies, ongoing staff training and environmental strategies.
Outcomes	• Ratio ratio • Odds of falling (per admission) • Number of fractures • Adverse events

Duration of the study	50 weeks. After a 10-week control period, two units started the intervention-this procedure continued at 10-week intervals until all eight units had crossed over into the intervention period
Notes	Outcomes reported for subgroups by level of cognition. Stable median site control falls rate and absence of interaction effect of time and falls outcomes indicates confounding by seasonal effects unlikely

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "computer-generated, random allocation sequences." Judgement comment: computer generated.
Allocation concealment (selection bias)	Low risk	Judgement comment: allocation concealed, no individual participant recruitment required
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls collected by staff who are blinded, but entered into hospital report systems by unit staff who were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: no loss to follow-up.
Selective reporting (reporting bias)	Low risk	Judgement comment: protocol available, outcome measures consistent with final report
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and collected concurrently
Baseline imbalance	Low risk	Judgement comment: imbalances at baseline adjusted for in analyses
Other bias	Unclear risk	Judgement comment: stepped-wedge design means there is a potential for bias due to systematic influence of other external factors during the conduct of the trial. Possible influence of seasonal trends addressed by pre-specified statistical analysis

Houghton 2014

Methods	RCT (cluster randomised)
Participants	**Setting:** 31 residential care facilities, mixed-level care, UK N = 953 residents. 31 clusters. Sample: 76% women Age (years): 87 **Baseline Characteristics** **Medication review** • *N*: 381 • *Age - mean (SD)*: 88.4 (6.5) • *Female - N (%)*: 303 (79.5%) • *Medical status defined? - Y/N* : N • *Falls risk defined? - Y/N (at baseline with validated tool)*: N • *Dependency defined? - Y/N*: N • *Cognitive status defined? - Y/N*: Y **Control** • *N*: 445 • *Age - mean (SD)*: 86 (8.5) • *Female - N (%)*: 324 (72.8%) • *Medical status defined? - Y/N* : N • *Falls risk defined? - Y/N (at baseline with validated tool)*: N • *Dependency defined? - Y/N*: N • *Cognitive status defined? - Y/N*: Y **Inclusion criteria:** care homes: average age > 65, registered with GP in local area; registered with Care Quality Commission for at least 6 months **Exclusion criteria:** care homes specifically for people (of all ages) with learning disability, sensory impairment, mental health problems, physical disabilities and alcohol dependence; if have received a medication review service from the Primary Care Trust in the last 6 months; if they receive the services of a community geriatrician; or if they are subject to investigation of the safeguarding of vulnerable adults.Residents: those who self-medicate; those in respite care **Pretreatment differences:** nil significant
Interventions	• Medication review. Multi-professional medication review service (MMRS): a meeting involving a clinical pharmacist and pharmacy technician from the Primary Care Trust Medicines Management Team, care home staff and GP(s) responsible for the medical care of residents. Review conducted twice: at baseline (approx 1 month) and 6 months. Each meeting considers 15 residents on average and lasts up to 2 hours, multiple meetings as necessary. • Usual care (support from the NHS).
Outcomes	• Rate ratio
Duration of the study	6 months intervention, follow-up to 12 months.
Notes	ISRCTN90761620 CAREMED trial
Risk of bias	

Houghton 2014 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "For practical (i.e. workload) reasons, consenting homes will be allocated to intervention or control sequentially after consent is obtained using minimisation." Judgement comment: Sequential allocation by minimisation is equivalent to being random
Allocation concealment (selection bias)	Unclear risk	Judgement comment: insufficient information for judgement.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: staff were involved in medication review meetings so were not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by staff who were not blinded as they were involved in medication review meetings
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: 1 care home lost from intervention group, reason unclear, unclear if accounted for in analysis
Selective reporting (reporting bias)	Unclear risk	Method of analysis of falls data as provided by author unclear. Unsure if falls analysed using a linear mixed model as per published abstract, impact unclear
Method of ascertaining falls	Unclear risk	Insufficient information for judgement.
Baseline imbalance	High risk	Judgement comment: higher number of participants requiring nursing care in control group
Other bias	Low risk	Judgement comment: n0ne detected.

Huang 2016

Methods	RCT (individually randomised)
Participants	**Setting:** 6 residential care facilities, mixed-level care, Taiwan N = 80 Sample: 50% women Age (years): 79.4 **Baseline Characteristics:** **Cognitive behavioural alone** • *N*: 27 • *Age - mean (SD)* : 77.9 (7.3) • *Female N (%)*: 16 (59.3) • *Medical status defined? - Y/N*: Yes (medications, No chronic disease)

	• *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Yes • *Cognitive status defined? - Y/N*: Yes **Cognitive behavioural plus exercise** • *N*: 27 • *Age - mean (SD)* : 79.1 (6.9) • *Female N (%)*: 13 (48.1) • *Medical status defined? - Y/N*: Yes • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Yes • *Cognitive status defined? - Y/N*: Yes Usual care • *N*: 26 • *Age - mean (SD)* : 81.3 (5.4) • *Female N (%)*: 11 (42.3) • *Medical status defined? - Y/N*: Yes • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Yes • *Cognitive status defined? - Y/N*: Yes **Inclusion criteria:** 65 years or over; MMSE 13 or over; ability to communicate in Mandarin or Taiwanese; Ability to ambulate independently or with an assistive device; CB group needed to complete all 8 sessions **Exclusion criteria:** unstable physical condition or evidence of end stage terminal disease **Pretreatment differences:** no significant group differences
Interventions	• Cognitive-behavioural intervention adapted for a Fear of Falling Management Model, with a focus on falls risk reduction, conducted by trained facilitator. 8 weekly sessions of 20 to 25 minutes, in groups of 6 to 8. • Cognitive-behavioural intervention plus a supervised strength and balance exercise programme, twice a week for approx 30 minutes. • Usual care
Outcomes	• Number of falls.
Duration of the study	8-month trial: 8 weeks intervention, falls over monitored over 3 months pre-intervention and 3 months post-intervention
Notes	80 participants randomised, 5 withdrew during the study, final sample =75 participants

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "used a computer-developed table to randomise patient assignment to each of the three groups in each nursing home."
Allocation concealment (selection bias)	Low risk	Quote: "Allocation was concealed from the recruiting RA."

Huang 2016 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: unable to blind participants/ personnel.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote "To achieve greater accuracy in the number of falls during the study period, we collected data from chart record, accident report, in charge staff, and participants." Judgement comment: falls were recorded by participants and staff who were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: little missing data from randomisation, and are balanced across groups
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Low risk	Quote: "we collected data from chart record, accident report, in charge staff, and participants." Judgement comment: Quote " The number of falls was recorded using the Falls Record Checklist (Huang & Acton, 2004)" - this is a checklist for concurrent recording of falls by participants
Baseline imbalance	Low risk	Judgement comment: no imbalance at baseline
Other bias	Low risk	Judgement comment: none detected

Imaoka 2016

Methods	RCT (individually randomised.)
Participants	**Setting:** residential care facility, high-level care, Japan N = 91 Sample: 76% women Age (years): 84.8 (SD 8.8) **Baseline Characteristics** **Usual care group** • *N phase 1*: 23 • *Age: mean (SD)* : 82.5 (10.9) • *Female (%)*: 15 (65%) • *Medical status defined? (Y/N)*: N • *Falls risk defined?(Y/N)*: N • *Dependency defined? (Y/N)*: Y • *Cognitive status defined? (Y/N)*: Y Reduced exercise group • *N phase 1*: 22 • *Age: mean (SD)* : 82.6 (9.1)

	• *Female (%)*: 16 (73%) • *Medical status defined? (Y/N)*: N • *Falls risk defined?(Y/N)*: N • *Dependency defined? (Y/N)*: Y • *Cognitive status defined? (Y/N)*: Y **Nutrition group** • *N phase 1*: 23 • *Age: mean (SD)* : 84.6 (7.7) • *Female (%)*: 20 (87%) • *Medical status defined? (Y/N)*: N • *Falls risk defined?(Y/N)*: N • *Dependency defined? (Y/N)*: Y • *Cognitive status defined? (Y/N)*: Y **Multifactorial group** • *N phase 1*: 23 • *Age: mean (SD)* : 87.6 (6.5) • *Female (%)*: 18 (78%) • *Medical status defined? (Y/N)*: N • *Falls risk defined?(Y/N)*: N • *Dependency defined? (Y/N)*: Y • *Cognitive status defined? (Y/N)*: Y **Inclusion criteria:** residents of long-term health facility, not received any regular supplementation of vitamin D during the previous 12 months **Exclusion criteria:** receiving terminal care; with renal failure (chronic kidney disease [CDK] stage 3 or an estimated glomerular filtration rate [eGFR] of G2 or poorer); poor glycaemic control; a pacemaker **Pretreatment differences:** nil significant
Interventions	• Usual care: advice on environmental adaptations, falls prevention education for staff, care conference, selection of walking aids, plus undividualised exercise (gait, balance, strength, resistance) and group exercise (warm-up exercises, sit-to-stand, balance and resistance). Two sessions of individualised exercise for 20 minutes per week, and group resistance exercise for 30 minutes per week. 1-hour education to staff. Caregiver's conference. Assessment and trial of walking aid by physical therapist. Compared to nutrition vitamin D group, this is direct comparison of individual and group exercise to vitamin D. Compared to 'multifactorial group', this is direct comparison of vitamin D to group exercise alone. • Reduced exercise group. Same as usual care including individualised exercise only without group resistance exercise, plus other usual care interventions. Two sessions of individualised exercise for 20 minutes per week. 1 hour education to staff. Caregiver's conference. Assessment and trial of walking aid by physical therapist. Compared to usual care, inverting the ratios provides a evidence on effectiveness of additional group exercise. • Nutrition group. Administered oral vitamin D (900 IU/day) as Isocal jelly PCF (500 IU) and a supplement (400IU vitamin D3). Jelly vitamins were eaten at lunchtime and supplements were taken after dinner. • Multifactorial group. Low level of exercise (individualised but not group exercise) and vitamin D supplementation 900IU/day. Two sessions of individualised exercise for 20 minutes per week. 1-hour education to staff. Caregiver's conference. Assessment and

Imaoka 2016 (*Continued*)

	trial of walking aid by physical therapist.
Outcomes	• Number of fallers • Hazard ratio for falling
Duration of the study	3 months intervention, follow-up to 9 months. Outcomes data exclude the intervention period
Notes	Effect of group exercise presented by comparing 'usual care' to 'reduced exercise' group

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: envelope drawn
Allocation concealment (selection bias)	Low risk	Judgement comment: opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: outcome assessors not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss generally balanced between groups.
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified.
Method of ascertaining falls	Low risk	Quote: "Falls were carefully recorded by the staff who found a resident falling down." Quote: "Falls were defined according to the International Classification of Diseases."
Baseline imbalance	Low risk	Judgement comment: no significant differences at baseline.
Other bias	Low risk	Judgement comment: none identified

Irez 2011

Methods	RCT
Participants	**Setting:** 1 residential care facility in Ankara, Turkey, intermediate-level care N = 60 Sample: 100% women Age (years): 75.4 **Baseline Characteristics** **Exercise - Pilates** • *N*: 30 • *Age - mean (SD)* : 72.8 (6.7) • *Female - N (%)*: 30 (100%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 30 • *Age - mean (SD)* : 78.0 (5.7) • *Female - N (%)*: 30 (100%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Inclusion criteria:** female, healthy, over 65 years of age, and have been relatively sedentary (undertaking no leisure time physical activity or less than 30 minutes of physical activity per day) for at least a year **Exclusion criteria:** male, significant general health problem or orthopaedic problem that would keep them from fully participating in the intervention protocol and/or the inability to attend at least 80% of the training sessions **Pretreatment differences:** intervention group younger. Falls risk factors not reported
Interventions	• Exercise - Pilates. The first part (4 weeks) consisted of mat exercises (Pilates, 2001), in the second part, Thera-Band elastic resistance exercises were added, and in the third part, the participants performed Pilates ball exercises for beginners. Classes led by certified Pilates instructor.Sessions 60 minutes, 3 days per week • Usual care. Instructed to refrain from beginning a new exercise programme or changing their current activity levels during this time period.
Outcomes	• Mean number of falls
Duration of the study	12 weeks
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Judgement comment: method of randomisation not described

Irez 2011 (*Continued*)

Allocation concealment (selection bias)	Unclear risk	Judgement comment: no information provided.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by study participants who could not be blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: loss to follow-up not reported
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified.
Method of ascertaining falls	High risk	Judgement comment: patient reported falls, calendars collected monthly
Baseline imbalance	High risk	
Other bias	Low risk	Judgement comment:nNone identified

Jarvis 2007

Methods	RCT (individually randomised)
Participants	**Setting:** 1 elderly care rehabilitation ward (subacute), Leicester, UK N = 29 Sample: 100% women Age (years): not stated **Inclusion criteria:** female patients admitted for rehabilitation **Exclusion criteria:** acute stroke; Parkinson's disease; Abbreviated Mental Test Score ≤ 5; severe cardiac, lung or kidney disease; severe osteoarthritis or rheumatoid arthritis
Interventions	• Intervention group: physiotherapy x 10 sessions per week. Once a week physiotherapy treatment at home after discharge. 8-week intervention • Control group: physiotherapy x 3 sessions per week. Some seen 1 x per week in day hospital or no treatment after discharge. 8-week intervention Physiotherapy consisted of stretches, lower limb exercises, and balance and gait activities in both groups
Outcomes	• Rate of falls • Number of people falling
Duration of the study	8 weeks
Notes	

Jarvis 2007 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "... randomly assigned, using sealed envelopes ..." Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'
Allocation concealment (selection bias)	Unclear risk	Quote: "The elderly women fallers were randomly assigned, using sealed envelopes, to either a control group or intervention group." Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Physiotherapy team responsible for measurement of outcomes reported to be blinded of intervention. Some chance of unblinding of assessors
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: large loss to follow-up; 28.6% dropout in intervention arm
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified

Jensen 2002

Methods	RCT (cluster randomised)
Participants	**Setting:** 9 residential care facilities (intermediate care), Umeå, Sweden N = 402. 9 clusters. Sample: 72% women Age (years): mean (range) intervention group 83 (65 to 97), control group 84 (65 to100) **Inclusion criteria:** facilities with ≥ 25 residents; residents aged ≥ 65 **Exclusion criteria:** none stated
Interventions	• Multidisciplinary programme including general and resident-specific tailored interventions for 11 weeks: supervised exercises, medication review, modifying

	environmental hazards, supplying and repairing aids, hip protectors, education of staff, post fall problem solving conferences and staff guidance. Individually tailored supervised exercises (gait, balance, coordination and functional + strength/resistance) 2 to 3 x per week. Intervention delivered by registered nurses, physician and physiotherapists • Usual care. Physiotherapist tasks unchanged, no hip protectors provided, no systematic fall-related problem-solving conferences or major fall-related environmental modifications
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (hip fracture)
Duration of the study	34-week follow-up
Notes	Eight extra physiotherapists employed for intervention period (a total of 200 hours/week) and three during the follow-up period (total of 10 hours/week)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Cluster-randomised study in nine facilities, divided into groups A and B (control or intervention). Quote: "Two sealed, dark envelopes" were used. Carried out by a person not connected with the study. Insufficient information to permit judgement
Allocation concealment (selection bias)	Low risk	Randomisation achieved by using by sealed dark envelopes by a person with no knowledge of study. Particiating individuals underwent baseline assessment prior to the randomisation of facilities
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Large loss to follow-up but loss is balanced and all patients included in fall analysis until lost
Selective reporting (reporting bias)	Low risk	No protocol identified but all expected falls outcomes reported: Falls, fallers, IRR and injuries reported and adjusted for clustering
Method of ascertaining falls	Low risk	Falls recorded concurrently with clear definition

Baseline imbalance	Low risk	Baseline differences adjusted for in analysis
Other bias	Low risk	None identified

Juola 2015

Methods	RCT (cluster randomised).
Participants	**Setting:** 20 wards of assisted living facilities in Helsinki, residential care, mixed-level care, Finland. N = 227 residents. 20 clusters. Sample: 71% women Age (years): 83 **Baseline Characteristics** 93% of population had dementia diagnosis. **Nursing educational intervention** • *N* : 118 • *Age - mean (SD)*: 82.9 (7.5) • *Female - N (%)*: 77 (65.3) • *Medical status defined? - Y/N*: Y • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y **Usual care** • *N* : 109 • *Age - mean (SD)*: 83.5 (6.9) • *Female - N (%)*: 84 (77.1) • *Medical status defined? - Y/N*: Y • *Falls risk defined? - Y/N*: N • *Dependency defined? - Y/N*: Y • *Cognitive status defined? - Y/N*: Y **Inclusion criteria:** age 65 years or older; living permanently in an assisted living facility; Finnish speaking; using at least one medication; having an estimated life expectancy of > 6 months; being able to provide written informed consent (or have a proxy who is able to provide written informed consent in the case of cognitive impairment) **Exclusion criteria:** none provided **Pretreatment differences:** significant baseline differences in Chalsons comorbidity index,dependence in mobility, prior stroke or transient ischaemic attack (TIA), 15D quality of life score; PRN dug use; proportion of sample using harmful medications; and borderline significant difference between groups in gender (P = 0.05). NOTE - some of these reported in Pitkala paper, some in Joula paper
Interventions	• Nursing educational intervention on harmful medications. Education based on constructive learning theory to recognise harmful medications and adverse drug events. Two x four-hour interactive training sessions • Usual care. Nurses were free to participate in any other continuing education, including programmes relating to medication use

Juola 2015 (*Continued*)

Outcomes	• Incidence rate ratio • Number of fallers • Number with multiple falls
Duration of the study	12 months
Notes	ACTRN12611001078943.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: dyads of matched facilities, then random number generator
Allocation concealment (selection bias)	Low risk	Judgement comment: person independent of assessment procedures telephoned another person not familiar with wards or residents to receive allocation
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Judgement comment: nursing staff were not aware that falls data was being analysed as part of the study, however, there is no explanation of whether attempts were made to keep participants and personnel blinded
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Judgement comment: nursing staff recorded falls as part of routine care - not aware that data was being analysed (main study outcome / focus was change in medications)
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: imbalance in group losses, in addition to 3 intervention and 5 control participants not accounted for
Selective reporting (reporting bias)	Low risk	Falls outcomes additional in secondary analysis. Describes all outcomes reported as per methods in the paper
Method of ascertaining falls	Unclear risk	Judgement comment: no definition of falls provided
Baseline imbalance	High risk	Judgement comment: significant baseline differences on mobility and Charlson comorbidity index, no adjustments reported
Other bias	Low risk	None detected

Kennedy 2015

Methods	RCT (cluster-randomised, pilot study)
Participants	**Setting:** 40 residential care facilities, mixed-level care, Canada. Mean 137 beds N = 5478. 40 clusters. Sample: 71% women Age (years): 84.4 (SD 10.9) **Baseline Characteristics** **ViDOS multifaceted KT intervention** • *N*: 2185 • *Age - mean (SD)* : 84.0 (11.1) • *Female - N (%)*: 1,532/2,175 (70.4%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Control** • *N*: 3293 • *Age - mean (SD)* : 84.6 (10.7) • *Female - N (%)*: 2329/3277 (71.1%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Inclusion criteria:** facilities: long-term care facilities - serviced by a particular pharmacy provider; have more than one prescribing physician; residents: none **Exclusion criteria:** residents: none **Pretreatment differences:** mean facility size was larger in control (157 beds, SD 80.2) versus intervention homes (115 beds, SD 67.9); however, both study arms had a similar proportion of small (< 100 beds) and large (> 250 beds) homes.In the control arm, there was a higher prevalence of hip fractures; osteoporosis diagnoses; and baseline use of vitamin D≥ 800 IU/day, calcium≥ 500 mg/day, and osteoporosis medications
Interventions	• ViDOS multifaceted KT intervention. Interactive educational sessions for an interdisciplinary team (comprising the Administrator, Medical Director, Director of Care, Consultant Pharmacist, Director of Food Services/Dietician, and other nursing, medical or rehabilitation staff) delivered via webinar with onsite study co-ordinator, aimed at increasing prescription of adequate levels of vitamin D, calcium and osteoporosis medication. Includes presentation by expert opinion leaders, action planning for quality improvement, audit and feedback review. Quarterly meetings. 3 sessions, approx 6 months apart. First 2 45 to 60 minutes, third 30 minutes. • Usual care - no additional information except fracture prevention toolkits (provided to all homes in the province)
Outcomes	• Number of falls • Number of fallers • Number with multiple falls • Number with fracture falls
Duration of the study	12.2 months; final follow-up 16 months

Notes		

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Judgement comment: computer-generated random allocation sequence
Allocation concealment (selection bias)	Low risk	Judgement comment: allocation adequately concealed at unit level and individual residents not recruited
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: staff recording falls were not blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	Reports ITT for all participating facilities and also separately for the facilities in the intervention group who were "active". Large loss of intervention facilities from recruitment to active participation (7 of the 19 intervention facilities recruited did not proceed to implement the intervention). Baseline data are reported for all intervention and control facilities (i.e. all 19 intervention facilities), but there is no comparison between those who participated (n = 12) and those who were recruited but did not participate in the intervention (n = 7 facilities) to ensure the remaining sample were not biased in any way relative to the recruited and randomised intervention sample
Selective reporting (reporting bias)	Unclear risk	Protocol indicated outcomes as per adjusted analyses would be reported but absolute number of falls and fractures only reported. Impact of this unclear
Method of ascertaining falls	High risk	Quote: "Researchers provided the homes with a standardized data collection sheet and homes completed the information using various sources including electronic/paper-based charts, internal monitoring systems, Resident Assessment Instrument - Minimum Data Set 2.0 (RAI-MDS 2.0), and critical incident re-

		ports." Judgement comment: falls data collected for 3 month blocks from various data sources - different homes had different reporting systems. This is acknowledged as a limitation
Baseline imbalance	High risk	Judgement comment: there were imbalances in baseline characteristics that may impact on falls rates (e.g. hip fractures), the protocol indicated adjustment in analyses (with generalised estimating equations) but adjusted analyses not reported for falls outcomes. P = 0.002 for hip fracture
Other bias	Low risk	Judgement comment: none detected

Kerse 2004

Methods	RCT (cluster randomised)
Participants	**Setting:** 14 mixed-level dependency residential care homes (intermediate- and high-level care), New Zealand N = 617 residents. 14 clusters. Sample: 72% women Age (years): mean 83.2 (SD 10.6) **Inclusion criteria:** resident in one of the included residential care homes **Exclusion criteria:** none stated but data excluded if enrolled in the study for < 2 days and had > 2 falls in one of those days
Interventions	• Falls risk management programme of 12 months duration ○ Falls co-ordinator in each home (carried out fall-risk assessment of all residents using tool, developed specific recommendations and care plans, co-ordinated with other healthcare professionals, and ensured that recommendations were followed) ○ Evidence-based risk assessment tool + detailed management strategies relating to mobility impairments, mental impairments, medications, continence, sensory impairments ○ Tailored care plan based on assessment + OT, PT, medical and specialist referrals ○ Logo on high-risk residents walls + colour-coded dots showing fall-prevention strategies ○ Manual containing the risk assessment form, information for strategies, high-risk fall logos, all forms, and educational information for nurses, doctors, physiotherapists and OTs • Usual care
Outcomes	• Rate of falls • Number of people falling

Kerse 2004 (*Continued*)

Duration of the study	12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "... homes were stratified by type, and an independent researcher, not involved in the study, block randomized them into intervention or control group using computer-generated random numbers."
Allocation concealment (selection bias)	Low risk	See above, and allocation of all cluster units performed at the start of the study AND individual participant recruitment was completed prior to assignment of the cluster, and the same participants were followed up over time
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	High risk	All falls data included in analysis, but large imbalance in those transferred or discharged (15 vs 35
Selective reporting (reporting bias)	Low risk	Falls, fallers, injurious falls and rates of falls reported and appropriately adjusted
Method of ascertaining falls	Low risk	Falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Baseline differences accounted for in analysis
Other bias	Low risk	None identified.

Kerse 2008

Methods	RCT (cluster randomised)
Participants	**Setting:** 41 low-level dependency residential care homes (intermediate-level care), New Zealand. N = 682 residents. 41 clusters. Sample: 74% women Age (years): mean 84.3 (SD 7.2) **Inclusion criteria:** able to engage in conversation about a goal; remember the goal;

	participate in a programme to achieve the goal **Exclusion criteria:** unable to communicate to complete the study measures; anxiety as main diagnosis; acutely unwell; terminally ill
Interventions	• Promoting independence in residential care (PIRC) intervention ○ Goal setting: resident + gerontology nurse (GN) set meaningful goal to promote progressive increase in activity. New goals set when one achieved ○ Functional assessment by GN and individualised programme developed to improve physical function. Physical activities based on repetitions of ADL, e.g. rising from a chair, additional walking, or repeated transfers. Exercise activities at least once a day. Physiotherapist and OT available to help achieve goal. Presciptive plan to increase independence in patient's file and above bed ○ GN trained health care assistants who helped implement programme, supervised by nursing staff ○ GN provided weekly staff support for 1 month, then monthly support ○ Six month intervention but staff expected to continue encouraging residents to activate after that. • Control: usual care + 2 social visits
Outcomes	• Rate of falls • Number of people falling • Adverse events
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "After recruitment of all homes and residents and collection of baseline data, a biostatistician not involved in recruitment randomised homes to the intervention or control group by using computer generated random numbers. "
Allocation concealment (selection bias)	Low risk	Allocation of all cluster units performed at the start of the study. Individual participant recruitment was completed prior to assignment of the cluster, and the same participants were followed up over time
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities who recorded falls were likely to be aware of their facility's allocation status

Kerse 2008 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses balanced between groups. Falls data for 310/330 and 329/352
Selective reporting (reporting bias)	Unclear risk	Judgement comment: details of falls outcomes not reported in trials registration
Method of ascertaining falls	Unclear risk	Judgement comment: falls clearly defined, but method of ascertainment unclear
Baseline imbalance	High risk	Judgement comment: difference in antidepressants at baseline between groups
Other bias	Low risk	Judgement comment: none identified

Klages 2011

Methods	RCT (individually randomised)
Participants	**Setting:** 1 long-term care home (appears to be high- and intermediate-level care), Ontario, Canada N = 24 Sample: 68% women in the analysis Age (years): mean (SD) intervention group 84 (6.6), control group 89 (3.2) **Inclusion criteria:** cognitively impaired (MMSE score < 25); able to follow simple walking instructions; able to walk with minimal assistance; no Snoezelen room attendance in 3 months prior to study **Exclusion criteria:** history of seizures; legal blindness; profound hearing loss; history of limb fractures; extrapyramidal system disruptions (inability to remain motionless or to initiate movement)
Interventions	• Multisensory stimulation in a Snoezelen room: individual 30-minute sessions of stimulation and relaxation, 2 x per week for 6 weeks, with at least 2 days between sessions • Control: individual visits from volunteers (same frequency and duration): listening to readings of the newspaper, looking at magazines, playing cards or a board game, and talking
Outcomes	• Number of falls
Duration of the study	3 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**

Klages 2011 (*Continued*)

Random sequence generation (selection bias)	Low risk	Quote: "A total of 24 eligible residents were recruited. Prior to the commencement of the study a computer-based random number generator was used to randomly select 12 numbers out of 24. These numbers were assigned to the intervention group. The remaining 12 numbers were allotted to participants in the control group."
Allocation concealment (selection bias)	Low risk	Quote: "As multiple recruitment packages were sent out simultaneously, and the participants were assigned a number in chronological order when a signed consent document was received, recruitment order and group allocation were unpredictable."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Nursing staff recording falls were not blind to group allocation and "The investigator [reviewing charts] . .. was not blind to group allocation."
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: one frequent faller excluded from the analysis
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	High risk	Judgement comment: falls recorded concurrently, but falls definition not reported
Baseline imbalance	High risk	Judgement comment: significant difference in age between groups
Other bias	Low risk	Judgement comment: none detected

Koh 2009

Methods	RCT (cluster randomised)
Participants	**Setting:** two acute care hospitals, Singapore N = 1122 patients. 2 clusters. Sample: 641 nurses in medical, surgical and geriatric units in the two hospitals (% female patients not stated) Age (years) patients: mean 68 **Inclusion criteria:** all patients **Exclusion criteria:** none stated

Koh 2009 (*Continued*)

Interventions	• Multifaceted strategy for implementation of Ministry of Health Fall Prevention Clinical Practice Guideline (CPG) • Revision of hospital's fall prevention policy in line with CPG • Identification of change champions from within staff • Educational sessions for staff aimed at promoting and supporting the adoption of the recommendations • Reminders and identification systems, e.g. mandatory fall risk-assessment tool in nursing assessment notes, posters in ward toilets, high-risk patients identified by pink name card above the bed, pink stickers on clinical/nursing notes, and pink identification bracelets • Audit and feedback on incidence of falls and compliance with use of risk assessment tool • Control: routine dissemination strategies for implementation of CPG
Outcomes	• Rate of falls
Duration of the study	6 months
Notes	Intervention targeted nursing staff. Age of patients not stated in Koh 2009. Obtained by personal communication with author

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The two study hospitals were randomly allocated either to the "intervention" site... or the "control" site". Author states carried out by supervised coin toss; heads gets the intervention
Allocation concealment (selection bias)	High risk	No concealment. After first site randomised, second site automatically becomes the control group
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Fall incidence and fall-associated injury rates were obtained from the hospitals' fall incidence database"
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: falls data for a random sample of medical records used. How representative these are of all patients and what proportion unknown
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified

Koh 2009 (*Continued*)

Method of ascertaining falls	High risk	Judgement comment: falls determined through audits of hospital records. Definitions and practices may vary between hospitals
Baseline imbalance	Unclear risk	Judgement comment: baseline characteristics of patients not reported
Other bias	Low risk	Judgement comment: none detected

Kovacs 2012

Methods	Pilot RCT (individually randomised).
Participants	**Setting:** One residential care facility, intermediate-level care, Hungary N = 41 Sample: 100% women Age (years): 69.2 **Baseline Characteristics** **Multimodal exercise plus osteoporosis exercise** • *N*: 21 • *Age - mean (SD)* : 68.7 (6.9) • *Female - N (%)*: 21 (100%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y (Berg Balance Scale) • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: N **Osteoporosis exercise programme** • *N*: 20 • *Age - mean (SD)* : 69.7 (6.5) • *Female - N (%)*: 20 (100%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: N **Inclusion criteria:** living in the National Institution for Blind People (partial sightedness or blindness); aged 60 years or over; being female **Exclusion criteria:** being totally blind; had lived in the nursing home for less than 2 months; being unable to walk around their own residence; having progressive neurological, and unstable cardiovascular diseases that would limit participation in exercise programme; planned moving away from the nursing home during the study period and; participated in an exercise programme including balance exercise within 6 months **Pretreatment differences:** nil
Interventions	• Multimodal exercise including strength, balance and progressive resistance based on Otago Exercise Programme, modified for visual impairment, plus walking programme plus standard osteoporosis exercise programme. Strength exercises were directed to major lower limb muscle groups playing roles in postural control, balance exercises were closely related to everyday activity. Group training in groups 3 to 6

	supervised by physiotherapist. Plus flexibility warm-up and cool-down. 2 x weekly 30-minute multimodal exercise plus 2 x weekly 30-minute osteoporosis exercise, plus 20 to 30 minutes walking. • Osteoporosis exercise programme. Standard osteoporosis exercise programme alone with strength and flexibility exercises. Not progressive or individually tailored. Plus flexibility warm-up and cool-down. 30 minutes, 4 times per week.
Outcomes	• Relative risk for falling • Adverse events
Duration of the study	6 months
Notes	Visually impaired participants.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Judgement comment: details of sequence generation not reported
Allocation concealment (selection bias)	Low risk	Judgement comment: numbered opaque identical sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls extracted from staff records (medical and nursing documentation) , blinding of staff not feasible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: no loss to follow-up
Selective reporting (reporting bias)	High risk	Judgement comment: number of falls in the follow-up period not reported
Method of ascertaining falls	Unclear risk	Judgement comment: falls clearly defined but details of documentation inadequate for judgement
Baseline imbalance	Low risk	Quote: "There were no significant differences between groups on any baseline characteristics. "
Other bias	Low risk	Judgement comment: none detected

Kovacs 2013

Methods	**Study design:** RCT (individually randomised)
Participants	**Setting:** One residential care facility, mixed-level care, Hungary. N = 86 Sample: 81% women Age (years): 77.9 **Baseline Characteristics** **Multimodal exercises programme** • *N*: 43 • *Age - mean (SD)* : 76.4 (9.6) • *Female - N (%)*: 36 (83%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Control** • *N*: 43 • *Age - mean (SD)* : 79.3 (12.7) • *Female - N (%)*: 34 (79%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** cognitive impairment (MMSE < 24), residents of nursing home, 60 years or over **Exclusion criteria:** living in nursing home < 2 months, < 60 years of age, unable to walk 6 metres with or without walking aid, unable to follow simple verbal exercise instructions, unstable cardiovascular or pulmonary diseases that would limit participation in exercise programme, terminal illness, planned moving from the nursing home during the study, no consent **Pretreatment differences:** using a frame (20.9% int, 41.9% con)
Interventions	• A multimodal exercise programme based on Otago Exercise Programme consisting of strength, balance exercises plus 10 minutes flexibility warm-up and cool down, with progressive resistance supervised by physiotherapist and group based (2 to 4 participants), and supervised walking training. Exercise programme twice weekly, walking once a week • Usual care: no exercise programme, participation in social activities
Outcomes	• Rate ratio • Risk of falling • Number with multiple falls
Duration of the study	12 months
Notes	Compliance reported. Cognitively impaired participants.
Risk of bias	

Kovacs 2013 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Judgement comment: insufficient detail for judgement
Allocation concealment (selection bias)	Low risk	Quote: "Consecutively numbered opaque identical sealed envelopes were used for allocation."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: carers recorded falls not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: attrition numbers and reasons balanced between groups
Selective reporting (reporting bias)	Low risk	Judgement comment: although no protocol was identified, falls outcomes were reported clearly and as multiple measures (fallers, falls, recurrent fallers, as n and RR)
Method of ascertaining falls	Low risk	Judgement comment: falls recorded concurrently on calendar using clear definition
Baseline imbalance	High risk	Judgement comment: significant difference between groups in the proportion using a frame, not adjusted for in analysis
Other bias	Low risk	Judgement comment: none detected

Lapane 2011

Methods	RCT (cluster randomised)
Participants	**Setting:** 25 nursing homes (appear to be high- and intermediate-level care), Ohio, USA N = 3321 residents. 25 clusters. Sample: 73% women Age (years): no overall age available **Inclusion criteria** (facilities): facilities serviced by one of two Omincare pharmacies and with stable contracts; Medicare and Medicaid certified; ≥ 50 geriatric beds; few short-stay residents **Exclusion criteria:** none stated

Interventions	• Clinical informatics tool (Geriatric Risk Assessment MedGuide (GRAM)) to assist consultant pharmacists and nursing staff identify residents at risk for delirium and falls based on prescribed medications, implement proactive monitoring plans as appropriate, and provide reports to assist consultant pharmacists conducting monthly medication review. Detailed instruction of staff on medications implicated in falls and delirium, use of reports, care plans and flow charts etc. Detailed instruction of consultant pharmacists providing targeted medication review for all high-risk residents. Reports within 24 hours of admission for new admissions and used during monthly review, in addition to generation at time of Minimum Data Set reports or when falls or delirium triggered resident assessment protocols. • Control: usual care including monthly medication review by consultant pharmacist.
Outcomes	• Number of people falling
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "Homes were randomised ..." Insufficient information to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Insufficient information to permit judgement, although clinical staff recording falls would have been aware of allocation of the nursing home
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Loss to follow-up not clearly reported
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified. Data not suitable for use of rate of falls or injurious falls in meta-analysis as per review Appendix 6
Method of ascertaining falls	High risk	Judgement comment: no definition of falls provided, only states "MDS data", may vary between sites
Baseline imbalance	High risk	Judgement comment: number of falls in past 30 days was much higher in intervention group

Lapane 2011 (*Continued*)

Other bias	Low risk	Judgement comment: none detected

Law 2006

Methods	RCT (cluster randomised by unit).
Participants	**Setting:** 118 homes for elderly people, 223 units (intermediate- and high-level care), throughout the UK N = 3717 residents. 223 clusters. Sample: 76% women Age (years): mean 85 **Inclusion criteria:** facility resident; aged ≥ 60 **Exclusion criteria:** temporary residents; taking vitamin D or calcium supplements or medications to increase bone density; sarcoidosis; malignancy; life threatening illness
Interventions	• 2.5 mg oral ergocalciferol (vitamin D2) every 3 months (equivalent to 1100 IU/day) • Usual care (no placebo)
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (non vertebral fractures)
Duration of the study	Median length of follow-up 10 months (interquartile range 7 to 14)
Notes	Mean baseline serum vitamin D level collected from 1% of the intervention group; mean 59 nmol/L

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Cluster randomisation by computer. No further information provided
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not possible to blind participants but personnel recorded the fall data were blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses balanced between groups. 2.8% left care homes in intervention

Law 2006 (*Continued*)

		group, 3.3% control group, other losses due to death (p484 first para, text)
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified, but fractures, fallers and falls reported, adjusted for clustering
Method of ascertaining falls	High risk	Judgement comment: falls not clearly defined. Study conducted across 118 homes which may have variations in reporting practice and definitions. Falls recorded daily
Baseline imbalance	Unclear risk	Judgement comment: similar at baseline for demographic characteristics (e.g. age, gender), but did not discuss prognostic factors e.g. falls rate/ medical status
Other bias	Low risk	Judgement comment: none identified

Mador 2004

Methods	RCT (individually randomised)
Participants	**Setting:** two metropolitan acute hospitals, South Australia N = 71 Sample: 48% women Age (years): mean 82.5 **Inclusion criteria:** inpatients on medical and surgical wards; aged ≥ 60; confusion due to either dementia or delirium; problematic behaviour **Exclusion criteria:** primary psychiatric illness; no next of kin available to give consent
Interventions	• Participants assessed for causes of confusion and behavioural disturbance by extended practice nurse within 24 hours of referral. Management plan formulated with respect to non pharmacological strategies to help manage problematic behaviour which was discussed with nursing staff. Ongoing support and education provided to carry out strategies • Usual care
Outcomes	• Number of people falling
Duration of the study	11 months. Median length of stay 12 days for intervention group and 9 days for control group
Notes	Potential contamination as staff receiving training were also caring for controls

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**

Mador 2004 (*Continued*)

Random sequence generation (selection bias)	Low risk	Quote: "Sequential sealed opaque envelopes were prepared by a person who was external to the study in blocks of ten stratified for the two hospitals, using a computer-generated table of random numbers."
Allocation concealment (selection bias)	Low risk	Quote: "Sequential sealed opaque envelopes were prepared by a person who was external to the study..." Randomised by the Repatriation Hospital Pharmacy Department
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcome assessors were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: little loss and ITT analysis conducted
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	High risk	Quote: "Number of falls for each patient was extracted from the hospital's database of critical incidents." Judgement comment: no falls definition reported
Baseline imbalance	High risk	Quote: "There was a significant difference for prior residence, with more control participants entering hospital from home com- pared with the intervention group (p ¼ 0.035). The number of participants under the care of a geriatrician was greater in the intervention than in the control group (p ¼ 0.006)."
Other bias	Low risk	Judgement comment: none identified

Mayo 1994

Methods	RCT (individually randomised)
Participants	**Setting:** rehabilitation (subacute) hospital, Canada N = 134 Sample: 46% women Age (years): mean (SD) intervention 70.9 (12.6), control 72.9 (11.8) **Inclusion criteria:** one or more of the following: admission diagnosis of stroke or ataxia; an episode of incontinence; a history of multiple falls; aged ≥ 80; using topical eye medication, anticonvulsants, vitamin supplements or anti-ulcer medications

Mayo 1994 (*Continued*)

	Exclusion criteria: unable to understand what was being asked of them; participated in this study during a previous admission
Interventions	All participants selected as being high risk of falling • Blue identification bracelet. Told to use bracelet as reminder to be careful when moving around hospital • Usual care: no blue bracelet
Outcomes	• Rate of falls • Number of people falling
Duration of the study	12 months. Median lengths of stay 75 days (intervention group), 65 days (control group)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "Subjects were interviewed to obtain baseline information ... and were then randomly assigned to either the intervention group or the control group."
Allocation concealment (selection bias)	Unclear risk	Insufficient information on process of allocation to permit judgement of 'Low risk' or 'High risk'
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Falls ascertained through incident reports. Staff completing incident reports would have been aware of whether or not participant was wearing a blue bracelet
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis, rate of falls, all patients appear to have been included
Selective reporting (reporting bias)	Low risk	Judgement comment: data presented for number of falls, fallers and rate of falls as per methods
Method of ascertaining falls	Unclear risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	High risk	Judgement comment: some imbalance in ability to walk independently at baseline
Other bias	Low risk	Judgement comment: none identified

McMurdo 2000

Methods	RCT (cluster randomised)
Participants	**Setting: 9** residential care facilities (intermediate-level care), Dundee, Scotland, UK N = 133 residents. 9 clusters. Sample: 81% women Age (years): mean 84 (SD 7) **Inclusion criteria:** aged ≥ 70 **Exclusion criteria:** MMSE score < 12

Interventions	• Multifactorial, multidisciplinary intervention
	  ○ Falls risk assessment and modification performed for each participant including medication review. Recommendations sent to participant's GP, optometrist review if indicated, and review of lighting levels
	  ○ Supervised exercises to improve balance, strength and flexibility; 30 minutes 2 x per week for 6 months. Performed seated because of frailty of participants; not individually tailored. Not specified who delivered the exercise intervention
	• Control: reminiscence therapy

Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures) • Adverse events
Duration of the study	12 months. 6 month intervention + 6 months follow-up
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "... allocated at random ..." Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'
Allocation concealment (selection bias)	Unclear risk	Insufficient information on process of allocation to permit judgement of 'Low risk' or 'High risk'
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities recording falls in calendar were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	High risk	Large difference in dropout rates between arms

McMurdo 2000 (*Continued*)

Selective reporting (reporting bias)	Unclear risk	Number of falls, fallers, multiple fallers and fracture falls reported. No adjustment for clustering. NIHR link broken
Method of ascertaining falls	Low risk	Falls clearly defined and recorded daily on a falls calendar by staff
Baseline imbalance	Low risk	Groups balanced at baseline
Other bias	Low risk	None identified

Meyer 2009

Methods	RCT (cluster randomised)
Participants	**Setting:** 58 nursing homes (high-level nursing care), Hamburg, Germany N = 1125 residents. 58 clusters. Sample: 85% women Age (years): mean (SD) intervention group 86 (6), control group 87 (6) **Inclusion criteria** (facilities): ≥ 30 residents; not using a fall risk assessment tool or willing to stop using a tool. Inclusion criteria (residents): ≥ 70 years; able to walk with or without assistance; living in the nursing home for > 3 months **Exclusion criteria:** none stated
Interventions	• Use of one fall risk assessment tool (Downton Index) by ward staff • Control: no fall risk assessment tool (nurses judgement of risk)
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures)
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Computer-generated randomisation lists were prepared by the biostatistician for concealed allocation of clusters by external central telephone."
Allocation concealment (selection bias)	Low risk	See above
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded

Meyer 2009 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Nursing staff recorded falls (presumably not blind). External investigator verified completeness of falls data - not clear if blind to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: fall data reported for all participants for time in study
Selective reporting (reporting bias)	Low risk	Judgement comment: outcomes consistent with protocol and adjusted for clustering
Method of ascertaining falls	Low risk	Judgement comment: falls recorded concurrently and clearly defined
Baseline imbalance	Low risk	Judgement comment: balanced at baseline
Other bias	Low risk	Judgement comment: none detected

Michalek 2014

Methods	Pilot RCT (pseudo-randomised to one of two clusters)
Participants	**Setting:** Subacute hospital setting. Median length of stay 20 days. Germany N = 114. 2 clusters. Sample: 79% women Age (years): Mean NR **Baseline Characteristics** **FORTA** • *N*: 58Table 4 • *Age - MEDIAN (IQR)*: 84 (81-87) • *Female - N (%)*: 42 (75%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Usual care** • *N*: 56 • *Age - MEDIAN (IQR)*: 83 (79-87) • *Female - N (%)*: 48 (83%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** aged >70 years; stable health condition defined as no need for intermediate or intensive care unit treatment; at least three diseases in need for drug treatment; at least three medical prescriptions; admitted during the first 3 days of the week because of staff availability; patients or proxies had to give written informed consent **Exclusion criteria:** critical or terminal illness; dementia (MMSE <25); refusal to partic-

	ipate **Pretreatment differences:** nil significant reported at baseline, BMI Borderline (P = 0. 052)
Interventions	• FORTA. Drugs were classified according to the FORTA list, combining positive and negative labelling of drugs, ranging from A (indispensable), B (beneficial), C (questionable), D (avoid). Drugs were changed in first week of hospitalisation as guided by FORTA. Weekly meetings of drug evaluation and need encompassing patient disease, functional status, prognosis and need for drugs with decisions based on FORTA suggestions. Drugs were continued despite unfavourable FORTA labelling if patients insisted. Overprescription and under prescription were identified and corrected according to FORTA recommendations. Weekly meetings. • Usual geriatric hospital care
Outcomes	• Falls rate • Number of fallers • Number with multiple falls
Duration of the study	Until discharge (median hospital stay 20 days)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Judgement comment: patients were assigned randomly by number of entrance to one of two wards
Allocation concealment (selection bias)	High risk	Judgement comment: quasi randomised to one of two wards - high risk of bias. Individuals randomised by number of entrance, sequence predictable
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible, patients admitted to intervention or control wards
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded in hospital recording system by staff who will know ward allocation of patients
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: no reported loss to follow-up during study period, attrition after enrolment unlikely in acute hospital setting, however falls data reported for 178 patients in Frohnhofen 2013 abstract

Michalek 2014 (*Continued*)

Selective reporting (reporting bias)	Low risk	Judgement comment: falls not recorded as outcome on trials registry but falls outcomes seems to be completely reported in multiple ways (fallers, falls rate)
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and likely concurrent through established hospital reporting system
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	High risk	Judgement comment: analysis was by individual but quasi randomisation and it was to one of two clusters (although not specifically cluster randomised), which should have been addressed in the analysis

Mulrow 1994

Methods	RCT (individually randomised)
Participants	**Setting:** 9 nursing homes (high-level nursing care), USA N = 194 Sample: 71% women Age (years): mean (SD) intervention group 79.7 (8.5), control group 81.4 (7.9) **Inclusion criteria:** aged > 60; resident in nursing home for ≥ 3 months; dependant in ≥ 2 ADLs **Exclusion criteria:** terminal illness; acute medical condition; MMSE score < 50%, unable to follow two-step command; assaultive behaviour; received physiotherapy within last 2 months
Interventions	• Tailored exercises 3 x per week for 30 to 45 minutes, 4 months duration. Exercises comprised gait, balance and co-ordination + strength/resistance + flexibility exercises. Intervention delivered by physical therapists (one on one) • Friendly visit
Outcomes	• Rate of falls • Number of people falling • Adverse events
Duration of the study	4 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement

Mulrow 1994 (*Continued*)

Random sequence generation (selection bias)	Low risk	Quote: "Randomization was performed after baseline assessments by calling a central number. Randomization was blocked in groups of four and stratified by nursing home site."
Allocation concealment (selection bias)	Low risk	Randomisation was performed after baseline assessments by calling a central number. No further description
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Falls recorded in charts and incident reports. Staff recording falls likely to be aware of allocation status. Research assistants examining charts and incident reports were reported to be blinded to allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: 14 dropouts, 12 due to death, other 2 unexplained but unlikely to be related to outcome
Selective reporting (reporting bias)	Low risk	Judgement comment: falls and fallers outcomes reported
Method of ascertaining falls	High risk	Judgement comment: no falls definition reported and may vary between sites
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified

Neyens 2009

Methods	RCT (cluster randomised by ward)
Participants	**Setting:** 12 nursing homes, psychogeriatric wards (high-level nursing care), the Netherlands (6 wards in intervention group and 6 in control group). N = 518 residents. 12 clusters. Sample: 68% women Age (years): mean (SD) intervention group 82.1 (7.7), control group 83.3 (7.7) **Inclusion criteria** (wards): ≥ 25 beds; not using a fall prevention protocol; having the largest number of mobile patients **Exclusion criteria:** none stated
Interventions	• Multifactorial, multidisciplinary intervention: ○ General medical assessment by medical staff (at start of trial, on admission, if change in medical condition)

	○ Assessment with fall risk evaluation tool (fall history, medication intake, mobility, use of assistive and protective aids) by multidisciplinary team (physician, 2 nurses, physiotherapist, OT) at start of trial, on admission, after a fall, at request of ward staff, 2 x per year for all residents) ○ Team decisions about individually-tailored fall-prevention activities, e.g. medication review, individually-designed exercise programmes, assessing and providing assistive and protective aids. Fortnightly conferences discussing each assessed resident ○ Environmental hazard check on each ward by OT ○ Team could implement general fall prevention activities, e.g. staff training • Control: usual care, no insight on fall prevention programme
Outcomes	• Rate of falls
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "At random, using computer techniques, two intervention homes and two control homes were selected from each group [groups based on the mean fall incidence rate of psychogeriatric patients per psychogeriatric bed], resulting in a total of six intervention homes and six control homes. "
Allocation concealment (selection bias)	High risk	One ward per home was chosen after randomisation, based on inclusion criteria
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study was cluster randomised and nursing staff recorded falls
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data analysed by ITT
Selective reporting (reporting bias)	High risk	No protocol identified, fallers not reported
Method of ascertaining falls	Unclear risk	Not stated whether falls clearly defined
Baseline imbalance	Low risk	Reasonable comparability. More falls pre-trial in intervention arm, but adjusted for in analyses

Other bias	Low risk	None identified

Nowalk 2001

Methods	RCT (individually randomised)
Participants	**Setting:** 2 long-term care facilities (combined high-level nursing care and independent living), USA N = 110 participants Sample: 86% women Age (years): mean 84 **Inclusion criteria:** aged ≥ 65; cognitively able to be tested; able to ambulate with or without assistive device; able to follow simple directions; co-operative; capable of participating in group sessions **Exclusion criteria:** unwilling or unable to complete baseline assessments
Interventions	• "Fit NB Free" (FNBF): supervised exercises consisting of progressive strength training, flexibility, and endurance (treadmill and bicycling exercises), 3 x per week for 13 to 28 months. Duration of sessions not specified. Exercises were delivered by exercise physiologists. Exercises individually-tailored based on exercise capacity of participants • "Living and Learning/Tai Chi (LL/TC): Tai Chi 3 x per week for 13 to 28 months + psychotherapeutic and behavioural methods to reduce fear of falling. Exercises not individually-tailored. Tai Chi was delivered by professional instructor. Individualised assessment of participants not part of intervention • Usual routine activities Note: all groups also exposed to educational activities
Outcomes	• Number of people falling
Duration of the study	24 months
Notes	True N for each group unknown and data discrepancies within published manuscript

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Following completion of all assessments, participants were randomly assigned to one of three groups ... using permuted blocks ..."
Allocation concealment (selection bias)	Unclear risk	Insufficient information on process of allocation to permit judgement of 'Low risk' or 'High risk'
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible

Nowalk 2001 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls on incident report forms were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: attrition by allocation group unclear, but overall 41/112 lost, died or not followed for full time period
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified. Number of falls not reported
Method of ascertaining falls	Low risk	Judgement comment: falls defined and reliant on facility incident reports
Baseline imbalance	Low risk	Judgement comment: no important differences at baseline
Other bias	Low risk	Judgement comment: none identified

Patterson 2010

Methods	RCT (cluster-randomised matched pairs of nursing homes)
Participants	**Setting:** 22 nursing homes (high- and intermediate-level care), Northern Ireland N = 334 residents. 22 clusters. Sample: 73% women Age (years): mean 82.7 (SD 8.4) **Inclusion criteria** (facilities): > 30 resident beds (including homes for general nursing category residents and for elderly mentally infirm people). Inclusion criteria (residents): aged ≥ 65 **Exclusion criteria** (facilities): caring exclusively for terminally ill people. Exclusion criteria (residents): terminally ill; attending day care only
Interventions	• Pharmacists visited intervention facilities monthly for 12 months. Reviewed residents' clinical and prescribing information, applied an algorithm to assess appropriateness of psychoactive medication, worked with nurses and prescribers to improve the prescribing of these drugs • Usual care
Outcomes	• Rate of falls
Duration of the study	12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement

Patterson 2010 (*Continued*)

Random sequence generation (selection bias)	Low risk	Quote: "randomly assigned ... using a computer generated table of random numbers"
Allocation concealment (selection bias)	Low risk	An independent researcher blind to the identity of the homes carried out the randomisation (after consent obtained from the homes)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Routinely collected falls data were used. Staff not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss to follow-up entirely due to death, with similar percentages of deaths in each group. ITT analysis
Selective reporting (reporting bias)	High risk	Judgement comment: trial registry record indicates outcome as number of people falling, but only rate of falls reported
Method of ascertaining falls	High risk	Judgement comment: falls definition not reported and reliant on falls reporting within each home which may vary
Baseline imbalance	Low risk	Judgement comment: no major differences at baseline. Similar for falls risk factors. Main difference is more urban nursing homes in control group than in intervention group
Other bias	Low risk	Judgement comment: none identified

Peyro Saint Paul 2013

Methods	RCT (individually randomised)
Participants	**Setting:** hospital acute and residential care facility setting (92% residential care), France. N = 19 residents Sample: 58% women Age (years): 89.9 **Baseline Characteristics** **Changing drug therapy** • *N*: 9 • *Age - mean (SD)* : 90.8 (3.7)

	• *Female - N (%)*: 5 (56%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 10 • *Age - mean (SD)* : 89.0 (7.3) • *Female - N (%)*: 6 (60%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Inclusion criteria:** aged 65 and over; chronic moderate hyponatraemia (serum sodium 123 mEq/L to 134 mEq/L) detected using a biological control routine; in acute care unit or retirement home **Exclusion criteria:** **Pretreatment differences:** age and sex same, Nz level same, renal clearance worse in control
Interventions	• Changing drug therapy. Review by pharmacist of drugs that may cause hyponatraemia. • Usual care. Routine management with no drug review
Outcomes	• Number of falls • Number of fallers • Number with multiple falls • Adverse events
Duration of the study	3 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Judgement comment: Author correspondence, quote: "Random sequence was managed as a single randomization list managed by the sponsor". It is unclear how the randomisation sequence was generated
Allocation concealment (selection bias)	Low risk	Judgement comment: Author correspondence, quote: "Random sequence was managed as a single randomization list managed by the sponsor. Allocation was concealed using masking envelope."

Peyro Saint Paul 2013 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: author correspondence: "Staff were not blind to group allocation. Residents were not blind to group allocation."
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: author correspondence, quote: "The fall was recorded as soon in the patient file by the first caregiver who noted: carer, nurse or doctor. Caregiver were not blind to group allocation."
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: the proportion of missing data is considered high enough to potentially have a relevant effect on the effect estimate: falls data only available for 9/19 randomised patients. Response to enquiry received 19/7 from Peyro Saint Paul - participant flow chart still unclear
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Unclear risk	Judgement comment: author correspondence, quote: "the software allows to record falls in the patient file. " Methods of ascertaining falls not reported
Baseline imbalance	Unclear risk	Judgement comment: characteristics for key baseline factors (falls risk, medical status, dependency, cognitive status) relevant to falls are not reported
Other bias	Low risk	Judgement comment: none detected. Main publication not in English

Potter 2016

Methods	RCT (individually randomised)
Participants	**Setting:** 4 care facilities, mixed level of care, rural Australia. N = 95 participants randomised; 93 in analysis Sample: 52% women Age (years): mean 84.3 (SD 6.9) **Baseline Characteristics** **Deprescribing intervention** • *N*: 47 • *Age - mean (SD)* : 84 (6) • *Female - N (%)*: 26 (55%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Usual care**

	• *N*: 48 • *Age - mean (SD)* : 84 (8) • *Female - N (%)*: 23 (48%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** residents of residential aged care facilities aged 65 years or older **Exclusion criteria:** taking no regular medicines; were in the final terminal stages of an illness; or if their usual general practitioner (GP) or the RACF nurse manager did not agree to their participation **Pretreatment differences:** control participants had lower mean blood pressure.
Interventions	• Deprescribing intervention. An individualised medicine review followed by the planned cessation of non-beneficial medicines. The intention of deprescribing was to reduce the total number of unique medicines consumed. The review was led by a GP and a geriatrician who was also a clinical pharmacologist of older people. The medicine withdrawal plan, amended to reflect changes requested by participant, next-of-kin, or GP, was implemented over several months. The GP reviewed participants weekly during deprescribing. • Usual care. Medication review as per the control arm with no deprescribing (medication review plan not passed on to GPs).
Outcomes	• Number of falls • Number of fallers • Number of multiple fallers • Number with fracture fall • Number with an adverse event.
Duration of the study	12 months
Notes	After 12 months, 59% of targeted medicines were deprescribed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: using a digital random number generator
Allocation concealment (selection bias)	Low risk	Judgement comment: sealed opaque envelopes opened after the medication review, withdrawal plan and baseline assessments
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: no blinding with reference to falls outcome assessment possible

Potter 2016 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls outcomes were assessed by persons who would know the treatment allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: no evidence of significant incompleteness of falls outcome data
Selective reporting (reporting bias)	Low risk	Judgement comment: protocol available, outcomes reported as per protocol
Method of ascertaining falls	Low risk	Judgement comment: used routine data collection plus recall from relatives. Clear definition used
Baseline imbalance	Low risk	Judgement comment: difference in systolic blood pressure, however, deemed unlikely to significantly affect outcome
Other bias	Low risk	None detected

Ray 1997

Methods	RCT (cluster randomised)	
Participants	**Setting:** 14 nursing homes (high-level nursing care), USA N = 499 participants. 14 clusters. Sample: 78% women Age (years): mean 83 **Inclusion criteria:** high risk of falls with potential problem in a safety domain; likely to remain in nursing home **Exclusion criteria:** age < 65; anticipated stay < 6 months; bed bound; no fall in previous year	
Interventions	• Consultation service with individual assessment and recommendations targeting environmental and personal safety, wheelchair use, psychotropic medication use, transferring, and ambulation. Falls co-ordinator at each site. Intervention delivered by study team • Usual care	
Outcomes	• Number having 2 or more falls	
Duration of the study	12 months	
Notes	No published data on numbers of falls or fallers who had a single fall	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Ray 1997 (*Continued*)

Random sequence generation (selection bias)	Low risk	Seven "matched" pairs of facilities participated. Quote: "The statistician ... generated sealed-envelope random assignments for each pair from the SAS function RANUNI (using the clock for the seed)."
Allocation concealment (selection bias)	Low risk	Study author (statistician) generated sealed envelope random number assignments for each pair using the SAS function from RANUNI using the clock for the seed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff at the facilities who recorded falls were likely to be aware of their facility's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	ITT analysis, little loss to follow-up, reasons balanced
Selective reporting (reporting bias)	Unclear risk	No protocol identified
Method of ascertaining falls	Unclear risk	Falls clearly defined, relies on incidence reports, trial guidance on concurrent reporting unclear
Baseline imbalance	Low risk	No major differences. Difference in BMI, life space diameter, multivariate regression conducted, no differences in main falls risk factors
Other bias	Low risk	None identified

Rosendahl 2008

Methods	RCT (cluster randomised)
Participants	**Setting:** 9 residential care facilities (intermediate- and high-level nursing care), Sweden N = 191. 34 clusters. Sample: 73% women in 34 clusters (cluster equals 3 to 9 participants living on the same floor, wing, or unit) Age (years): mean 84.7 (SD 6.5) **Inclusion criteria:** aged ≥ 65; dependent in ≥ 1 personal ADLs; able to stand from armchair with help from 1 person; MMSE score ≥ 10; physician approval **Exclusion criteria:** none stated
Interventions	• Functional exercise programme: weight-bearing exercises challenging leg strength, postural stability, and gait ability. Physiotherapists selected exercises for each participant according to their functional deficits. High intensity and increasing load encouraged (5 sessions of 45 minutes every fortnight; total of 29 sessions)

	• Control: seated programme developed by OT, e.g. watching films, reading, singing (5 sessions of 45 minutes every fortnight)
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (hip fractures) • Adverse events
Duration of the study	6 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Researchers not involved in the study performed the randomization by using lots in sealed non-transparent envelopes."
Allocation concealment (selection bias)	Low risk	Randomisation by cluster was performed after the inclusion of participants and baseline assessments using sealed nontransparent envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses balanced and unlikely to affect outcome
Selective reporting (reporting bias)	Low risk	Judgement comment: outcomes reported consistently with trial registration. All expected outcomes reported
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: no major differences at baseline. Difference in self-perceived health but no differences in diagnoses, functional assessments, falls or drugs
Other bias	Low risk	Judgement comment: none identified.

Rubenstein 1990

Methods	RCT (individually randomised)
Participants	**Setting:** long-term care facility (intermediate- and high-level nursing care), Los Angeles, USA N = 160 Sample: 85% women Age (years): mean (SD) intervention group 86.8 (0.6), control group 87.9 (0.7) **Inclusion criteria:** fall within 7 days of nurse receiving fall incident report **Exclusion criteria:** unable to walk; unable to be evaluated within 7 days of fall due to acute illness or hospitalisation; unable to understand English
Interventions	• Comprehensive post fall assessment within 7 days of fall. Intervention delivered by nurse: physical examination including visual screening, extended pulse and blood pressure assessments with attention to postural changes, assessment of footwear and foot problems, a quantified gait and balance assessment, laboratory tests, ECG, 24 hours Holter monitoring, environmental assessment to identify potential hazards. Once only assessment with recommendations given to patient's primary care physician • Usual care. Control group did not receive the assessment and no recommendations were transmitted. "Less than half of the control group received no more than a brief check for injury after they fell."
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures
Duration of the study	24 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Eligible fallers were ... randomly assigned to either the intervention or control group, using computer generated, randomly sequenced cards in sealed envelopes."
Allocation concealment (selection bias)	Unclear risk	Insufficient information on process of allocation to permit judgement of 'Low risk' or 'High risk'. It is unclear who conducted the randomisation and envelopes not described as opaque and sequentially numbered
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded

Rubenstein 1990 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls after intervention were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data balanced between arms
Selective reporting (reporting bias)	Low risk	No protocol identified (1990 study) but expected falls outcomes reported as number of falls and fallers reported
Method of ascertaining falls	Low risk	Falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	No major differences between groups at baseline
Other bias	Low risk	None identified

Sakamoto 2006

Methods	RCT (individually randomised)
Participants	**Setting:** nursing care facilities and rehabilitation outpatient departments (intermediate care), Japan N = 553 Sample: 74% women Age (years): mean 81.6 (SD 9.0) **Inclusion criteria:** able to stand on their own while holding on to a bar **Exclusion criteria:** severe dementia
Interventions	• Single leg stance practice both legs for 1 minute each leg, 3 times daily • Usual care (without exercise)
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (hip fractures)
Duration of the study	6 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Randomization of the subjects into an exercise group or a control group was performed by the Department of Information Science of our university. " using a "table of random numbers"

Sakamoto 2006 (*Continued*)

Allocation concealment (selection bias)	Unclear risk	Randomisation by Department of Information Science. Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: greater loss from intervention group, 22 vs 4
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified, but number of falls, and fallers reported
Method of ascertaining falls	Unclear risk	Judgement comment: no definition of falls. Method of ascertaining falls not described
Baseline imbalance	Unclear risk	Judgement comment: baseline characteristics by group allocation unclear
Other bias	Low risk	Judgement comment: none identified

Sakamoto 2012

Methods	RCT (individually randomised)
Participants	**Setting:** 3 nursing homes (intermediate-level care), Aomori, Japan N = 145 Sample: 81% women Age (years): mean (SD) intervention group 84.2 (7.8), control group 84.1 (7.7) **Inclusion criteria:** aged ≥ 65; able to transfer independently with or without assistive devices **Exclusion criteria:** non consenting; pica disorder (the desire to eat "unnatural" things) in case they ate the patches
Interventions	• Lavender olfactory stimulation: commercially available white patch (1 cm x 2 cm, Aromaseal Lavender; Hakujuji Co., Tokyo, Japan) attached to inside of resident's clothing near the neck: continuous olfactory exposure for 24 hours. Patches replaced daily for 1 year. Odour can only be sensed by person wearing the patch • Control: placebo patch (1 cm x 2 cm, unscented Aromaseal) replaced daily for 1 year
Outcomes	• Rate of falls • Number of people falling • Adverse events

Duration of the study	12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "An independent statistician performed resident allocations using computer-generated randomization of numbers at each nursing home."
Allocation concealment (selection bias)	Low risk	Quote: "An independent statistician performed resident allocations ... at each nursing home. Treatment allocation status was delivered to the head nurse at each nursing home, and patches were prepared accordingly."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: placebo patch used but as was olfactory stimulation is a reasonable chance of unblinding
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Although the staff recording falls were blind to group allocation, the head nurse who "supervised the recording of falls regularly", was not
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: approx 30% withdrawal due to death and discharge, balanced between study arms. ITT analysis performed
Selective reporting (reporting bias)	Low risk	Judgement comment: falls, falls rate, fallers and recurrent falls reported unadjusted and adjusted. Falls outcomes thoroughly and completely reported
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: no differences between groups at baseline
Other bias	Low risk	Judgement comment: none identified

Salvà 2016

Methods	RCT (cluster randomised).
Participants	**Setting:** 16 residential care facilities, mixed-level care, Spain. N = 16 clusters randomised, 12 clusters in analysis. Sample: 72% women Age (years): 84.4 **Baseline Characteristics** **Multifactorial falls prevention programme** • *N*: 193 • *Age - mean (SD)* : 84.2 (6.8) • *Female - N (%)*: 141 (73.1) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y (MMSE 17(7)) **Control** • *N*: 137 • *Age - mean (SD)* : 84.5 (6.6) • *Female - N (%)*: 98 (71.5) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y (MMSE 18(8) **Inclusion criteria:** 65 years or more; People with or without cognitive impairment living indefinitely in a nursing home place; Able to walk with or without any kind of help or able to self transfer (as defined in category d420 of the WHO International Classification of Functioning, Disability and Health) without help; Give their consent (or the legal guardian in case of cognitive impairment) **Exclusion criteria:** terminal illness; occupying temporarily a nursing home place (convalescence period) or another kind of place (day centre, long-term care, etc) **Pretreatment differences:** nil
Interventions	• Multifactorial falls prevention programme. Mini Falls Assessment Instrument and implementation of a multifactorial tailored programme to prevent falls. Interventions provided to address individual risk factors including: gait and balance impairment, cognitive impairment, polypharmacy, assistance with ADLs, lower limb pain, urinary incontinence, weakness, symptomatic heart disease, fear of falling, neuroleptics/ psychotropic drugs, problems in feet, dizziness, visual impairment, depressive symptoms. 3 sessions weekly of 45 minutes • Control. Falls risk assessment, without intervention actions and usual care.
Outcomes	• Rate ratio • Odds ratio for falling • Number of fractures
Duration of the study	12 months
Notes	Additional information provided by author correspondence
Risk of bias	

Salvà 2016 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Judgement comment: random draw with opaque envelopes
Allocation concealment (selection bias)	High risk	Judgement comment: allocation not concealed from the person performing recruitment, as per author correspondence
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: open-label trial
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by staff who were not blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: large loss to follow-up after randomisation which is greater in the control arm (41%); 2 centres in control arm left the study (65 participants); 1 centre in each arm provided no falls data (14 participants in intervention group, 32 participants in control group)
Selective reporting (reporting bias)	Low risk	Judgement comment: all outcomes reported as specified in trial record, some by author correspondence
Method of ascertaining falls	Unclear risk	Judgement comment: falls recording concurrent, unclear if a definition of falls was provided
Baseline imbalance	Low risk	Judgement comment: no major imbalances. Imbalance in those with depression in dementia, however numbers are small
Other bias	Low risk	Judgement comment: none identified

Sambrook 2012

Methods	RCT (cluster randomised by facility).
Participants	**Setting:** 51 aged care facilities (intermediate care), North Sydney, Australia N = 602 residents. 51 clusters. Sample: 71% women Age (years): mean 86.4 (SD 6.6) **Inclusion criteria:** aged ≥ 70; ambulant; likely to survive for ≥ 12 months **Exclusion criteria:** taking vitamin D or calcium supplements; history of skin cancer in previous 3 years

Interventions	• UV: increased sunlight exposure to face, hands and arms, 30 to 40 minutes, 5 days per week • UV+: increased sunlight exposure (as above) + calcium carbonate 600 mg daily • Control: usual care + brochure on vitamin D deficiency and its treatment
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures) • Adverse events
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The random allocation sequence ... was generated by a statistician who was not involved in the recruitment"
Allocation concealment (selection bias)	Low risk	Quote: "... it was concealed from the study coordinators until after randomisation."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study was cluster randomised and nursing staff reported falls. Researchers visited each home every two months to record falls
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: low loss to follow-up and ITT analysis
Selective reporting (reporting bias)	Low risk	Judgement comment: falls, fallers, risk ratio and rate ratio reported, adjusted for clustering
Method of ascertaining falls	Low risk	Judgement comment: clear definition, falls documented concurrently (in nursing notes and incident reports) and recorded by research staff monthly
Baseline imbalance	Low risk	Judgement comment: only significant difference in cognition at baseline adjusted for in analysis

Other bias	Low risk	Judgement comment: none identified

Saravanakumar 2014

Methods	RCT (individually randomised)
Participants	**Setting:** residential care facilities, mixed-level care, Australia **Baseline Characteristics** **Tai chi group** • *N*: 9 • *Age - mean (SD)* : 81.1 (8.0) • *Female - N (%)*: 8 (72.7%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Yoga group** • *N*: 9 • *Age - mean (SD)* : 84.9 (6.7) • *Female - N (%)*: 10 (90.9%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Usual care** • *N*: 11 • *Age - mean (SD)* : 85.4 (9.1) • *Female - N (%)*: 6 (54.5%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** aged 60 and over; able to stand with support; able to understand English; able to understand and follow simple instructions and demonstrations **Exclusion criteria:** severe debilitating illness; severe cognitive impairment; severe hearing or visual impairment (as determined by the RCF staff) **Pretreatment differences:** nil significant
Interventions	• Tai chi. Modified Tai Chi programme beginning with warm-up exercises of different joints and progressing through 18 individual Tai chi and qigong movement patterns, with repetitions for each pattern, using imagery, breathing and posture control. The movements were slow, controlled and circular using functional patterns and engaging the mind. Modifications were made for functional capacity. 30-minute classes twice weekly. • Yoga. Modified traditional yoga exercises (asanas), breathing (pranayama), synchronising movements with breathing and yoga nidra, a type of relaxation. To make it suitable for frail residents, more seated exercises and preparatory movements were included. 30-minute classes twice weekly.

	• Usual care. The care facility encouraged all residents to access the Staying Active programme with weekly half-hour seated exercise sessions; physical culture, games and group activities like bingo, group reading, story-telling, etc.; a gym with bicycles, pulleys and massage by trained staff; assisted and independent activities such as walking, gardening.
Outcomes	• Mean number of falls • Number of adverse events
Duration of the study	14 weeks
Notes	ACTRN12612000103864

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Permuted block randomisation with a block size of 6 was generated using MS Office Excel."
Allocation concealment (selection bias)	Low risk	Quote: "After baseline assessments, participants were randomly allocated to tai chi, yoga or usual care groups by a researcher not involved in recruitment who prepared the randomised list in sealed envelopes that were given to the facility staff a day before the commencement of the interventions."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls would recorded by care home staff in RCF records, who would not be blinded to the intervention
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: loss from groups Control 9%, Tai Chi 18%, Yoga 27%. Given small trial numbers losses may have impacted on outcomes
Selective reporting (reporting bias)	Low risk	Judgement comment: falls outcomes reported as per trial registration
Method of ascertaining falls	Low risk	Quote: "Falls were defined as 'events that resulted in a per- son coming to rest inadvertently on the ground or floor or other lower level, excluding intentional change in position to rest in furniture, wall or other objects' (WHO, 2007). Fall incidence information was collected from the records main-tained at the RCF. The data was collected for the period of 6

		months pre-intervention, intervention period and 6 months post-intervention period."
Baseline imbalance	Low risk	Judgement comment: no significant differences at baseline
Other bias	Low risk	Judgement comment: none identified

Schnelle 2003

Methods	RCT (individually randomised)
Participants	**Setting:** 4 nursing homes (high-level nursing care), USA N = 190 Sample: 85% women Age (years): mean (SD) intervention group 87.3 (8.0), control group 88.6 (6.7) **Inclusion criteria:** incontinent; no in-dwelling catheter; follows one stage commands; not Medicare Part A for post acute care or terminal; occupying long stay bed **Exclusion criteria:** none stated
Interventions	• "FIT": incontinence care and functional exercises delivered by research staff. Every 2 hours from 08.00 to 16.00, 5 days a week, for 8 months. At each session patients prompted to toilet and changed if wet; encouraged to walk (or mobilise in wheel chair if not ambulatory); carried out sit-to-stand exercises with minimal assistance; offered fluids to drink before and after each episode. Upper body resistance training (arm curls and arm raises) at one episode per day. Individually tailored to meet weekly goals (up to 8 sit-to-stands, and up to 10 minutes walking (wheeling) per episode) • Control: usual care
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures)
Duration of the study	8 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "... subjects were randomized within NHs by computerized programs into intervention and control groups."
Allocation concealment (selection bias)	Unclear risk	Insufficient information on process of allocation to permit judgement of 'Low risk' or 'High risk'

Schnelle 2003 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Falls recorded in medical records. Staff recording falls were likely to be aware of allocation status. Researchers examining records were blinded to allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: ITT analysis
Selective reporting (reporting bias)	Low risk	Judgement comment: falls, fallers, injurious falls, fracture falls and falls incidence reported
Method of ascertaining falls	High risk	Judgement comment: no falls definition reported
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified

Schoenfelder 2000

Methods	RCT (individually randomised)
Participants	**Setting:** 2 nursing homes (high-level nursing care), USA N = 16 Sample: 75% women Age (years): mean 82.8 (range 66 to 95) **Inclusion criteria:** aged ≥ 65; ambulating independently with or without assistive device; understand English; MMSE score > 20 **Exclusion criteria:** unstable physical condition; terminal illness; history of acting out or abusive behaviour
Interventions	• Supervised ankle strengthening exercises followed by up to 10 minutes of walking, total time 20 minutes, 3 x per wk for 3 months. Exercises individually tailored. Intervention delivered by research member • Control: usual care
Outcomes	• Rate of falls • Number of people falling • Adverse events
Duration of the study	6 months
Notes	
Risk of bias	

Schoenfelder 2000 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'. Quote: "... subjects were matched in pairs and assigned randomly within each pair to the intervention or control group."
Allocation concealment (selection bias)	High risk	Allocation concealment not described and researchers changed group allocation of one participant after randomisation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls after intervention were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: no loss to follow-up
Selective reporting (reporting bias)	High risk	Judgement comment: no fallers data reported
Method of ascertaining falls	Unclear risk	Judgement comment: methods of collecting falls data unclear, no definition provided
Baseline imbalance	High risk	Judgement comment: differences in gender and falls efficacy at baseline
Other bias	Low risk	Judgement comment: none identified

Serra-Rexach 2011

Methods	RCT (individually randomised)
Participants	**Setting:** 1 geriatric nursing home (intermediate-level care), Madrid, Spain N = 40 Sample: 80% women Age (years): mean 92 (SD 2) **Inclusion criteria:** aged ≥ 90; planning to stay in the same nursing home during the study; able to ambulate with or without cane, walker, or parallel bars); able to communicate; able and willing to consent **Exclusion criteria:** acute or terminal illness; myocardial infarction in previous 3 months; unstable medical condition; upper or lower extremity fracture in previous 3 months; severe dementia; neuromuscular disease; using drugs affecting neuromuscular function

Interventions	• Training group: training sessions 45 to 50 minutes per day, 3 days per week for 8 weeks (stretching exercises to warm up and cool down + aerobic training on cycle ergometer (up to 15 minutes), strength training with leg press with variable resistance (2 to 3 sets of 8 to 10 repetitions with rests between), + upper limb resistance training with weights or resistance bands. Also received usual care physiotherapy (mobility exercises, i.e. passive and active stretching of joints, 40 to 45 minutes per day, 2 days per week) • Control: usual care physiotherapy (mobility exercises, i.e. passive and active stretching of joints, 40 to 45 minutes per day, 5 days per week)
Outcomes	• Number of falls • Adverse events
Duration of the study	12 weeks (8 weeks intervention and further 4 weeks follow-up)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "computer generated randomization sequence"
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "The assessment staff was blinded to participant randomization assignment. Participants were... reminded not to discuss their randomization assignment with assessment staff." "An independent researcher was in charge of auditing all nursing and medical records to record the number of falls in each participant over the study period"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss to follow-up low and reasons balanced
Selective reporting (reporting bias)	Unclear risk	Judgement comment: falls a secondary outcome. Falls defined as adverse event in published protocol but not final publication
Method of ascertaining falls	Low risk	Quote: "In our study, we will define falls as "unexpected event in which the participants come to rest on the ground, floor, or other lower level" [61,62].

Serra-Rexach 2011 (*Continued*)

		An independent researcher will be in charge of auditing all nursing and medical records to record all falls in the participants over the study period." Judgement comment: falls defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: groups similar at baseline
Other bias	Low risk	Judgement comment: none identified

Shaw 2003

Methods	RCT (individually randomised)
Participants	**Setting:** 2 accident and emergency (A&E) departments, Newcastle, UK N = 308 Sample: 79% of participants lived in high and intermediate nursing care facilities (personal communication), (80% women) Age (years): mean 84 (range 71 to 97) **Inclusion criteria:** presenting to A&E after a fall; age ≥ 65; MMSE score < 24; consent from patient; immediate carer and next of kin **Exclusion criteria:** unable to walk; medical diagnosis likely to have caused index fall, e.g. stroke; unfit for investigation within 4 months; unable to communicate for reasons other than dementia; living outside of a 15-mile radius of recruitment site; no major informant
Interventions	• Multifactorial, multidisciplinary assessment and intervention to identify and manage risk factors. ○ Assessment of feet and footwear, gait and balance (physiotherapist): provision of walking aids and footwear, chiropody referral if required. Home-based tailored exercise programme supervised by physiotherapist (gait training, balance, transfer and mobility interventions, functional limb strengthening and flexibility exercises) for 3 months ○ Medical intervention comprised investigation and management of untreated medical problems, medication review, vision assessment and referral if indicated and psychogeriatric review if indicated ○ Cardiovascular review and advice and/or treatment of identified cardiac risk factors for falls ○ OT assessment of environmental fall hazards using a standard checklist, and hazard modification if indicated • Multifactorial, multidisciplinary assessment without intervention + usual care
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (hip fractures)
Duration of the study	12 months

Shaw 2003 (*Continued*)

Notes		

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "We randomised patients by block randomisation using computer generated random numbers"
Allocation concealment (selection bias)	Low risk	Quote: "Group allocation was performed by a researcher who was independent of the recruitment process and blind to baseline interview data"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Data from postcards processed and coded off site by researcher blind to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Most losses due to death, withdrawals low and balanced
Selective reporting (reporting bias)	Unclear risk	No protocol identified
Method of ascertaining falls	Low risk	Falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Balanced at baseline
Other bias	Low risk	None identified

Shimada 2004

Methods	RCT (individually randomised)
Participants	**Setting:** 1 long-term care facility (intermediate-level care), Japan N = 32 Sample: 78% women Age (years): mean (SD) intervention group 81.8 (5.9), control group 83.1 (6.4) **Inclusion criteria:** none stated **Exclusion criteria:** not able to walk more than 3 minutes on treadmill at greater than 0.5 km/hour; unable to participate because of recognisable dementia; unspecified health problems
Interventions	• Supervised perturbed gait exercises on a treadmill (individually tailored) for 6 months (gait, balance and co-ordination + endurance) in addition to usual exercise. Complete programme of 600 minutes over 6 months, 1 to 3 x per week. Intervention

Shimada 2004 (*Continued*)

	delivered by physical therapists • Usual exercise. Programs consisting of stretching, resistance training, group training, and outdoor gait training.
Outcomes	• Rate of falls • Number of people falling
Duration of the study	6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The 32 subjects were randomly divided into two groups ..." Insufficient information to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible for participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	Collection of falls data not described but states "This study ... was carried out without blinding." Staff who recorded falls were likely to be aware of individual's allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses similar between groups
Selective reporting (reporting bias)	Low risk	Judgement comment: fall rates, number of falls and time to first fall reported
Method of ascertaining falls	Unclear risk	Judgement comment: falls ascertainment not reported
Baseline imbalance	Low risk	Judgement comment: groups similar at baseline
Other bias	Low risk	Judgement comment: one detected

Shorr 2012

Methods	RCT (cluster randomised).
Participants	**Setting:** 16 nursing units in an urban community hospital, acute care, USA N = 27,672 participants. 16 clusters. Sample: not stated. Age (years): not stated. **Baseline Characteristics** **Automated tele-vigilance system** • *N*: 11,115 participants • *Age - mean (SD)* : NR • *Female - N (%)*: NR • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 17,436 participants • *Age - mean (SD)* : NR • *Female - N (%)*: NR • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Inclusion criteria:** admission to one of 16 general medical-surgical nursing units in Methodist Healthcare-University Hospital, Memphis, Tennessee, during period 1 May 2006 to 30 Oct 2007 **Exclusion criteria:** nil. **Pretreatment differences:** baseline characteristics of patients unknown. Staffing hours significantly differ between groups, but controlled for in analysis
Interventions	• Automated tele-vigilance system. Education, training, and technical support to promote use of a standard bed alarm system which uses 1 to 2 weight-sensitive sensor pads applied to the bed, chair or commode. When contact is broken this activates alarm in patient's room and call at nurses' station. Automated tele-vigilance system cameras installed, cameras can work in visible or infrared range, physically linked to a server that will store encrypted video and analyse images data in real-time, sending an alert to the care staff via their computers and personal pagers if it detects a fall. Physician can also watch images in order to determine the cause of the incident and then act preventively and induce treatment/care strategies. • Usual care
Outcomes	• Rate of falls • Risk of falling • Injurious falls • Number of adverse events
Duration of the study	Admission period. Trials recruitment over 18 months.
Notes	Additional data provided by author.

Shorr 2012 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated randomisation on the basis of baseline fall rates
Allocation concealment (selection bias)	High risk	Allocation of clusters unblinded and recruitment of participants in acute hospital wards occurred over May 2006 - Oct 2007 after cluster allocation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Blinding of staff not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcome assessors were likely to be unblinded due to the cluster randomisation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Loss of two clusters due to closure of the units, but reason for loss not related to outcome and appropriately accounted for in analysis
Selective reporting (reporting bias)	Low risk	Falls, fallers and injurious falls reported
Method of ascertaining falls	Low risk	Falls clearly defined and recorded concurrently
Baseline imbalance	Unclear risk	Staffing hour for all 3 staff types significantly differ between groups, but controlled for in analysis. However baseline characteristics at patient level not known
Other bias	Low risk	Allocation of clusters occurred in pairs of units with similar falls rates within one hospital which may allow the randomisation sequence to be predicted. However this issue already considered under allocation concealment. No other risk identified

Sihvonen 2004

Methods	RCT (individually randomised)
Participants	**Setting:** 2 residential care homes (intermediate-level care), Finland N = 28 Sample: 100% women Age (years): mean (SD) intervention group 80.7 (6.1), control group 82.9 (4.2) **Inclusion criteria:** aged ≥ 70; able to stand without walking aid; able to visualise feedback from a computer; able to follow instructions **Exclusion criteria:** acute illness; dementia; impending hip surgery

Sihvonen 2004 (*Continued*)

Interventions	• Balance training using computerised visual feedback and a force platform (gait, balance and co-ordination exercises), 20 to 30-minute sessions, 3 x per week, for 4 weeks. Exercises individually tailored. Intervention delivered by the research team • Control: usual care
Outcomes	• Rate of falls • Number of people falling
Duration of the study	12 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The subjects … were randomly assigned to an exercise group or a control group ... Since the study was carried out in two separate places, the randomization was done in blocks." "Randomisation was carried out by drawing lots."
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Falls recorded by participants who were aware of group allocation. No mention of blinding of researchers contacting participants for details or if no diary returned
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: low loss to follow-up, unlikely to affect outcome
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified. Falls and fallers reported, falls rate calculable
Method of ascertaining falls	Unclear risk	Judgement comment: falls recorded by participants
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified

Sitja Rabert 2015

Methods	RCT (individually-randomised, multicentre trial)
Participants	**Setting:** 10 residential care facilities, mixed-level care, Spain N = Sample: Age (years): **Baseline Characteristics** **WBV + exercise** • *N*: 81 • *Age - mean (SD)* : 82.30 (7.75) • *Female - N (%)*: 53 (65%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Exercise (control)** • *N*: 78 • *Age - mean (SD)* : 82.55 (7.12) • *Female - N (%)*: 54 (69%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** volunteers of either sex aged older than 65 years; resident in a nursing home; and able to adopt a squat position on a vibrating platform **Exclusion criteria:** acute illness (not resolved within 10 days); epilepsy; severe heart disease; use of a pacemaker; high risk of thromboembolism; a hip or knee replacement; musculoskeletal disorders; cognitive or physical disorders that could interfere with training methods **Pretreatment differences:** nil significant
Interventions	• Whole body vibration + exercise: static/dynamic exercises (balance and resistance training) performed on a vibratory platform (frequency: 30-35 Hz; Amplitude: 2 mm to 4 mm). 3 x per week for 6 weeks.Warm-up and cool down exercises performed at each session. 30-minute sessions, 3 sessions per week, training volume increased progressively. • Exercise alone: same exercise programme with no whole body vibration. Group-based progressive static and dynamic exercise programme, involving balance and strength training. Warm up and cool down exercises performed at each session without vibration platform.
Outcomes	• Number of fallers • Number with multiple falls • Number with fracture fall • Adverse events
Duration of the study	6 weeks, total follow-up 6 months
Notes	NCT01375790

Sitja Rabert 2015 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "computer generated randomisation list will be generated for participants at each nursing-home using the statistical software SPSS17."
Allocation concealment (selection bias)	Low risk	Quote: "Allocation to treatment will be centralized by telephone. All the researchers will be blinded to the randomisation sequence list."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: open-label trial
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls collected from nursing home staff or relatives who were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses to follow-up balanced between groups, reasons balanced
Selective reporting (reporting bias)	High risk	Judgement comment: number of falls not reported by group allocation
Method of ascertaining falls	Low risk	Judgement comment: fall definition provided in Clinical Trial Registry (NCT01375790).Quote: "Fall: an unexpected event in which the participants come to rest on the ground, floor, or lower level". Concurrently recorded. Additional information from author 11/7: Report calendar: During the study, every falls was registered in a register falls specially created by the study and data concerning falls were regularly collected from each nursing home or from relatives if a participant had moved to a different address. During the follow-up period, systematically every week the two blinded physiotherapists registered the falls occurred
Baseline imbalance	Low risk	Judgement comment: groups well balanced at baseline

Sitja Rabert 2015 (*Continued*)

Other bias	Low risk	Judgement comment: none identified

Stenvall 2007

Methods	RCT (individually randomised)
Participants	**Setting:** acute hospital wards (geriatric and orthopaedic), Umeå, Sweden N = 199 Sample: 74% women Age (years): mean 82.2 (SD 6.3) **Inclusion criteria:** admitted to hospital with femoral neck fracture; aged ≥ 70 **Exclusion criteria:** severe rheumatoid arthritis; severe hip osteoarthritis; pathological fracture of the femoral neck; severe renal failure; bedridden prior to the fracture
Interventions	• Post-operative care in a geriatric orthopaedic service in a geriatric ward: multidisciplinary team providing comprehensive geriatric assessment, management, and rehabilitation • Control: usual care in an orthopaedic ward
Outcomes	• Rate of falls • Number of people falling • Number sustaining a fracture (all fractures)
Duration of the study	32 months. Follow-up time was until participants were discharged from hospital
Notes	Dementia subgroup analysis published in Stenvall 2012.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Patients were randomized ... in opaque sealed envelopes. The lots in the envelopes were sequentially numbered ... Persons not involved in the study performed these procedures."
Allocation concealment (selection bias)	Low risk	Used sequentially numbered, opaque sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: "The staffs on the intervention and control wards were not aware of the nature of the present study."

Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: all patients included in analysis (ITT)
Selective reporting (reporting bias)	Low risk	Judgement comment: no protocol identified but falls outcomes thoroughly reported. Falls, fallers, falls incidence and fracture falls reported, plus data by dementia subgroup
Method of ascertaining falls	Low risk	Judgement comment: falls clearly defined and recorded concurrently
Baseline imbalance	Low risk	Judgement comment: significant imbalance in depression and non-significant imbalance in dementia at baseline adjusted for in analyses
Other bias	Low risk	Judgement comment: none detected

Streim 2012

Methods	RCT
Participants	**Setting:** residents in nursing homes and assisted living facilities within 30 miles of Philadelphia, USA. Mixed levels of care N = 94 (36 randomised, 56 in a non-randomised patient preference arm) Sample: NR Age (years): NR **Baseline Characteristics** Age (years): range 60 to 95. Baseline characteristics not provided **Inclusion criteria:** 65 years and older; ambulatory; cognitively intact or with mild-moderate impairment but capable of self-reporting depression symptoms; receiving antidepressant treatment for a single episode of depression; in full remission for at least six months **Exclusion criteria:** bedridden; severe cognitive impairment **Pretreatment differences**: no differences in race and gender. Differences in medication use at baseline (benzodiazepines P = 0.034, serotonin norepinephrine reuptake inhibitors P = 0.0004, Lexapro P < 0.0001)
Interventions	• Discontinue taking antidepressants • Control: continue taking antidepressants A third non-randomised arm of people choosing to discontinue antidepressants
Outcomes	• Number of falls per week Other outcomes not included in this review, e.g. depression and cognition
Duration of the study	• Odds of fall

Streim 2012 (*Continued*)

Notes	Trial identified as an abstract only, with no falls results reported. Excerpt from unpublished manuscript provided by author correspondence

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Judgement comment: insufficient information for judgement
Allocation concealment (selection bias)	Unclear risk	Judgement comment: insufficient information for judgement
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Judgement comment: insufficient information for judgement
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Judgement comment: insufficient information for judgement
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: insufficient information for judgement
Selective reporting (reporting bias)	Unclear risk	No details of falls outcomes provided in trial registry
Method of ascertaining falls	Unclear risk	Judgement comment: insufficient information for judgement
Baseline imbalance	High risk	Judgement comment: differences in medication use between randomised groups at baseline
Other bias	Unclear risk	Judgement comment: imbalance in randomisation due to high number of patients choosing third 'preference' arm of study

Tideiksaar 1993

Methods	RCT (individually randomised)
Participants	**Setting:** acute geriatric care hospital ward, New York city, USA N = 70 Sample: 86% women Age (years): mean 84 (range 67 to 97) **Inclusion criteria:** one or more abnormal factors on a 9 point performance orientated environmental mobility screen (indicating impaired bed mobility) **Exclusion criteria:** none stated
Interventions	• Bed alarm system to alert staff when patient leaves their bed. Intervention delivered by nurses • Control: usual care

Tideiksaar 1993 (*Continued*)

Outcomes	• Rate of falls • Adverse events
Duration of the study	9 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients … were randomly assigned to either the experimental group … or the control group". Insufficient information to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff who recorded falls not blinded to individual participants' allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: no loss to follow-up, acute setting
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Unclear risk	Judgement comment: definition of falls provided but not clearly. Falls recorded concurrently
Baseline imbalance	Unclear risk	Judgement comment: baseline characteristics not reported
Other bias	Low risk	Judgement comment: none identified

Toulotte 2003

Methods	RCT (individually randomised)
Participants	**Setting:** nursing care facility, France. Published data implies residents receiving mixed high and intermediate levels of care N = 20 Sample: % women not stated Age (years): mean 81.4 (SD 4.7) **Inclusion criteria:** dementia (MMSE score < 21); history of ≥ 2 falls (not involving an

	environmental hazard) in previous 3 months; able to walk 10 metres without human assistance **Exclusion criteria:** none stated
Interventions	• Supervised exercises 1 hour, 2 x per week for 16 weeks in groups of 5. Exercises incorporated gait, balance and co-ordination, strength/resistance, and flexibility. Exercises not individually tailored. Two physicians delivered intervention in each group. Individualised assessment of participants not part of intervention • Usual care
Outcomes	• Rate of falls
Duration of the study	4 months follow-up
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "A randomised cross-over design was used." Insufficient information about the sequence generation process to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Physician conducting tests was blinded to allocation status. Unlikely that these tests included recording of falls. Staff who recorded falls likely to be aware of individual participants' allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss to follow-up unclear
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Unclear risk	Judgement comment: falls clearly defined but method of recording falls unclear
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified

Treacy 2015

Methods	RCT (individually randomised)
Participants	**Setting:** general rehabilitation ward (subacute) at one hospital, Australia N = Sample: Age (years): **Baseline Characteristics** **Standing balance circuit classes** • *N*: 81 • *Age - mean (SD)* : 82.6 (7.3) • *Female - N (%)*: 51 (62%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Usual care** • *N*: 81 • *Age - mean (SD)* : 81.4 (7.8) • *Female - N (%)*: 53 (65%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** 18 years or over; admission to the general rehabilitation ward a Bankstown-Lidcombe Hospital, NSW; no medical contraindications to exercise; able to: fully weight bear; stand unaided independently for at least 30 seconds; and participate in group therapy sessions with minimal supervision **Exclusion criteria:** 1. unable to fully weight bear as ordered by a medical officer (i.e. non, partial or touch weight bearing status through one or both legs).2. Have a medical condition precluding exercise, e.g. unstable cardiac disease, uncontrolled hypertension, uncontrolled metabolic diseases, large abdominal aortic aneurysm. 3. Have an identified multi-resistant organism infection or other infection that would pose a significant risk to others in a group setting **Pretreatment differences:** no imbalances. See online appendix.
Interventions	• Standing balance circuit classes. Group training, supervised by 2 physiotherapists standing balance circuit class programme focused on posture whilst standing and stepping. Involving 7 exercise stations, with 3 levels of difficulty, each with a specific balance exercise, plus standard rehabilitation. Six 1-hour classes over 2 weeks. • Usual care. Assessment and treatment by the multidisciplinary ward team. Patients are predominately treated within a group setting in physiotherapy with additional one-to-one sessions as required with the focus being on weight bearing exercises. Outpatient therapy, as required. Once or twice per day at least two hours per day.
Outcomes	• Rate ratio for falls • Adverse events
Duration of the study	2 weeks

Notes		
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "The allocation schedule was computer generated using randomly ordered blocks of four and six. "
Allocation concealment (selection bias)	Low risk	Quote: "A concealed allocation procedure (numbered sealed opaque envelopes)" Quote: "Randomisation schedule and envelopes were prepared and held by a staff member not involved in study recruitment or intervention. Participants and therapists were made aware of group allocation once the envelopes had been opened." Judgement comment: allocation adequately concealed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls reliant on self-report, person responsible and hospital incident reporting system. Not possible to blind staff
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: little loss to follow-up at 2 weeks
Selective reporting (reporting bias)	Low risk	Judgement comment: protocol available and outcomes reported as planned
Method of ascertaining falls	Low risk	Judgement comment: Trial registry: "Fall incidence will be measured by participant and/or 'person responsible' self-report. In-patient fall data will also be collected via the hospital Incident Information Management System (incident reporting system)."Hospital system will have clear definition and concurrent recording of falls"
Baseline imbalance	Low risk	Judgement comment: groups balanced on range of demographic variables at baseline
Other bias	Low risk	Judgement comment: none detected

Tuunainen 2013

Methods	RCT (individually randomised)
Participants	**Setting:** residential care facility, mixed-level care, Finland N = Sample: Age (years): **Baseline Characteristics** **Strength training** • *N*: 18 • *Age - mean (SD)* : 84.7 (5.5) • *Female - N (%)*: 12 (67%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Balance and strength training** • *N*: 18 • *Age - mean (SD)* : 85 (4.2) • *Female - N (%)*: 16 (89%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Self-administered training** • *N*: 19 • *Age - mean (SD)* : 86.1 (7.3) • *Female - N (%)*: 14 (74%) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** participant's ability to raise himself/herself from a chair without using hands or arms for support.Willingness to participate. Prof Pyykko stated in email 31/8/16 additional inclusion criteria: could move independently, arise from a chair 5 times in a row, follow instructions **Exclusion criteria:** nil stated. **Pretreatment differences:** the Strength training group had 33% male, The Balance and Muscle Training had 11% male, The self-administered group had 26% male.In the Strength training group, 39% were prescribed sleeping medications. In the balance and muscle training group, 56% were prescribed sleeping medications. In the self-administered group, 68% were prescribed sleeping medications
Interventions	• Strength training. groups of 5, under supervision by 2 physiotherapists. Progressively graded strengthening exercises for hip and other postural muscles using 1.2 kg weights attached to ankles from 6th session and using stairs from the 19th session. Twice-weekly for approx 1 hour. • Balance and strength training. Groups of 5, under supervision by 2 physiotherapists. Progressively challenging balance tasks. Strength training similar to strength training group but ankle weights not used. Twice weekly for approx 1 hour. • Self-administered training. Nurses provided encouragement to keep to self-guided

	training tasks. Written exercise instructions provided by physiotherapists, comprising stretching from a sitting position, crouching and rising. Twice-weekly for approx 1 hour.
Outcomes	• Number of falls • Falls rate • Number of fallers • Number with multiple falls • Compliance Other outcomes not included in this review
Duration of the study	13 weeks. Follow-up 3 years.
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: drawing of envelopes
Allocation concealment (selection bias)	Unclear risk	Judgement comment: details of allocation concealment not reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by ward nurses who are unlikely to be blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: it appears that the residents in the intervention groups who stopped training were not included in the analysis
Selective reporting (reporting bias)	Unclear risk	Judgement comment: study protocol not available. In correspondence, author states data on fracture falls data was collected but not included
Method of ascertaining falls	Unclear risk	Judgement comment: insufficient information to enable judgement
Baseline imbalance	High risk	Judgement comment: larger proportion of prescribing of sleeping medications in the Self administered group may have contributed to that group's higher falls rate

Other bias	Low risk	Judgement comment: none identified

Van de Ven 2014

Methods	RCT (cluster randomised).
Participants	**Setting:** 34 units from 11 residential care facilities, high-level care, the Netherlands N = 318. 11 clusters. Sample: 75% women Age (years): 84.7 **Baseline Characteristics** **Dementia care mapping** • *N*: 154 • *Age - mean (SD)* : 84.8 (6.0) • *Female - N (%)*: 118 (76.6) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Usual care** • *N*: 164 • *Age - mean (SD)* : 84.59 (6.6) • *Female - N (%)*: 121 (73.8) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Inclusion criteria** (facilities): those with Dementia-Special Care Units (DSCUs).Residents: Age of 65 years or more;Dementia diagnosed by an elderly care physician according to the Diagnostic and statistical manual of mental disorders-IV criteria for dementia; Approval of the elderly care physician for inclusion; At least one of the following neuropsychiatric symptoms: aggression, motor or verbal agitation,psychosis, depression, and apathy; Informed consent given by the residents themselves, their families, or their legal guardians; The resident must use the common areas, such as the shared living room, at least 4 hours a day **Exclusion criteria:** residents: an estimated life expectancy of 6 weeks; those who are physically unable to spend time in common areas of the facility; withdrawal of consent **Pretreatment differences:** the intervention and control groups differed in terms of the proportions of staff in permanent positions. There were no other statistically significant differences at baseline between the intervention and control groups
Interventions	• Dementia Care Mapping (DCM) based on principles of person-centred care, involving action plans based on systematic observations of care. Nurses received DCM training, a DCM organisational briefing day and conducted the 4-months DCM-intervention twice during the study. single DCM cycle consists of observation, feedback to the staff, and action plans for the residents. 10 staff members attended basic and advanced training to become certified DCM mappers, then attended an organisational briefing day. Intervention delivered twice.

	• Usual care without DCM training.
Outcomes	• Number of falls • Falls rate • Costs
Duration of the study	18 months
Notes	Author contact: Geertje van de Ven, Radboud University, G.vandeVen@elg.umcn.nl. Author clari fied study details by email. Dutch Trials Registry NTR2314http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2314

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Judgement comment: computer-generated sequence "soft-ware"
Allocation concealment (selection bias)	Low risk	Quote: "Randomisation will take place after the study sample has been recruited and informed consent has been given,"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by staff who are not blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: attrition rates due to no medical file higher in the intervention group (44% 68/154) vs control group (21% 35/164). (Fig 1). Unclear if medical file is source of falls data
Selective reporting (reporting bias)	Low risk	Judgement comment: outcomes reported as per protocol
Method of ascertaining falls	Unclear risk	Judgement comment: insufficient information for judgement
Baseline imbalance	High risk	Judgement comment: large difference in baseline fall rates. Baseline data for many potential confounders for falls outcomes not recorded
Other bias	Low risk	Judgement comment: none identified

Van Gaal 2011a

Methods	RCT (cluster randomised by ward)
Participants	**Setting:** 6 nursing homes, 10 wards (high-level nursing care), the Netherlands N = 392 participants included in study. 10 clusters. Sample: 66% women Age (years): mean (SD) intervention group 78 (9.9), control group 78 (11.7) **Inclusion criteria** (facilities): 2 or 4 more or less comparable wards. Inclusion criteria (residents): none stated **Exclusion criteria:** none stated
Interventions	• Implementation of 3 guidelines (falls, urinary tract infection, pressure ulcers) targeting ward nursing staff ○ Educational meetings for all nurses (90 minutes) on the causes of 3 adverse events, assessment of patients at risk and prevention ○ Two case discussions on every ward (30 minutes) covering these topics ○ CD-ROM with education material issued to every ward (information, test and feedback) ○ Information leaflets and oral information regarding prevention of pressure ulcers, urinary tract infection and falls issued to at-risk patients ○ Nurses recorded presence or absence of adverse events in a computerised registration system daily. This programme generated feedback on process and outcome indicators to the nurses • Control: usual care
Outcomes	• Rate of falls
Duration of the study	23 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	States randomised after stratification. Insufficient information to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding of staff not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff recording falls would be aware of allocation. Cluster randomised trial so likely the person collecting data from patient files would be aware also
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: high loss to follow-up for Van Gaal 2011a (nursing home setting)

Van Gaal 2011a (*Continued*)

Selective reporting (reporting bias)	Low risk	Judgement comment: falls reported as per protocol and adjusted for clustering
Method of ascertaining falls	High risk	Judgement comment: falls clearly defined but reliant on existing reporting systems in patient records which may vary between sites
Baseline imbalance	High risk	Judgement comment: 2011a (NH): nurse characteristics balanced at baseline but significant difference in physically impaired patients (reviewer $P < 0.001$ Chi^2), rehabilitation patients (reviewer Chi^{-2} $P < 0.001$)
Other bias	Low risk	Judgement comment: none detected

Van Gaal 2011b

Methods	RCT (cluster randomised).
Participants	**Sample:** 4 hospitals (acute care), 10 wards, the Netherlands N = 2201 participants included in study. 10 clusters. Sample: 55% women Age (years): mean (SD) intervention group 66 (14.5), control group 64 (16.9) **Inclusion criteria** (hospitals): 2 or 4 more or less comparable wards. Inclusion criteria (patients): expected length of stay of ≥ 5 days **Exclusion criteria:** none stated
Interventions	• Implementation of 3 guidelines (falls, urinary tract infection, pressure ulcers) targeting ward nursing staff ○ Educational meetings for all nurses (90 minutes) on the causes of 3 adverse events, assessment of patients at risk and prevention ○ Two case discussions on every ward (30 minutes) covering these topics ○ CD-ROM with education material issued to every ward (information, test and feedback) ○ Information leaflets and oral information regarding prevention of pressure ulcers, urinary tract infection and falls issued to at-risk patients ○ Nurses recorded presence or absence of adverse events in a computerised registration system daily. This programme generated feedback on process and outcome indicators to the nurses • Control: usual care
Outcomes	• Rate of falls
Duration of the study	23 months
Notes	
Risk of bias	

Van Gaal 2011b (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	States randomised after stratification. Insufficient information to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding of staff not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff recording falls would be aware of allocation. Cluster-randomised trial so likely the person collecting data from patient files would be aware also
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: high loss to follow-up for van Gaal 2011b (hospitals)
Selective reporting (reporting bias)	Low risk	Judgement comment: falls reported as per protocol and adjusted for clustering
Method of ascertaining falls	High risk	Judgement comment: falls clearly defined but reliant on existing reporting systems in patient records which may vary between sites
Baseline imbalance	High risk	Judgement comment: 2011a (NH): nurse characteristics balanced at baseline but significant difference in physically impaired patients (reviewer $P < 0.001$ Chi-^{2}), rehabilitation patients (reviewer Chi-^{2} $P < 0.001$).
Other bias	Low risk	Judgement comment: none detected

Van het Reve 2014

Methods	RCT (individually randomised)
Participants	**Setting:** residential care, intermediate-level care, Switzerland (13 care facilities) and Germany (1 facility) N = 182 Sample: 55% women Age (years): 81.5 (SD 7.3) **Baseline Characteristics** **Strength-balance-cognitive training** • *N*: 88 • *Age - mean (SD)* : 81.1 (8.3) • *Female - N (%)*: 49 (58.3) • *Medical status defined? - Y/N*: Y

	• *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Strength-balance training** • *N*: 94 • *Age - mean (SD)* : 81.9 (6.3) • *Female - N (%)*: 52 (53.1) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: Y **Inclusion criteria:** older than 65 years; able to walk 20 meters with or without aids; signed informed consent statement **Exclusion criteria:** severe cognitive impairment (Mini-Mental State Examination below 22 points); rapidly progressive or terminal illness, acute illness or unstable chronic illness **Pretreatment differences:** nil significant
Interventions	• Multiple intervention: strength-balance-cognitive training. Same exercise programme as strength-balance training group plus a computer-based cognitive training programme, with a focus on improving attention. Cognitive intervention: 10 minutes, 3 times per week. Exercise programme: 30 minutes resistance and 10 minutes balance training, 2 times per week. • Exercise: strength-balance training. Exercise programme consisting of progressive resistance training on age-adapted machines and balance training. Flexibility exercises followed each training session. 30 minutes resistance and 10 minutes balance training, 2 times per week
Outcomes	• Falls rate • Number of falls • Number of fallers • Compliance
Duration of the study	15 months comprising 12 weeks intervention and 12 months post-intervention follow-up period
Notes	ISRCTN75134517

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "using simple (unrestricted) randomisation [70] based on a table of random numbers."
Allocation concealment (selection bias)	High risk	Judgement comment: an "assessor" performed the randomisation and group allocation

Van het Reve 2014 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: unable to blind participants and personnel
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls calendars filled in by staff. "Blinding of investigator was not possible."
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: approximately 25% missing data for falls
Selective reporting (reporting bias)	Low risk	Judgement comment: protocol available (ISRCTN75134517) with falls reported as per protocol
Method of ascertaining falls	Low risk	Quote: "Falls, defined as 'unexpected events in which the participant comes to rest on the ground, floor or lower level, were assessed from 6 months retrospectively to 12 months prospectively using a fall calendar."
Baseline imbalance	Low risk	Judgement comment: nil significant
Other bias	Low risk	Judgement comment: four participants with vision impairment reallocated to control group, however, this number is small relative to intervention group sizes

Wald 2011

Methods	CCT (odd vs even medical record number)
Participants	**Setting:** acute medical units in 1 hospital, Colorado, USA N = 217 Sample: 55% women Age (years): mean (SD) intervention group 80.5 (6.5), control group 80.7 (7.0) **Inclusion criteria:** aged ≥ 70 **Exclusion criteria:** patients admitted to medical subspecialty service (cardiology, pulmonary, oncology)
Interventions	• Hospitalist-run acute care for the elderly service (ACE) (interdisciplinary team approach): admitted to 12-bed medical unit when beds available, attendance of patients by doctor with additional training in geriatrics, standardised geriatric assessment, daily (Monday to Friday), interdisciplinary rounds focusing on geriatric syndromes, standardised geriatric screens, clinical focus on mitigating harm and discharge planning; novel inpatient geriatrics training curriculum • Control: usual care. Admitted to general internal medicine unit with general medical teams with daily discharge planning rounds with social worker and discharge

Wald 2011 (*Continued*)

	planner
Outcomes	• Rate of falls
Duration of the study	22 weeks
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	A systematic non-random method was used (odd /even case record number)
Allocation concealment (selection bias)	High risk	Not possible to blind prior to allocation (see above)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Falls from hospital event reports. Last digit of medical record number was used for group allocation. Allocation not concealed
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: all patients included in analyses of other outcomes. Falls incidence per patient days reported
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Unclear risk	Judgement comment: falls definition not reported. Falls determined from standard reporting system which will be concurrent
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none identified.

Walker 2015

Methods	RCT (cluster randomised).
Participants	**Setting:** 6 residential care facilities, mixed-level care, UK N = 52 residents. 6 clusters. Sample: 67% women Age (years): 83 **Baseline Characteristics** **Implementation of the Guide to Action Care Home tool** • *N*: 25 (3 sites) • *Age - mean (SD)* : 84 (14.8) • *Female - N (%)*: 18 (72%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 27 (3 sites) • *Age - mean (SD)* : 82 (13.4) • *Female - N (%)*: 17 (63%) • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: N • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: N **Inclusion criteria:** care homes: listed on the Care Quality Commission database, long stay, old age, dementia or learning disability registration, nursing/residential registration, over 10 residents, no prior experience of Guide to Action Care Home. Care homes were purposively selected from those who replied expressing interest, to reflect a range of ownership, size and registration. Residents: (high risk): aged over 50 years, fallen at least once in the past year **Exclusion criteria:** bed-bound, hoist-dependent or terminally ill **Pretreatment differences:** nil
Interventions	• Implementation of the Guide to Action Care Home tool. Training in Guide to Action Care Home tool (a checklist of falls risk factors with suggested actions), with reference manual and certificate on training completion. Plus standard care. Intervention takes 15 to 20 minutes, can lead to interventions which take an average of 2 hours to complete. • Usual care. Access to standard care, but no Guide to Action Care Home training or manual.
Outcomes	• Falls rate • Injurious falls rate
Duration of the study	6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement

Walker 2015 (*Continued*)

Random sequence generation (selection bias)	Low risk	Judgement comment: randomisation done
Allocation concealment (selection bias)	Low risk	Judgement comment: allocation concealed according to standard operating procedure
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: 26% missing data (7/27) from control arm vs 12% intervention arm
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls would have been recorded by staff who would not be blinded to the intervention (staff training)
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: 26% missing data (7/27) from control arm vs 12% intervention arm
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Low risk	Judgement comment: no information on most potential confounders (e.g. medical status, dependency)
Baseline imbalance	Unclear risk	Judgement comment: none detected
Other bias	Low risk	Judgement comment: blinding not feasible

Ward 2010

Methods	RCT (cluster randomised by facility).
Participants	**Setting:** 88 residential aged care facilities (high-care, low-care and dementia-specific), New South Wales, Australia N = 5391 residents. 88 clusters. Sample: 73% women Age (years): median age 86 **Inclusion criteria** (facilities): ≥ 20 beds **Exclusion criteria:** none stated
Interventions	• Intervention: full-time project nurse to assist facilities in using evidence-based approaches to falls injury prevention relating to risk assessment; mobility assessment; use of hip protectors; calcium and vitamin D supplementation; continence management; exercise programs; appropriate footwear; medication review; and post-fall management review. Project nurse provided intervention facilities with information and resources on preventing falls and fractures. Initial training session followed by 3-monthly network meetings. Intervention staff also could attend workshop on planning and running exercise programs • Control: usual care. Staff attended a workshop where data collection procedures were explained

Ward 2010 (*Continued*)

Outcomes	• Number of falls • Number sustaining a fracture (hip fractures)
Duration of the study	17 months
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "randomly allocated within strata into intervention or control groups by the statistician ... using the procedure "surveyselect" in SAS statistical software"
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Staff recording falls and carrying out monthly record audit were aware of group allocation. Failure to produce monthly data followed up by project nurse (also aware of group allocation)
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement comment: losses balanced but large loss of 3 facilities/arm of study
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified, fallers data not reported
Method of ascertaining falls	High risk	Judgement comment: no definition of falls. Fall data retrieved by facility through record audit - likely to be variable reporting between facilities
Baseline imbalance	High risk	Judgement comment: although data in table 1 (limited participant variables) show reasonable balance between groups, there was moderate difference (2 falls /month) between groups in the 7-month pre-intervention falls data
Other bias	Low risk	Judgement comment: none detected

Whitney 2017

Methods	Cluster RCT (pilot, cross-over study).
Participants	**Setting:** Four nursing homes and five residential homes in London, UK, mixed-level care, 97% cognitively impaired. 9 clusters: 5 intervention, 4 usual care N = 191 participants. 9 clusters. Sample: 69% women Age (years): mean 83.5 (SD 8.8) **Baseline Characteristics** **Individualised fall prevention programme** • *N*: 103 • *Age - mean (SD)* : 84.6 (5.6) • *Female - N (%)*: 92 (46.0) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: N • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 88 • *Age - mean (SD)* : 84.1 (7.7) • *Female - N (%)*: 173 (56.1) • *Medical status defined? - Y/N*: Yes • *Falls risk defined (with valid tool at baseline)? -Y/N*: No • *Dependency defined? Y/N*: No • *Cognitive status defined? Y/N*: N **Inclusion criteria:** over 65 years; admitted to rehabilitation ward **Exclusion criteria:** restricted to bed; refused to participate **Pretreatment differences** (phase 1): longer stay in the control group patients (P <0.001) ; higher percentage of females in the control group (P =0.03)
Interventions	• Multifactorial intervention (exercise, dementia related behaviour management, comprehensive geriatric assessment including medication review, staff training, movement sensors). Falls risk assessment and management: including medical interventions, environmental modifications, equipment modifications, cognitive and behavioural treatment, family guidance. Mobility restrictions and optimising location on ward instituted in high risk patients. For moderate-risk patients mobility (transfers, walking, toilets usage, etc.) was done only under supervision and/or assistance of a professional staff member. High-risk patients had permanent personal supervision. Weekly assessment. • Usual care. Any activities undertaken by the participants recommended or administered by their treating team
Outcomes	• Rate ratio • Risk ratio • Numbers on injurious falls and fractures • Adverse events
Duration of the study	6 months
Notes	Costs of the programme to be reported. Other outcomes not included in this review

Whitney 2017 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Judgement comment: computer-generated randomisation, stratified by nursing home beds
Allocation concealment (selection bias)	Low risk	Judgement comment: randomisation conducted by separate clinical trials unit. Allocation concealed and no recruitment after allocation revealed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: blinding not feasible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by care home staff who were not blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	Judgement comment: 17.8% loss to follow-up. Large amounts of missing data on many outcomes (up to 60%). Not clear what loss to follow-up/missing data for falls outcome
Selective reporting (reporting bias)	Low risk	Judgement comment: falls, fallers, injury and fracture falls data reported
Method of ascertaining falls	Unclear risk	Judgement comment: falls definition used. Facilities used their "usual reporting mechanisms" for falls - no detail of what these mechanisms were or if they varied substantially between facilities
Baseline imbalance	High risk	Judgement comment: significant baseline differences in number of medical conditions, time to complete Timed Up and Go, and likelihood on being in nursing home bed. Although analysis involved some adjustments (for the baseline score on the outcome being investigated) it does not appear these baseline differences were adjusted for across the outcome measures
Other bias	Low risk	None detected

Wolf 2013

Methods	RCT (individually randomised)
Participants	**Setting:** Subacute hospital setting, single geriatric ward, Germany N = 98 Sample: 65% women Age (years): 76.1 **Baseline Characteristics** **Bed-exit alarm** • *N*: 48 • *Age - mean (SD)* : NR • *Female - N (%)*: NR • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: N **Usual care** • *N*: 50 • *Age - mean (SD)* : NR • *Female - N (%)*: NR • *Medical status defined? - Y/N*: N • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: N **Inclusion criteria:** patients at high risk of falls defined by a score of 3 or more in STRATIFY; requirement for assistance with mobilization during resting time **Exclusion criteria:** immobility; participation in another trial **Pretreatment differences:** NR
Interventions	**Intervention Characteristics** • Bed-exit alarm. Patients fitted with sensors to upper leg at rest time. Based on Wireless Sensing Triple Axis Reference Design. Sensors worn during rest periods 1 to 3 pm and 8 pm to 6 am. • Usual care
Outcomes	• Number of falls • Number of fallers
Duration of the study	During admission period, total trial period 13 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Judgement comment: method of sequence generation not described in adequate detail
Allocation concealment (selection bias)	Unclear risk	Judgement comment: no information provided about allocation methods

Wolf 2013 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Judgement comment: falls recorded by nurses who were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: author correspondence indicated no loss to follow-up
Selective reporting (reporting bias)	Unclear risk	Judgement comment: no protocol identified
Method of ascertaining falls	Unclear risk	Judgement comment: no falls definition provided, standard reporting mechanisms used
Baseline imbalance	Unclear risk	Judgement comment: inadequate details on baseline characteristics of patients to make a judgement
Other bias	Low risk	Judgement comment: none detected

Yokoi 2015

Methods	RCT (cluster randomised).
Participants	**Setting:** 5 residential care facilities, intermediate-level care, Japan N = 105 participants. 5 clusters. Sample: 60% women Age (years): 79.4 **Baseline Characteristics** **Short stick exercises** • *N*: 51 • *Age - mean (SD)* : 80.2 (7.9) • *Female - N (%)*: 33 (64.7) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y (TUG) • *Dependency defined? Y/N*: Y (independent for inclusion & SF-8 performed) • *Cognitive status defined? Y/N*: Y (MMSE) **Usual care** • *N*: 54 • *Age - mean (SD)* : 78.5 (5.2) • *Female - N (%)*: 30 (55.6) • *Medical status defined? - Y/N*: Y • *Falls risk defined (with valid tool at baseline)? -Y/N*: Y • *Dependency defined? Y/N*: Y • *Cognitive status defined? Y/N*: Y **Inclusion criteria** (facilities): with 50 beds in the Kinki area in Japan; where no in-

	tervention for fall prevention was conducted. Residents: able to walk without assistive devices and take care of themselves without assistance; had sufficient cognition to follow directions; had never performed an SSE before; were allowed by their chief physician to exercise **Exclusion criteria:** residents: with dementia or severe cardiac, pulmonary or musculoskeletal disorders that are associated with a higher fall risk **Pretreatment differences:** BMI significantly less in the Intervention group, but as both groups were in normal range, probably would not have had impact on outcome
Interventions	**Intervention Characteristics** • Short stick exercises. Group-based supervised short stick exercises, performed in a seated position, and performing 6 activities with a rolled Japanese newspaper as the stick (warm up included). 25 minute sessions, twice weekly. • Usual care. Daily housekeeping, hobbies, work and 10-minute group stretching exercises were continued.
Outcomes	• Time to first fall • Number of falls • Number of fallers • Compliance
Duration of the study	12 months, 6 months intervention period.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Judgement comment: randomisation of the 5 facilities was by lottery using envelopes by a researcher not involved with study. Insufficient information but reason for not using sequence generation not really valid despite only 5 facilities, so some risk of bias
Allocation concealment (selection bias)	Unclear risk	Judgement comment: allocation of facilities probably adequate, assuming envelopes were sealed and opaque. It does not appear that individual participant recruitment was completed prior to cluster allocation. The study states that research assistants were not informed of the results of randomisation, but it appears that the research assistants were involved with falls data collection, not with recruitment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not possible to blind participants. Highly unlikely that personnel could be blinded

Yokoi 2015 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: "Research assistants were not informed of the results of the randomization. The staff was asked not to tell the research assistants about which group was undergoing the intervention." Judgement comment: unblinding is likely
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: losses to follow-up balanced between groups with similar reasons
Selective reporting (reporting bias)	High risk	Judgement comment: the number of falls were not reported
Method of ascertaining falls	Unclear risk	Falls determined by interviews with staff and medical records. Not clear whether staff were asked to recall for periods longer than one month. Unclear whether the method and reliability of staff recording falls in patient records were the same in all the facilities
Baseline imbalance	Unclear risk	Judgement comment: baseline characteristics of individuals in the facilities appear to be reasonably balanced although BMI significantly different but both groups within normal range for BMI so not likely to be important. Baseline characteristics of the facilities were not compared - in particular the rates of falling in each of the facilities prior to the intervention
Other bias	Low risk	None detected

Zermansky 2006

Methods	RCT (individually randomised)
Participants	**Setting:** 65 care homes for the elderly (high, intermediate and mixed levels of care), UK N = 661 Sample: 77% women Age (years): mean 85 (interquartile range 80 to 90) **Inclusion criteria:** aged ≥ 65; resident in a care home with ≥ 6 residents **Exclusion criteria:** participating in another trial; terminally ill; already receiving clinical medication review; at GP request
Interventions	• Clinical medication review by a pharmacist comprising a review of the GP record and consultation with the participant and their carer. Written recommendations forwarded to participant GPs • Control: usual care
Outcomes	• Rate of falls • Number of people falling

Zermansky 2006 (*Continued*)

Duration of the study	6 months	
Notes		
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "After collection of baseline data, patients were randomised in randomly sized blocks of two to eight patients using an algorithm written in Visual Basic in Microsoft Access."
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement comment: not possible to blind the intervention
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Falls data collected from accident book. Unclear whether staff recording falls in accident book would have been aware of allocation status
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement comment: loss to follow-up similar in both groups, as was main reason for loss (death)
Selective reporting (reporting bias)	Unclear risk	Judgement comment: falls not reported as an outcome in trial registration
Method of ascertaining falls	High risk	Judgement comment: no falls definition reported
Baseline imbalance	Low risk	Judgement comment: groups balanced at baseline
Other bias	Low risk	Judgement comment: none detected

A&E: emergency department
ADLs: activities of daily living
AMTS: Abbreviated Mental Test Score
BMI: body mass index
CPG: clinical practice guideline
DCM: dementia care mapping
GCS: Glasgow Coma Score
GP: general practitioner
IQR: interquartile range
ITT: intention-to-treat
IU: international unit
MMSE: Mini Mental State Examination

N: No
NR: not reported
OT: occupational therapist
RCT: randomised controlled trial
SD: standard deviation
STOPP/START: Screening Tool of Older Persons potentially inappropriate Prescriptions/Screening Tool to Alert doctors to Right Treatment
vs: versus
Y: Yes

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Barreca 2004	RCT. Falls outcomes. Supervised exercises in older people post stroke
Bernhardt 2008	RCT. Falls recorded as adverse events. Early rehabilitation post stroke
Bosner 2012	Not randomised. Five nursing homes agreed to participate; three were assigned sequentially for the intervention and two for the control group
Bouwen 2008	RCT (cluster randomised). Nursing homes. Outcome of the study was a subgroup of falls only (falls with medical consequences)
Capezuti 1998	RCT (cluster randomised). Nursing homes. The intervention was designed to minimise restraints, not to reduce falls. Falls reported as adverse events
Crotty 2002	RCT. Accelerated discharge after hip fracture and home based rehabilitation in the community. Not designed to reduce falls. Falls recorded as adverse events
Cucca 2017	Falls recorded as adverse events
Cummings 2015	Falls recorded as adverse events
Dattalo 2015	Wrong setting, participants likely to be in the community. Attempts to contact authors unsuccessful
Davison 2005	RCT. Post-fall intervention with falls outcomes. Only one participant in residential/nursing care
de Morton 2007	CCT. The primary outcome was discharge destination. Falls were recorded as adverse events
de Souto 2016	Falls recorded as adverse events
de Souto 2017	Falls recorded as adverse events
DeSure 2013	RCT, cross-over trial. Exercise Program to Prevent Falls in Institutionalized Elderly with Cognitive Deficits. Falls data for phase 1 not clearly reported, falls data for phase 2 has contamination of intervention group. Attempts to contact author unsuccessful. Available falls data considered invalid

(*Continued*)

Donat 2007	RCT. Exercise interventions in nursing homes. No falls outcomes
Drahota 2013	Pilot RCT. Intervention was intended to reduce fall injuries, rather than falls
Fiatarone 1994	RCT. Boston FICSIT study in nursing home residents. No falls outcomes
Forster 2017	Falls recorded as adverse events
Fossey 2006	RCT. Nursing homes. Intervention to reduce antipsychotics in people with severe dementia. Falls were recorded as adverse events
Galik 2014	Falls reported as adverse events
Galik 2015	Falls recorded as adverse events
Gianturco 2013	Wrong setting, RCT conducted at a geriatric day service with community-dwelling participants
Ginde 2017	Falls recorded as adverse events
Graafmans 1996	Wrong setting. 49% residing in homes for the elderly, included in community review
Grant 2005	RCT. Participants recruited in hospital after a hip fracture. Preventing falls in older people living in the community
Greenspan 2013	Wrong study design, not an RCT.
Greenspan 2015	Falls recorded as adverse events
Gruber-Baldini 2011	RCT. Intervention to motivate nursing assistants to actively engage nursing home residents in functional and physical activities. Falls recorded as adverse events
Gu 2006	Non-randomised controlled trial of exercise intervention in nursing homes. Experimental group was a convenience sample from two nursing homes; matched control group selected from another nursing home [personal communication]
Hardin 2013	Wrong patient population. Hospital setting. Author confirmed age of participants unknown
Harwood 2004	RCT. Participants recruited at the end of ward rehabilitation post hip fracture. Preventing falls in older people living in the community
Hauer 2001	RCT. Exercise intervention. Recruited at the end of ward rehabilitation. Majority were community-dwelling (4% living in nursing homes)
Heiberg 2017	Falls recorded as adverse events
Herrmann 2016	Falls reported as adverse events

(*Continued*)

Hopman-Rock 1999	RCT. Participants with dementia in homes for the elderly. Falls recorded as safety issue, i.e. as adverse events
Huang 2005	RCT. Discharge planning intervention to prevent falls in older people living in the community
Il'nitskii 2014	Wrong study design, not an RCT
Ilfeld 2010	Falls recorded as adverse events
Jarret 2015	RCT. Intervention delivered in a rehabilitation setting, patients admitted from community, no falls in hospitals, falls outcomes recorded post-discharge. Included in community review
Jeon 2015	Only injurious falls reported
JPRN-UMIN0000167	Wrong setting: likely community. Attempts to contact author unsuccessful. Trial discontinued
Kato 2006	Not RCT. "Prospective clinical trial" of an exercise programme in a long-term care facility with falls outcomes. Nurses volunteered their ward to be an intervention ward (personal communication from authors)
Katz 2004	RCT in residential care population. Intervention: three doses of risperidone in people with dementia and psychosis or agitation. Post hoc subgroup analysis of falls based on 85.9% of those randomised. Falls reported as adverse events
Katz 2005	This study was not primarily a falls prevention intervention. Falls reported as adverse events
Kenny 2001	RCT. Follow-up of falls outcomes appears to be primarily in the community
Koczy 2011	The intervention was designed to minimise restraints, not to reduce falls. Falls reported as adverse events
Kopke 2012	RCT (cluster randomised). Nursing homes. The intervention was designed to minimise restraints, not to reduce falls. Falls reported as adverse events
Kwok 2006	RCT. Intervention to determine whether bed-chair pressure sensors reduced physical restraint use. Falls reported as adverse events
Lackner 2008	RCT in cognitively impaired nursing home residents with urge urinary incontinence. Falls reported as adverse events
Li 2017	Falls recorded as adverse events
Lord 2003b	RCT. Majority of participants community-dwelling. Only 121/551 participants were residents of an intermediate level nursing care facility
Mailhot 2012	Falls recorded as adverse events
Mailhot 2014	Falls reported as adverse events
Mak 2016	Wrong setting. Intervention delivered in hospital, falls recorded in the community

(*Continued*)

Mansfield 2015	Falls recorded as adverse events
McRae 1996	Not RCT. Falls and fallers were not primary outcomes but were monitored as possible adverse events
Mudge 2008	Non-randomised controlled study. Patients admitted to an intervention ward or control ward
NCT00973297	Wrong population: Patients post-stroke
NCT01054287	Author correspondence confirmed that study unpublished and unlikely to be published as primary author has left the institution. Trial discontinued (results unavailable)
NCT01523600	Trial discontinued due to lack of funding.
NCT01618786	Intervention intended to reduce injuries not falls.
NCT02686515	Wrong population: Patients post-stroke
Nyaruhirira 2013	Wrong setting. Setting unclear, attempts to contact author unsuccessful
Ouslander 2005	RCT testing 'Functional Incidental Training' in nursing homes. Not designed to reduce falls. Falls recorded as adverse events
Parasurum 2011	Wrong patient population. Hospital mental health setting, patient age unknown, attempt to contact authors unsuccessful, participants unlikely to be elderly
Pedreira 2014	Wrong population: Patients post-stroke
Peng 2014	Falls recorded as adverse events
Peri 2008	RCT (cluster). Pilot for Kerse 2008 (same intervention). Excluded because falls were recorded as possible adverse effects of the intervention
Rantz 2001	RCT. Quality improvement intervention in nursing care facilities targeting 29 quality indicators, of which falls was one. Only included 87/113 homes in the analysis (23% loss). Insufficient information provided on falls outcomes to use in this review
Ray 2005	RCT. Study of falls related injuries. No data provided on falls or fallers
Reinhardt 2014	Falls reported as adverse events
Resnick 2002	Participants resident in continuing care retirement community but all living independently
Resnick 2012	RCT in assisted living facilities. Testing changing model of care to function-focused care. Falls monitored as a safety issue, i.e. adverse events. Hypothesised that the intervention might increase the likelihood of falling
Richter 2015	Falls recorded as adverse events

(*Continued*)

Rolland 2007	RCT. Exercise programme to improve ability to perform ADL for people with Alzheimer's disease in nursing homes. Falls monitored as a safety issue, i.e. adverse events
Sackley 2009	RCT. Falls described as an outcome at trial registration but not mentioned as an outcome in the published paper
Sahota 2014	Specific type of falls only, reported bedside and injurious falls, not total falls
Said 2012	Falls recorded as adverse events
Said 2015	RCT. Falls recorded as adverse events
Sato 2000	RCT. Etidronate versus placebo in older people with post stroke hemiplegia. Falls outcomes. Wrong population; article subsequently retracted
Sato 2005a	RCT. Vitamin D vs placebo in older people with post stroke hemiplegia. Falls outcomes. Wrong population; article subsequently retracted
Sato 2005b	RCT. Folate and mecobalamine (vitamin B12) vs placebo in older people with post stroke hemiplegia. Falls outcomes. Wrong population; article subsequently retracted
Sato 2011	RCT. Aledronate versus alphacalcidol in older people post-stroke. Falls outcomes. Wrong population; scientific misconduct also likely
Schneider 2006	The objective of this study was to determine the effectiveness of atypical antipsychotic medications. Falls were monitored as a potential adverse effect
Schwendimann 2006	Not RCT. Described as quasi-randomised in abstract but author confirmed that all consecutively admitted patients were allocated at non-random order either to nursing unit A or B whenever a free hospital bed was available (1 to 5 admissions/discharges per day). Nurse-led fall prevention programme
Sherrington 2016a	Wrong setting, correspondence with the author indicated 3% participants were in care - excluded as majority living in a community setting
Shimada 2003	RCT. Majority of participants community-dwelling (62%)
Shimada 2009	Not RCT. Exercise intervention versus control in a residential-care facility. Falls outcomes. Intervention on 2 days per week and 2 other days randomly selected to be control days
Siddiqi 2016	No falls outcomes
Sjoberg 2013	Wrong setting. Intervention partly in hospital and partly in community. Author confirmed that < 50% residing in nursing homes at 6 and 12 months
Smith 2017	Falls data not reported separately to slips and trips. Not an RCT
Sola 2014	RCT. Setting unclear, likely to be in the community. Attempts to contact author unsuccessful

(*Continued*)

Southard 2006	RCT with no falls outcomes. Balance and confidence were the primary outcomes of this study
Steadman 2003	RCT. Participants were attendees of a hospital-based falls clinic. "Prevously living in the community" [personal communication]. Not preventing falls in hospital or nursing care facility
Tanikawa 2014	Falls recorded as adverse events
Tariot 2004	RCT. Trial testing effectiveness of memantine in people with Alzheimer's disease already receiving donepezil. Falls were monitored as a potential adverse effect of the intervention
Tariot 2005	RCT. Trial testing effectiveness of divalproex sodium in nursing home residents with possible or probable Alzheimer disease. Falls were monitored as a potential adverse effect of the intervention
Teresi 2013	Wrong study design. Not an RCT, random selection for data collection, rather than allocation
Underwood 2011	Ongoing RCT (cluster randomised). Exercise intervention in residential and nursing homes Primary outcome depression. No falls outcomes. Recording peripheral fractures and fear of falling
van Ooijen 2013	Wrong setting. Intervention delivered in hospital, author confirmed falls recorded post dischage and the majority of participants were in the community
Vassallo 2004	Non-randomised controlled trial of a multidisciplinary fall-prevention programme in hospital. Falls outcomes
Visvanathan 2015	Not an RCT
Von Koch 2001	RCT. Intervention: rehabilitation at home after a stroke. Not intervention to prevent falls; falls recorded as adverse events
Wolf 2003	RCT. Participants in independent living facilities or congregate living facilities, i.e. not nursing care facilities. Community-dwelling
Zhong 2007	RCT. Institutionalised participants with dementia randomised to quetiapine 200 mg per day, 100 mg per day, or placebo. Falls monitored as a potential adverse effect of the intervention

ADL: activities of daily living
CCT: controlled clinical trial
RCT: randomised controlled trial

Characteristics of studies awaiting assessment *[ordered by study ID]*

Dever 2016

Methods	RCT
Participants	Setting and sample: 26 care facilities, N = 150, Canada Age (years): mean 86
Interventions	• Falls-risk assessment • Usual care
Outcomes	• Number of falls
Notes	Article located in search update (August 2017); pending processing Falls reported as medians with range

Frohnhofen 2013

Methods	RCT
Participants	Setting and sample: Hospital setting, N = 178 geriatric patients during rehabilitation
Interventions	FORTA (Fit-for-the-Aged) medication review
Outcomes	• Number of falls.
Notes	Conference abstract only. Publication likely to be same study as Michalek 2014, attempts to contact author unsuccessful 6/7/16 and 26/10/16

Hewitt 2014

Methods	RCT (cluster randomised)
Participants	Inclusion criteria: residential-aged care facilities: have a mix of high-care residents and/or low-care residents; likely to have 15 residents willing to participate; the facility manager consents to participation in the trial and to the allocation of staff time.Participants: permanently reside in residential-aged care Exclusion criteria: terminal or unstable illness; significant advanced cognitive decline (Mini Mental State Examination </= 15); physical symptoms that preclude the safe use of exercise equipment in a group setting (e.g.. Parkinson's disease or hemiplegia); permanently wheelchair- or bed-bound; performed a similar balance and/or resistance training programme within the previous 12 months
Interventions	SUNBEAM program (Strength and Balance Exercise in Aged Care) conducted in group settings; comprising progressive resistance training and balance exercises from 0-6 months; then maintenance exercises for 7-12 months Usual care
Outcomes	• Number of falls • Falls rate
Notes	ACTRN12613000179730

MacRitchie 2001

Methods	RCT
Participants	Setting and sample: two nursing homes, Connecticut, USA N = 88 Age (years): mean 84 (SD 6.9), range 65 to 98 Inclusion criteria: none stated
Interventions	• Standing-exercise Functional Maintenance programme of 4 months duration • Control
Outcomes	• Incidence of falls
Notes	Thesis identified in the Cochrane Library (CENTRAL). No usable falls data in abstract. No published papers identified

Raymond 2017

Methods	RCT
Participants	Setting and sample: Hospital setting, sub-acute, N = 468, Australia ≥ 65 years.
Interventions	• Standing high-intensity functional group exercise 3x week plus individual physiotherapy 2x week. • Daily individual physiotherapy exercises.
Outcomes	• Number of falls • Number of fallers
Notes	Article located in search update (August 2017); pending processing Few falls (total 12), not reported by group allocation.

Tallon 2013

Methods	RCT
Participants	Setting and sample: residential care Inclusion criteria: living in nursing home, able to walk, no contra-indication to whole body vibration
Interventions	• Exercise with whole body vibration, 3 times weekly for 20 minutes • Standard exercise: same exercises on a non-vibrating platform
Outcomes	• Number of falls • Risk of falling
Notes	Study completed. Conference abstract available. Author indicated study completed but analysis ongoing, study unpublished [email 11/7/16]. No response received to follow-up email 31/1/2017

Van der Linden 2017

Methods	Consecutive allocation, prospective controlled trial.
Participants	Setting and sample: hospital setting, sub-acute, N = 172, Belgium
Interventions	• Medication review using RASP (Rationalization of home medication by an Adjusted STOPP in older Patients) list and pharmacist led review. • Control.
Outcomes	• Number of falls • Number of participants falling.
Notes	Article located in search update (August 2017); pending processing NCT01513265

Wylie 2017

Methods	Pilot RCT
Participants	Setting and sample: 6 care facilities, N = 468, UK, East Scotland
Interventions	• 3-month podiatry intervention comprising core podiatry care, foot and ankle exercises, orthoses and footwear provision • Usual care
Outcomes	• Number of falls • Time to first fall
Notes	Article located in search update (August 2017); pending processing NCT02178527

RCT: randomised controlled trial
SD: standard deviation

Characteristics of ongoing studies *[ordered by study ID]*

ACTRN12613000228785

Trial name or title	Preventing falls and fractures in low-level aged-care residents by increasing dairy food intake by two serves per day
Methods	RCT
Participants	Low-level aged care residents with dietary calcium intake below 600 mg/day
Interventions	• Additional 2 serves of dairy foods per day • Usual diet

ACTRN12613000228785 (*Continued*)

Outcomes	• Falls • Fractures
Starting date	Not commenced.
Contact information	Dr Sandra Iuliano Endocrinology, Level 2 Centaur Building Heidelberg Repatriation Hospital Waterdale Rd, West Heidelberg, VIC, 3081 Australia +61394963216 sandraib@unimelb.edu.au
Notes	

ACTRN12615000817549

Trial name or title	Establishing the effectiveness, cost-effectiveness and student experience of simulation training for the prevention of falls amongst hospitalised inpatients
Methods	RCT
Participants	**Inclusion criteria:** patients admitted to intervention wards within a public hospital **Group 1** • All health professional undergraduate students from Monash University attending placement at Peninsula Health for at least two weeks or more. • Placement on wards which have been randomised to the intervention or control. **Group 2** Patients admitted to intervention wards within PH
Interventions	• Health professional students attend a four hour simulation training session • Usual care
Outcomes	• Rate of falls • Proportion of fallers • Rate of injurious falls
Starting date	17/08/2015
Contact information	Dr Cylie Williams Peninsula Health 2 Hastings Rd Frankston, VIC, 3199, Australia cyliewilliams@phcn.vic.gov.au +61 3 97848125
Notes	

ACTRN12617000314325

Trial name or title	Does abbreviating patient falls risk screening in documentation impact on falls in hospital inpatients: a stepped wedge cluster randomised control trial
Methods	RCT (stepped wedge)
Participants	All patients who are admitted to intervention wards at Peninsula Health, Non inclusion of paediatric and maternity wards
Interventions	• The short Falls Risk Assessment Tool (FRAT) - a template which guides the user to falls intervention strategies only. • Patient focused falls interventions will be documented on a Short FRAT based on observed and personalised need rather then the risk level. • Control: the traditional FRAT
Outcomes	• Rate of falls
Starting date	March 2017.
Contact information	Dr Cylie Williams Peninsula Health Level 3 - Office for Research 2 Hastings Rd, Frankston VIC 3199 Australia cyliewilliams@monash.edu
Notes	Trial may be eligible depending on mean age of patients on trial completion

Dal Bello-Haas 2012

Trial name or title	The effects of a long-term care walking programme on balance, falls and well-being
Methods	RCT
Participants	**Inclusion criteria:** 60 years or older; living in long-term care facility; able to follow simple instructions; able to ambulate with or without an aid for at least 10 m; available Monday to Friday; willing to participate in a 5 days per week walking programme over a 4-month period **Exclusion criteria:** recent cardiovascular event; uncontrolled hypertension; uncontrolled epilepsy; recent fracture; unable to satisfactorily comply with the protocol requirements; recent admission into an acute care facility (past 4 months); scheduled for surgery or hospitalisation in the next 6 months; participating in another regular exercise programme (half an hour or more, three or more times per week) aimed at improving balance or strength
Interventions	• Individualised; progressive; one-to-one supervised walking programme provided by study personnel and supervised by a licensed physiotherapist • Usual care
Outcomes	1. Falls incidence

Dal Bello-Haas 2012 (*Continued*)

Starting date	December 2010 Estimated completion December 2016
Contact information	Vanina PM Dal Bello-Haas School of Rehabilitation Sciences, McMaster University, 1400 Main Street West, 403/E, Hamilton, Ontario L8S 1C7, Canada vdalbel@mcmaster.ca
Notes	CT.gov NCT01277809

Hassett 2016

Trial name or title	Activity and MObility UsiNg Technology (AMOUNT) rehabilitation trial
Methods	RCT
Participants	**Inclusion criteria:** admitted for rehabilitation or assessment at one of the 3 study sites with: reduced mobility (Short Physical Performance Battery score of less than 12); clinician-assessed capacity for improvement in mobility; likely life expectancy of more than 12 months; anticipated length of stay of greater than or equal to 10 days; ability to maintain a standing position with 1 person assist as a minimum standard **Exclusion criteria**: marked cognitive impairment; insufficient English language skills to participate in rehabilitation and no available interpreter; inadequate vision to use the devices; medical condition precluding exercise (unstable cardiac disease, uncontrolled hypertension, uncontrolled metabolic diseases, large abdominal aortic aneurysm or a weight-bearing restriction); lack of interest in the use of the technologies; anticipated discharge to nursing home; discharge location too far from study site to complete home visits and follow-up assessments
Interventions	• Tailored technology use (video and computer games/exercises and tablet applications as well as activity monitors) to promote physical activity in addition to usual care • Usual care
Outcomes	1. Number of falls.
Starting date	September 2014. Data collection completed.
Contact information	Prof Cathie Sherrington The George Institute for Global Health PO Box M201, Missenden Road Sydney NSW 2050 Australia Phone: +61280524300 Email: csherrington@georgeinstitute.org.au
Notes	ANZCTR. ACTRN12614000936628

ISRCTN34353836

Trial name or title	Finch: Falls in care homes study
Methods	RCT (cluster randomised)
Participants	**Inclusion criteria: Care Home inclusion criteria** • Long stay with old age and/or dementia registration • 10 or more potentially eligible residents • Routinely record falls in resident personal records and on incident sheets • Consent of care home manager to comply with the protocol and identify a care home fall champion **Resident inclusion criteria** • All long-term care home residents **Staff Inclusion Criteria (Process Evaluation Only)** • Employed by a Care Home participating in FINCH and selected for participation in the Process Evaluation • Employed in a caring role **Exclusion criteria: Care Home exclusion criteria** • Participated in GtACH pilot/feasibility studies • Homes exclusively providing care for those with learning difficulties or substance dependency • Homes with contracts under suspension with health or social providers, or that are currently subject to safeguarding investigations or homes under CQC special measures • Homes with a significant proportion of beds taken up by health-service commissioned intermediate-care services • Trained and routinely using a systematic falls prevention programme **Resident exclusion criteria** • Residents on short-term care (e.g. respite) • Residents identified to be in the last few days of life **Staff Exclusion Criteria (Process Evaluation Only)** • Have a significant proportion of time caring for residents in health-service commissioned intermediate-care services funded beds
Interventions	• Guide to Action Care Home (GtACH) fall-prevention programme • Usual care
Outcomes	• Rate of falls • Rate of fractures • Injurious falls
Starting date	1 November 2016 No longer recruiting
Contact information	Pip Logan B108a Div Rehab and Ageing Medical School Queen's Medical Centre NG7 2UH Nottingham United Kingdom pip.logan@nottingham.ac.uk
Notes	

ISRCTN42003273

Trial name or title	Polypharmacy reduction in patients treated for chronic diseases (POLITE-RCT)
Methods	RCT (cluster randomised)
Participants	Hospital (acute care) inpatients aged 65 and over
Interventions	• Medication review • Usual care
Outcomes	• Number of falls • Fall related injury • Fall related fractures
Starting date	1 November 2013. Completed October 2016.
Contact information	Prof Attila Altiner Rostock University Medical Center Institute of General Practice (Universitätsmedizin Rostock Institut für Allgemeinmedizin) POB 100888 Rostock 18055 Germany +49 (0)381 4942481 altiner@med.uni-rostock.de
Notes	

JPRN-UMIN000000555

Trial name or title	The effects of whole body vibration for the prevention of falls in elderly
Methods	RCT
Participants	≥ 65 years, resident of senior citizen institution **Excluded criteria:** bedridden
Interventions	• Whole Body Vibration; 3 minutes twice weekly for 12 months
Outcomes	• Rate of falls
Starting date	Study registered 25/12/2006 Study completed. Analysis completed 1/6/2009.
Contact information	Tatsuya Koike, Osaka City University Medical School, Abenoku Asahimachi 1-4-3, Osaka, 545-8585, Japan
Notes	Trials registry page last updated on 28/11/2012 .Attempt to contact author 3/7/16 unsuccessful

JPRN-UMIN000008361

Trial name or title	Multicenter, randomised, double-blind, placebo controlled, parallel group trial to evaluate the effect of Vitamin D supplementation for fall prevention
Methods	RCT (double-blind)
Participants	Residents in the social welfare corporation kensyokai associated facilities
Interventions	• Beverage contained Vitamin D supplement (liquid), 1 drop/day (1,000 IU) for 1 year • Placebo beverage without Vitamin D supplement for 1 year
Outcomes	• Falls • Fracture incidence
Starting date	Start: 20 Jan 2013. Data analysis completed 31/12/2014. No publication identified
Contact information	Tetsuya Enishi Division of Rehabilitation, Tokushima University Hospital, Tokushima University enishi.tetsuya@tokushima-u.ac.jp
Notes	Authors contacted 16/5/16, no response received. Last modified 17/8/2017, status indicates unpublished

McCullagh 2016

Trial name or title	A twice-daily individual targeted exercise program in frail hospitalised older medical in-patients (APEP)
Methods	RCT
Participants	**Inclusion criteria:** ≥65 years, medical patients, anticipated length of stay greater than 2 days, planned for discharge home, mobility aid and /or assistance required on admission **Exclusion criteria:** contraindications to exercise, unable to follow commands in the English language, unable to exercise with the assistance of one person only, when active palliative care is required, when full isolation for containment of a contagious infection is required
Interventions	• Twice-daily, individual, targeted, strengthening, balance and endurance exercise sessions • Twice-daily, individual, stretching and relaxation exercise sessions (sham exercise)
Outcomes	• Number of falls • Number of falls injuries
Starting date	March 2015. Estimated completion May 2017.
Contact information	Dr Suzanne Timmons, Senior Lecturer in Gerontology and Rehabilitation, University College Cork
Notes	NCT02463864

Mestres 2017

Trial name or title	Supporting Clinical Rules Engine in the Adjustment of Medication (SCREAM)
Methods	RCT (cluster randomised)
Participants	**Inclusion criteria:** residents living in a nursing home in the Netherlands. The nursing homes are able to deliver the medication and lab data electronically
Interventions	• Medication review. A clinical decision support system, the CRR (clinical rule reporter) will be used to weekly screen medication list, laboratory values and medical history in order to obtain potential clinical relevant remarks that will be sent to the correspondent physician with an advice on how to improve/solve the situation. • Usual care.
Outcomes	• Number of falls (as part of composite measure)
Starting date	June 2013. Planned completion June 2016.
Contact information	Dr. PHM van der Kuy
Notes	NTR5165

Mudge 2017

Trial name or title	CHERISH (Collaborative for Hospitalised Elders: Reducing the Impact of Stays in Hospital)
Methods	RCT (cluster randomised)
Participants	**Inclusion criteria:** ≥ 65 years, admitted to hospital for 3 or more days, with admission to nominated intervention or control ward **Exclusion criteria:** discharged from hospital within 2 days; palliative intent of care
Interventions	• "Eat Walk Engage", a quality improvement programme designed to enhance uptake of evidence-based processes of care for older inpatients.The target processes are early mobility; adequate oral nutritional intake; and meaningful, cognitively stimulating activities. • Usual care, including any facility based improvement programmes.
Outcomes	• Number of falls (as part of composite measure)
Starting date	October 2015.
Contact information	Prof Alison Mudge Building C28 Level 1 Royal Brisbane and Women's Hospitals Herston Queensland 4029 Australia Email Alison.Mudge@health.qld.gov.au
Notes	ACTRN12615000879561

NCT00636675

Trial name or title	CONNECT
Methods	RCT (cluster randomised by nursing home)
Participants	16 nursing homes (560 residents and 576 staff members)
Interventions	• CONNECT plus standard FALLS quality improvement programme. CONNECT is a multi-component intervention that helps staff; learn new strategies to improve day-to-day interactions; establish relationship networks for creative problem-solving; and sustain newly acquired interaction behaviours through mentorship • FALLS quality improvement programme
Outcomes	• Fall rates (secondary outcome)
Starting date	September 2009. Estimated completion September 2016.
Contact information	Ruth Anderson, RN, PhD Duke University School of Nursing Durhan, North Carolina, USA, 27710 Email: ruth.anderson@duke.edu
Notes	Included study (Colon-Emeric 2013) is a pilot study including 8 care facilities, this study includes 16 sites

NCT01483456

Trial name or title	Impact of multidisciplinary program on falls in elderly inpatients (IPR)
Methods	RCT (stepped wedge)
Participants	**Setting:** hospitals (rehabilitation wards and geriatric acute wards), France N = 1680 (target sample size) **Inclusion Criteria:** aged ≥ 65; admitted during study; consenting **Exclusion Criteria:** cognitively impaired (MMSE < 10); psychiatric pathology; bedridden
Interventions	• Multifactorial intervention; identification of patient's fall risk; multifactorial fall-prevention programme (integrated actions targeted on risk factors; exercise programs and review of the hospital environment); "Get up" workshop and morbidity and mortality conferences related to fall cases • Usual care
Outcomes	• Incidence of falls • Incidence of fall-related injury
Starting date	July 2011
Contact information	P Krolak-Salmon Hospices Civils de Lyon Email: pierre.krolak-salmon@chu-lyon.fr

NCT01483456 (*Continued*)

Notes	IPR (in French "Identifier, Prévenir, Relever"). Study design described as "Intervention model: single group assignment" no mention of a control group. Contact person has confirmed that this is an RCT Author correspondence confirmed trial design. Enquired about study completion 13 Jan 2017, no response received

NCT01551121

Trial name or title	Assessment of an automated telesurveillance system on the incidence of serious falls in nursing homes (TELEHPAD)
Methods	RCT (individually randomised)
Participants	**Settting:** 3 Nursing homes in the Limousin region **Target sample size:** N = 216 **Sample:** people admitted to Limoges or Gueret nursing homes I**nclusion criteria:** aged 75; consenting; able to understand the study and complete evaluations; able to stand up from the bed; covered by French health insurance **Exclusion criteria:** short-term prognosis; in multiple bed room and one co-occupant does consent to participate
Interventions	• Installation of automated telesurveillance system (camera installed in room) • Usual care
Outcomes	Duration: 1 year • Number of people falling
Starting date	March 2012.
Contact information	Thierry Dantoine, MD University Hospital Limoges Email: thierry.dantoine@chu-limoges.fr
Notes	Correspondence with T Dantoine confirmed study ongoing 10 August 2016. Study listed as recruiting as at 10 November 2017

NCT01561872

Trial name or title	Assessment of an automated telesurveillance system on serious falls prevention in an elderly suffering from dementia specialized care unit: the URCC (GET-BETTER)
Methods	RCT (individually randomised)
Participants	**Setting:** Limoges and Brive's URCC Target sample size = 350 **Inclusion Criteria:** men and women aged > 65; admitted to Limoges or Brive's URCC (dementia care unit) ; consenting; covered by French health insurance

NCT01561872 (*Continued*)

	Exclusion Criteria: short-term prognosis
Interventions	• Automated telesurveillance system (camera installed) • Control: usual care (no telesurveillance)
Outcomes	Duration of study: 6 months • Rate of falls • Rate of injurious falls
Starting date	April 2012.Completed 2016
Contact information	Dr T Dantoine University Hospital Limoges France Email: thierry.dantoine@chu-limoges.fr
Notes	URCC: Unité de Réadaptation Cogintico-Comportementale (Unit for demented patients' rehabilitation) (Dantoine T, personal communication Oct 20 2012). Correspondence with T Dantoine confirmed study completed, analysis ongoing as at 10 August 2016

NCT01735682

Trial name or title	Whole body vibration exercise training for institutionalized elderly
Methods	RCT (single blind)
Participants	**Inclusion Criteria** • ≥ 65 years • Functional Ambulation Category 1 to 4 • able to understand simple verbal commands • able to tolerate intermittent physical activity for at least 45 minutes • able to perform knee flexion > 45 degree • able to stand with or without support for 1 minute or more
Interventions	• Whole body vibration • Conventional exercise • Upper limb exercise
Outcomes	• Falls incidence
Starting date	Estimted study completion October 2015. Last verified May 2014
Contact information	The Hong Kong Polytechnic, University Shatin Hospital, Hong Kong
Notes	Enquiry sent to author about study completion 3 July 2016. No response received

NCT01876095

Trial name or title	Discontinuing Inappropriate Medication in Nursing Home Residents (DIM-NHR)
Methods	RCT
Participants	**Inclusion criteria** **Wards** • Long-stay ward • Capability and commitment to perform a multidisciplinary multi-step medication review. **Participants** • A life expectancy of > 4 weeks as judged by the treating elderly care physician. • IC provided by patients themselves or provided by a legal representative for incapacitated patients.
Interventions	• Multidisciplinary medication review • Usual care
Outcomes	1. Falling
Starting date	Study completed April 2016
Contact information	Dr Katja Taxis University of Groningen ZonMw: The Netherlands Organisation for Health, Research Development,
Notes	Author enquiry sent 3 July 2016, 14 Oct 2016, no response received

NCT02295462

Trial name or title	Effect of person-centred-care on antipsychotic drug use in nursing homes: a cluster-randomised trial
Methods	RCT
Participants	**Inclusion Criteria** • Nursing homes with at least 50 residents • All residents within a cluster are eligible to participate in the study
Interventions	• Medication review + person-centred care • Medication review only
Outcomes	• Falls
Starting date	Start date December 2014. Estimated completion March 2017
Contact information	Prof. Dr.Gabriele Meyer, Martin-Luther-Universität Halle-Wittenberg Halle (Saale), Sachsen-Anhalt, Germany, 06110 +49 ext 3455574498gabriele.meyer@medizin.uni-halle.de

NCT02295462 (*Continued*)

Notes	

NCT02570945

Trial name or title	Trial of a pharmacist-physician intervention model to reduce high-risk drug use by hospitalised elderly patients
Methods	RCT
Participants	**Inclusion Criteria** • Patients 65 and older admitted at the Centre hospitalier universitaire de Sherbrooke
Interventions	• Pharmacist-physician medication review to reduce high-risk medication use by elderly inpatients • Control
Outcomes	• Number of falls
Starting date	Study completed. Completion December 2015
Contact information	Benoit Cossette, Principal investigator, Centre de recherche du Centre hospitalier universitaire de Sherbrooke
Notes	

NCT02604056

Trial name or title	Pragmatic cluster trial for nursing home antipsychotic prescribing
Methods	RCT (cluster randomised)
Participants	**Inclusion Criteria** • Nursing homes within pre-determined regions of Ontario that expressed an interest in the full intervention (the regions; or hubs; contain a wide variety of nursing home types within a reasonable travel distance [i.e. < 100 km]) • Nursing homes within the hubs in which the medical and administrative leads agree to and support the project **Exclusion Criteria** • Nursing homes with a previous or ongoing involvement in externally supported quality improvement initiatives focusing on antipsychotic medications • Nursing homes without any prescribers caring for at least 10 residents routinelyNursing homes with fewer than 30 residents
Interventions	• Audit & feedback & educational outreach. Educational Outreach offered to each prescriber and team members in the home • Usual care: Audit & feedback. Standard quality improvement supports (including online Audit and Feedback reports for each prescriber in the home)
Outcomes	• Falls

NCT02604056 (*Continued*)

Starting date	September 2015. Estimated completion December 2017
Contact information	Women's College Hospital, Ontario Ministry of Health, Long Term Care, Ontario Medical Association, Health Quality Ontario, Centre for Effective Practice
Notes	

NCT02702037

Trial name or title	Older Person's Exercise and Nutrition study (OPEN): a simple physical exercise combined with protein supplement - effects on functional status and independence among older people: a cluster randomised controlled trial
Methods	RCT (individually randomised)
Participants	**Inclusion Criteria** • 75 Years and older • Able to rise independently from a seated position to standing • Nursing home setting **Exclusion Criteria** • BMI >30 • Severe dysphagia • Tube feeding • Severe kidney failure • Bedridden people • Terminal stage of life • Lack of informed consent
Interventions	• The participants will be supported to perform the sit-to-stand exercise at least four times per day during 12 weeks (7 days/week). The participants will also be offered an oral protein-rich supplement (125 mL, 18 g protein (24% of RDI), 300 kcal) twice a day in conjunction with two of the four sit-to-stand exercises during 12 weeks (7 days/week) • Usual care
Outcomes	• Falls
Starting date	March 2016
Contact information	Karolinska Institute, Nutricia Foundation
Notes	Anne-Marie Bostrom, PhD Karolinska Institutet Stockholm, Sweden anne-marie.bostrom@ki.se

NCT02714257

Trial name or title	Seniors avoiding falls through exercise study
Methods	RCT
Participants	**Inclusion Criteria** • ≥ 65 years old • Previous Fragility Fracture (FF) in past 5 years • Able to speak and understand English • Participants will need to be willing to try exercising and agree to annual follow-up measurements
Interventions	• Enhanced Usual Care plus Exercise Coaching Intervention. Participants will receive the three printed pamphlets on fall risks and exercising in groups (same as the controls) plus; (1) an exercise programme that includes strength, balance, and aerobic exercises; (2) an exercise coach that provides in-person and telephone support/feedbacks to enhance participation in the exercise programme; and (3) regular progress reports sent by coaches by fax/Electronic Health Records every 12 weeks, to communicate the patient's progress • 2. Usual care. Enhanced usual care by reviewing three printed pamphlets on fall risks and recommendation to exercise. In addition, to maximise patient safety, the investigators will communicate the baseline bone density results (measured by Dual-energy X-ray absorptiometry, DXA) to the patient's primary care provider, and any critical values of a baseline measure
Outcomes	• Number of falls • Injurious falls
Starting date	September 2016 Estimated study completion August 2020
Contact information	Sol M Rodriguez-Colon Penn State Hershey Medical Center Hershey, Pennsylvania, USA, 17033 smr359@psu.edu
Notes	The intervention will be held in churches, community centres, and senior residential facilities. Study may be eligible depending on proportion of participants in aged-care facilities

NCT02714582

Trial name or title	Feasibility, appropriateness, meaningfulness and effectiveness of bedside shift reporting
Methods	RCT
Participants	**Inclusion Criteria** • Admitted on a participating hospital ward • Be conscious • Speak Dutch • Participated in at least 3 bedside shift reports **Exclusion Criteria** • Dementia or other severe cognitive/mental disorders

NCT02714582 (*Continued*)

Interventions	• Bedside Shift Reporting (BSR). The experimental group (nurses and patients) will: ○ develop a tailored BSR-intervention by use of co-design; diagnostic interviews and pilot testing ○ use the tailored BSR-intervention, with participation of the patient; instead of the regular nurse shift report • Usual care. No bedside shift report
Outcomes	• Falls incidence
Starting date	March 2016 Estimated completion February 2018
Contact information	Ann Van Hecke, MSc, PhD Ghent University/Ghent University Hospital
Notes	

NCT02757131

Trial name or title	Dedicated ambulator-assisted physical activity to improve hospital outcome measures in elderly patients
Methods	RCT
Participants	Hospital setting. **Inclusion Criteria** • Men or women 60 years of age or older admitted as inpatients to participating wards in the Medicine Institute, Cleveland Clinic Main Campus during the study time period • Hospitalised for a medical illness • Complete history and physical examination on file • Physical therapy consult and 6-Clicks score between 16-20 a. This is based on a usual care assessment ordered by a physician that will happen prior to any study recruitment - it is entirely independent of the study **Exclusion Criteria** • Observational status • Admission to ICU • Surgical patients • Patients diagnosed with: decompensated heart failure, unstable angina, other medical conditions precluding participation in exercise/ambulation • Comfort care measures only
Interventions	• Ambulator-assisted physical therapy, Ambulation protocol as directed by physical therapist and three times daily under supervision of dedicated ambulator patient care nursing assistant • Usual care
Outcomes	• Number of falls.
Starting date	July 2016. Completed July 2017.

NCT02757131 (*Continued*)

Contact information	Aaron Hamilton, MD Cleveland Clinic Foundation Cleveland, Ohio, USA, 44195
Notes	

NCT02969343

Trial name or title	Patient safety learning laboratory: making acute care more patient-centered
Methods	RCT (stepped wedge)
Participants	**Hospital setting** Estimated enrolment 21,000 participants. **Inclusion Criteria**: patients 18-99 years of age on hospital care units where the PSLL patient safety health information technology tools are implemented
Interventions	• Implementation of three Patient Safety Learning Laboratory (PSLL) toolkits. 1) Patient-centered fall-prevention toolkit 2) Patient safety checklist tool and 3) MySafeCare Patient Safety Reporting System • Usual care
Outcomes	• Rate of falls • Rate of injurious falls
Starting date	April 2015. Estimated completion September 2018.
Contact information	Principal Investigator: David W. Bates, MD, MSc, Brigham and Women's Hospital Contact: Alexandra C Businger 617-732-7063 abusinger@partners.org Contact: Patricia Dykes, RN PhD 617-732-8925 pdykes@partners.org Boston, Massachusetts, USA, 02115
Notes	Trial may be eligible depending on age of patients on trial completion

NCT03014570

Trial name or title	Testing iImplementation of EIT-4-BPSD.
Methods	RCT
Participants	**Inclusion Criteria** • Living in the nursing home • 55 years of age or older • score 0-12 on the Brief Interview of Mental Status **Exclusion Criteria** • Enrolled in hospice • in the nursing home for short-stay rehabilitation

NCT03014570 (*Continued*)

Interventions	• 4-step intervention: a. Assessment of the environment and policies; b. Education of staff; c. Establishing person-centered care plans; and d. Mentoring and motivating staff. • Education-only control
Outcomes	• Number of falls.
Starting date	April 2016.
Contact information	Barbara Resnick, Professor, University of Maryland Baltimore, Maryland, USA, 21201
Notes	

NCT03019211

Trial name or title	Feasibility aquatic physical exercise to reduce falls in institutionalized elderly (PrePhysFalls)
Methods	RCT
Participants	**Inclusion criteria** • Be institutionalised in a care centre • Participate voluntarily and sign the informed consent • Have a punctuation of 2 or more in The Downton Fall Risk Index **Exclusion criteria** • Suffer from a condition that can be affected or hinder exercise • Acute disease unresolved in 10 days • Not controlled hypertension • Contagious skin disorder • Urinary or faecal incontinence
Interventions	• Hydrotherapy. Static/dynamic exercises (balance and resistance training) in an aquatic environment • Control. Exercises out of a water environment
Outcomes	• Number of falls.
Starting date	Started April 2016. Completed March 2017.
Contact information	Mercè Sitjà Rabert Universitat Ramon Llull, Barcelona, Spain
Notes	

NCT03192384

Trial name or title	A service intervention to reduce falls in hospital
Methods	RCT (stepped wedge, cluster randomised)
Participants	**Inclusion Criteria** • All patients on ward
Interventions	• Implementation of educational programme intervention • Usual care
Outcomes	• Rate of falls.
Starting date	May 2017
Contact information	Professor Richard Lilford, University of Warwick Coventry, Warwickshire, United Kingdom, CV2 2DX
Notes	

NTR5015

Trial name or title	Randomized controlled intervention trial on falling and functional decline in the hospitalised elderly
Methods	RCT (cross-over trial)
Participants	**Inclusion criteria:** patients >70 years; stay in hospital > 3 days; agreement by the attending doctor; informed consent; ability to read and write Dutch **Exclusion criteria:** patients in isolation precautions; patients who can not go to the room where the activity programme is given; patients participating in another study
Interventions	• A daily two hours activities of daily life programme with occupational therapy by volunteers; physiotherapy and ergotherapy to improve the physical and mental condition
Outcomes	• Incidence of falls
Starting date	5 January 2015
Contact information	Sandra Koster s.koster@mst.nl
Notes	Author correspondence indicated that quote: “we can inform you that the main group of participants can be defined as elderly patients (> 65 year)”

Scheffers-Barnhoorn 2017

Trial name or title	FIT-HIP. Fear of falling intervention in hip fracture geriatric rehabilitation: a cluster randomised controlled trial
Methods	RCT (cluster randomised)
Participants	**Inclusion criteria** • Aged 65 years or older • Admitted to a geriatric rehabilitation (GR) unit for rehabilitation due to a hip fracture • Concerned to fall. This is measured by the one item fear of falling question (answering 'positively' in the category sometimes, often or very often) **Exclusion criteria** The patient has a condition interfering with learning ability, such as: • A diagnosis of dementia or score on the 'hetero-anamnesis list cognition' > 1, suggesting pre-morbid cognitive problems • A major psychiatric disease • Insufficient mastery of Dutch language • The patient has a limited life expectancy • The patient has a pathological hip fracture • Pre-fracture Barthel-index score < 15 (as a measure of ADL dependency)
Interventions	• Treatment of fear of falling. The FIT-HIP intervention consists of various elements of cognitive-behaviour therapy (guided exposure, psycho-education, cognitive restructuring, relapse prevention). This will be combined with exercise training in the physiotherapy sessions • Usual care
Outcomes	• Number of falls
Starting date	March 2016.
Contact information	Maaike Scheffers-Barnhoorn Leiden University Medical Center (LUMC), Department of Public Health and Primary Care The Netherlands.
Notes	NTR5695

ADL: activities of daily living
BMI: body mass index
IC: informed consent
ICU: intensive care unit
IU: international unit
RCT: randomised controlled trial

DATA AND ANALYSES

Comparison 1. Care facilities: Exercise vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	10	2002	Rate ratio (Random, 95% CI)	0.93 [0.72, 1.20]
2 Rate of falls and number of fallers: trials with incomplete data			Other data	No numeric data
3 Number of fallers	10	2090	Risk Ratio (Random, 95% CI)	1.02 [0.88, 1.18]
4 Number of people sustaining a fracture	1		Risk Ratio (Fixed, 95% CI)	Subtotals only
4.1 Hip fractures	1	183	Risk Ratio (Fixed, 95% CI)	0.16 [0.01, 2.81]
4.2 All fractures	1	183	Risk Ratio (Fixed, 95% CI)	0.88 [0.25, 3.14]
5 Rate of falls, excluding studies with ≤20 participants in each arm	8	1959	Rate ratio (Random, 95% CI)	0.91 [0.72, 1.15]
6 Number of fallers, excluding studies with ≤20 participants in each arm	9		Risk Ratio (Random, 95% CI)	1.04 [0.89, 1.21]
7 Adverse events: aches and pains	1	582	Risk Ratio (M-H, Fixed, 95% CI)	1.23 [0.61, 2.48]
7.1 Severe soreness	1	194	Risk Ratio (M-H, Fixed, 95% CI)	0.91 [0.40, 2.04]
7.2 Severe bruises	1	194	Risk Ratio (M-H, Fixed, 95% CI)	2.0 [0.18, 21.69]
7.3 Severe fatigue	1	194	Risk Ratio (M-H, Fixed, 95% CI)	4.0 [0.46, 35.14]

Comparison 2. Care facilities: Exercises vs usual care (grouped by type of exercise)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	10		Rate Ratio (Random, 95% CI)	Subtotals only
1.1 Gait, balance, functional training	4	1347	Rate Ratio (Random, 95% CI)	0.96 [0.69, 1.33]
1.2 Whole body vibration	1	62	Rate Ratio (Random, 95% CI)	0.96 [0.58, 1.60]
1.3 Combination of exercise categories (*see* Appendix 4 for categories in each trial)	6	683	Rate Ratio (Random, 95% CI)	0.94 [0.60, 1.47]
2 Number of fallers	10		Risk Ratio (Random, 95% CI)	Subtotals only
2.1 Gait, balance, and functional training	5	1452	Risk Ratio (Random, 95% CI)	1.02 [0.80, 1.31]
2.2 3D (Tai Chi)	1	59	Risk Ratio (Random, 95% CI)	0.60 [0.19, 1.87]
2.3 Whole body vibration vs usual care	1	62	Risk Ratio (Random, 95% CI)	0.88 [0.54, 1.43]

2.4 Combination of exercise categories (*see* Appendix 4 for categories in each trial)	4	607	Risk Ratio (Random, 95% CI)	1.07 [0.88, 1.29]

Comparison 3. Care facilities: Exercise vs usual care (grouped by level of care)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	10		Rate ratio (Random, 95% CI)	Subtotals only
1.1 High level nursing care facilities	2	210	Rate ratio (Random, 95% CI)	1.79 [0.89, 3.60]
1.2 Intermediate level care facilities	5	1315	Rate ratio (Random, 95% CI)	0.70 [0.47, 1.04]
1.3 Facilities providing mixed levels of care	3	477	Rate ratio (Random, 95% CI)	1.08 [0.92, 1.28]
2 Number of fallers	10	2090	Risk Ratio (Random, 95% CI)	1.02 [0.88, 1.18]
2.1 High level nursing care facilities	1	194	Risk Ratio (Random, 95% CI)	1.16 [0.83, 1.62]
2.2 Intermediate level care facilities	6	1419	Risk Ratio (Random, 95% CI)	0.94 [0.75, 1.17]
2.3 Mixed level care facilities	3	477	Risk Ratio (Random, 95% CI)	1.05 [0.76, 1.47]

Comparison 4. Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	5		Rate Ratio (Fixed, 95% CI)	Subtotals only
1.1 Additional gait, balance, functional training	2	56	Rate Ratio (Fixed, 95% CI)	0.62 [0.40, 0.96]
1.2 Strength/resistance vs self-training	1	34	Rate Ratio (Fixed, 95% CI)	0.74 [0.50, 1.10]
1.3 Balance and strength vs self-training	1	32	Rate Ratio (Fixed, 95% CI)	0.48 [0.30, 0.77]
1.4 Flexibility (Yoga) vs 'Staying active' program	1	20	Rate Ratio (Fixed, 95% CI)	0.47 [0.24, 0.91]
1.5 3D (Tai Chi) vs 'Staying active' program	1	20	Rate Ratio (Fixed, 95% CI)	0.52 [0.28, 0.98]
1.6 Flexibility (Yoga) vs 3D (Tai Chi)	1	18	Rate Ratio (Fixed, 95% CI)	1.11 [0.51, 2.37]
1.7 3D exercises ("In balance") vs Functional balance, strength & mobility	1	142	Rate Ratio (Fixed, 95% CI)	0.73 [0.60, 0.89]
1.8 Wii balance board vs Otago balance program	1	60	Rate Ratio (Fixed, 95% CI)	0.35 [0.19, 0.63]

2 Rate of falls and number of fallers: trials with incomplete data			Other data	No numeric data
3 Number of fallers	5		Risk Ratio (Fixed, 95% CI)	Subtotals only
3.1 Additional gait, balance, and functional training	2	56	Risk Ratio (Fixed, 95% CI)	0.79 [0.43, 1.45]
3.2 Strength/resistance vs self-training	1	34	Risk Ratio (Fixed, 95% CI)	0.56 [0.30, 1.03]
3.3 Balance and strength vs self-training	1	32	Risk Ratio (Fixed, 95% CI)	0.55 [0.29, 1.05]
3.4 Additional whole body vibration	1	159	Risk Ratio (Fixed, 95% CI)	1.28 [0.71, 2.31]
3.5 3D exercises ("In balance") vs Functional balance, strength & mobility	1	142	Risk Ratio (Fixed, 95% CI)	0.92 [0.70, 1.21]
3.6 Comparison of combination exercise programmes	1	41	Risk Ratio (Fixed, 95% CI)	0.54 [0.29, 1.01]
4 Number of people sustaining a fracture	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
4.1 Total fractures	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 5. Care facilities: Medication review vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	7		Rate ratio (Random, 95% CI)	Subtotals only
1.1 General medication reviews vs usual care	6	2409	Rate ratio (Random, 95% CI)	0.93 [0.64, 1.35]
1.2 Medication review for hyponatraemia	1	9	Rate ratio (Random, 95% CI)	0.63 [0.16, 2.49]
2 Number of fallers	7		Risk Ratio (Random, 95% CI)	Subtotals only
2.1 General medication review vs usual care	6	5139	Risk Ratio (Random, 95% CI)	0.93 [0.80, 1.09]
2.2 Medication review for hyponatraemia	1	9	Risk Ratio (Random, 95% CI)	0.42 [0.07, 2.59]
3 Number of people sustaining a fracture	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
3.1 General medication review vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
4 Rate of falls post-hoc sensitivity analysis (excluding Potter 2016)	5		Rate ratio (Random, 95% CI)	Subtotals only
4.1 General medication reviews vs usual care	5		Rate ratio (Random, 95% CI)	0.82 [0.60, 1.11]
5 Serious adverse events	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected

5.1 General medication review vs usual care	1	Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 6. Care facilities: Vitamin D supplementation vs no vitamin D supplementation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	6		Rate ratio (Random, 95% CI)	Subtotals only
1.1 Additional Vitamin D supplementation	4	4512	Rate ratio (Random, 95% CI)	0.72 [0.55, 0.95]
1.2 Multivitamins (including vitamin D3 + calcium) vs placebo	1	91	Rate ratio (Random, 95% CI)	0.38 [0.20, 0.71]
1.3 Education on Vitamin D + calcium + osteoporosis medications vs usual care	1	4017	Rate ratio (Random, 95% CI)	1.03 [0.85, 1.25]
2 Number of fallers	7		Risk Ratio (Random, 95% CI)	Subtotals only
2.1 Vitamin D supplementation	4	4512	Risk Ratio (Random, 95% CI)	0.92 [0.76, 1.12]
2.2 Vitamin D + calcium supplementation vs placebo	1	583	Risk Ratio (Random, 95% CI)	1.03 [0.90, 1.18]
2.3 Multivitamins (including vitamin D3 + calcium) vs usual care or placebo	1	91	Risk Ratio (Random, 95% CI)	0.82 [0.40, 1.66]
2.4 Education on Vitamin D + calcium + osteoporosis medications vs usual care	1	4017	Risk Ratio (Random, 95% CI)	1.05 [0.90, 1.23]
3 Number of people sustaining a fracture	4		Risk Ratio (Random, 95% CI)	Subtotals only
3.1 Vitamin D supplementation	3	4464	Risk Ratio (Random, 95% CI)	1.09 [0.58, 2.03]
3.2 Vitamin D3 + calcium vs placebo	1	583	Risk Ratio (Random, 95% CI)	0.62 [0.36, 1.07]
4 Adverse events	4		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
4.1 Multivitamins (including vitamin D3 + calcium) vs usual care or placebo	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
4.2 Vitamin D + calcium supplementation	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
4.3 Vitamin D supplementation	2		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 7. Care facilities: Environmental interventions vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	1		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Wireless position-monitoring patch vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 8. Care facilities: Social environment vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	4		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Staff education on fracture prevention vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Guideline implementation programme vs control	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.3 Risk assessment tool vs nurses' judgement	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.4 Dementia care mapping vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of fallers	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
2.1 Risk assessment tool vs nurses' judgement	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Number of people sustaining a fracture	2		Risk Ratio (Fixed, 95% CI)	Subtotals only
3.1 Risk assessment tool vs nurses' judgement	1	1125	Risk Ratio (Fixed, 95% CI)	0.96 [0.57, 1.63]
3.2 Project nurse facilitating best-practice falls injury prevention strategies vs usual care	1	5391	Risk Ratio (Fixed, 95% CI)	0.95 [0.63, 1.44]

Comparison 9. Care facilities: Psychological interventions vs control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	1		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Exercise + cognitive training vs exercise	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of fallers	1		Risk Ratio (Fixed, 95% CI)	Totals not selected

2.1 Exercise + cognitive training vs exercise	1	Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 10. Care facilities: Other single interventions vs control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	2		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Lavender patch vs placebo	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Sunlight exposure vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of fallers	2		Risk Ratio (Fixed, 95% CI)	Totals not selected
2.1 Lavender patch vs placebo	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2.2 Sunlight exposure vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Number of people sustaining a fracture	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
3.1 Sunlight exposure vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 11. Care facilities: Multiple interventions vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	2		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Exercise + management of urinary incontinence + fluid therapy vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Sunlight exposure + calcium vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of fallers	2		Risk Ratio (Fixed, 95% CI)	Totals not selected
2.1 Exercise + management of urinary incontinence + fluid therapy vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2.2 Sunlight exposure + calcium vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Number of people sustaining a fracture	2		Risk Ratio (Fixed, 95% CI)	Totals not selected
3.1 Exercise + management of urinary incontinence + fluid therapy vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
3.2 Sunlight exposure + calcium vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 12. Care facilities: Multifactorial interventions vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	10	3439	Rate ratio (Random, 95% CI)	0.88 [0.66, 1.18]
2 Number of fallers	9	3153	Risk Ratio (Random, 95% CI)	0.92 [0.81, 1.05]
3 Number of people sustaining a fracture	5	2160	Risk Ratio (Random, 95% CI)	0.79 [0.30, 2.07]

Comparison 13. Care facilities: Multifactorial interventions vs usual care (grouped by level of care)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	10	3439	Rate Ratio (Random, 95% CI)	0.88 [0.66, 1.18]
1.1 High level nursing care facilities	2	1499	Rate Ratio (Random, 95% CI)	0.59 [0.44, 0.79]
1.2 Intermediate level care facilities	3	670	Rate Ratio (Random, 95% CI)	0.64 [0.50, 0.83]
1.3 Mixed level care facilities	5	1270	Rate Ratio (Random, 95% CI)	1.23 [0.85, 1.77]
2 Number of fallers	9		Risk Ratio (Random, 95% CI)	Subtotals only
2.1 High level nursing care facilities	1	981	Risk Ratio (Random, 95% CI)	0.75 [0.57, 0.98]
2.2 Intermediate level care facilities	3	670	Risk Ratio (Random, 95% CI)	0.75 [0.60, 0.94]
2.3 Mixed level care facilities	5	1502	Risk Ratio (Random, 95% CI)	1.01 [0.88, 1.15]

Comparison 14. Care facilities: Multifactorial interventions vs usual care (grouped by level of cognition)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	10		Rate ratio (Random, 95% CI)	Subtotals only
1.1 Participants with cognitive impairment	4	1199	Rate ratio (Random, 95% CI)	0.83 [0.49, 1.40]
1.2 Participants with no cognitive impairment or mixed sample	8	1805	Rate ratio (Random, 95% CI)	0.84 [0.62, 1.13]
2 Number of fallers	10		Risk Ratio (Random, 95% CI)	Subtotals only
2.1 Participants with cognitive impairment	4	955	Risk Ratio (Random, 95% CI)	0.79 [0.57, 1.12]
2.2 Participants with no cognitive impairment or mixed sample	8	1805	Risk Ratio (Random, 95% CI)	0.94 [0.78, 1.12]

Comparison 15. Hospitals: Additional exercises vs usual physiotherapy

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	2	215	Rate Ratio (Fixed, 95% CI)	0.59 [0.26, 1.34]
2 Number of fallers	2	83	Risk Ratio (Fixed, 95% CI)	0.36 [0.14, 0.93]

Comparison 16. Hospitals: Medication review vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	1		Rate Ratio (Fixed, 95% CI)	Totals not selected
2 Number of fallers	1		Risk Ratio (Fixed, 95% CI)	Totals not selected

Comparison 17. Hospitals: Vitamin D supplements vs no vitamin D supplements

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Number of fallers	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
1.1 Vitamin D + calcium vs calcium	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of people sustaining a fracture	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
2.1 Vitamin D + calcium vs calcium	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Adverse events	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
3.1 Gastrointestinal complaints (nausea, vomiting, diarrhoea)	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 18. Hospitals: Environmental interventions vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	5		Rate Ratio (Random, 95% CI)	Subtotals only
1.1 Carpet flooring vs vinyl flooring	1	54	Rate Ratio (Random, 95% CI)	14.73 [1.88, 115.35]
1.2 Low-low beds vs usual care	1	11099	Rate Ratio (Random, 95% CI)	1.39 [0.22, 8.78]

1.3 Blue identification bracelet vs usual care (no bracelet)	1	134	Rate Ratio (Random, 95% CI)	1.15 [0.72, 1.84]
1.4 Bed alarms vs usual care	2	28649	Rate Ratio (Random, 95% CI)	0.60 [0.27, 1.34]
2 Number of fallers	4		Risk Ratio (Fixed, 95% CI)	Subtotals only
2.1 Carpet flooring vs vinyl flooring	1	54	Risk Ratio (Fixed, 95% CI)	8.33 [0.95, 73.37]
2.2 Blue identification bracelet vs usual care (no bracelet)	1	134	Risk Ratio (Fixed, 95% CI)	1.34 [0.76, 2.36]
2.3 Bed alarms vs usual care	2	28649	Risk Ratio (Fixed, 95% CI)	0.93 [0.38, 2.24]

Comparison 19. Hospitals: Social environment vs control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	5		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Organisational service model change (fall prevention guideline implementation)	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Organisation service model change (falls prevention, incontinence and ulcer guideline implementation)	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.3 Organisational service model change (fall prevention toolkit software)	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.4 Acute care service for elderly patients vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.5 Post-operative orthogeriatric service after hip fracture	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of fallers	3		Risk Ratio (Fixed, 95% CI)	Totals not selected
2.1 Fall prevention tool kit software vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2.2 Behaviour advisory service vs usual care	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2.3 Post-operative orthogeriatric service after hip fracture	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Number of people sustaining a fracture	1		Risk Ratio (Fixed, 95% CI)	Totals not selected
3.1 Post-operative orthogeriatric service after hip fracture	1		Risk Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 20. Hospitals: Knowledge/education interventions vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	1		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Educational materials + health professional follow-up vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Educational materials only vs usual care	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of fallers	2		Risk Ratio (Random, 95% CI)	Totals not selected
2.1 Individualised educational session vs usual care	1		Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
2.2 Educational materials + health professional follow-up vs usual care	1		Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
2.3 Educational materials only vs usual care	1		Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Comparison 21. Hospitals: Multifactorial interventions vs usual care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	5	44664	Rate ratio (Random, 95% CI)	0.80 [0.64, 1.01]
2 Number of fallers	3	39889	Risk Ratio (Random, 95% CI)	0.82 [0.62, 1.09]
3 Number of people sustaining a fracture	2		Risk Ratio (Fixed, 95% CI)	0.76 [0.14, 4.10]

Comparison 22. Hospitals: Multifactorial interventions vs usual care (grouped by type of care)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of falls	5	44664	Rate Ratio (Random, 95% CI)	0.80 [0.64, 1.01]
1.1 Acute level of care	1	35264	Rate Ratio (Random, 95% CI)	1.04 [0.79, 1.37]
1.2 Subacute or acute (mixed) levels of care	2	5653	Rate Ratio (Random, 95% CI)	0.88 [0.61, 1.27]
1.3 Subacute level of care	2	3747	Rate Ratio (Random, 95% CI)	0.67 [0.54, 0.83]
2 Number of fallers	3		Risk Ratio (Random, 95% CI)	0.82 [0.62, 1.09]
2.1 Acute level care	1		Risk Ratio (Random, 95% CI)	0.99 [0.33, 3.00]
2.2 Subacute or acute (mixed) levels of care	1		Risk Ratio (Random, 95% CI)	1.04 [0.48, 2.28]
2.3 Subacute level of care	1		Risk Ratio (Random, 95% CI)	0.78 [0.57, 1.07]

3 Number of people sustaining a fracture	2	Risk Ratio (Fixed, 95% CI)	0.76 [0.14, 4.10]
3.1 Subacute or acute (mixed) levels of care	1	Risk Ratio (Fixed, 95% CI)	0.32 [0.01, 8.95]
3.2 Subacute level of care	1	Risk Ratio (Fixed, 95% CI)	1.02 [0.14, 7.24]

Analysis 1.1. Comparison 1 Care facilities: Exercise vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 1 Care facilities: Exercise vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Exercise N	Usual care N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
Buckinx 2014 (1)	31	31	-0.04 (0.26)		9.7 %	0.96 [0.58, 1.60]
Faber 2006 (2)	142	90	0.12 (0.09)		14.6 %	1.13 [0.95, 1.35]
Irez 2011	30	30	-1.27 (0.33)		7.8 %	0.28 [0.15, 0.54]
Kerse 2008	310	329	0.1 (0.14)		13.3 %	1.11 [0.84, 1.45]
Kovacs 2013	32	30	-0.26 (0.38)		6.7 %	0.77 [0.37, 1.62]
Mulrow 1994	97	97	0.28 (0.17)		12.4 %	1.32 [0.95, 1.85]
Rosendahl 2008 (3)	87	96	-0.2 (0.32)		8.1 %	0.82 [0.44, 1.53]
Sakamoto 2006	315	212	-0.2 (0.12)		13.9 %	0.82 [0.65, 1.04]
Schoenfelder 2000	9	7	1 (0.33)		7.8 %	2.72 [1.42, 5.19]
Sihvonen 2004	20	7	-0.92 (0.43)		5.8 %	0.40 [0.17, 0.93]
Total (95% CI)	**1073**	**929**			**100.0 %**	**0.93 [0.72, 1.20]**

Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 37.19$, df = 9 (P = 0.00002); $I^2 = 76\%$

Test for overall effect: Z = 0.57 (P = 0.57)

Test for subgroup differences: Not applicable

(1) 12 months follow-up

(2) Functional Walking (FW) and In Balance groups (IB) combined vs control

(3) Functional exercise programme vs seated activities

Analysis 1.2. Comparison 1 Care facilities: Exercise vs usual care, Outcome 2 Rate of falls and number of fallers: trials with incomplete data.

Rate of falls and number of fallers: trials with incomplete data

Study	**Intervention**	**Comparator**	**Participants (N)**	**Study findings (NR = not reported)**
Buettner 2002	Exercise: Supervised group exercises, combination exercises.	Usual care	27	Rate of falls: Falls were reduced but the treatment effect estimate and confidence interval were not reported in the published study or research monograph Risk of falling: NR
Cadore 2014	Exercise: Multicomponent exercise programme including gait/balance and strength/resistance training	Usual care including mobility exercises	24	Rate of falls: Over 12 weeks there were no falls in the multicomponent arm in comparison to a rate of falls of 0.8 falls per patient per month in the mobility exercises arm of the study (P < 0.001). Participants were aged > 85 years. Risk of falling: NR
da Silva Borges 2014	Exercise: Ballroom dancing (3D exercises; EG)	No regular physical activity (CG)	59	Rate of falls: The authors reported " fewer falls in the EG post-test compared to the CG post-test (p<0.0001)." Risk of falling: NR
Nowalk 2001	Exercise: 1. "Fit NB Free" Individually tailored combination exercises 2. "Living and Learning/ Tai Chi"	Usual routine activities	110	Rate of falls: NR Risk of falling: No significant difference in risk of falling (time to first fall) between either intervention group and the usual care group (P = 0.29)
Toulotte 2003	Exercise: Supervised exercises, combination exercises.	Usual care	20	Rate of falls: The authors reported that falls were reduced but a falls rate could not be determined from the published data Risk of falling: NR

Analysis 1.3. Comparison 1 Care facilities: Exercise vs usual care, Outcome 3 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 1 Care facilities: Exercise vs usual care

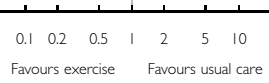
Outcome: 3 Number of fallers

Study or subgroup	Exercise N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Buckinx 2014 (1)	31	31	-0.13 (0.25)		7.8 %	0.88 [0.54, 1.43]
Choi 2005	29	30	-0.51 (0.58)		1.7 %	0.60 [0.19, 1.87]
Faber 2006 (2)	142	90	0.31 (0.19)		12.0 %	1.36 [0.94, 1.98]
Kerse 2008	310	329	0.17 (0.12)		21.8 %	1.19 [0.94, 1.50]
Kovacs 2013	32	30	-0.4 (0.31)		5.4 %	0.67 [0.37, 1.23]
Mulrow 1994	97	97	0.15 (0.17)		14.1 %	1.16 [0.83, 1.62]
Rosendahl 2008	87	96	0.05 (0.16)		15.4 %	1.05 [0.77, 1.44]
Sakamoto 2006	315	212	-0.11 (0.16)		15.4 %	0.90 [0.65, 1.23]
Sihvonen 2004	20	7	-0.26 (0.31)		5.4 %	0.77 [0.42, 1.42]
Yokoi 2015 (3)	51	54	-1.2 (0.74)		1.0 %	0.30 [0.07, 1.28]
Total (95% CI)	**1114**	**976**			**100.0 %**	**1.02 [0.88, 1.18]**

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 11.61$, df = 9 (P = 0.24); $I^2 = 23\%$

Test for overall effect: Z = 0.26 (P = 0.80)

Test for subgroup differences: Not applicable



(1) 12 months follow-up

(2) Functional Walking (FW) and In Balance (IB) groups combined vs control

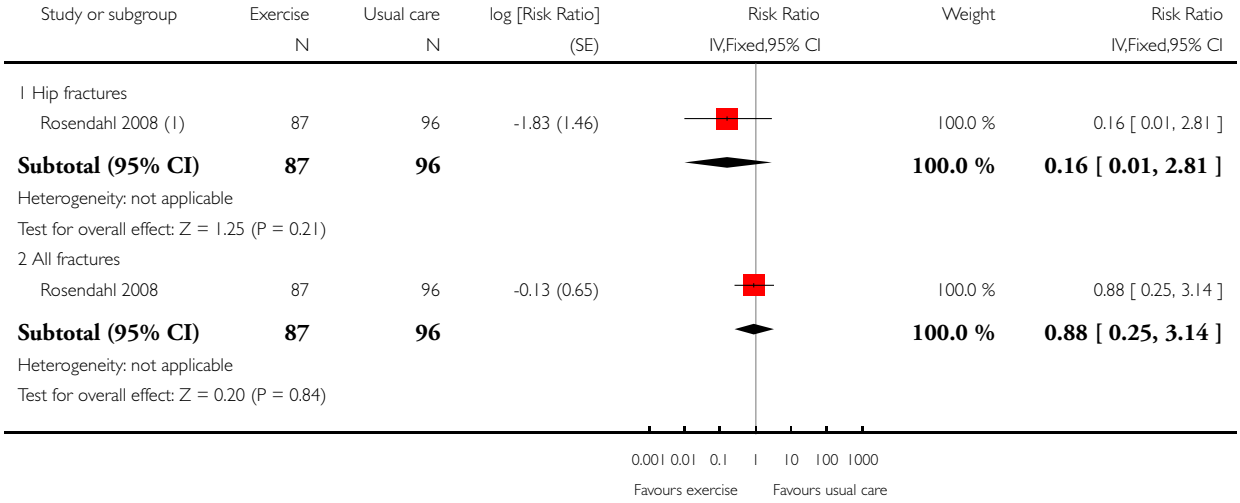
(3) 12 month outcomes

Analysis 1.4. Comparison 1 Care facilities: Exercise vs usual care, Outcome 4 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 1 Care facilities: Exercise vs usual care

Outcome: 4 Number of people sustaining a fracture



(1) Functional exercise programme vs seated activities; mixed levels of care

Analysis 1.5. Comparison 1 Care facilities: Exercise vs usual care, Outcome 5 Rate of falls, excluding studies with ≤20 participants in each arm.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 1 Care facilities: Exercise vs usual care

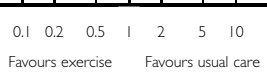
Outcome: 5 Rate of falls, excluding studies with ≤20 participants in each arm

Study or subgroup	Exercise N	Usual care N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
Buckinx 2014	31	31	-0.04 (0.26)		10.4 %	0.96 [0.58, 1.60]
Faber 2006 (1)	142	90	0.12 (0.09)		18.6 %	1.13 [0.95, 1.35]
Irez 2011	30	30	-1.27 (0.33)		8.0 %	0.28 [0.15, 0.54]
Kerse 2008	310	329	0.1 (0.14)		16.2 %	1.11 [0.84, 1.45]
Kovacs 2013	32	30	-0.26 (0.38)		6.7 %	0.77 [0.37, 1.62]
Mulrow 1994	97	97	0.28 (0.17)		14.6 %	1.32 [0.95, 1.85]
Rosendahl 2008	87	96	-0.2 (0.32)		8.3 %	0.82 [0.44, 1.53]
Sakamoto 2006	315	212	-0.2 (0.12)		17.2 %	0.82 [0.65, 1.04]
Total (95% CI)	**1044**	**915**			**100.0 %**	**0.91 [0.72, 1.15]**

Heterogeneity: $Tau^2 = 0.07$; $Chi^2 = 23.47$, df = 7 (P = 0.001); $I^2 = 70\%$

Test for overall effect: Z = 0.80 (P = 0.43)

Test for subgroup differences: Not applicable



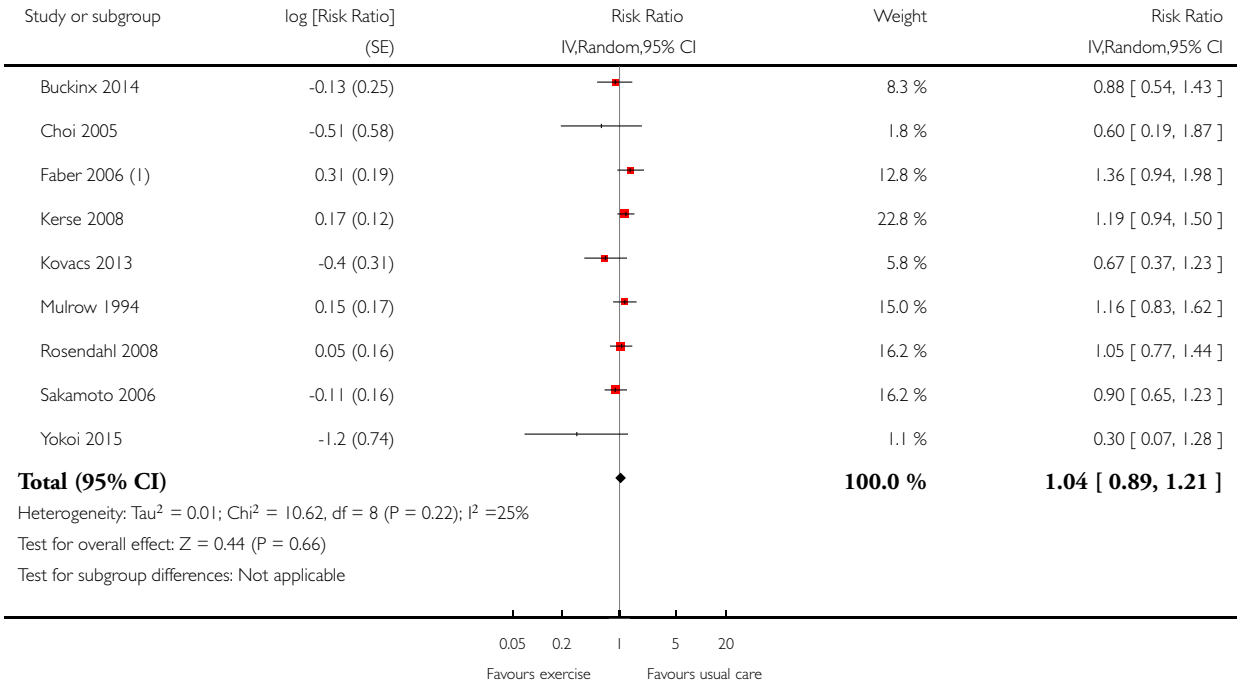
(1) Functional Walking (FW) and In Balance groups (IB) combined vs control

Analysis 1.6. Comparison 1 Care facilities: Exercise vs usual care, Outcome 6 Number of fallers, excluding studies with ≤20 participants in each arm.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 1 Care facilities: Exercise vs usual care

Outcome: 6 Number of fallers, excluding studies with ≤20 participants in each arm



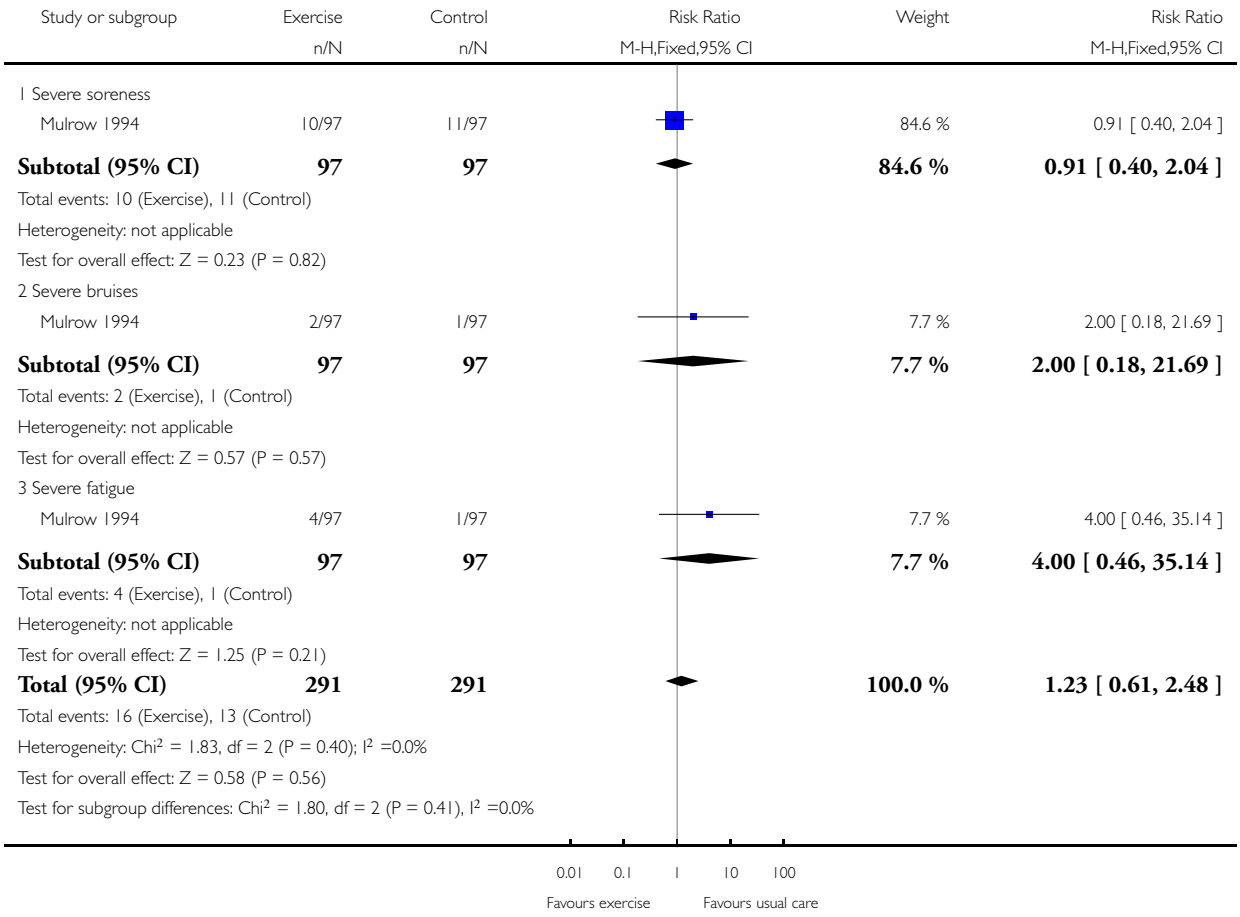
(1) Functional Walking (FW) and In Balance (IB) groups combined vs control

Analysis 1.7. Comparison 1 Care facilities: Exercise vs usual care, Outcome 7 Adverse events: aches and pains.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 1 Care facilities: Exercise vs usual care

Outcome: 7 Adverse events: aches and pains

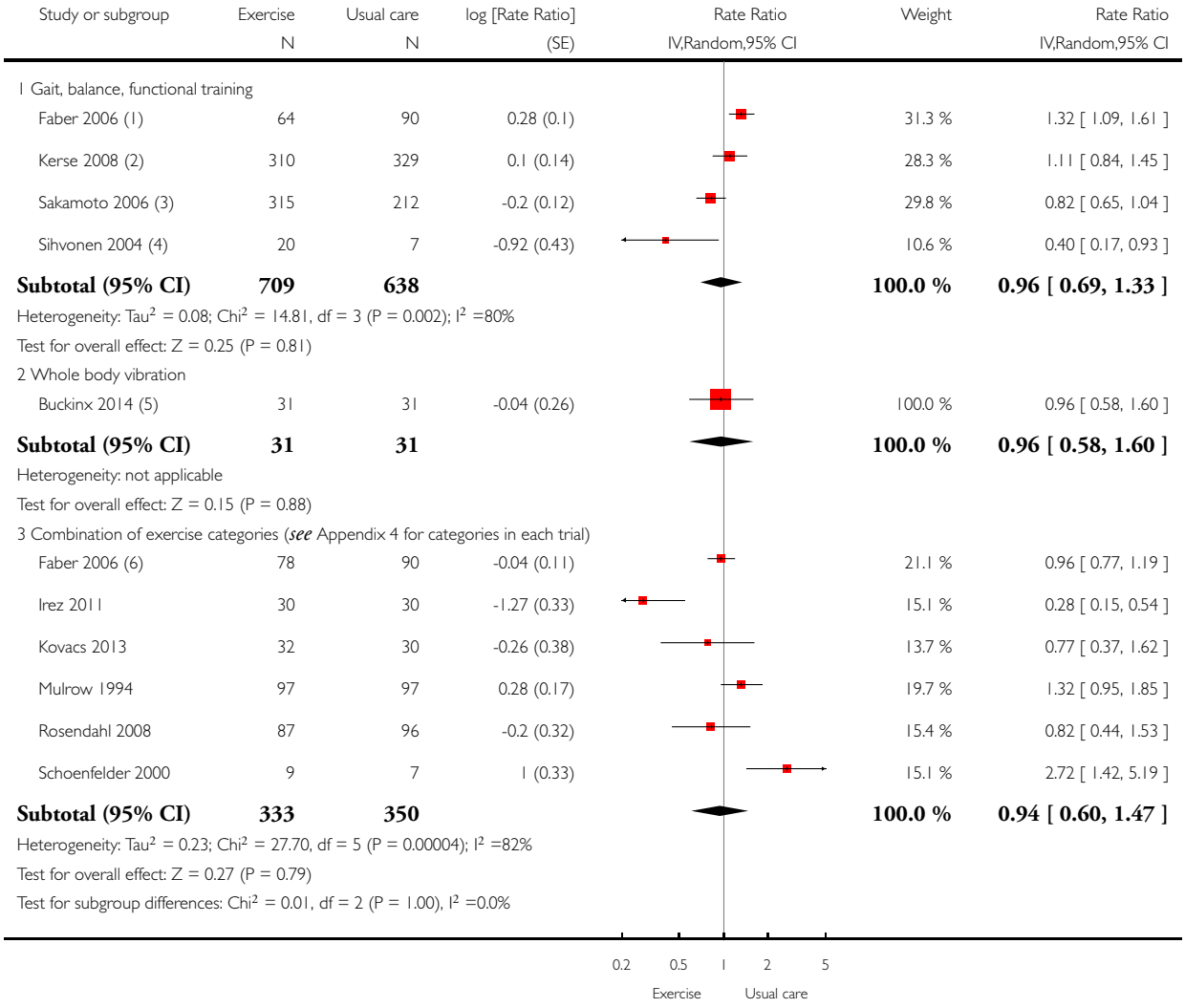


Analysis 2.1. Comparison 2 Care facilities: Exercises vs usual care (grouped by type of exercise), Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 2 Care facilities: Exercises vs usual care (grouped by type of exercise)

Outcome: 1 Rate of falls



(1) Functional Walking (FW) group vs control

(2) goal-setting physical activity programme

(3) balance training: one-leg standing

(4) balance training: mechanical apparatus

(5) Whole body vibration vs usual care (12 months)

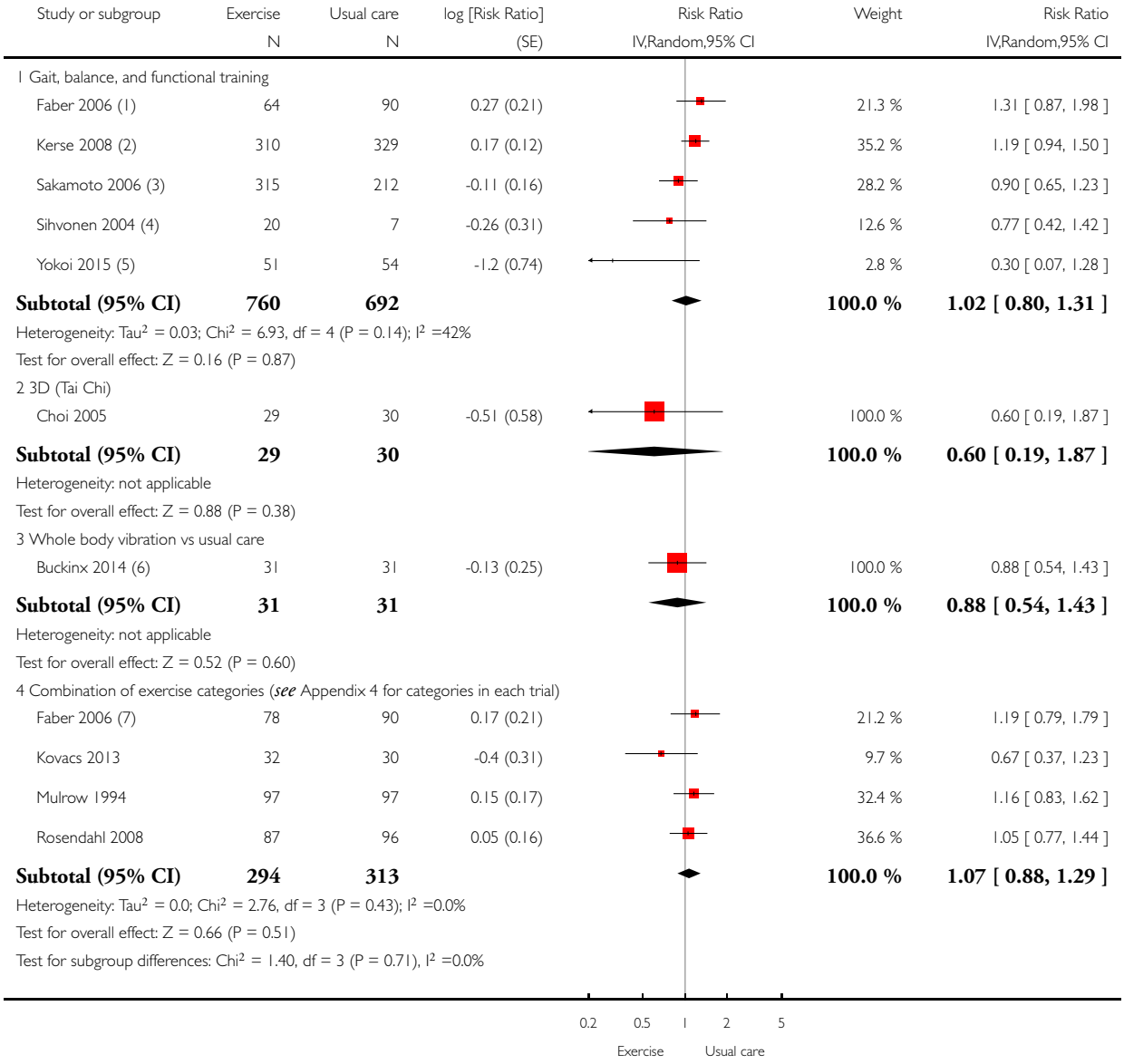
(6) In Balance (IB) group vs control

Analysis 2.2. Comparison 2 Care facilities: Exercises vs usual care (grouped by type of exercise), Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 2 Care facilities: Exercises vs usual care (grouped by type of exercise)

Outcome: 2 Number of fallers



(1) Functional Walking (FW) group vs control

(2) goal-setting physical activity programme

(3) balance training: one-leg standing

(4) balance training: mechanical apparatus

(5) short stick exercises, 12 month outcomes

(6) Whole body vibration vs usual care (12 months)

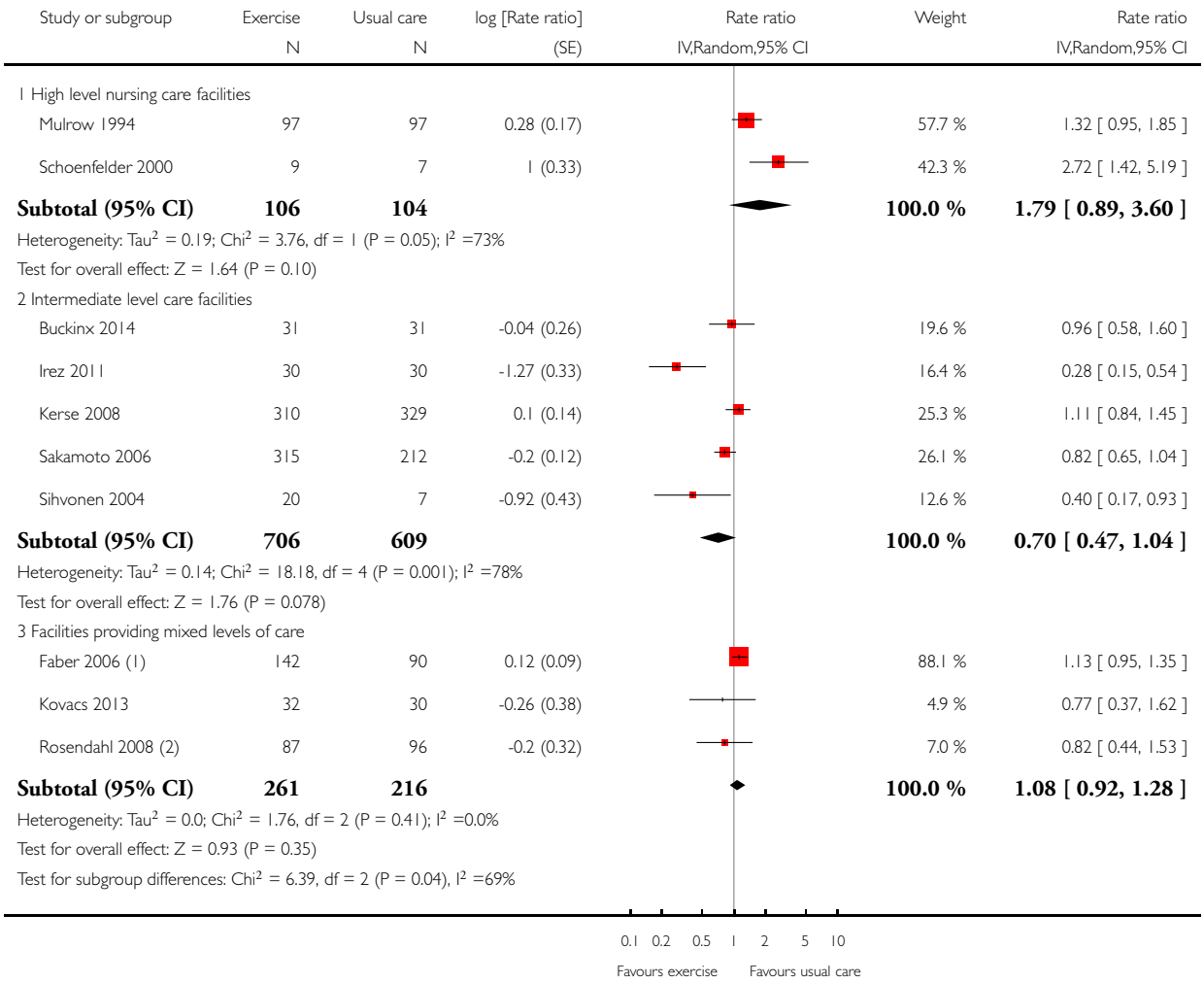
(7) In Balance (IB) group vs control

Analysis 3.1. Comparison 3 Care facilities: Exercise vs usual care (grouped by level of care), Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 3 Care facilities: Exercise vs usual care (grouped by level of care)

Outcome: 1 Rate of falls



(1) Functional Walking (FW) and In Balance groups (IB) combined vs control

(2) Functional exercise programme vs seated activities

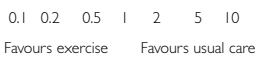
Analysis 3.2. Comparison 3 Care facilities: Exercise vs usual care (grouped by level of care), Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 3 Care facilities: Exercise vs usual care (grouped by level of care)

Outcome: 2 Number of fallers

Study or subgroup	Exercise N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 High level nursing care facilities						
Mulrow 1994	97	97	0.15 (0.17)		14.1 %	1.16 [0.83, 1.62]
Subtotal (95% CI)	**97**	**97**			**14.1 %**	**1.16 [0.83, 1.62]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.88 (P = 0.38)						
2 Intermediate level care facilities						
Buckinx 2014	31	31	-0.13 (0.25)		7.8 %	0.88 [0.54, 1.43]
Choi 2005	29	30	-0.51 (0.58)		1.7 %	0.60 [0.19, 1.87]
Kerse 2008	310	329	0.17 (0.12)		21.8 %	1.19 [0.94, 1.50]
Sakamoto 2006	315	212	-0.11 (0.16)		15.4 %	0.90 [0.65, 1.23]
Sihvonen 2004	20	7	-0.26 (0.31)		5.4 %	0.77 [0.42, 1.42]
Yokoi 2015 (1)	51	54	-1.2 (0.74)		1.0 %	0.30 [0.07, 1.28]
Subtotal (95% CI)	**756**	**663**			**53.1 %**	**0.94 [0.75, 1.17]**
Heterogeneity: Tau^2 = 0.02; Chi^2 = 6.85, df = 5 (P = 0.23); I^2 =27%						
Test for overall effect: Z = 0.56 (P = 0.58)						
3 Mixed level care facilities						
Faber 2006 (2)	142	90	0.31 (0.19)		12.0 %	1.36 [0.94, 1.98]
Kovacs 2013	32	30	-0.4 (0.31)		5.4 %	0.67 [0.37, 1.23]
Rosendahl 2008	87	96	0.05 (0.16)		15.4 %	1.05 [0.77, 1.44]
Subtotal (95% CI)	**261**	**216**			**32.8 %**	**1.05 [0.76, 1.47]**
Heterogeneity: Tau^2 = 0.04; Chi^2 = 3.90, df = 2 (P = 0.14); I^2 =49%						
Test for overall effect: Z = 0.31 (P = 0.76)						
Total (95% CI)	**1114**	**976**			**100.0 %**	**1.02 [0.88, 1.18]**
Heterogeneity: Tau^2 = 0.01; Chi^2 = 11.61, df = 9 (P = 0.24); I^2 =23%						
Test for overall effect: Z = 0.26 (P = 0.80)						
Test for subgroup differences: Chi^2 = 1.15, df = 2 (P = 0.56), I^2 =0.0%						



(1) 12 month outcomes

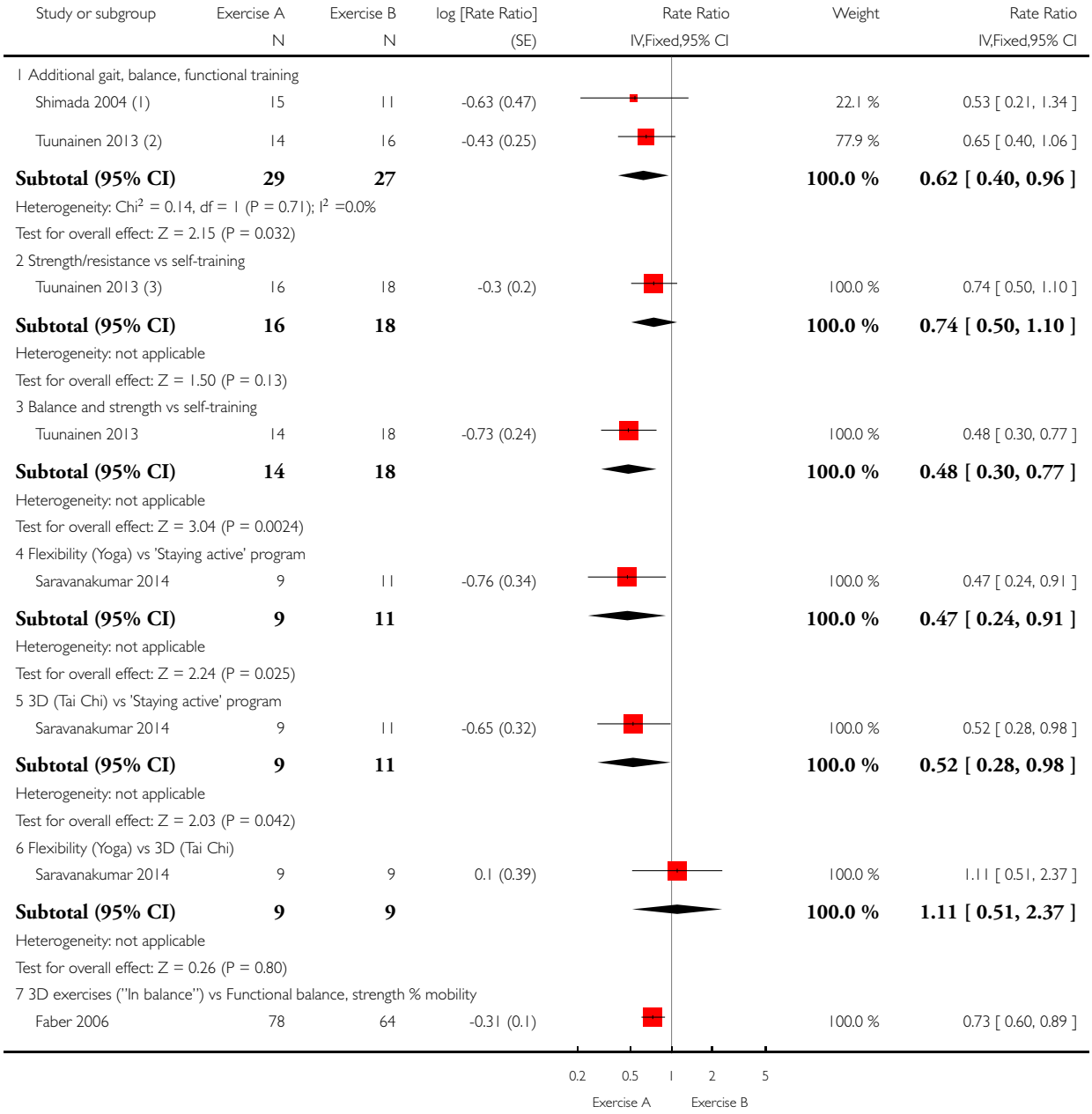
(2) Functional Walking (FW) and In Balance (IB) groups combined vs control

Analysis 4.1. Comparison 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details), Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

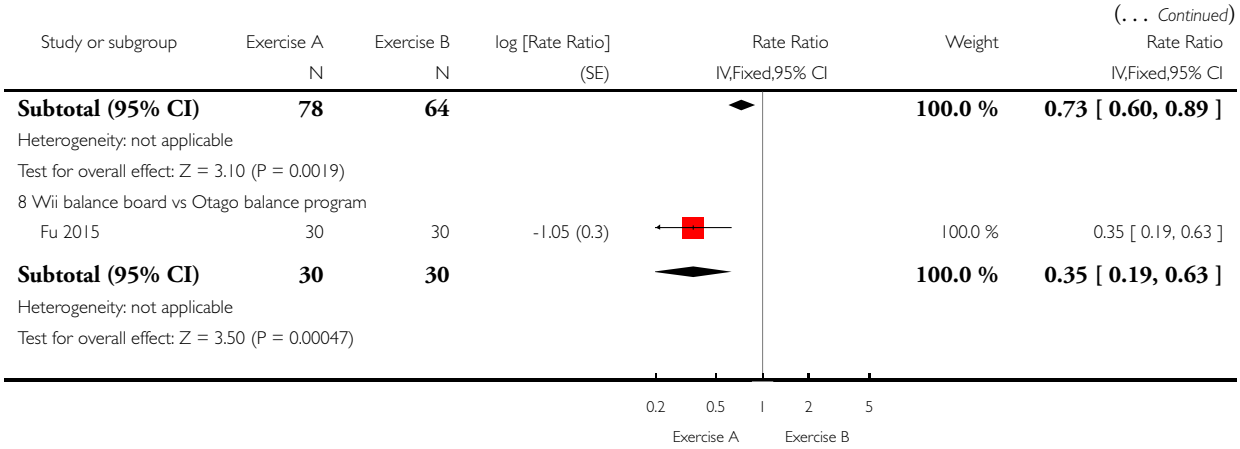
Comparison: 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details)

Outcome: 1 Rate of falls



(*Continued* . . .)

(. . . *Continued*)



(1) balance training: mechanical apparatus + combination exercises vs combination exercises

(2) Balance and strength training vs strength training

(3) Progressive resistance group training vs self-training

Analysis 4.2. Comparison 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details), Outcome 2 Rate of falls and number of fallers: trials with incomplete data.

Rate of falls and number of fallers: trials with incomplete data

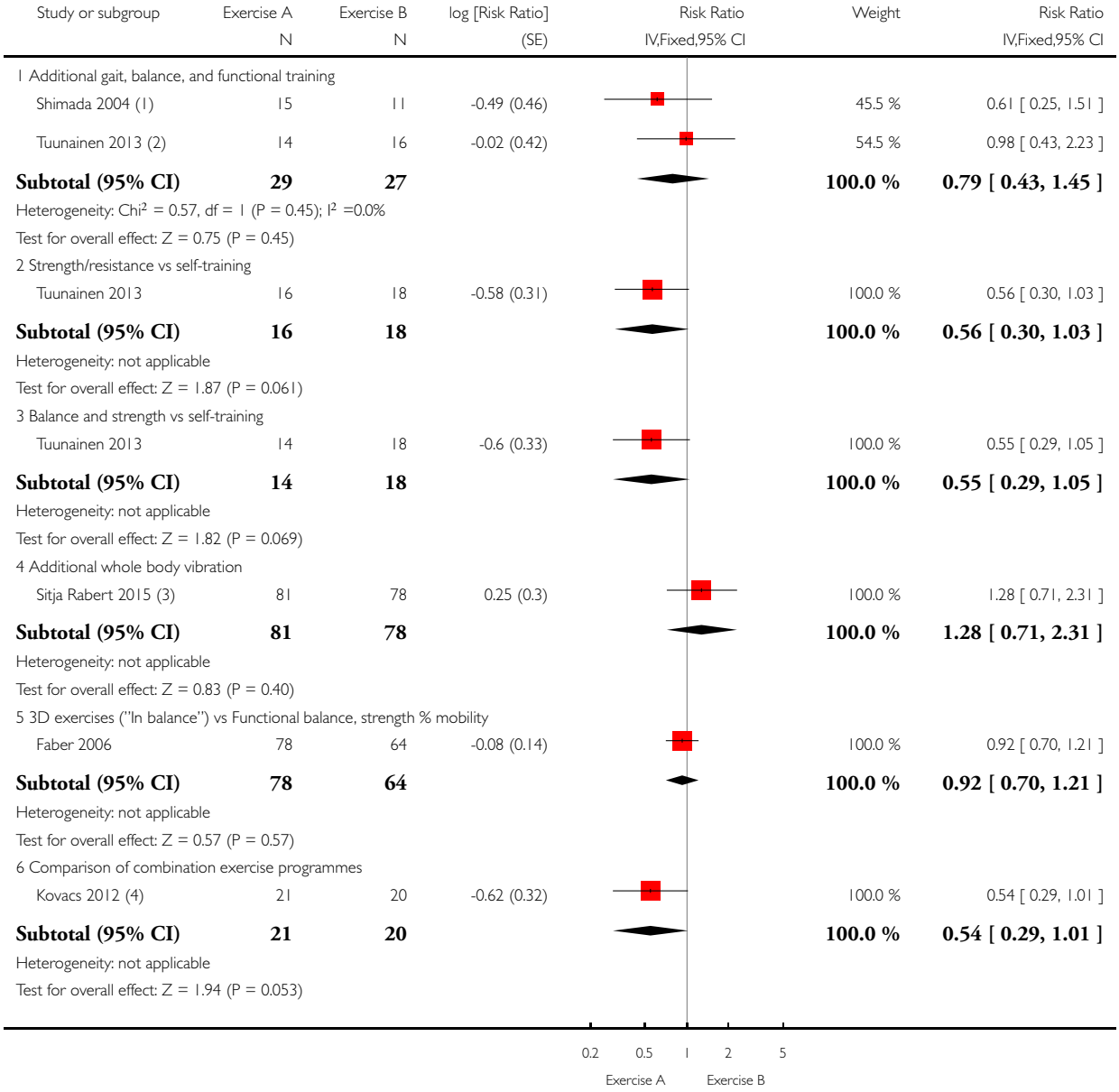
Study	Intervention	Comparator	Participants (N)	Study findings
Imaoka 2016	Exercise: Additional group exercise (described by author as "Usual care": combination group exercises plus individualised exercise)	Individualised exercise (described by author as "reduced exercise")	39	Rate of falls: Not reported Risk of falling: No strong evidence for a reduction in the risk of falling in the post-intervention period with additional group exercise (RR 0.48, 95% CI 0.17 to 1. 3). The falls data are not presented in the forest plot as they exclude the intervention period
Serra-Rexach 2011	Exercise: Training sessions (combination exercises) plus usual care physiotherapy	Usual care physiotherapy (40-45 min / day 5 x weekly)	40	Rate of falls: "The mean number of falls per participant recorded over the study period was 1.2 fewer in the intervention group than in the control group (95% CI = 0.0-3.0, P =.03)." Risk of falling: not reported

Analysis 4.3. Comparison 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details), Outcome 3 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details)

Outcome: 3 Number of fallers



(1) balance training: mechanical apparatus + combination exercises vs combination exercises

(2) Balance and strength training vs strength training

(3) Whole body vibration balance % strength training vs balance % strength training

(4) Multimodel exercise programme based on Otago plus oesteoporosis exercises vs osteoporosis exercises

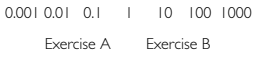
Analysis 4.4. Comparison 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details), Outcome 4 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 4 Care facilities: Comparisons of different exercise programs (*see* Appendix 4 for details)

Outcome: 4 Number of people sustaining a fracture

Study or subgroup	Exercise A N	Exercise B N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Total fractures					
Sitja Rabert 2015 (1)	81	78	1.06 (1.62)		2.89 [0.12, 69.07]



(1) Whole body vibration balance % strength training vs balance % strength training

Analysis 5.1. Comparison 5 Care facilities: Medication review vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 5 Care facilities: Medication review vs usual care

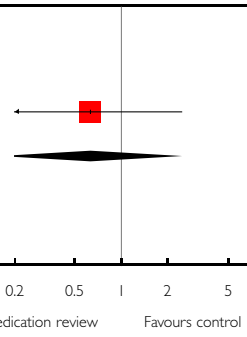
Outcome: 1 Rate of falls

Study or subgroup	Medication review N	Control N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
1 General medication reviews vs usual care						
Frankenthal 2014 (1)	160	146	-0.49 (0.12)		16.8 %	0.61 [0.48, 0.78]
Houghton 2014 (2)	381	445	0.01 (0.16)		16.0 %	1.01 [0.74, 1.38]
Juola 2015 (3)	93	96	-0.33 (0.14)		16.4 %	0.72 [0.55, 0.95]
Patterson 2010 (4)	173	161	0.36 (0.15)		16.2 %	1.43 [1.07, 1.92]
Potter 2016 (5)	45	48	0.51 (0.11)		17.0 %	1.67 [1.34, 2.07]
Zermansky 2006 (6)	331	330	-0.48 (0.08)		17.5 %	0.62 [0.53, 0.72]
Subtotal (95% CI)	**1183**	**1226**			**100.0 %**	**0.93 [0.64, 1.35]**

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Medication review N	Control N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
Heterogeneity: Tau^2 = 0.20; Chi^2 = 75.25, df = 5 (P<0.00001); I^2 =93%						
Test for overall effect: Z = 0.39 (P = 0.70)						
2 Medication review for hyponatraemia						
Peyro Saint Paul 2013 (7)	4	5	-0.46 (0.7)		100.0 %	0.63 [0.16, 2.49]
Subtotal (95% CI)	**4**	**5**			**100.0 %**	**0.63 [0.16, 2.49]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.66 (P = 0.51)						

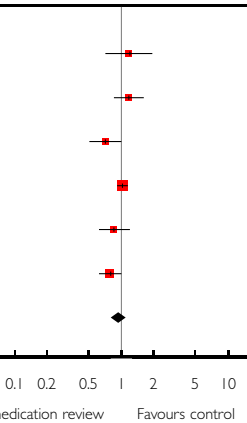


(1) Medication review with recommendations to chief physician based on STOPP/START criteria
(2) Medication review meeting involving a meeting involving clinical pharmacist, pharmacy technician, care home staff and GP(s)
(3) Nurse education on harmful medications in older people, adjusted for age, sex, comorbidities
(4) Monthly review targeting psychoactive medication prescribing for 12 months
(5) Medication review with desprescribing vs medication review without deprescribing
(6) One review of GP record + consultation with patient and carer
(7) Pharmacist review of medications of patients identified with hyponatremia

Analysis 5.2. Comparison 5 Care facilities: Medication review vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals
Comparison: 5 Care facilities: Medication review vs usual care
Outcome: 2 Number of fallers

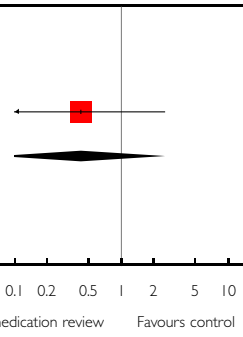
Study or subgroup	Medication review N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 General medication review vs usual care						
Crotty 2004a (1)	56	54	0.17 (0.26)		7.4 %	1.19 [0.71, 1.97]
Crotty 2004b (2)	381	384	0.16 (0.16)		14.8 %	1.17 [0.86, 1.61]
Juola 2015 (3)	93	96	-0.342 (0.177)		13.0 %	0.71 [0.50, 1.00]
Lapane 2011 (4)	1769	1552	0.03 (0.06)		31.0 %	1.03 [0.92, 1.16]
Potter 2016 (5)	45	48	-0.15 (0.17)		13.7 %	0.86 [0.62, 1.20]
Zermansky 2006 (6)	331	330	-0.24 (0.12)		20.1 %	0.79 [0.62, 1.00]
Subtotal (95% CI)	**2675**	**2464**			**100.0 %**	**0.93 [0.80, 1.09]**



(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Medication review N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Heterogeneity: Tau^2 = 0.02; Chi^2 = 9.66, df = 5 (P = 0.09); I^2 =48%						
Test for overall effect: Z = 0.86 (P = 0.39)						
2 Medication review for hyponatraemia						
Peyro Saint Paul 2013 (7)	4	5	-0.87 (0.93)		100.0 %	0.42 [0.07, 2.59]
Subtotal (95% CI)	**4**	**5**			**100.0 %**	**0.42 [0.07, 2.59]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.94 (P = 0.35)						



(1) Pharmacist transition coordinator for patients discharged from hospital to nursing care facilities for the first time

(2) Pharmacist-led outreach programme (audit + feedback + education of staff regarding medications and falls risk)

(3) Nurse education on harmful medications in older people

(4) GRAM software for decision support for prescribing practices vs monthly medication review

(5) A GP and a geriatrician/pharmacologist independently identified deprescribing targets using a list of potentially inappropriate medicines

(6) One review of GP record + consultation with patient and carer

(7) Pharmacist review of medications of patients identified with hyponatremia

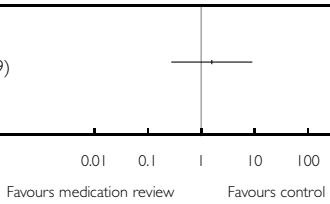
Analysis 5.3. Comparison 5 Care facilities: Medication review vs usual care, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 5 Care facilities: Medication review vs usual care

Outcome: 3 Number of people sustaining a fracture

Study or subgroup	Medication review N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 General medication review vs usual care					
Potter 2016	45	48	0.47 (0.89)		1.60 [0.28, 9.16]

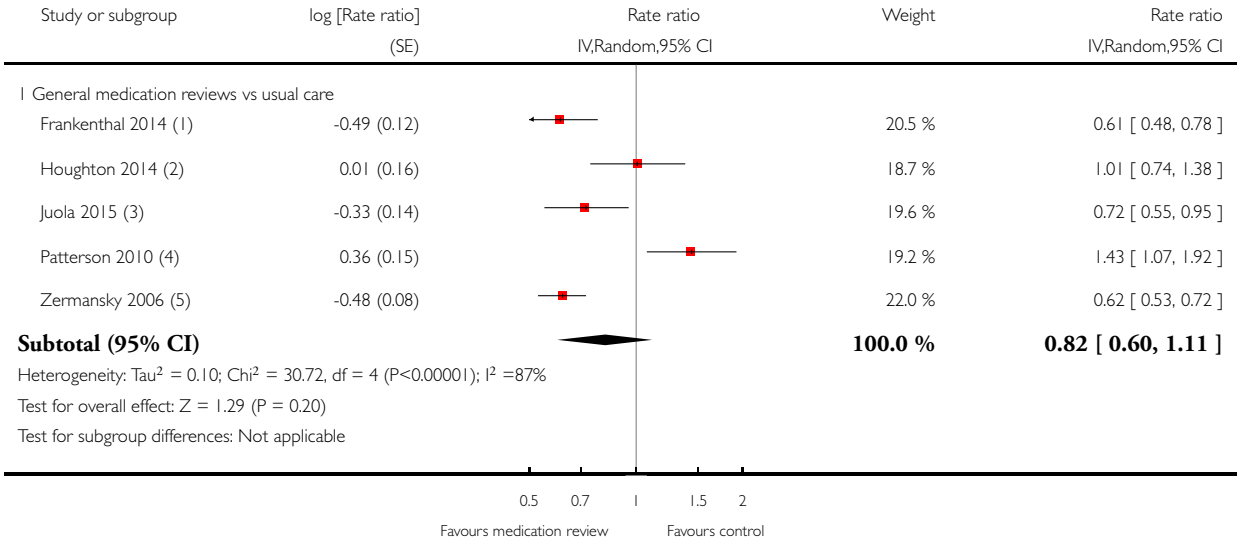


Analysis 5.4. Comparison 5 Care facilities: Medication review vs usual care, Outcome 4 Rate of falls post-hoc sensitivity analysis (excluding Potter 2016).

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 5 Care facilities: Medication review vs usual care

Outcome: 4 Rate of falls post-hoc sensitivity analysis (excluding Potter 2016)



(1) Medication review with recommendations to chief physician based on STOPP/START criteria

(2) Medication review meeting involving a meeting involving clinical pharmacist, pharmacy technician, care home staff and GP(s)

(3) Nurse education on harmful medications in older people, adjusted for age, sex, comorbidities

(4) Monthly review targeting psychoactive medication prescribing for 12 months

(5) One review of GP record + consultation with patient and carer

Analysis 5.5. Comparison 5 Care facilities: Medication review vs usual care, Outcome 5 Serious adverse events.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 5 Care facilities: Medication review vs usual care

Outcome: 5 Serious adverse events

Study or subgroup	Medication review n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 General medication review vs usual care				
Potter 2016	3/45	3/48		1.07 [0.23, 5.01]

Analysis 6.1. Comparison 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation

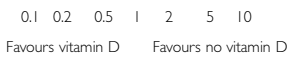
Outcome: 1 Rate of falls

Study or subgroup	Vitamin D N	No vitamin D N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
1 Additional Vitamin D supplementation						
Bischoff 2003 (1)	62	60	-0.67 (0.41)		9.5 %	0.51 [0.23, 1.14]
Broe 2007 (2)	23	25	-1.27 (0.51)		6.5 %	0.28 [0.10, 0.76]
Flicker 2005 (3)	313	312	-0.31 (0.13)		35.3 %	0.73 [0.57, 0.95]
Law 2006 (4)	1762	1955	-0.14 (0.04)		48.7 %	0.87 [0.80, 0.94]
Subtotal (95% CI)	**2160**	**2352**			**100.0 %**	**0.72 [0.55, 0.95]**
Heterogeneity: Tau^2 = 0.04; Chi^2 = 7.85, df = 3 (P = 0.05); I^2 =62%						
Test for overall effect: Z = 2.32 (P = 0.020)						
2 Multivitamins (including vitamin D3 + calcium) vs placebo						
Grieger 2009	48	43	-0.97 (0.32)		100.0 %	0.38 [0.20, 0.71]
Subtotal (95% CI)	**48**	**43**			**100.0 %**	**0.38 [0.20, 0.71]**
Heterogeneity: not applicable						
Test for overall effect: Z = 3.03 (P = 0.0024)						

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Vitamin D N	No vitamin D N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
3 Education on Vitamin D + calcium + osteoporosis medications vs usual care						
Kennedy 2015	1290	2727	0.03 (0.1)		100.0 %	1.03 [0.85, 1.25]
Subtotal (95% CI)	**1290**	**2727**			**100.0 %**	**1.03 [0.85, 1.25]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.30 (P = 0.76)						
Test for subgroup differences: Chi^2 = 11.35, df = 2 (P = 0.00), I^2 =82%						



(1) Vitamin D3 + calcium vs calcium
(2) 800 IU vitamin D group only vs placebo
(3) Vitamin D3 + calcium vs calcium
(4) Vitamin D2 vs usual care

Analysis 6.2. Comparison 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation, Outcome 2 Number of fallers.

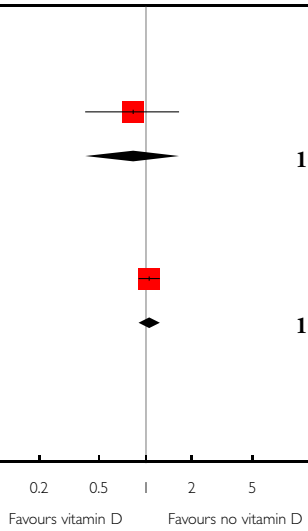
Review: Interventions for preventing falls in older people in care facilities and hospitals
Comparison: 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation
Outcome: 2 Number of fallers

Study or subgroup	Vitamin D N	No vitamin D N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 Vitamin D supplementation						
Bischoff 2003 (1)	62	60	-0.36 (0.41)		5.1 %	0.70 [0.31, 1.56]
Broe 2007 (2)	23	25	-0.82 (0.55)		3.0 %	0.44 [0.15, 1.29]
Flicker 2005 (3)	313	312	-0.15 (0.11)		35.7 %	0.86 [0.69, 1.07]
Law 2006 (4)	1762	1955	0.03 (0.05)		56.2 %	1.03 [0.93, 1.14]
Subtotal (95% CI)	**2160**	**2352**			**100.0 %**	**0.92 [0.76, 1.12]**
Heterogeneity: Tau^2 = 0.01; Chi^2 = 5.16, df = 3 (P = 0.16); I^2 =42%						
Test for overall effect: Z = 0.82 (P = 0.41)						
2 Vitamin D + calcium supplementation vs placebo						
Chapuy 2002 (5)	393	190	0.03 (0.07)		100.0 %	1.03 [0.90, 1.18]
Subtotal (95% CI)	**393**	**190**			**100.0 %**	**1.03 [0.90, 1.18]**

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Vitamin D N	No vitamin D N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Heterogeneity: not applicable						
Test for overall effect: Z = 0.43 (P = 0.67)						
3 Multivitamins (including vitamin D3 + calcium) vs usual care or placebo						
Grieger 2009	48	43	-0.2 (0.36)		100.0 %	0.82 [0.40, 1.66]
Subtotal (95% CI)	**48**	**43**			**100.0 %**	**0.82 [0.40, 1.66]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.56 (P = 0.58)						
4 Education on Vitamin D + calcium + osteoporosis medications vs usual care						
Kennedy 2015	1290	2727	0.05 (0.08)		100.0 %	1.05 [0.90, 1.23]
Subtotal (95% CI)	**1290**	**2727**			**100.0 %**	**1.05 [0.90, 1.23]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.63 (P = 0.53)						
Test for subgroup differences: Chi² = 1.52, df = 3 (P = 0.68), I² =0.0%						



(1) Vitamin D3 + calcium vs calcium

(2) 800 IU vitamin D group only vs placebo

(3) Vitamin D3 + calcium vs calcium

(4) Vitamin D2 vs usual care

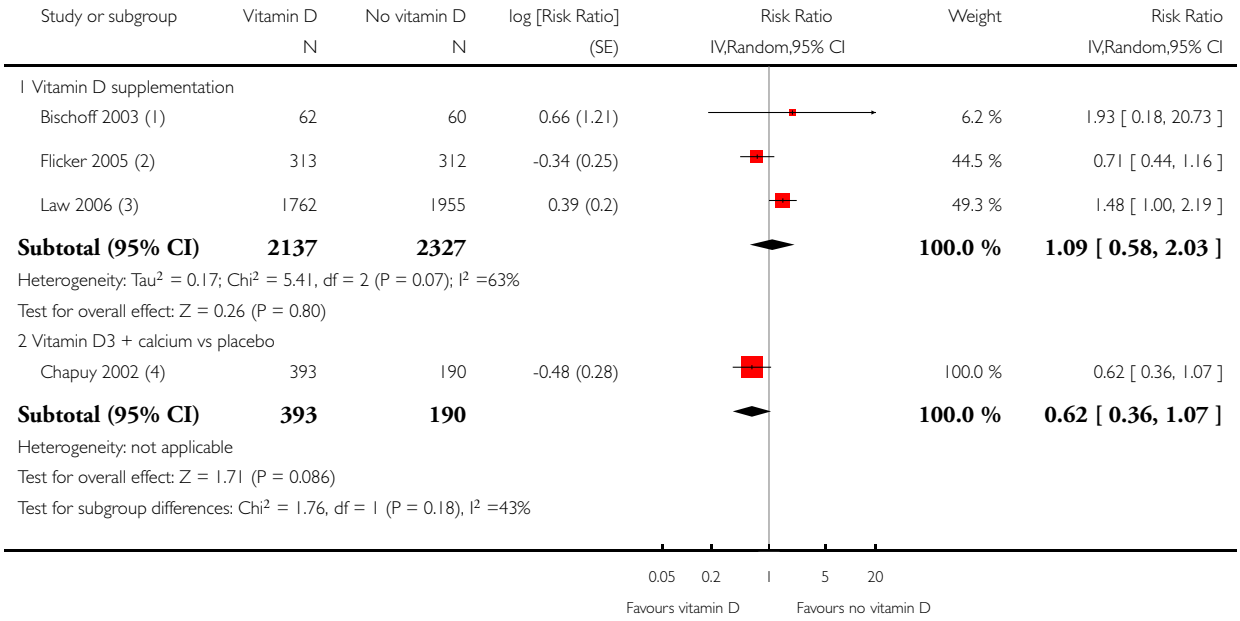
(5) Vitamin D3 + calcium vs placebo

Analysis 6.3. Comparison 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation

Outcome: 3 Number of people sustaining a fracture



(1) Hip fracture; Vitamin D3 + calcium vs calcium

(2) All fractures; Vitamin D3 + calcium vs calcium

(3) Non vertebral fractures; Vitamin D2 vs usual care

(4) Hip fracture; Vitamin D3 + calcium vs placebo

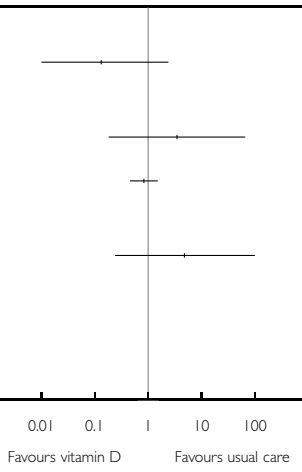
Analysis 6.4. Comparison 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation, Outcome 4 Adverse events.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 6 Care facilities: Vitamin D supplementation vs no vitamin D supplementation

Outcome: 4 Adverse events

Study or subgroup	Vitamin D n/N	No vitamin D n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Multivitamins (including vitamin D3 + calcium) vs usual care or placebo				
Grieger 2009 (1)	0/48	3/43		0.13 [0.01, 2.41]
2 Vitamin D + calcium supplementation				
Chapuy 2002 (2)	3/393	0/190		3.39 [0.18, 65.36]
Chapuy 2002 (3)	27/393	16/190		0.82 [0.45, 1.48]
3 Vitamin D supplementation				
Bischoff 2003 (4)	2/62	0/60		4.84 [0.24, 98.80]
Bischoff 2003 (5)	0/62	0/60		Not estimable
Flicker 2005 (6)	0/313	0/312		Not estimable



(1) rash/vertigo, behavioural issues, indigestion

(2) Hypercalcaemia

(3) Gastrointestinal disorders

(4) constipation

(5) Hypercalcaemia

(6) No adverse events

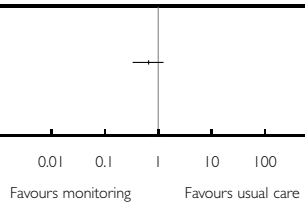
Analysis 7.1. Comparison 7 Care facilities: Environmental interventions vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 7 Care facilities: Environmental interventions vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Wireless monitoring N	Usual care N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Wireless position-monitoring patch vs usual care					
Clifton 2009	33	39	-0.43 (0.34)		0.65 [0.33, 1.27]



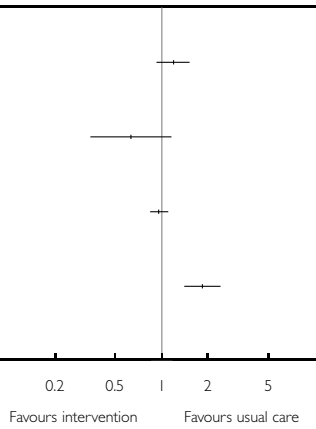
Analysis 8.1. Comparison 8 Care facilities: Social environment vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 8 Care facilities: Social environment vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Social environment change N	Usual care N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Staff education on fracture prevention vs usual care					
Cox 2008	3315	2322	0.17 (0.13)		1.19 [0.92, 1.53]
2 Guideline implementation programme vs control					
Van Gaal 2011a	196	196	-0.46 (0.31)		0.63 [0.34, 1.16]
3 Risk assessment tool vs nurses' judgement					
Meyer 2009	574	551	-0.04 (0.07)		0.96 [0.84, 1.10]
4 Dementia care mapping vs usual care					
Van de Ven 2014	137	156	0.61 (0.14)		1.84 [1.40, 2.42]

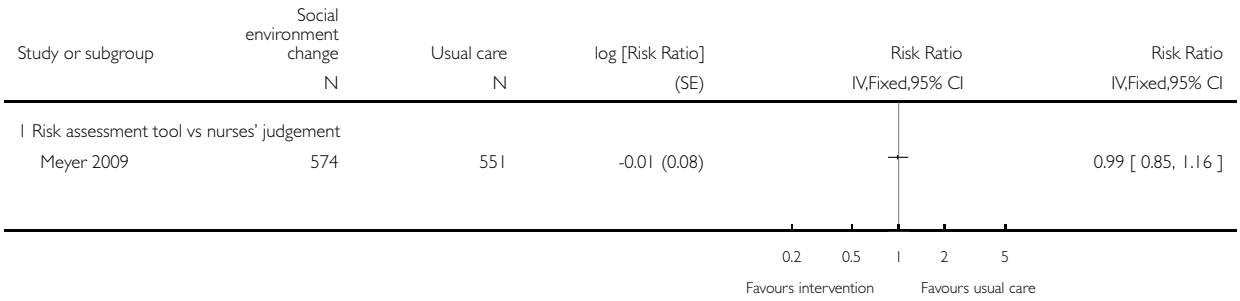


Analysis 8.2. Comparison 8 Care facilities: Social environment vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 8 Care facilities: Social environment vs usual care

Outcome: 2 Number of fallers

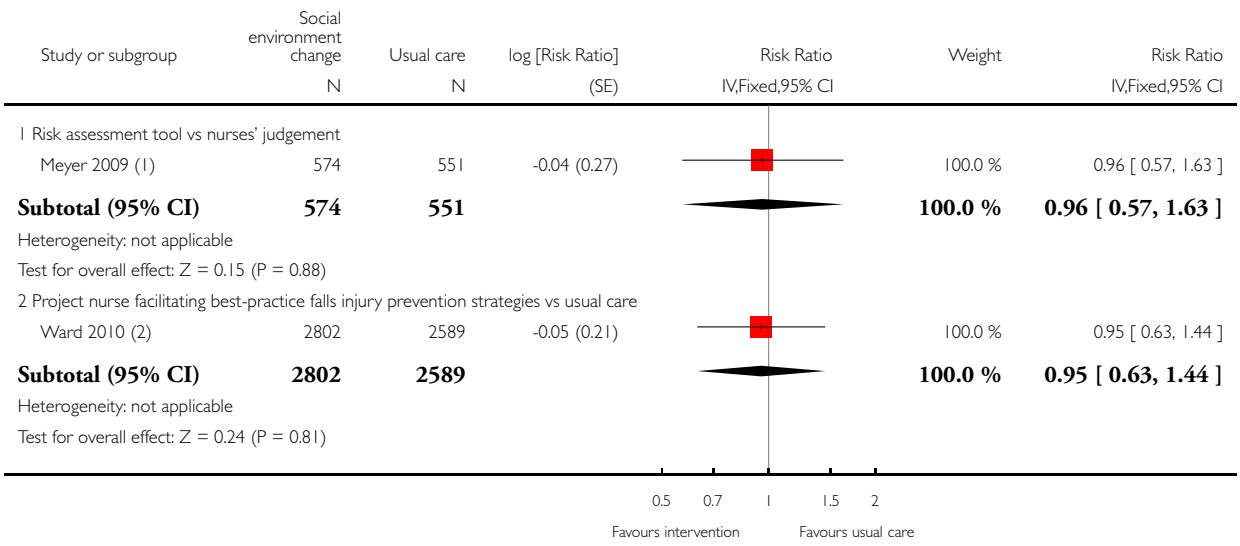


Analysis 8.3. Comparison 8 Care facilities: Social environment vs usual care, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 8 Care facilities: Social environment vs usual care

Outcome: 3 Number of people sustaining a fracture

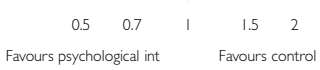


(1) All fractures
(2) Hip fracture

Analysis 9.1. Comparison 9 Care facilities: Psychological interventions vs control, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals
Comparison: 9 Care facilities: Psychological interventions vs control
Outcome: 1 Rate of falls

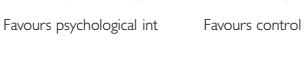
Study or subgroup	Psychological Int N	Control N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Exercise + cognitive training vs exercise					
Van het Reve 2014	60	54	0.2 (0.23)		1.22 [0.78, 1.92]



Analysis 9.2. Comparison 9 Care facilities: Psychological interventions vs control, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals
Comparison: 9 Care facilities: Psychological interventions vs control
Outcome: 2 Number of fallers

Study or subgroup	Psychological Int N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Exercise + cognitive training vs exercise					
Van het Reve 2014	60	54	0.3 (0.9)		1.35 [0.23, 7.88]



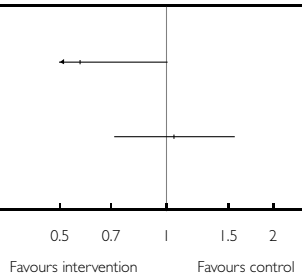
Analysis 10.1. Comparison 10 Care facilities: Other single interventions vs control, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 10 Care facilities: Other single interventions vs control

Outcome: 1 Rate of falls

Study or subgroup	Intervention N	Control N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Lavender patch vs placebo					
Sakamoto 2012	73	72	-0.56 (0.29)		0.57 [0.32, 1.01]
2 Sunlight exposure vs usual care					
Sambrook 2012	190	205	0.05 (0.2)		1.05 [0.71, 1.56]



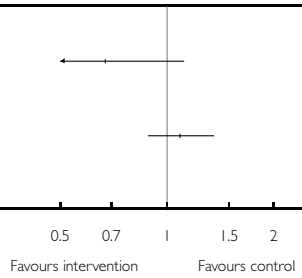
Analysis 10.2. Comparison 10 Care facilities: Other single interventions vs control, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 10 Care facilities: Other single interventions vs control

Outcome: 2 Number of fallers

Study or subgroup	Intervention N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Lavender patch vs placebo					
Sakamoto 2012	73	72	-0.4 (0.26)		0.67 [0.40, 1.12]
2 Sunlight exposure vs usual care					
Sambrook 2012	190	205	0.09 (0.11)		1.09 [0.88, 1.36]



Analysis 10.3. Comparison 10 Care facilities: Other single interventions vs control, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 10 Care facilities: Other single interventions vs control

Outcome: 3 Number of people sustaining a fracture

Study or subgroup	Intervention N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Sunlight exposure vs usual care					
Sambrook 2012	190	205	0.07 (0.36)		1.07 [0.53, 2.17]

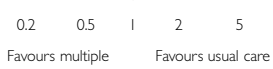
Analysis 11.1. Comparison 11 Care facilities: Multiple interventions vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 11 Care facilities: Multiple interventions vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Multiple Intervention N	Usual care N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Exercise + management of urinary incontinence + fluid therapy vs usual care					
Schnelle 2003	92	98	-0.48 (0.25)		0.62 [0.38, 1.01]
2 Sunlight exposure + calcium vs usual care					
Sambrook 2012	207	205	0.03 (0.1)		1.03 [0.85, 1.25]



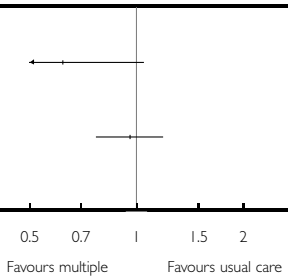
Analysis 11.2. Comparison 11 Care facilities: Multiple interventions vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 11 Care facilities: Multiple interventions vs usual care

Outcome: 2 Number of fallers

Study or subgroup	Multiple Intervention N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Exercise + management of urinary incontinence + fluid therapy vs usual care					
Schnelle 2003	92	98	-0.48 (0.27)		0.62 [0.36, 1.05]
2 Sunlight exposure + calcium vs usual care					
Sambrook 2012	207	205	-0.04 (0.11)		0.96 [0.77, 1.19]



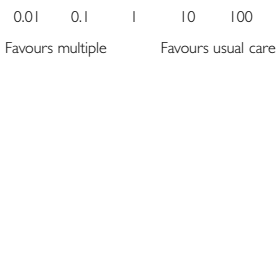
Analysis 11.3. Comparison 11 Care facilities: Multiple interventions vs usual care, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 11 Care facilities: Multiple interventions vs usual care

Outcome: 3 Number of people sustaining a fracture

Study or subgroup	Multiple Intervention N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Exercise + management of urinary incontinence + fluid therapy vs usual care					
Schnelle 2003	92	98	1.45 (1.11)		4.26 [0.48, 37.55]
2 Sunlight exposure + calcium vs usual care					
Sambrook 2012	207	205	-0.25 (0.39)		0.78 [0.36, 1.67]



Analysis 12.1. Comparison 12 Care facilities: Multifactorial interventions vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 12 Care facilities: Multifactorial interventions vs usual care

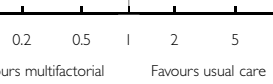
Outcome: 1 Rate of falls

Study or subgroup	Multifactorial N	Usual care N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
Becker 2003	509	472	-0.6 (0.15)		11.7 %	0.55 [0.41, 0.74]
Dyer 2004	102	94	-0.62 (0.13)		12.1 %	0.54 [0.42, 0.69]
Jensen 2002	188	196	-0.29 (0.2)		10.7 %	0.75 [0.51, 1.11]
Kerse 2004	309	238	0.29 (0.12)		12.3 %	1.34 [1.06, 1.69]
McMurdo 2000	52	38	-0.25 (0.24)		9.9 %	0.78 [0.49, 1.25]
Neyens 2009	249	269	-0.24 (0.31)		8.4 %	0.79 [0.43, 1.44]
Rubenstein 1990	79	81	-0.05 (0.1)		12.6 %	0.95 [0.78, 1.16]
Salv 2016	193	137	0.8 (0.23)		10.1 %	2.23 [1.42, 3.49]
Walker 2015	22	20	-0.73 (0.45)		6.0 %	0.48 [0.20, 1.16]
Whitney 2017	103	88	0.46 (0.44)		6.2 %	1.58 [0.67, 3.75]
Total (95% CI)	**1806**	**1633**			**100.0 %**	**0.88 [0.66, 1.18]**

Heterogeneity: $Tau^2 = 0.17$; $Chi^2 = 57.84$, df = 9 (P<0.00001); $I^2 = 84\%$

Test for overall effect: Z = 0.85 (P = 0.39)

Test for subgroup differences: Not applicable



Analysis 12.2. Comparison 12 Care facilities: Multifactorial interventions vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 12 Care facilities: Multifactorial interventions vs usual care

Outcome: 2 Number of fallers

Study or subgroup	Multifactorial N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Becker 2003	509	472	-0.29 (0.14)		13.4 %	0.75 [0.57, 0.98]
Dyer 2004	102	94	0.03 (0.28)		5.0 %	1.03 [0.60, 1.78]
Jensen 2002	188	196	-0.34 (0.14)		13.4 %	0.71 [0.54, 0.94]
Kerse 2004	309	238	0.25 (0.17)		10.6 %	1.28 [0.92, 1.79]
McMurdo 2000	52	38	-0.42 (0.3)		4.4 %	0.66 [0.36, 1.18]
Rubenstein 1990	79	81	-0.06 (0.1)		18.5 %	0.94 [0.77, 1.15]
Salv 2016	193	137	0.27 (0.22)		7.4 %	1.31 [0.85, 2.02]
Shaw 2003	130	144	-0.08 (0.07)		23.2 %	0.92 [0.80, 1.06]
Whitney 2017	103	88	0.09 (0.32)		4.0 %	1.09 [0.58, 2.05]
Total (95% CI)	**1665**	**1488**			**100.0 %**	**0.92 [0.81, 1.05]**

Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 13.73$, df = 8 (P = 0.09); $I^2 = 42\%$

Test for overall effect: Z = 1.19 (P = 0.23)

Test for subgroup differences: Not applicable

Analysis 12.3. Comparison 12 Care facilities: Multifactorial interventions vs usual care, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 12 Care facilities: Multifactorial interventions vs usual care

Outcome: 3 Number of people sustaining a fracture

Study or subgroup	Multifactorial N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Becker 2003 (1)	509	472	-0.07 (0.81)		20.8 %	0.93 [0.19, 4.56]
Jensen 2002 (2)	188	196	-1.47 (0.7)		24.2 %	0.23 [0.06, 0.91]
Salv 2016 (3)	193	137	1.49 (1.09)		14.2 %	4.44 [0.52, 37.58]
Shaw 2003 (4)	130	144	-0.6 (0.49)		32.4 %	0.55 [0.21, 1.43]
Whitney 2017 (5)	103	88	1.45 (1.54)		8.4 %	4.26 [0.21, 87.21]
Total (95% CI)	**1123**	**1037**			**100.0 %**	**0.79 [0.30, 2.07]**

Heterogeneity: $Tau^2 = 0.50$; $Chi^2 = 7.12$, df = 4 (P = 0.13); $I^2 = 44\%$

Test for overall effect: Z = 0.48 (P = 0.63)

Test for subgroup differences: Not applicable

(1) Hip fracture

(2) Hip fracture

(3) Total fractures

(4) Hip fracture

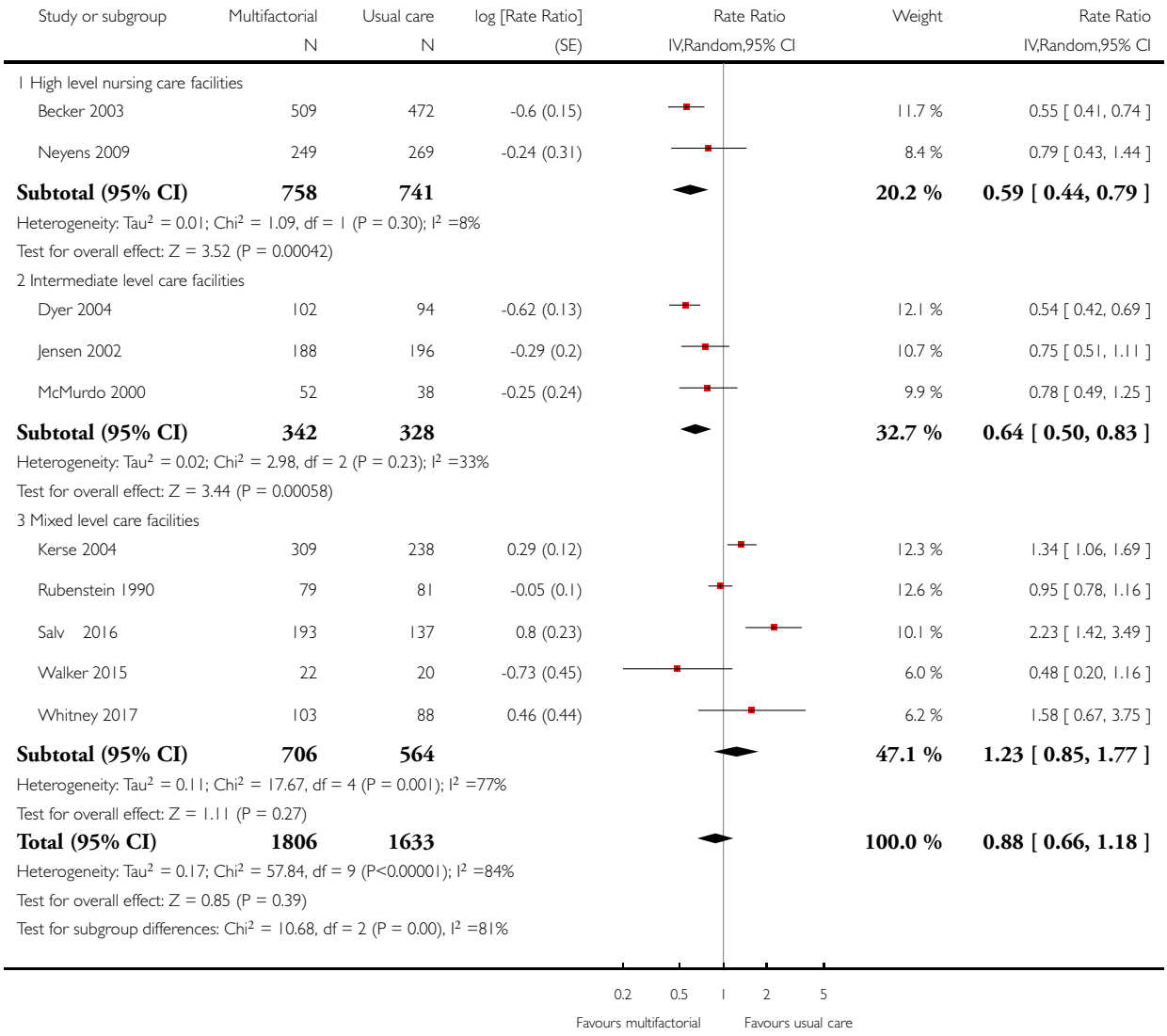
(5) Total fractures

Analysis 13.1. Comparison 13 Care facilities: Multifactorial interventions vs usual care (grouped by level of care), Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 13 Care facilities: Multifactorial interventions vs usual care (grouped by level of care)

Outcome: 1 Rate of falls



Analysis 13.2. Comparison 13 Care facilities: Multifactorial interventions vs usual care (grouped by level of care), Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 13 Care facilities: Multifactorial interventions vs usual care (grouped by level of care)

Outcome: 2 Number of fallers

Study or subgroup	Multifactorial N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 High level nursing care facilities						
Becker 2003	509	472	-0.29 (0.14)		100.0 %	0.75 [0.57, 0.98]
Subtotal (95% CI)	**509**	**472**			**100.0 %**	**0.75 [0.57, 0.98]**
Heterogeneity: not applicable						
Test for overall effect: Z = 2.07 (P = 0.038)						
2 Intermediate level care facilities						
Dyer 2004	102	94	0.03 (0.28)		17.0 %	1.03 [0.60, 1.78]
Jensen 2002	188	196	-0.34 (0.14)		68.1 %	0.71 [0.54, 0.94]
McMurdo 2000	52	38	-0.42 (0.3)		14.8 %	0.66 [0.36, 1.18]
Subtotal (95% CI)	**342**	**328**			**100.0 %**	**0.75 [0.60, 0.94]**
Heterogeneity: Tau^2 = 0.0; Chi^2 = 1.62, df = 2 (P = 0.44); I^2 =0.0%						
Test for overall effect: Z = 2.50 (P = 0.012)						
3 Mixed level care facilities						
Kerse 2004	309	238	0.25 (0.17)		13.4 %	1.28 [0.92, 1.79]
Rubenstein 1990	79	81	-0.06 (0.1)		29.7 %	0.94 [0.77, 1.15]
Salv 2016	193	137	0.27 (0.22)		8.6 %	1.31 [0.85, 2.02]
Shaw 2003	130	144	-0.08 (0.07)		44.0 %	0.92 [0.80, 1.06]
Whitney 2017	103	88	0.09 (0.32)		4.3 %	1.09 [0.58, 2.05]
Subtotal (95% CI)	**814**	**688**			**100.0 %**	**1.01 [0.88, 1.15]**
Heterogeneity: Tau^2 = 0.01; Chi^2 = 5.27, df = 4 (P = 0.26); I^2 =24%						
Test for overall effect: Z = 0.11 (P = 0.91)						
Test for subgroup differences: Chi^2 = 7.01, df = 2 (P = 0.03), I^2 =71%						

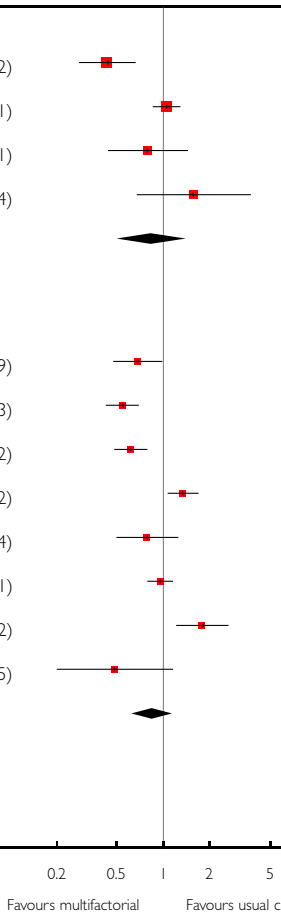
Analysis 14.1. Comparison 14 Care facilities: Multifactorial interventions vs usual care (grouped by level of cognition), Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 14 Care facilities: Multifactorial interventions vs usual care (grouped by level of cognition)

Outcome: 1 Rate of falls

Study or subgroup	Multifactorial N	Usual care N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
1 Participants with cognitive impairment						
Becker 2003 (1)	150	169	-0.84 (0.22)		27.4 %	0.43 [0.28, 0.66]
Jensen 2002 (2)	69	102	0.05 (0.11)		31.9 %	1.05 [0.85, 1.30]
Neyens 2009 (3)	249	269	-0.24 (0.31)		23.2 %	0.79 [0.43, 1.44]
Whitney 2017 (4)	103	88	0.46 (0.44)		17.6 %	1.58 [0.67, 3.75]
Subtotal (95% CI)	**571**	**628**			**100.0 %**	**0.83 [0.49, 1.40]**
Heterogeneity: Tau^2 = 0.21; Chi^2 = 14.98, df = 3 (P = 0.002); I^2 =80%						
Test for overall effect: Z = 0.71 (P = 0.48)						
2 Participants with no cognitive impairment or mixed sample						
Becker 2003 (5)	215	191	-0.39 (0.19)		12.6 %	0.68 [0.47, 0.98]
Dyer 2004	102	94	-0.62 (0.13)		14.0 %	0.54 [0.42, 0.69]
Jensen 2002 (6)	112	79	-0.49 (0.12)		14.3 %	0.61 [0.48, 0.78]
Kerse 2004	309	238	0.29 (0.12)		14.3 %	1.34 [1.06, 1.69]
McMurdo 2000	52	38	-0.25 (0.24)		11.3 %	0.78 [0.49, 1.25]
Rubenstein 1990	79	81	-0.05 (0.1)		14.6 %	0.95 [0.78, 1.16]
Salv 2016 (7)	96	77	0.59 (0.2)		12.3 %	1.80 [1.22, 2.67]
Walker 2015	22	20	-0.73 (0.45)		6.6 %	0.48 [0.20, 1.16]
Subtotal (95% CI)	**987**	**818**			**100.0 %**	**0.84 [0.62, 1.13]**
Heterogeneity: Tau^2 = 0.15; Chi^2 = 52.63, df = 7 (P<0.00001); I^2 =87%						
Test for overall effect: Z = 1.15 (P = 0.25)						
Test for subgroup differences: Chi^2 = 0.00, df = 1 (P = 0.97), I^2 =0.0%						



(1) At least one sign of cognitive impairment or depression based on Minimum Data Set of the Resident Assessment Instrument (MDS RAI 2.0)

(2) Subgroup with MMSE score <19

(3) Psychogeriatric patients

(4) 97% Addenbrooke' s Cognitive Examination (ACE-R) score <80

(5) No sign of cognitive impairment or depression based on Minimum Data Set of the Resident Assessment Instrument (MDS RAI 2.0)

(6) Subgroup with MMSE score ≥19

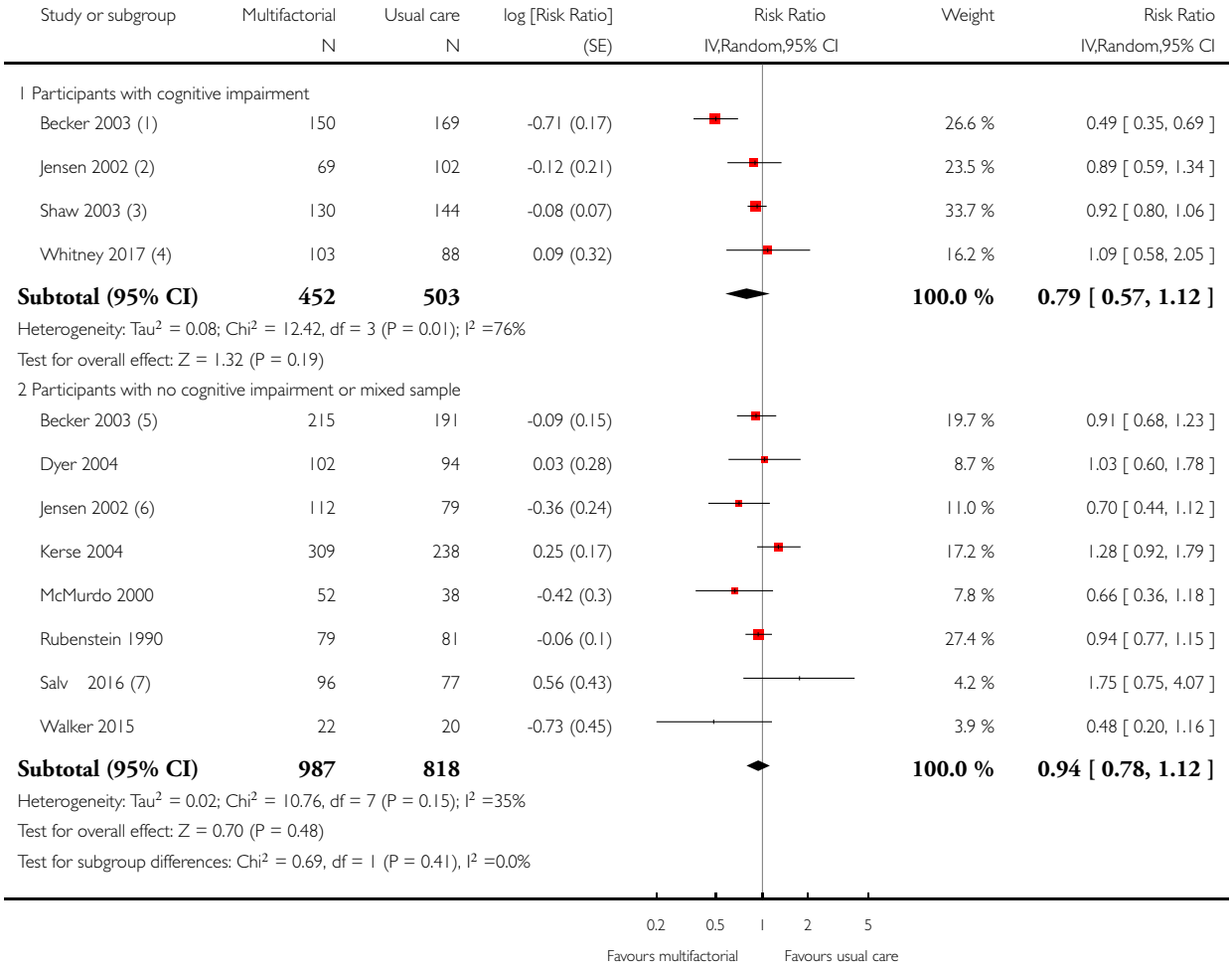
(7) Higher cognition subgroup (excluding those with dementia)

Analysis 14.2. Comparison 14 Care facilities: Multifactorial interventions vs usual care (grouped by level of cognition), Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 14 Care facilities: Multifactorial interventions vs usual care (grouped by level of cognition)

Outcome: 2 Number of fallers



(1) At least one sign of cognitive impairment or depression based on Minimum Data Set of the Resident Assessment Instrument (MDS RAI 2.0)

(2) Subgroup with MMSE score <19

(3) All participants had an MMSE score <24

(4) 97% Addenbrooke' s Cognitive Examination (ACE-R) score <80

(5) No sign of cognitive impairment or depression based on Minimum Data Set of the Resident Assessment Instrument (MDS RAI 2.0)

(6) Subgroup with MMSE score ≥19

(7) Higher cognition subgroup (excluding those with dementia)

Analysis 15.1. Comparison 15 Hospitals: Additional exercises vs usual physiotherapy, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 15 Hospitals: Additional exercises vs usual physiotherapy

Outcome: 1 Rate of falls

Study or subgroup	Additional exercise N	Usual physiotherapy N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Weight	Rate Ratio IV,Fixed,95% CI
Donald 2000 (1)	30	24	-0.62 (0.62)		45.8 %	0.54 [0.16, 1.81]
Treacy 2015	80	81	-0.45 (0.57)		54.2 %	0.64 [0.21, 1.95]
Total (95% CI)	**110**	**105**			**100.0 %**	**0.59 [0.26, 1.34]**

Heterogeneity: $Chi^2 = 0.04$, df = 1 (P = 0.84); $I^2 = 0.0\%$

Test for overall effect: Z = 1.26 (P = 0.21)

Test for subgroup differences: Not applicable

(1) Factorial design: additional exercises with carpet or vinyl flooring vs conventional physiotherapy with carpet or vinyl flooring

Analysis 15.2. Comparison 15 Hospitals: Additional exercises vs usual physiotherapy, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 15 Hospitals: Additional exercises vs usual physiotherapy

Outcome: 2 Number of fallers

Study or subgroup	Additional exercise N	Usual physiotherapy N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Weight	Risk Ratio IV,Fixed,95% CI
Donald 2000 (1)	30	24	-1.56 (0.87)		30.8 %	0.21 [0.04, 1.16]
Jarvis 2007	14	15	-0.78 (0.58)		69.2 %	0.46 [0.15, 1.43]
Total (95% CI)	**44**	**39**			**100.0 %**	**0.36 [0.14, 0.93]**

Heterogeneity: $Chi^2 = 0.56$, df = 1 (P = 0.46); $I^2 = 0.0\%$

Test for overall effect: Z = 2.11 (P = 0.035)

Test for subgroup differences: Not applicable

(1) Factorial design: additional exercises with carpet or vinyl flooring vs conventional physiotherapy with carpet or vinyl flooring

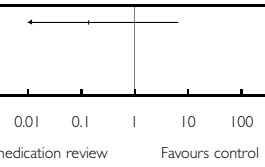
Analysis 16.1. Comparison 16 Hospitals: Medication review vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 16 Hospitals: Medication review vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Medication review N	Control N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
Michalek 2014	58	56	-1.97 (1.97)		0.14 [0.00, 6.63]



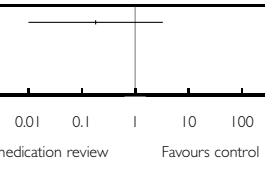
Analysis 16.2. Comparison 16 Hospitals: Medication review vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 16 Hospitals: Medication review vs usual care

Outcome: 2 Number of fallers

Study or subgroup	Medication review N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
Michalek 2014	58	56	-1.71 (1.49)		0.18 [0.01, 3.35]



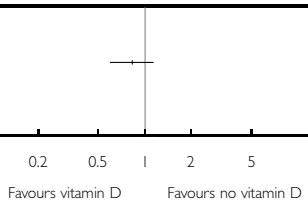
Analysis 17.1. Comparison 17 Hospitals: Vitamin D supplements vs no vitamin D supplements, Outcome 1 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 17 Hospitals: Vitamin D supplements vs no vitamin D supplements

Outcome: 1 Number of fallers

Study or subgroup	Vitamin D N	No vitamin N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Vitamin D + calcium vs calcium					
Burleigh 2007	100	103	-0.2 (0.17)		0.82 [0.59, 1.14]



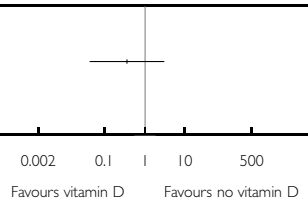
Analysis 17.2. Comparison 17 Hospitals: Vitamin D supplements vs no vitamin D supplements, Outcome 2 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 17 Hospitals: Vitamin D supplements vs no vitamin D supplements

Outcome: 2 Number of people sustaining a fracture

Study or subgroup	Vitamin D N	No vitamin N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Vitamin D + calcium vs calcium					
Burleigh 2007	100	103	-1.08 (1.12)		0.34 [0.04, 3.05]

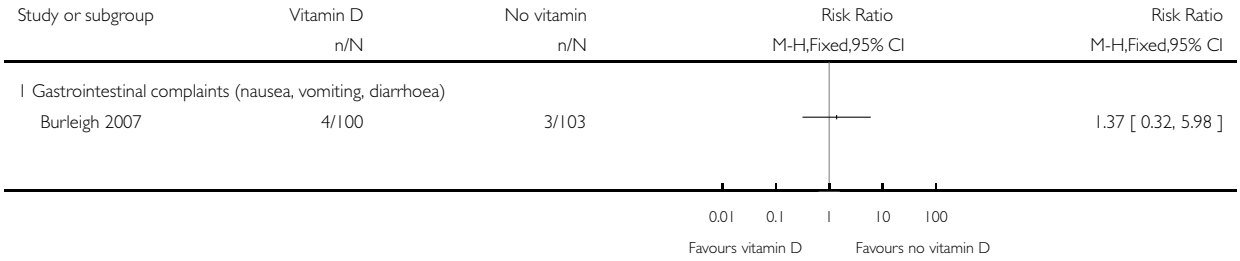


Analysis 17.3. Comparison 17 Hospitals: Vitamin D supplements vs no vitamin D supplements, Outcome 3 Adverse events.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 17 Hospitals: Vitamin D supplements vs no vitamin D supplements

Outcome: 3 Adverse events

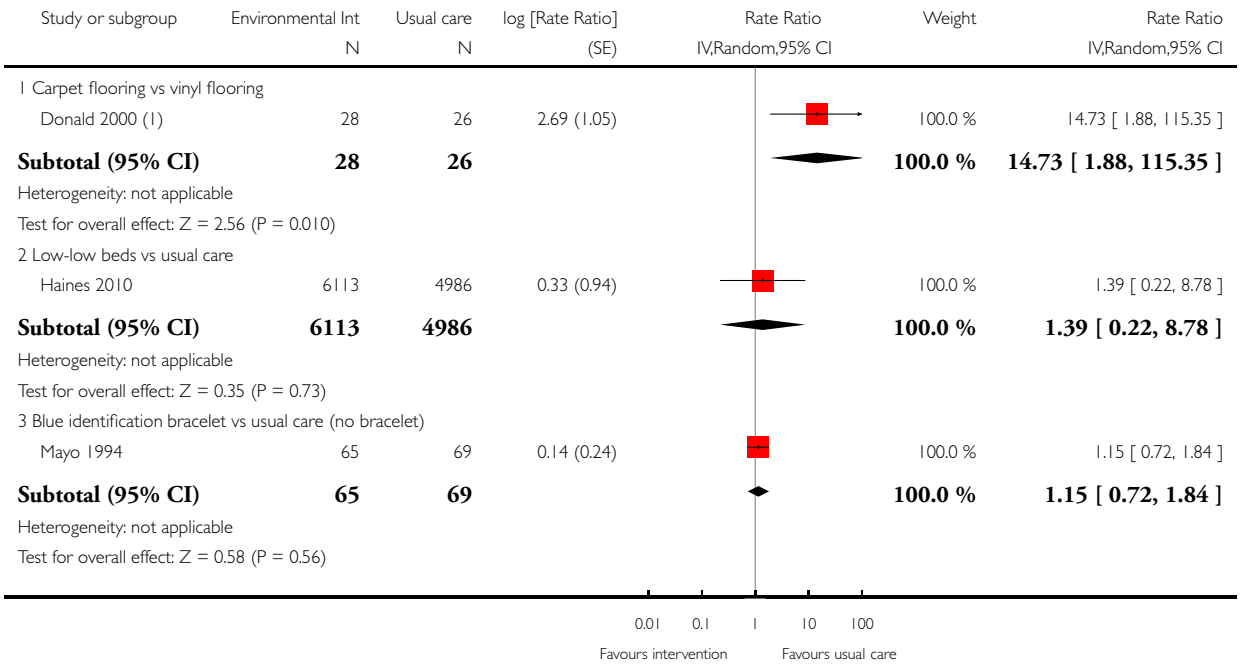


Analysis 18.1. Comparison 18 Hospitals: Environmental interventions vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 18 Hospitals: Environmental interventions vs usual care

Outcome: 1 Rate of falls



(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Environmental Int N	Usual care N	log [Rate Ratio] (SE)	Rate Ratio IV,Random,95% CI	Weight	Rate Ratio IV,Random,95% CI
4 Bed alarms vs usual care						
Shorr 2012 (2)	11115	17436	0.09 (1.02)		16.3 %	1.09 [0.15, 8.08]
Wolf 2013 (3)	48	50	-0.63 (0.45)		83.7 %	0.53 [0.22, 1.29]
Subtotal (95% CI)	**11163**	**17486**			**100.0 %**	**0.60 [0.27, 1.34]**
Heterogeneity: Tau² = 0.0; Chi² = 0.42, df = 1 (P = 0.52); I² =0.0%						
Test for overall effect: Z = 1.25 (P = 0.21)						

(1) Factorial design: carpet flooring with or without additional exercises vs vinyl flooring with or without additional exercises

(2) Education and support on use of sensor pads applied to the bed, chair or commode

(3) Sensor alarm fitted to patients upper leg at rest time

Analysis 18.2. Comparison 18 Hospitals: Environmental interventions vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 18 Hospitals: Environmental interventions vs usual care

Outcome: 2 Number of fallers

Study or subgroup	Environmental Int N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Weight	Risk Ratio IV,Fixed,95% CI
1 Carpet flooring vs vinyl flooring						
Donald 2000	28	26	2.12 (1.11)		100.0 %	8.33 [0.95, 73.37]
Subtotal (95% CI)	**28**	**26**			**100.0 %**	**8.33 [0.95, 73.37]**
Heterogeneity: not applicable						
Test for overall effect: Z = 1.91 (P = 0.056)						
2 Blue identification bracelet vs usual care (no bracelet)						
Mayo 1994 (1)	65	69	0.29 (0.29)		100.0 %	1.34 [0.76, 2.36]
Subtotal (95% CI)	**65**	**69**			**100.0 %**	**1.34 [0.76, 2.36]**
Heterogeneity: not applicable						
Test for overall effect: Z = 1.00 (P = 0.32)						
3 Bed alarms vs usual care						
Shorr 2012 (2)	11115	17436	0.05 (0.9)		25.0 %	1.05 [0.18, 6.13]
Wolf 2013 (3)	48	50	-0.12 (0.52)		75.0 %	0.89 [0.32, 2.46]

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Environmental Int N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Weight	Risk Ratio IV,Fixed,95% CI
Subtotal (95% CI)	**11163**	**17486**			**100.0 %**	**0.93 [0.38, 2.24]**
Heterogeneity: Chi^2 = 0.03, df = 1 (P = 0.87); I^2 =0.0%						
Test for overall effect: Z = 0.17 (P = 0.86)						

(1) Blue identification bracelet vs usual care (no bracelet)

(2) Education and support on bed alarm use

(3) Bec and chair sensor alarm

Analysis 19.1. Comparison 19 Hospitals: Social environment vs control, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 19 Hospitals: Social environment vs control

Outcome: 1 Rate of falls

Study or subgroup	Social environment N	Control N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Organisational service model change (fall prevention guideline implementation)					
Koh 2009 (1)	612	510	0.6 (1.06)		1.82 [0.23, 14.55]
2 Organisation service model change (falls prevention, incontinence and ulcer guideline implementation)					
Van Gaal 2011b (2)	1081	1120	-0.4 (0.69)		0.67 [0.17, 2.59]
3 Organisational service model change (fall prevention toolkit software)					
Dykes 2010 (3)	2755	2509	-0.6 (1.73)		0.55 [0.02, 16.29]
4 Acute care service for elderly patients vs usual care					
Wald 2011	122	95	-0.33 (1)		0.72 [0.10, 5.10]
5 Post-operative orthogeriatric service after hip fracture					
Stenvall 2007 (4)	102	97	-0.97 (0.34)		0.38 [0.19, 0.74]

(1) Multifaceted fall prevention guideline implementation vs routine dissemination

(2) Guideline implementation (falls, urinary tract infection, pressure ulcers) programme vs control

(3) Fall prevention tool kit software vs usual care

(4) Acute care: unit specialising in geriatric orthopaedic care versus conventional orthopaedic care after proximal femoral fracture surgery

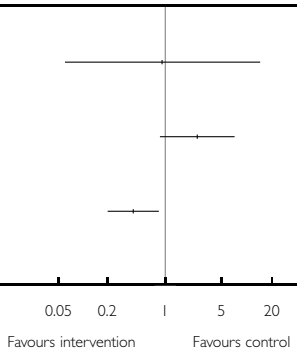
Analysis 19.2. Comparison 19 Hospitals: Social environment vs control, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 19 Hospitals: Social environment vs control

Outcome: 2 Number of fallers

Study or subgroup	Social environment N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Fall prevention tool kit software vs usual care					
Dykes 2010	2755	2509	-0.09 (1.4)		0.91 [0.06, 14.21]
2 Behaviour advisory service vs usual care					
Mador 2004	36	35	0.89 (0.54)		2.44 [0.85, 7.02]
3 Post-operative orthogeriatric service after hip fracture					
Stenvall 2007 (1)	102	97	-0.89 (0.36)		0.41 [0.20, 0.83]



(1) Acute care: unit specialising in geriatric orthopaedic care versus conventional orthopaedic care after proximal femoral fracture surgery

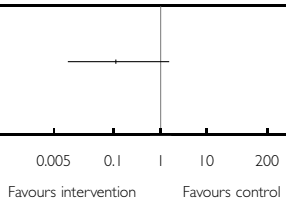
Analysis 19.3. Comparison 19 Hospitals: Social environment vs control, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 19 Hospitals: Social environment vs control

Outcome: 3 Number of people sustaining a fracture

Study or subgroup	Social environment N	Control N	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Risk Ratio IV,Fixed,95% CI
1 Post-operative orthogeriatric service after hip fracture					
Stenvall 2007 (1)	102	97	-2.21 (1.34)		0.11 [0.01, 1.52]



(1) Acute care: unit specialising in geriatric orthopaedic care versus conventional orthopaedic care after proximal femoral fracture surgery

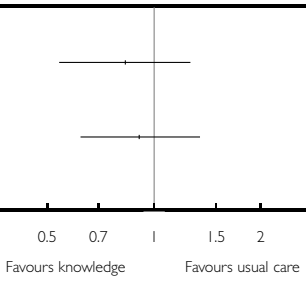
Analysis 20.1. Comparison 20 Hospitals: Knowledge/education interventions vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 20 Hospitals: Knowledge/education interventions vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Knowledge/education N	Usual care N	log [Rate Ratio] (SE)	Rate Ratio IV,Fixed,95% CI	Rate Ratio IV,Fixed,95% CI
1 Educational materials + health professional follow-up vs usual care					
Haines 2011	401	381	-0.19 (0.22)		0.83 [0.54, 1.27]
2 Educational materials only vs usual care					
Haines 2011	424	381	-0.09 (0.2)		0.91 [0.62, 1.35]



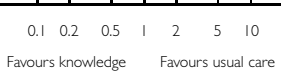
Analysis 20.2. Comparison 20 Hospitals: Knowledge/education interventions vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 20 Hospitals: Knowledge/education interventions vs usual care

Outcome: 2 Number of fallers

Study or subgroup	Knowledge/education N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Risk Ratio IV,Random,95% CI
1 Individualised educational session vs usual care					
Ang 2011	910	912	-1.24 (0.48)		0.29 [0.11, 0.74]
2 Educational materials + health professional follow-up vs usual care					
Haines 2011	401	381	-0.3 (0.22)		0.74 [0.48, 1.14]
3 Educational materials only vs usual care					
Haines 2011	424	381	-0.17 (0.21)		0.84 [0.56, 1.27]



Analysis 21.1. Comparison 21 Hospitals: Multifactorial interventions vs usual care, Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 21 Hospitals: Multifactorial interventions vs usual care

Outcome: 1 Rate of falls

Study or subgroup	Multifactorial N	Usual care N	log [Rate ratio] (SE)	Rate ratio IV,Random,95% CI	Weight	Rate ratio IV,Random,95% CI
Barker 2016 (1)	17698	17566	0.04 (0.14)		25.4 %	1.04 [0.79, 1.37]
Cumming 2008 (2)	2047	1952	-0.04 (0.15)		24.1 %	0.96 [0.72, 1.29]
Haines 2004 (3)	310	316	-0.36 (0.13)		26.8 %	0.70 [0.54, 0.90]
Healey 2004 (4)	749	905	-0.53 (0.42)		6.3 %	0.59 [0.26, 1.34]
Hill 2015 (5)	1402	1719	-0.51 (0.21)		17.3 %	0.60 [0.40, 0.91]
Total (95% CI)	**22206**	**22458**			**100.0 %**	**0.80 [0.64, 1.01]**

Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 8.38$, df = 4 (P = 0.08); $I^2 = 52\%$

Test for overall effect: Z = 1.90 (P = 0.058)

Test for subgroup differences: Not applicable

(1) Acute care: risk assessment and up to 6 interventions for high risk patients, plus staff education vs usual care

(2) Acute and subacute care: risk assessment, staff and patient education, drug review, environmental modifications, exercise vs usual care

(3) Subacute: risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) vs usual care

(4) Acute and subacute care: risk factor screening and targeted care plan in at-risk patients vs usual care

(5) Subacute care: Multimedia falls education with follow-up for patients plus staff education and feedback.

Analysis 21.2. Comparison 21 Hospitals: Multifactorial interventions vs usual care, Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 21 Hospitals: Multifactorial interventions vs usual care

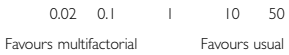
Outcome: 2 Number of fallers

Study or subgroup	Multifactorial N	Usual care N	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
Barker 2016 (1)	17698	17566	-0.01 (0.565)		6.5 %	0.99 [0.33, 3.00]
Cumming 2008 (2)	2047	1952	0.04 (0.4)		12.9 %	1.04 [0.48, 2.28]
Haines 2004 (3)	310	316	-0.25 (0.16)		80.6 %	0.78 [0.57, 1.07]
Total (95% CI)	**20055**	**19834**			**100.0 %**	**0.82 [0.62, 1.09]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.57$, df = 2 (P = 0.75); $I^2 = 0.0\%$

Test for overall effect: Z = 1.37 (P = 0.17)

Test for subgroup differences: Not applicable



(1) Acute care: risk assessment and up to 6 interventions for high risk patients, plus staff education vs usual care

(2) Acute and subacute care: risk assessment, staff and patient education, drug review, environmental modifications, exercise vs usual care

(3) Subacute: risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) vs usual care

Analysis 21.3. Comparison 21 Hospitals: Multifactorial interventions vs usual care, Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 21 Hospitals: Multifactorial interventions vs usual care

Outcome: 3 Number of people sustaining a fracture

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Fixed,95% CI	Weight	Risk Ratio IV,Fixed,95% CI
Cumming 2008 (1)	-1.14 (1.7)		25.7 %	0.32 [0.01, 8.95]
Haines 2004 (2)	0.02 (1)		74.3 %	1.02 [0.14, 7.24]
Total (95% CI)			**100.0 %**	**0.76 [0.14, 4.10]**

Heterogeneity: $Chi^2 = 0.35$, df = 1 (P = 0.56); $I^2 = 0.0\%$

Test for overall effect: Z = 0.32 (P = 0.75)

Test for subgroup differences: Not applicable

(1) Acute and subacute care: risk assessment, staff and patient education, drug review, environmental modifications, exercise vs usual care

(2) Subacute: risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) vs usual care

Analysis 22.1. Comparison 22 Hospitals: Multifactorial interventions vs usual care (grouped by type of care), Outcome 1 Rate of falls.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 22 Hospitals: Multifactorial interventions vs usual care (grouped by type of care)

Outcome: 1 Rate of falls

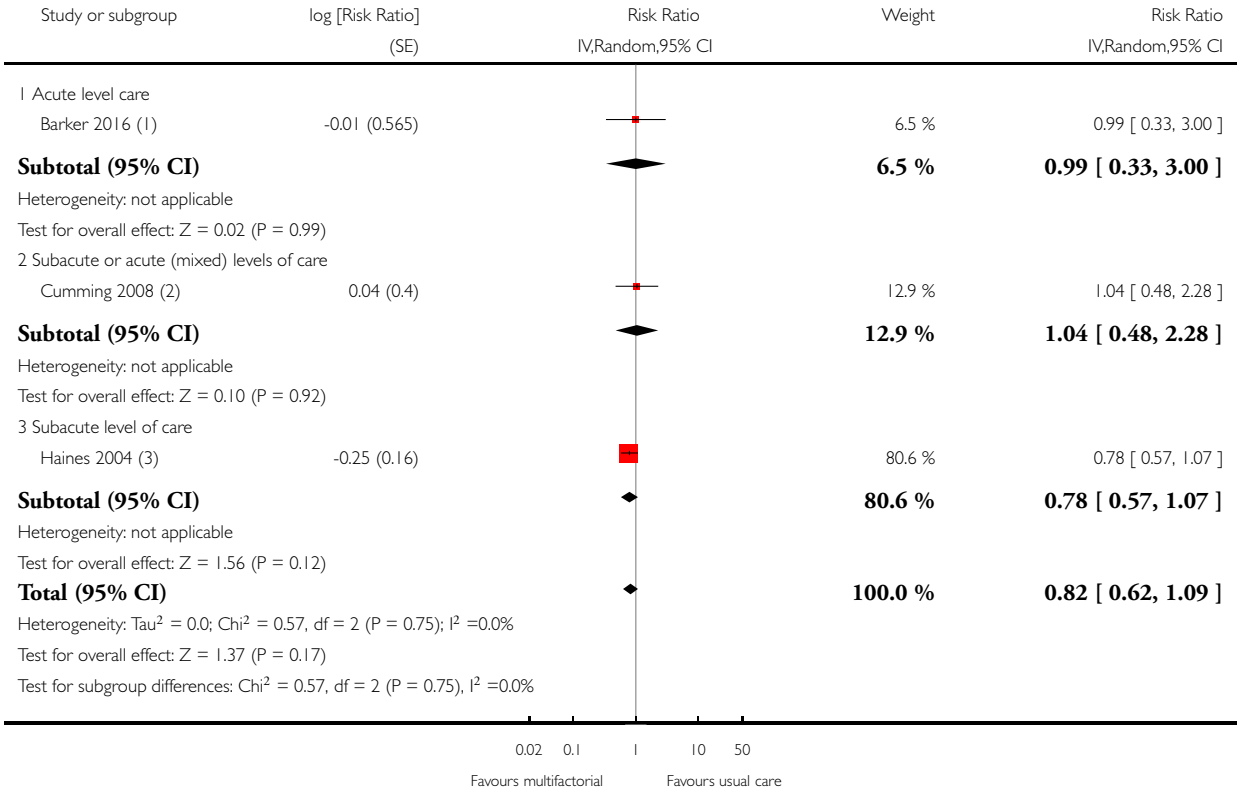
Study or subgroup	Multifactorial N	Usual care N	log [Rate Ratio] (SE)	Rate Ratio IV,Random,95% CI	Weight	Rate Ratio IV,Random,95% CI
1 Acute level of care						
Barker 2016	17698	17566	0.04 (0.14)		25.4 %	1.04 [0.79, 1.37]
Subtotal (95% CI)	**17698**	**17566**			**25.4 %**	**1.04 [0.79, 1.37]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.29 (P = 0.78)						
2 Subacute or acute (mixed) levels of care						
Cumming 2008	2047	1952	-0.04 (0.15)		24.1 %	0.96 [0.72, 1.29]
Healey 2004	749	905	-0.53 (0.42)		6.3 %	0.59 [0.26, 1.34]
Subtotal (95% CI)	**2796**	**2857**			**30.4 %**	**0.88 [0.61, 1.27]**
Heterogeneity: Tau^2 = 0.02; Chi^2 = 1.21, df = 1 (P = 0.27); I^2 =17%						
Test for overall effect: Z = 0.68 (P = 0.50)						
3 Subacute level of care						
Haines 2004	310	316	-0.36 (0.13)		26.8 %	0.70 [0.54, 0.90]
Hill 2015	1402	1719	-0.51 (0.21)		17.3 %	0.60 [0.40, 0.91]
Subtotal (95% CI)	**1712**	**2035**			**44.1 %**	**0.67 [0.54, 0.83]**
Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.37, df = 1 (P = 0.54); I^2 =0.0%						
Test for overall effect: Z = 3.63 (P = 0.00028)						
Total (95% CI)	**22206**	**22458**			**100.0 %**	**0.80 [0.64, 1.01]**
Heterogeneity: Tau^2 = 0.03; Chi^2 = 8.38, df = 4 (P = 0.08); I^2 =52%						
Test for overall effect: Z = 1.90 (P = 0.058)						
Test for subgroup differences: Chi^2 = 6.38, df = 2 (P = 0.04), I^2 =69%						

Analysis 22.2. Comparison 22 Hospitals: Multifactorial interventions vs usual care (grouped by type of care), Outcome 2 Number of fallers.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 22 Hospitals: Multifactorial interventions vs usual care (grouped by type of care)

Outcome: 2 Number of fallers



(1) Acute care: risk assessment and up to 6 interventions for high risk patients, plus staff education vs usual care

(2) Acute and subacute care: risk assessment, staff and patient education, drug review, environmental modifications, exercise vs usual care

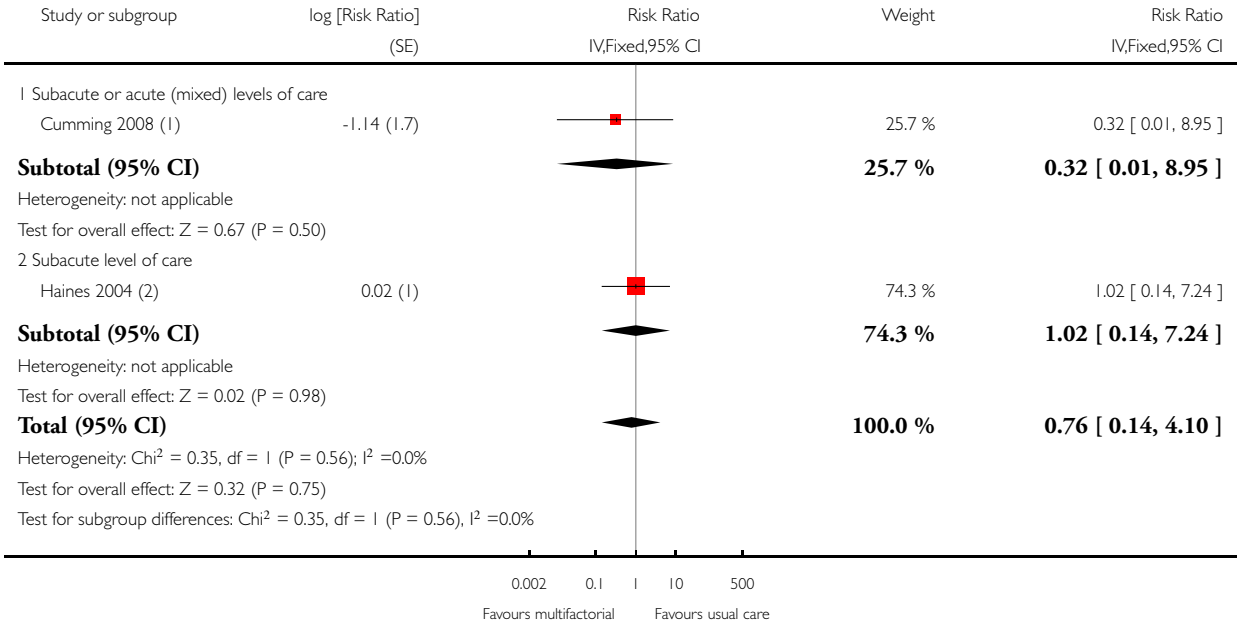
(3) Subacute: risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) vs usual care

Analysis 22.3. Comparison 22 Hospitals: Multifactorial interventions vs usual care (grouped by type of care), Outcome 3 Number of people sustaining a fracture.

Review: Interventions for preventing falls in older people in care facilities and hospitals

Comparison: 22 Hospitals: Multifactorial interventions vs usual care (grouped by type of care)

Outcome: 3 Number of people sustaining a fracture



(1) Acute and subacute care: risk assessment, staff and patient education, drug review, environmental modifications, exercise vs usual care

(2) Subacute: risk assessment and targeted interventions (exercise, educational sessions from OT, hip protectors) vs usual care

ADDITIONAL TABLES

Table 1. Description of included studies: reference links

Study description	Links to references
Additional studies included in this update	**Care facilities** N = 28: Beck 2016; Buckinx 2014; Cadore 2014; Colon-Emeric 2013; da Silva Borges 2014; Houghton 2014; Frankenthal 2014; Fu 2015; Garcia Gollarte 2014; Huang 2016; Imaoka 2016; Irez 2011; Juola 2015; Kennedy 2015; Kovacs 2012; Kovacs 2013; Peyro Saint Paul 2013; Potter 2016; Salvà 2016; Saravanakumar 2014; Sitja Rabert 2015; Streim 2012; Tuunainen 2013; Van de Ven 2014; Van het Reve 2014; Walker 2015; Whitney 2017; Yokoi 2015 **Hospitals** N = 7: Aizen 2015; Barker 2016; Hill 2015; Michalek 2014; Shorr 2012; Treacy 2015; Wolf 2013

Table 1. Description of included studies: reference links (*Continued*)

Design	**Cluster randomised** N = 42: Aizen 2015; Barker 2016; Beck 2016; Becker 2003; Chenoweth 2009; Choi 2005; Colon-Emeric 2013; Cox 2008; Crotty 2004b; Cumming 2008; Dyer 2004; Dykes 2010; Garcia Gollarte 2014; Haines 2010; Healey 2004; Hill 2015; Houghton 2014; Jensen 2002; Juola 2015; Kennedy 2015; Kerse 2004; Kerse 2008; Koh 2009; Lapane 2011; Law 2006; McMurdo 2000; Meyer 2009; Michalek 2014; Neyens 2009; Patterson 2010; Ray 1997; Rosendahl 2008; Salvà 2016; Sambrook 2012; Shorr 2012; Van de Ven 2014; Van Gaal 2011a; Van Gaal 2011b; Walker 2015; Ward 2010; Whitney 2017; Yokoi 2015
Setting (country)	**Australia** (N = 17): Barker 2016; Chenoweth 2009; Crotty 2004a; Crotty 2004b; Cumming 2008; Flicker 2005; Grieger 2009; Haines 2004; Haines 2010; Haines 2011; Hill 2015; Mador 2004; Potter 2016; Sambrook 2012; Saravanakumar 2014; Treacy 2015; Ward 2010 **Belgium** (N = 1): Buckinx 2014 **Brazil** (N = 1): da Silva Borges 2014 **Denmark** (N = 1): Beck 2016 **Canada** (N = 3): Kennedy 2015; Klages 2011; Mayo 1994 **China** (N = 1): Fu 2015 **Finland** (N = 3): Juola 2015; Sihvonen 2004; Tuunainen 2013 **France** (N = 3): Chapuy 2002; Peyro Saint Paul 2013; Toulotte 2003 **Germany** (N = 4): Becker 2003; Meyer 2009; Michalek 2014; Wolf 2013 **Hungary** (N = 2): Kovacs 2012; Kovacs 2013 **Israel** (N = 2): Aizen 2015; Frankenthal 2014 **Korea** (N = 1): Choi 2005 **Japan** (N = 5): Imaoka 2016; Sakamoto 2006; Sakamoto 2012; Shimada 2004; Yokoi 2015 **The Netherlands** (N = 5): Faber 2006; Neyens 2009; Van de Ven 2014; Van Gaal 2011a; Van Gaal 2011b **New Zealand** (N = 2): Kerse 2004; Kerse 2008 **Singapore** (N = 2): Ang 2011; Koh 2009 **Spain** (N = 5): Cadore 2014; Garcia Gollarte 2014; Salvà 2016; Serra-Rexach 2011; Sitja Rabert 2015 **Sweden** (N = 3): Jensen 2002; Rosendahl 2008; Stenvall 2007 **Switzerland** (N = 2): Bischoff 2003; Van het Reve 2014 **Taiwan** (N = 1): Huang 2016 **Turkey** (N = 1): Irez 2011 **United Kingdom** (N = 14): Burleigh 2007; Cox 2008; Houghton 2014; Donald 2000; Dyer 2004; Healey 2004; Jarvis 2007; Law 2006; McMurdo 2000; Patterson 2010; Shaw 2003; Walker 2015; Whitney 2017; Zermansky 2006 **USA** (N = 16): Broe 2007; Buettner 2002; Clifton 2009; Colon-Emeric 2013; Dykes 2010; Lapane 2011; Mulrow 1994; Nowalk 2001; Ray 1997; Rubenstein 1990; Schnelle 2003; Schoenfelder 2000; Shorr 2012; Streim 2012; Tideiksaar 1993; Wald 2011
Setting	**Care facilities** N = 71 ***High level nursing care*** N = 17: Beck 2016; Becker 2003; Bischoff 2003; Broe 2007; Chenoweth 2009; Clifton 2009; Crotty 2004a; Fu 2015; Imaoka 2016; Meyer 2009; Mulrow 1994; Neyens 2009; Ray 1997; Schnelle 2003; Schoenfelder 2000; Van de Ven 2014; Van Gaal 2011a;

Table 1. Description of included studies: reference links (*Continued*)

	Intermediate level care N = 17: Buckinx 2014; Chapuy 2002; Choi 2005; da Silva Borges 2014; Dyer 2004; Irez 2011; Jensen 2002; Kerse 2008; Kovacs 2012; McMurdo 2000; Sakamoto 2006; Sakamoto 2012; Sambrook 2012; Serra-Rexach 2011; Sihvonen 2004; Van het Reve 2014; Yokoi 2015 ***Mixed levels of care*** N = 37: Buettner 2002; Cadore 2014; Colon-Emeric 2013; Cox 2008; Crotty 2004b; Houghton 2014; Faber 2006; Flicker 2005; Frankenthal 2014; Garcia Gollarte 2014; Grieger 2009; Huang 2016; Juola 2015; Kennedy 2015; Kerse 2004; Klages 2011; Kovacs 2013; Lapane 2011; Law 2006; Nowalk 2001; Patterson 2010; Peyro Saint Paul 2013; Potter 2016; Rosendahl 2008; Rubenstein 1990; Salvà 2016; Saravanakumar 2014; Shaw 2003; Shimada 2004; Sitja Rabert 2015; Streim 2012; Toulotte 2003; Tuunainen 2013; Walker 2015; Ward 2010; Whitney 2017; Zermansky 2006 **Hospitals** N = 24 ***Acute care*** N = 10: Ang 2011; Barker 2016; Dykes 2010; Koh 2009; Mador 2004; Shorr 2012; Stenvall 2007; Tideiksaar 1993; Van Gaal 2011b; Wald 2011 ***Subacute care*** N = 12: Aizen 2015; Burleigh 2007; Donald 2000; Haines 2004; Haines 2010; Healey 2004; Hill 2015; Jarvis 2007; Mayo 1994; Michalek 2014; Treacy 2015; Wolf 2013 ***Acute and subacute care*** N = 2: Cumming 2008; Haines 2011
Care facilities	**Exercises** N = 23: Buckinx 2014; Buettner 2002; Cadore 2014; Choi 2005; da Silva Borges 2014; Faber 2006; Fu 2015; Irez 2011; Kerse 2008; Kovacs 2012; Kovacs 2013; Mulrow 1994; Nowalk 2001; Rosendahl 2008; Sakamoto 2006; Saravanakumar 2014; Schoenfelder 2000; Serra-Rexach 2011; Shimada 2004; Sihvonen 2004; Sitja Rabert 2015; Toulotte 2003; Tuunainen 2013; Yokoi 2015

Table 2. Description of interventions in studies of exercise interventions in care facilities

Trial	Intervention	Control	Comment
Buckinx 2014	Whole body vibration exercise programme	Usual care: no change to lifestyle	
Buettner 2002	Supervised group exercises	Usual care	
Cadore 2014	Multicomponent exercises. Twice-weekly, 40-minute duration	Usual care: "mobility" exercises (30 minutes per day at least 4 days per week), small active and passive movements applied as stretches in a rhythmic fashion	
Choi 2005	Tai Chi	Usual care: routine activities, without participation in any regular exercise classes	
da Silva Borges 2014	Ballroom dancing	Usual care: agreed not to engage in any regular physical activity	

Table 2. Description of interventions in studies of exercise interventions in care facilities (*Continued*)

Faber 2006	• A functional balance, strength and mobility programme • 3D (balance)	Usual care: no change usual pattern of activity	Both comparisons of interventions vs usual care considered under exercise vs usual care. Comparisons of interventions arms considered under comparisons of different exercise categories
Fu 2015	Wii balance training (1 hour 3 x week)	Different exercise: Balance training (Otago) (1 hour, 3 x week)	
Imaoka 2016	Reduced exercise - individualised exercise only.	Different exercise: groups plus individualised exercises (described by study authors as usual care)	
Irez 2011	Combination exercises: Pilates	Usual care: no Pilates, instructed not to change current activity levels	
Kerse 2008	Activity programme	Usual care	
Kovacs 2012	Multimodal exercise - Otago Exercise programme	Different exercise: Osteoporosis exercise programme, includes balance and strengthening exercises	
Kovacs 2013	Multimodal exercise - Otago Exercise programme	Usual care: social activities such as board games, listening to music	
Mulrow 1994	Tailored exercises	Usual care: friendly visit, usually involved reading to participant, avoided physical activity	
Nowalk 2001	• Supervised exercise • Tai Chi Plus control (basic enhanced programme)	Usual care: basic enhanced programme including falls-prevention programme with 3 education sessions and a walking programme	Results for interventions vs usual care as reported by study authors presented in Analysis 1.2 as data not suitable for calculation of RaR or RR.
Rosendahl 2008	Functional exercise programme	Usual care: Seated activities, including watching films, reading, singing	
Sakamoto 2006	Single leg practice 1 min / leg, 3 x daily	Usual care: no details	
Saravanakumar 2014	• Tai Chi • Flexibility (yoga)	Different exercise: "staying active": includes games, group activities, a gym with bike and activities such as walking and gardening	All comparisons presented under comparisons of different exercise categories

Table 2. Description of interventions in studies of exercise interventions in care facilities (*Continued*)

Schoenfelder 2000	Ankle-strengthening exercise	Usual care: little information	
Serra-Rexach 2011	Training sessions + usual care physiotherapy	Different exercise: usual care physiotherapy (40 to 45 minutes/day 5 x weekly)- stretches, aerobic exercise such as walking (though low intensity)	
Shimada 2004	Gait exercises + usual exercises	Different exercise: physiotherapy for pain, stretches, low- and high-intensity resistance training, gait training, stairs, lower limb function	
Sihvonen 2004	Balance training (visual feedback)	Usual care: little information	
Sitja Rabert 2015	Whole body vibration + exercise static and dynamic balance and strength exercise)	Different exercise: same exercise programme done on land	
Toulotte 2003	Supervised exercises	Usual care: continued daily routine	
Tuunainen 2013	• Group strength training: Progressive resistance, supervised group training, 1 hour, 2x weekly • Balance and strength training	Different exercise: self-administered training (1 hour, 2 x weekly) : Stretching, crouching and rising administered by nurses written instructions from physiotherapist	All comparisons presented under comparisons of different exercise categories
Yokoi 2015	Group supervised seated stick exercises 25 minutes, 2 x weekly (included daily house-keeping and hobbies for both exercise and control group)	Usual care: activities of daily living and 10-minute group stretching exercises continued. No other exercises were conducted	

Table 3. Description of interventions in the medication review trials

Study	Medication review	Control	Comment
Crotty 2004a	Additional pharmacist	Usual care	
Crotty 2004b	Additional pharmacist	Usual care	
Frankenthal 2014	Medication review	No interventional recommendations made by pharmacist to chief physician	

Table 3. Description of interventions in the medication review trials (*Continued*)

Garcia Gollarte 2014	Physician education on drug use in older people, plus medication review in 10%	No intervention or information about an educational intervention	Falls data excludes the intervention period; not suitable for pooling
Houghton 2014	Multiprofessional medication review	Usual care (support from the NHS)	
Juola 2015	Nursing education to reduce medication use	Usual care	
Lapane 2011	Clinical informatics tool for medication review: providing reports to pharmacists and nursing staff to assist identifying residents at risk for delirium and falls. Reports generated within 24 hours of admission, used during monthly medication review and at time of Minimum Data Set reporting or when falls or delirium triggered resident assessment protocols	Usual care (includes monthly medication review by pharmacist)	
Patterson 2010	Pharmacist review of psychoactive medications	Usual care	
Peyro Saint Paul 2013	Ceasing medication to avoid hyponatraemia	Usual care	Unusual study, not pooled with others
Potter 2016	Deprescribing	Medication review without deprescribing	
Streim 2012	Deprescribing antidepressants	Continue taking antidepressants	Data not suitable for pooling.
Zermansky 2006	Medication review by pharmacist	Usual care	

Table 4. Summary of 'Risk of bias' assessment of included studies

Risk of Bias	Low	High	Unclear
Sequence generation (selection bias)	69% (66/95)	2% (2/95)	28% (27/95)
Allocation (selection bias)	45% (43/95)	15% (14/95)	40% (38/95)
Blinding of participants and personnel (performance bias)	7% (7/95)	91% (86/95)	2% (2/95)

Table 4. Summary of 'Risk of bias' assessment of included studies (*Continued*)

Blinding of outcome assessors (detection bias)	11% (10/95)	68% (65/95)	21% (20/95)
Incomplete outcome data (attrition bias)	63% (60/95)	27% (26/95)	9% (9/95)
Selective reporting (reporting bias)	53% (50/95)	8% (8/95)	39% (37/95)
Method of ascertaining falls	47% (45/95)	28% (27/95)	24% (23/95)
Baseline imbalance	54% (51/95)	27% (26/95)	19% (18/95)
Other bias	92% (87/95)	2% (2/95)	6% (6/95)

APPENDICES

Appendix 1. Search strategies (2012 to February 2016)

For this update the searches were modified to broaden sensitivity and bring them in line with current Cochrane guidelines. Previous search strategies are given in Cameron 2012.

The search process was run in two stages: the first search was run in February 2016 and a second top-up search was run in August 2017.

CENTRAL 2016, Issue 2 (Cochrane Central Register of Studies Online)

Initial search (February 2016)

#1 MESH DESCRIPTOR Accidental Falls (945)
#2 MESH DESCRIPTOR Hip Fractures EXPLODE ALL TREES WITH QUALIFIERS PC (122)
#3 (falls or faller*):TI,AB,KY (2980)
#4 #1 OR #2 OR #3 (3051)
#5 MESH DESCRIPTOR Aged EXPLODE ALL TREES (863)
#6 (older or senior* or elderly):TI,AB,KY (35860)
#7 #5 OR #6 (36186)
#8 #4 AND #7 (1491)
#9 MESH DESCRIPTOR Residential Facilities EXPLODE ALL TREES (1269)
#10 MESH DESCRIPTOR Long-Term Care (989)
#11 MESH DESCRIPTOR Institutionalization (159)
#12 MESH DESCRIPTOR Hospitalization (3772)
#13 MESH DESCRIPTOR Subacute Care (9)
#14 MESH DESCRIPTOR Hospitals EXPLODE ALL TREES (2630)
#15 MESH DESCRIPTOR Hospital Units (173)
#16 MESH DESCRIPTOR Rehabilitation Centers (233)
#17 MESH DESCRIPTOR Inpatients (631)
#18 MESH DESCRIPTOR Geriatric Assessment (1117)

#19 ((long stay or long term or acute or sub-acute or subacute or residential or hospital) adj3 (care or ward* or hospital)):TI,AB,KY (9444)
#20 ((rehabilitation or geriatric) adj (ward* or hospital* or unit* or department*)):TI,AB,KY (2261)
#21 (hostel* or nursing home*):TI,AB,KY (2109)
#22 inpatient*:TI,AB,KY (7336)
#23 residen*:TI,AB,KY (7244)
#24 institution*:TI,AB,KY (8275)
#25 #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 (36817)
#26 #8 AND #25 (495)
#27 18/04/2012 TO 29/02/2016:DL (261267)
#28 #26 AND #2 (7214)
#29 * NOT INMEDLINE NOT INEMBASE AND 18/04/2012 TO 29/02/2016:DL (61657)
#30 #28 AND #29 (7)

Top-up search (August 2017)

#27 29/02/2016 TO 31/08/2017:DL (146249)
#28 #26 AND #27 (120)

MEDLINE (OvidSP)

Initial search (February 2016)

1 Accidental Falls/ or exp Hip Fractures/pc [Prevention & Control] (18380)
2 (falls or faller$).tw. (33218)
3 or/1-2 (42468)
4 exp Aged/ or Middle Aged/ (4118285)
5 (older or senior$ or elderly).tw. (473795)
6 or/4-5 (4287430)
7 and/3,6 (21348)
8 exp Residential Facilities/ (45187)
9 Long-Term Care/ (22760)
10 Institutionalization/ or Hospitalization/ (84278)
11 Subacute Care/ (757)
12 exp Hospitals/ (230464)
13 Hospital Units/ (9255)
14 Rehabilitation Centers/ (7271)
15 Inpatient/ (14941)
16 Geriatric Assessment/ (20228)
17 ((long stay or long term or acute or sub-acute or subacute or residential or hospital) adj3 (care or ward$1 or hospital)).tw. (744645)
18 ((rehabilitation or geriatric) adj (ward$1 or hospital$1 or unit$1 or department$1)).tw. (7183)
19 (hostel$1 or nursing home$).tw. (24258)
20 inpatient.tw. (51064)
21 residen$.tw. (170300)
22 institution$.tw. (191229)
23 or/8-22 (1281719)
24 and/7,23 (6980)
25 Randomized controlled trial.pt. (406953)
26 Controlled clinical trial.pt. (90108)
27 randomized.ab. (336211)
28 placebo.ab. (166425)
29 Drug therapy.fs. (1819658)
30 randomly.ab. (242642)
31 trial.ab. (347439)
32 groups.ab. (1517503)

33 or/25-32 (3659217)
34 exp Animals/ not Humans/ (4187037)
35 33 not 34 (3146945)
36 24 and 35 (1918)
37 (2012* or 2013* or 2014* or 2015* or 2016*).ed,dc. (4739332)
38 36 and 37 (660)
Top-up search (August 2017)
37 (2016* or 2017*).ed,dc,yr. (2902640)
38 36 and 37 (444)

Embase (OvidSP)

Initial search (February 2016)
1 Falling/ or exp Hip fracture/pc (30681)
2 (falls or faller$).tw. (42331)
3 or/1-2 (60124)
4 Aged/ or Middle Aged/ (2951209)
5 (older or senior$ or elderly).tw. (623077)
6 or/3-4 (2990799)
7 and/3,6 (60124)
8 Residential Home/ or Nursing Home/ or Assisted Living Facility/ (48670)
9 Halfway House/ or Long Term Care/ (102560)
10 Hospitalization/ (243942)
11 Institutional Care/ or Residential Care/ or Home For The Aged/ or Institutionalization/ (29979)
12 exp Hospital/ or Hospital Patient/ (893392)
13 Rehabilitation Center/ (10566)
14 ((long stay or long term or acute or sub-acute or subacute or residential or hospital) adj3 (care or ward$1 or hospital)).tw. (1054527)
15 ((rehabilitation or geriatric) adj (ward$1 or hospital$1 or unit$1 or department$1)).tw. (11032)
16 (hostel$1 or nursing home$).tw. (30080)
17 inpatient.tw. (78633)
18 residen$.tw. (208729)
19 institution$.tw. (287669)
20 or/8-19 (2160272)
21 and/7,20 (15557)
22 exp Randomized Controlled Trial/ or exp Single Blind Procedure/ or exp Double Blind Procedure/ or Crossover Procedure/ (443586)
23 (random* or RCT or placebo or allocat* or crossover* or 'cross over' or trial or (doubl* adj1 blind*) or (singl* adj1 blind*)).ti,ab. (1472662)
24 22 or 23 (1551624)
25 (exp Animal/ or animal.hw. or Nonhuman/) not (exp Human/ or Human cell/ or (human or humans).ti.) (5440113)
26 24 not 25 (1369711)
27 21 and 26 (1849)
28 (2012* or 2013* or 2014* or 2015* or 2016*).em,dd. (6468106)
29 27 and 28 (849)
Top-up search (August 2017)
28 (2016* or 2017*).dd,yr. (2947022)
29 27 and 28 (362)

CINAHL (EBSCOhost)

Initial search (February 2016)
S1 (MH "Accidental Falls") (14,702)
S2 TI ((falls or faller or fallers)) OR AB ((falls or faller or fallers)) (18,518)
S3 S1 or S2 (25,905)

S4 (MH "Aged+") (554,747)
S5 TI ((senior or seniors or elderly or older)) OR AB ((senior or seniors or elderly or older)) (154,950)
S6 S4 or S5 (606,645)
S7 S3 and S6 (12,500)
S8 (MH "Residential Facilities+") (24,586)
S9 (MH "Long Term Care") (20,495)
S10 MH Hospitalization OR MH institutionalisation (22,416)
S11 (MH "Subacute Care") (1,163)
S12 (MH "Hospitals+") (82,740)
S13 (MH "Hospital Units") (5,365)
S14 (MH "Rehabilitation Centers") (6,003)
S15 TX (long stay or acute or sub-acute or subacute or residential) N3 (care or ward or wards or hospital*) (42,572)
S16 TX (rehabilitation or geriatric) N1 (ward* or hospital* or unit* or department*) (27,626)
S17 TX hostel OR TX hostels (342)
S18 TI inpatient OR AB inpatient (23,497)
S19 TI residen* OR AB residen* (44,727)
S20 TI institution* OR AB institution* (42,946)
S21 TX nursing home (49,403)
S22 S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 (306,197)
S23 S7 AND S22 (3,504)
S24 PT Clinical Trial (79,124)
S25 (MH "Clinical Trials+") (196,188)
S26 TI clinical trial* OR AB clinical trial* (51,126)
S27 TI ((single blind* or double blind*)) OR AB ((single blind* or double blind*)) (23,585)
S28 TI random* OR AB random* (166,482)
S29 S24 OR S25 OR S26 OR S27 OR S28 (302,149)
S30 S23 AND S29 (496)
S31 EM 2012 OR EM 2013 OR EM 2014 OR EM 2015 OR EM 2016 (1,482,299)
S32 S30 AND S31 (145)

Top-up search (August 2017)

S31 EM 2016 OR EM 2017 (1,830,054)
S32 S30 AND S31 (169)

WHO ICTRP

Initial search (February 2016)

fall* AND prevent* OR fall AND reduc* (368 records for 361 trials)

Top-up search (August 2017)

89 additional records identified

ClinicalTrials.gov

Initial search (March 2016)

(fall OR falls OR falling) AND (prevention OR prevent OR reduce OR reduction)
Interventional Studies
received from 01/01/2012 to 22/03/2016
551 records

Top-up search (August 2017)

232 additional records identified

Appendix 2. 'Risk of bias' assessment criteria

Bias	Judgement of risk of bias: LOW, HIGH, or UNCLEAR
Random sequence generation Relating to selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence	According to recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions*.
Allocation concealment Relating to selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment	According to recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions*. In cluster randomised trials, if patients were recruited following allocation of the cluster, this was considered as high risk. The timing of recruitment of individuals to clusters was considered within this domain
Blinding of participants and personnel Relating to performance bias due to knowledge of the allocated interventions by participants and personnel carrying out the interventions	According to recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions*.
Blinding of outcome assessment Relating to detection bias due to knowledge of the allocated interventions by outcome assessors	According to recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions*.
Incomplete outcome data Relating to attrition bias due to amount, nature or handling of incomplete outcome data	According to recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions*. For cluster-randomised trials, potential bias due to loss of clusters was considered within this domain
Selective outcome reporting Relating to bias due to the selective reporting or non reporting of findings	According to recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions*. Where no protocol was identified, but all expected falls outcomes were reported and appropriate adjustments for clustering were performed, a 'low risk' rather than unclear judgement was made
Method of ascertaining falls Relating to bias in the recall of falls due to unreliable methods of ascertainment	All studies were assessed as follows. **Judgement of 'Low risk'** if the study used a clear definition of falls plus some form of concurrent collection of data about falling, e.g. staff recorded falls daily on a hospital register. **Judgement of 'High risk'** if ascertainment relied on participant recall at longer intervals than one month during the study or at its conclusion, or if there were important differences in the methods of ascertainment of falls between study arms, or falls were poorly defined **Judgement of 'Unclear'** if there was retrospective recall over a short period only, or a definition of falls was not described, or details of ascertainment were not described, i.e. insufficient information was provided to allow a judgement of 'Low risk' or 'High risk'

(*Continued*)

Bias resulting from major baseline imbalances Relating to bias resulting from major imbalances in key baseline characteristics	**Judgement of 'Low risk'** if good comparability of groups, or confounding adjusted for in analysis **Judgement of 'High risk'** if imbalance in characteristics likely to impact on falls rate (particularly age, previous falls/falls risk, medical status, dependency, cognitive function) and confounding not adjusted for in analysis **Judgement of 'Unclear'** if not discussed.

Appendix 3. Settings, combinations and categories of interventions (ProFaNE) for each included study

Setting/ Combination	Study ID	Exercises	Medication (drug target)	Management of urinary incontinence	Fluid or nutritional therapy	Environment/ assistive technology	Social environment	Knowledge	Other
CARE FACILITIES									
Single	Bischoff 2003		****						
	Broe 2007		****						
	Buckinx 2014	****							
	Buettner 2002	****							
	Cadore 2014	****							
	Chapuy 2002		****						
	Chenoweth 2009						****		
	Choi 2005	****							
	Clifton 2009					****			

(*Continued*)

	Colon-Emeric 2013						****		
	Cox 2008						****		
	Crotty 2004a		****						
	Crotty 2004b		****						
	da Silva Borges 2014	****							
	Houghton 2014		****						
	Faber 2006	****							
	Flicker 2005		****						
	Franken-thal 2014		****						
	Fu 2015	****							
	Garcia Gollarte 2014		****						
	Grieger 2009		****						
	Huang 2016 (CB)								**** Psycholog-ical
	Imaoka 2016 (RED EX)	****							
	Imaoka 2016 (Vit D)		****						
	Irez 2011	****							

(*Continued*)

	Juola 2015		****						
	Kennedy 2015		****						
	Kerse 2008	****							
	Klages 2011								**** Multisensory stimulation
	Kovacs 2012	****							
	Kovacs 2013	****							
	Lapane 2011		****				****		
	Law 2006		****						
	Meyer 2009						****		
	Mulrow 1994	****							
	Nowalk 2001	****							
	Patterson 2010		****						
	Peyro Saint Paul 2013		****						
	Potter 2016		****						
	Rosendahl 2008	****							
	Sakamoto 2006	****							
	Sakamoto 2012								**** Lavender patches

(*Continued*)

	Sambrook 2012 (UV)								**** Sunlight
	Sara-vanaku-mar 2014	****							
	Schoen-felder 2000	****							
	Serra-Rexach 2011	****							
	Shimada 2004	****							
	Sihvonen 2004	****							
	Sitja Rabert 2015	****							
	Streim 2012		****						
	Toulotte 2003	****							
	Tuunainen 2013	****					****		
	Van de Ven 2014								
	Van Gaal 2011a						****		
	Van het Reve 2014								**** Psycholog-ical
	Ward 2010						****		
	Yokoi 2015	****							

(*Continued*)

	Zermansky 2006		****						
Multiple	Huang 2016	****							****
	Imaoka 2016	****	****						
	Schnelle 2003	****		****	****				
	Sambrook 2012 (UV+)		****						**** Sunlight
Multifactorial	Beck 2016	****			****				
	Becker 2003	****				****	****	****	
	Dyer 2004	****	****			****	****		**** Podiatry referral
	Jensen 2002	****	****			****	****		
	Kerse 2004		****	****		****	****		
	McMurdo 2000	****	****			****		****	
	Neyens 2009	****	****			****	****		
	Ray 1997		****			****	****	****	
	Rubenstein 1990		****			****			
	Salvà 2016 [a]	****	****	****		****			

(*Continued*)

	Shaw 2003	****	****			****			
	Walker 2015	****	****	****	****	****		****	
	Whitney 2017	****	****			****	****		
HOSPITALS									
Single	Ang 2011							****	
	Burleigh 2007		****						
	Donald 2000 (2 x 2 factorial)	****				****			
	Dykes 2010						****		
	Haines 2010					****			
	Haines 2011							****	
	Jarvis 2007	****							
	Koh 2009						****		
	Mador 2004						****		
	Mayo 1994					****			
	Michalek 2014		****						
	Shorr 2012					****			
	Stenvall 2007						****		

(*Continued*)

	Tideiksaar 1993					****			
	Treacy 2015	****							
	Van Gaal 2011b						****		
	Wald 2011						****		
	Wolf 2013					****			
Multifactorial	Aizen 2015		****	****		****	****		Psychological
	Barker 2016			****		****	****		
	Cumming 2008	****	****			****	****	****	
	Haines 2004	****				****		****	
	Healey 2004		****			****			**** Opthalmology referral
	Hill 2015						****	****	

[a] ***Likely types of interventions based on falls risk factors assessed, actual interventions instigated unclear***

Abbreviations

CB: cognitive behavioural
RED EX: reduced exercise
UV: increased sunlight exposure group.
UV+: increased sunlight exposure + calcium supplementation group

Appendix 4. Categories of exercise (ProFaNE) by study setting and combination

Study setting/type	Study ID	Gait/balance/ functional training	Strength/ resistance training	Flexibility	3D (Tai Chi, dance etc)	General physical activity	Endurance	Other
CARE FACILITIES								
Single	Buckinx 2014	****	****					**** (WBV)
	Buettner 2002	****	****	****		****		****
	Cadore 2014	****	****	****				
	Choi 2005				****			
	da Silva Borges 2014			****	****			
	Faber 2006 (FW)	****						
	Faber 2006 (IB)	****	****	****	****			
	Fu 2015							****
	Imaoka 2016							****
	Irez 2011	****	****	****				
	Kerse 2008	****						
	Kovacs 2012	****	****	****		****		
	Kovacs 2013	****	****	****		****		
	Mulrow 1994	****	****	****				
	Nowalk 2001 (FNBF)		****	****				

(*Continued*)

	Nowalk 2001 (LL/TC)				****			
	Rosendahl 2008	****	****					
	Sakamoto 2006	****						
	Saravanakumar 2014 (Tai Chi)				****			
	Saravanakumar 2014 (Yoga)			****				
	Schoenfelder 2000		****			****		
	Serra-Rexach 2011		****	****			****	
	Shimada 2004	****						
	Sihvonen 2004	****						
	Sitja Rabert 2015	****	****					**** (WBV)
	Toulotte 2003	****	****	****				
	Tuunainen 2013 (MF)		****					
	Tuunainen 2013(MFB)	****	****					
	Yokoi 2015	****		****				

(*Continued*)

Multiple	Huang 2016	****	****					
	Imaoka 2016						****	
	Schnelle 2003		****			****		
Multifactorial	Beck 2016	****	****					
	Becker 2003	****	****					
	Dyer 2004	****	****	****		****		
	Jensen 2002	****	****					
	McMurdo 2000	****	****	****				
	Neyens 2009[a]							
	Salvà 2016	****	****	****		****		
	Shaw 2003	****	****	****				
	Walker 2015[a]							
	Whitney 2017	****						
HOSPITALS								
Single	Donald 2000 (EX)		****					
	Jarvis 2007	****	****	****				
	Treacy 2015	****						

(*Continued*)

Multifactorial	Cumming 2008	****						
	Haines 2004	****	****		****			

[a] No description of the exercise components

Abbreviations

EX: supplementary exercises
FNBF: 'Fit NB Free' group
FW: 'Functional Walking' group
IB: 'In Balance' group
LL/TC: 'Living and learning/Tai Chi' group
MF: muscle force
MFB: muscle force & balance

WBV: whole body vibration

Appendix 5. Categories of environment/assistive technology interventions (ProFaNE) by study setting and combination

Study setting/ type	Study ID	Furnishing/ adaptations	Personal mobility aids	Communication/signalling aids	Body worn care/protection aids	Other environmental
CARE FACILITIES						
Single	Clifton 2009			****		
Multifactorial	Becker 2003	****	****		****	
	Dyer 2004	****				
	Jensen 2002	****	****	****	****	
	Kerse 2004	****	****		****	
	McMurdo 2000	****				
	Neyens 2009	****	****			

(*Continued*)

	Ray 1997	****	****			
	Rubenstein 1990	****				
	Salvà 2016[a]					
	Shaw 2003	****	****		****	
	Walker 2015	****	****		****	
	Whitney 2017	****	****	****		
HOSPITALS						
Single	Donald 2000 (FL)	****				
	Mayo 1994			****		
	Haines 2010	****				
	Shorr 2012			****		
	Tideiksaar 1993			****		
	Wolf 2013			****		
Multifactorial	Aizen 2015	****	****		****	
	Barker 2016	****		****		
	Cumming 2008	****	****	****		
	Haines 2004			****	****	
	Healey 2004	****		****	****	****
	Stenvall 2007					**** Home visit by OT and/or PT

[a] ***No clear description of types of environment/assistive technology***

Abbreviations

FL: carpet flooring group
OT: occupational therapist
PT: physiotherapist

Appendix 6. Categories of medication (drug target, ProFaNE) interventions by study setting and combination

Setting/ Combination	Study ID	Vitamin D	Calcium	Other bone health medication	Antidepressants	Antipsychotics/ neuroleptics	Medication review	Other
CARE FACILITIES								
Single	Bischoff 2003	****						
	Broe 2007	****						
	Chapuy 2002	****	****					
	Crotty 2004a						**** Pharm	
	Crotty 2004b						**** Pharm	
	Houghton 2014						**** MultiP	
	Flicker 2005	****						
	Frankenthal 2014						**** Pharm	
	Garcia Gollarte 2014						**** Educ	
	Grieger 2009[a]	****	****					
	Imaoka 2016[a]	****	****					
	Juola 2015[b]							****

(*Continued*)

	Kennedy 2015[c]	****	****	****				
	Law 2006	****						
	Patterson 2010[d]					****		
	Peyro Saint Paul 2013[e]						****	
	Potter 2016							
	Sambrook 2012	**** (UV)						
	Streim 2012						**** Depresc	
	Zermansky 2006						**** Pharm	
Multiple	Imaoka 2016[a]							
	Sambrook 2012	**** (UV)	****					
Multifacto-rial	Dyer 2004						****	
	Jensen 2002						****	
	McMurdo 2000						****	
	Neyens 2009						****	
	Ray 1997					****		
	Rubenstein 1990						****	
	Salvà 2016						****	
	Shaw 2003						****	

(*Continued*)

	Walker 2015							
	Whitney 2017						****	
HOSPITALS								
	Burleigh 2007	****						
	Michalek 2014						****	
Multifactorial	Aizen 2015						??	
	Cumming 2008						****	
	Healey 2004						****	
	Stenvall 2007	****	****	****				

[a] Multivitamin
[b] Nurse education on harmful medications
[c] Training to increase appropriate prescription of vitamin D, calcium and osteoporosis medications
[d] Medication review of antipsychotics
[e] Review by a pharmacologist for patients with hyponatraemia
Abbreviations
Depresc: deprescribing
Educ: education on medication review
Multi P: multiprofessional review by clinical pharmacist, pharmacy technician, care home staff and GP
Pharm: pharmacist
UV: increased sunlight exposure group

Appendix 7. Source of data for generic inverse variance analysis (*see* footnotes for explanation of codes)

Study ID	Source for rate ratio (falls)	Source for risk ratio (fallers)	Source of risk ratio (number with fractures)
Aizen 2015	ND	ND	NA
Ang 2011	NA	4	NA
Barker 2016	1b	7c	ND
Beck 2016	ND	NA	NA
Becker 2003	1b	5b	7c
Becker 2003 (Cognitively impaired/not impaired subgroup analysis)	1	5	NA
Bischoff 2003	1a	5a	7
Broe 2007 (800 IU)	1a	4a	NA
Buckinx 2014	3	7	NA
Buettner 2002	ND	NA	NA
Burleigh 2007	ND	5	7
Cadore 2014	ND	NA	NA
Chapuy 2002	NA	7	7
Chenoweth 2009	NA	ND	NA
Choi 2005	NA	7c	NA
Clifton 2009	3	NA	NA
Colon-Emeric 2013	ND	NA	NA
Cox 2008	1ab	NA	ND
Crotty 2004a	NA	5	NA
Crotty 2004b	NA	5ab	NA
Cumming 2008	1ab	7c	7c

(*Continued*)

da Silva Borges 2014	ND	NA	NA
Donald 2000	3	5	NA
Dyer 2004	3c	6b	NA
Dykes 2010	3c	7c	NA
Faber 2006	3	4 (FW vs control and IB vs control) 4a (FW + IB vs control)	NA
Flicker 2005	1	4	7
Frankenthal 2014	3	NA	NA
Fu 2015	1a	NA	NA
Garcia Gollarte 2014	ND	ND	NA
Grieger 2009	3	7	NA
Haines 2004	3	5	7
Haines 2010	3c	NA	NA
Haines 2011	2a	6a	NA
Healey 2004	3c	NA	NA
Hill 2015	1ab	ND[a]	ND[a]
Houghton 2014	1b	NA	NA
Huang 2016	ND	ND	NA
Imaoka 2016	ND	ND	NA
Irez 2011	3	NA	NA
Jarvis 2007	ND	7	NA
Jensen 2002	1b	4b	6a
Jensen 2002 (MMSE < 19/ ≥ 19 subgroup analysis)	1b	7c	NA

(*Continued*)

Juola 2015	1ac	7c	NA
Juola 2015 (MMSE >15, 10-15, <10 subgroups)	3ac	NA	NA
Kennedy 2015	3c	7c	ND
Kerse 2004	1ab	7c	NA
Kerse 2008	2b	7c	NA
Klages 2011	ND	NA	NA
Koh 2009	3c	NA	NA
Kovacs 2012	NA	5	NA
Kovacs 2013	1	5	NA
Lapane 2011	NA	4b	NA
Law 2006	3c	7c	5ab
Mador 2004	NA	7	NA
Mayo 1994	3	4	NA
McMurdo 2000	3c	7c	7c
Meyer 2009	3c	7c	7c
Michalek 2014	3c	7c	NA
Mulrow 1994	3	7	NA
Neyens 2009	1b	NA	NA
Nowalk 2001	NA	ND	NA
Patterson 2010	3c	NA	NA
Peyro Saint Paul 2013	3	7	NA
Potter 2016	3	7	7
Ray 1997	NA	ND	NA
Rosendahl 2008	1c	7c	7c

(*Continued*)

Rubenstein 1990	3	7	7
Sakamoto 2006	3	7	7
Sakamoto 2012	1	4	NA
Salvà 2016	1ab	7c	7c
Salvà 2016 (subgroup excluding dementia)	1ab	6ab	NA
Sambrook 2012	1c	7c	7c
Saravanakumar 2014	3	NA	NA
Schnelle 2003	3	7	7
Schoenfelder 2000	3	NA	NA
Serra-Rexach 2011	ND	NA	NA
Shaw 2003	ND	5	5
Shimada 2004	3	7	NA
Shorr 2012	3ab	7c	NA
Sihvonen 2004	1a	7	NA
Sitja Rabert 2015	ND	7	7
Stenvall 2007	1	4	ND
Stenvall 2007 (dementia subgroup in Stenvall 2012)	1	7	7
Streim 2012	ND	NA	NA
Tideiksaar 1993	ND	NA	NA
Toulotte 2003	ND	NA	NA
Treacy 2015	1	NA	NA
Tuunainen 2013	3	7	NA
Van de Ven 2014	3c	NA	NA
Van Gaal 2011a	1c	NA	NA

(*Continued*)

Van Gaal 2011b	1c	NA	NA
Van het Reve 2014	3	7	NA
Wald 2011	3	NA	NA
Walker 2015	3c	NA	NA
Ward 2010	ND	NA	7c
Whitney 2017	1b	5a	7c
Wolf 2013	3	7	NA
Yokoi 2015	NA	7c	NA
Zermansky 2006	3	7	NA

[a] Data reported as admissions not patients

Abbreviations

FW: 'Functional Walking' group

IB: 'In Balance' group

MMSE: Mini Mental State Examination

800 IU: 800 International Units vitamin D group

Codes for source of rate ratio:

1: incidence rate ratio reported by trial authors

2: hazard ratio/relative hazard (multiple events) reported by trial authors

3: incidence rate ratio calculated by review authors

a: adjusted for confounders by trial authors

b: adjusted for clustering by trial authors

c: adjusted for clustering by review authors

Codes for source of risk ratio:

4: hazard ratio/relative hazard (first fall only) reported by trial authors

5: relative risk reported by trial authors

6: odds ratio reported by trial authors

7: relative risk calculated by review authors

a: adjusted for confounders by trial authors

b: adjusted for clustering by trial authors

c: adjusted for clustering by review authors

NA: not applicable. Falls (for rate ratio) or fallers (for risk ratio) or number of people sustaining a fracture (for risk ratio) not reported as an outcome in the trial

ND: outcomes relating to falls or fallers or fractures were reported, but there were no useable data; results from the paper reported in the text of the review

Appendix 8. Raw data for rate of falls and number of fallers when available

Study ID	Intervention group: falls per person year	Control group: falls per person year	Intervention group: number of fallers	Intervention group: number in analysis	Intervention group: proportion of fallers	Control group: number of fallers	Control group: number in analysis	Control group: proportion of fallers
CARE FACILITIES								
Beck 2016	0	0.43	---	9	---	---	22	
Becker 2003	1.40	2.56	188	509	0.37	247	472	0.52
Becker 2003 (Cognitively impaired)	1.10	2.71	50	150	0.33	98	169	0.58
Becker 2003 (Not cognitively impaired)	1.42	2.04	93	215	0.43	91	191	0.48
Bischoff 2003	---	---	14	62	0.23	18	60	0.30
Broe 2007 (800 IU)	0.28	1.00	5	23	0.22	11	25	0.44
Buckinx 2014	1.16	1.21	15	31	0.48	17	31	0.55
Buettner 2002	---	---	---	---	---	---	---	---
Cadore 2014	0	9.6	---	11	---	---	13	---
Chapuy 2002	---	---	251	393	0.64	118	190	0.62
Chenoweth 2009	---	---	---	---	---	---	---	---
Choi 2005	---	---	9	29	0.31	15	30	0.50
Clifton 2009	2.45	3.79	---	43	---	---	43	---

(*Continued*)

Colon-Emeric 2013	2.06[a]	2.64[a]	---	---	---	---	---	---
Cox 2008[b]	---	---	---	3315	---	---	2322	---
Crotty 2004a	---	---	19	44	0.43	16	44	0.36
Crotty 2004b	---	---	97	381	0.26	73	334	0.22
da Silva Borges 2014	---	---	---	30	---	---	29	---
Houghton 2014	3.32	3.0	---	381	---	---	445	---
Dyer 2004	2.17	4.02	56	102	0.55	51	94	0.54
Faber 2006 (FW)	3.3	2.5	40	64	0.63	48	90	0.53
Faber 2006 (IB)	2.4	2.5	45	78	0.58	48	90	0.53
Faber 2006 (FW + IB)	2.8	2.5	85	142	0.60	48	90	0.53
Flicker 2005	1.26	1.90	170	313	0.54	185	312	0.59
Frankenthal 2014	0.80	1.30	---	160	---	---	146	---
Fu 2015	0.54	1.52	---	30	---	---	30	---
Garcia Gollarte 2014	1.28	1.72	82	344	0.24	104	372	0.28
Grieger 2009	0.60	1.60	11	48	0.23	12	43	0.28
Huang 2016 (CB)	0.00	1.67	0	25	0.00	7	24	0.29
Huang 2016 (CB + EX)	0.00	1.67	0	24	0.00	7	24	0.29

(*Continued*)

Imaoka 2016 (RED EX)	---	---	7	22	0.32	9	17	0.53
Imaoka 2016 (Vit D)	---	---	6	17	0.35	9	17	0.53
Imaoka 2016 (multiple)	---	---	4	19	0.21	9	17	0.53
Irez 2011	1.60	5.63	---	30	---	---	30	---
Jensen 2002	2.45	3.03	82	188	0.44	109	196	0.56
Jensen 2002 (MMSE < 19)	3.50	3.34	37	69	0.54	62	102	0.61
Jensen 2002 (MMSE ≥ 19)	1.77	2.90	42	112	0.38	43	79	0.54
Juola 2015	2.25	3.25	42	93	0.45	60	96	0.63
Juola 2015 (MMSE >15)	3.90	3.08	---	45	---	---	50	---
Juola 2015 (MMSE 10-15)	1.12	4.22	---	23	---	---	22	---
Juola 2015 (MMSE <10)	0.61	2.70	---	25	---	---	24	---
Kennedy 2015	2.57	2.51	853	1290	0.66	1712	2727	0.63
Kerse 2004	4.1	2.3	173	309	0.56	103	238	0.43
Kerse 2008	---	---	162	310	0.52	146	329	0.44
Klages 2011	---	---	---	---	---	---	---	---
Koh 2009	0.40	0.22	---	612	---	---	510	---

(*Continued*)

Kovacs 2012	---	---	8	21	0.38	14	20	0.70
Kovacs 2013	0.69	0.97	16	32	0.50	20	30	0.67
Lapane 2011	---	---	---	1769	---	---	1552	---
Law 2006	2.01	2.31	770	1762	0.44	833	1955	0.43
McMurdo 2000	3.02	3.85	20	52	0.38	22	38	0.58
Meyer 2009	1.97	2.04	299	574	0.52	291	551	0.53
Mulrow 1994	1.86	2.44	44	97	0.45	38	97	0.39
Neyens 2009	2.09	2.54	---	249	---	---	269	---
Nowalk 2001 (LL/TC)	---	---	---	---	---	---	---	---
Nowalk 2001 (FNBF)	---	---	---	---	---	---	---	---
Patterson 2010	1.96	1.37	---	173	---	---	161	---
Peyro Saint Paul 2013	3.00	4.80	1	4	0.25	3	5	0.60
Potter 2016	4.91	2.96	25	45	0.56	31	48	0.65
Ray 1997	---	---	---	---	---	---	---	---
Rosendahl 2008	3.6	4.6	46	87	0.53	49	96	0.51
Rubenstein 1990	2.49	2.63	56	79	0.71	61	81	0.75
Sakamoto 2006	0.93	1.14	68	315	0.22	51	212	0.24
Sakamoto 2012	1.04	1.40	26	73	0.36	36	72	0.50

(*Continued*)

Salvà 2016	1.93	0.89	94	193	0.49	52	137	0.38
Salvà 2016 (excluding dementia)	---	---	---	---	---	---	---	---
Sambrook 2012 (UV)	---	---	111	190	0.58	111	205	0.54
Sambrook 2012 (UV+)	---	---	108	207	0.52	111	205	0.54
Saravanakumar 2014 (Tai Chi)	2.02	3.90	---	9	---	---	11	---
Saravanakumar 2014 (Yoga)	2.87	3.90	---	9	---	---	11	---
Schnelle 2003	0.68	1.09	17	92	0.18	29	98	0.30
Schoenfelder 2000	9.33	3.43	---	9	---	---	7	---
Serra-Rexach 2011	---	---	---	---	---	---	---	---
Shaw 2003	---	---	96	130	0.74	115	144	0.80
Shimada 2004	1.07	2.00	5	15	0.33	6	11	0.55
Sihvonen 2004	---	---	11	20	0.55	5	7	0.71
Sitja Rabert 2015	---	---	20	81	0.25	15	78	0.19
Streim 2012	---	---	---	---	---	---	---	---
Toulotte 2003	---	---	---	---	---	---	---	---

(*Continued*)

Tuunainen 2013 (MF)	0.88	1.19	7	16	0.44	14	18	0.78
Tuunainen 2013 (MFB)	0.57	1.19	6	14	0.43	14	18	0.78
Van de Ven 2014	1.81	3.33	---	137	---	---	156	---
Van Gaal 2011a	1.56	2.08	---	196	---	---	196	---
Van het Reve 2014	0.50	0.80	3	54	0.06	2	60	0.03
Walker 2015	4.00	1.90		22			20	
Ward 2010	---	---	---	---	---	---	---	---
Whitney 2017	1.51	0.93	31	103	0.30	25	88	0.28
Yokoi 2015	---	---	6	51	0.12	16	54	0.30
Zermansky 2006	1.60	2.60	84	331	0.25	106	330	0.32
HOSPITALS								
Aizen 2015	0.67	0.48	13.00	200	0.065	8.00	308	0.026
Ang 2011	---	---	4	910	0.004	14	912	0.02
Barker 2016	2.72	2.57	623	17698	0.035	646	17566	0.04
Burleigh 2007	---	---	36	100	0.36	45	103	0.44
Cumming 2008	3.36	3.39	157	2047	0.08	143	1952	0.07
Donald 2000 (FL)	5.75	0.39	7	28	0.25	1	26	0.04
Donald 2000 (EX)	2.22	2.10	2	30	0.07	6	24	0.25

(*Continued*)

Dykes 2010	1.01	1.84	34	2755	0.01	51	2509	0.02
Haines 2004	4.12	5.94	54	310	0.17	71	316	0.22
Haines 2010	1.91	1.37	---	6113	---	---	4986	---
Haines 2011 (ED)	3.14	3.39	56	424	0.13	54	381	0.14
Haines 2011 (ED+)	2.79	3.39	44	401	0.11	54	381	0.14
Healey 2004	4.12	7.03	---	749	---	---	905	---
Hill 2015	2.85	5.03	136	1623	0.08	248	1983	0.13
Jarvis 2007	---	---	3	14	0.21	7	15	0.47
Mador 2004	---	---	10	36	0.28	4	35	0.11
Mayo 1994	4.62	4.01	27	65	0.42	21	69	0.30
Michalek 2014	0.55	3.87	2	58	0.03	12	56	0.21
Shorr 2012	2.05	1.66	282	11115	0.03	359	17436	0.02
Stenvall 2007	2.30	5.95	12	102	0.12	26	97	0.27
Stenvall 2007 dementia subgroup (Stenvall, 2012)	0.65	10.67	1	28	0.04	11	36	0.31
Tideiksaar 1993	---	---	---	---	---	---	---	---
Treacy 2015	2.28	3.53	---	---	---	---	---	---
Van Gaal 2011b	1.04	1.04	---	1081	---	---	1120	---
Wald 2011	1.75	2.45	---	122	---	---	95	---
Wolf 2013	3.00	5.66	6	48	0.13	7	50	0.14

(*Continued*)

[a] Falls per bed year.
[b] Raw data not available, data reported by authors as rate ratios
Abbreviations
ED: educational materials only group
ED+: educational materials plus physiotherapist follow-up
EX: supplementary exercises group
FL: carpet flooring group
FNBF: 'Fit NB Free' group
FW: 'Functional Walking' group (a functional balance, strength & mobility programme)
IB: 'In Balance' group
MF: muscle force
MFB: muscle force & balance
MMSE: Mini Mental State Examination
LL/TC: 'Living and learning/Tai Chi' group
RED EX: reduced exercise
UV: increased sunlight exposure group.
UV+: increased sunlight exposure + calcium supplementation group
Vit D: Vitamin D3 & calcium in multivitamin supplement
800 IU: 800 International Units vitamin D group

Appendix 9. Raw data for number of fractures when available

Study ID	Intervention group: number of people with fractures	Intervention group: number in analysis	Intervention group: proportion of fracture fallers	Control group: number of people with fractures	Control group: number in analysis	Control group: proportion of fracture fallers
CARE FACILITIES						
Beck 2016	---	9	---	---	22	---
Becker 2003: hip	17	509	0.033	15	472	0.032
Bischoff 2003: hip	2	62	0.032	1	60	0.017
Broe 2007 (800 IU)	---	23	---	---	25	
Buckinx 2014	---	31	---	---	31	
Buettner 2002	---	---	---	---	---	
Cadore 2014	---	11	---	---	13	

(*Continued*)

Chapuy 2002: NV	70	393	0.178	34	190	0.179
Chapuy 2002: hip	27	393	0.069	21	190	0.111
Chenoweth 2009	---	---	---	---	---	
Choi 2005	---	29	---	---	30	
Clifton 2009	---	43	---	---	43	
Colon-Emeric 2013	---	---	---	---	---	
Cox 2008[a]	---	3315	---	---	2322	
Crotty 2004a	---	44	---	---	44	
Crotty 2004b	---	381	---	---	334	
da Silva Borges 2014	---	30	---	---	29	
Houghton 2014	---	381	---	---	445	
Dyer 2004	4	102	0.039	3	94	0.032
Faber 2006 (FW)	---	64	---		90	
Faber 2006 (IB)	---	78	---		90	
Faber 2006 (FW + IB)	---	142	---		90	
Flicker 2005	25	313	0.080	35	312	0.112
Frankenthal 2014	---	160	---		146	
Fu 2015	---	30	---		30	
Garcia Gollarte 2014	---	344	---		372	
Grieger 2009	---	48	---		43	

(*Continued*)

Huang 2016	---	51	---		24	
Imaoka 2016 (RED EX)	---	22	---		17	
Imaoka 2016 (Vit D)	---	17	---		17	
Imaoka 2016 (multiple)	---	19	---		17	
Irez 2011	---	30	---		30	
Jensen 2002	3	188	0.016	12	196	0.061
Juola 2015	---	93	---		96	
Kennedy 2015	---	1290	---		2727	
Kerse 2004	---	309	---		238	
Kerse 2008	---	310	---		329	
Klages 2011	---	---	---		---	
Koh 2009	---	612	---		510	
Kovacs 2012	---	21	---		20	
Kovacs 2013	---	32	---		30	
Lapane 2011	---	1769	---		1552	
Law 2006: NV	64	1762	0.036	51	1955	0.026
Law 2006: hip	24	1762	0.014	20	1955	0.010
McMurdo 2000	1	52	---	3	38	
Meyer 2009	39	574	0.068	38	551	0.069
Mulrow 1994	---	97			97	
Neyens 2009	---	249			269	
Nowalk 2001 (LL/TC)	---	---			---	

(*Continued*)

Nowalk 2001 (FNBF)	---	---			---	
Patterson 2010	---	173			161	
Peyro Saint Paul 2013	---	4			5	
Potter 2016	3	45	0.067	2	48	0.042
Ray 1997	---	---			---	
Rosendahl 2008	4	87	0.046	6	96	0.063
Rubenstein 1990	7	79	0.089	5	81	0.062
Sakamoto 2006: hip	1	315	0.003	1	212	0.005
Sakamoto 2012	---	73			72	
Salvà 2016	10	193	0.052	1	137	0.007
Salvà 2016 (excluding dementia)	---	---			---	
Sambrook 2012 (UV)	17	190	0.089	17	205	0.083
Sambrook 2012 (UV+)	13	207	0.063	17	205	0.083
Saravanakumar 2014	---					
Schnelle 2003	4	92	0.043	1	98	0.010
Schoenfelder 2000	---	9			7	
Serra-Rexach 2011	---	---			---	
Shaw 2003	6	130	0.046	12	144	0.083
Shimada 2004	---	15			11	

(*Continued*)

Sihvonen 2004	---	20			7	
Sitja Rabert 2015	1	81	0.012	0	78	0
Streim 2012	---	---			---	
Toulotte 2003	---	---			---	
Tuunainen 2013 (MF)	---	16			18	
Tuunainen 2013 (MFB)	---	14			18	
Van de Ven 2014	---	137			156	
Van Gaal 2011a	---	196			196	
Van het Reve 2014	---	54			60	
Walker 2015	---	22			20	
Ward 2010	109	2802	0.039	106	2589	0.041
Whitney 2017	3	103	0.029	0	88	0
Yokoi 2015	---	51			54	
Zermansky 2006	---	331			330	
HOSPITALS						
Aizen 2015	---	200			308	
Ang 2011	---	910			912	
Barker 2016	11	17698	0.0006	13	17566	0.0007
Burleigh 2007	1	100	0.010	3	103	0.029
Cumming 2008	2	2047	0.001	3	1952	0.002
Donald 2000 (FL)	---	28			26	
Donald 2000 (EX)	---	30			24	

(*Continued*)

Dykes 2010	---	2755			2509	
Haines 2004	2	310	0.006	2	316	0.006
Haines 2010	---	6113			4986	
Haines 2011 (ED)	---	424			381	
Haines 2011 (ED+)	---	401			381	
Healey 2004	---	749			905	
Hill 2015	4[b]	1623	---	6[b]	1983	---
Jarvis 2007	---	14			15	
Mador 2004	---	36			35	
Mayo 1994	---	65			69	
Michalek 2014	---	58			56	
Shorr 2012	---	11115			17436	
Stenvall 2007	0	102	0	4	97	0.041
Stenvall 2007 dementia subgroup (Stenvall 2012)	0	28	0	3	36	0.083
Tideiksaar 1993	---	---			---	
Treacy 2015	---	---			---	
Van Gaal 2011b	---	1081			1120	
Wald 2011	---	122			95	
Wolf 2013	---	48			50	

[a] Raw data not available, data reported by authors as rate ratios
[b] admissions
Abbreviations
ED: educational materials only group
ED+: educational materials plus physiotherapist follow-up
EX: supplementary exercises group

(*Continued*)

FL: carpet flooring group
FNBF: 'Fit NB Free' group
FW: 'Functional Walking' group (a functional balance, strength & mobility programme)
IB: 'In Balance' group
LL/TC: 'Living and learning/Tai Chi' group
MF: muscle force
MFB: muscle force & balance
NV: non-vertebral
RED EX: reduced exercise
UV: increased sunlight exposure group.
UV+: increased sunlight exposure + calcium supplementation group
Vit D: Vitamin D3 & calcium in multivitamin supplement
800 IU: 800 International Units vitamin D group

Appendix 10. Studies reporting cost-effectiveness or costs of the intervention and/or healthcare resource use

Study ID (source if not primary reference), sample, efficacy analyses, type of evaluation	Intervention(s) and comparator (N in analysis)	Perspective(s), type of currency, price year, time horizon	Cost items measured	Mean (SD) intervention cost per person	Healthcare service costs	Incremental cost per fall prevented/per QALY gained
•Buettner 2002 •Residents of 3 dementia care units (Oxford, Boston, and Palo Alto, USA) ≥ 2 falls in 1 month, mean age 83 (range 60 to 98) years •No effectiveness data available for analysis •Cost analysis	•Daily "graded" walking, "exercise for function" programme 3 x week, sensory air mat 2 x week (evenings) for 2 months vs usual care, number allocated to each group not reported (total N = 27)	•Not stated •US dollar •Not stated •2 months	•Therapist time (intervention only) •Cost of falls and injuries ("based on research data on falls")		•Treatment group USD 30,031, control group USD 79,535	
•Chenoweth 2009 (Norman 2008) •Residents from 15 dementia care	•Dementia care mapping (DCM) (N = 109, 5 sites) vs person centred	•Health service •Australian dollar •2008 •8 months	•Trainer time, post-training support, staff replacement (DCM, PCC)	•Not reported (annual total cost per residential care setting DCM AUD	•Annual pharmaceutical cost per resident AUD 545.55	•Not reported •Incremental cost per behaviour (CMAI point)

(*Continued*)

sites across Sydney, Australia, category 1 to 3 on Australian Resident Classification Scale (high level of care), mean age 84 (SD 7) years •No effectiveness data available for analysis •Cost-effectiveness analysis	care (PCC) (N = 98, 5 sites) vs usual care (N = 82, 5 sites) for 4 months		•Pharmaceutical use	10,034, PCC AUD 2250)		averted DCM vs usual care AUD 46.89, PCC vs usual care AUD 6.43
•Clifton 2009 •Skilled nursing care-facility residents, Eastern Washington State, USA, mean age 82 (SD 7) years •Analysis 7.1 •Analytic model	•Wear FallSaver monitor for 60 days (N = 33) vs no device for 60 days (N = 39), cross-over trial	•Not stated •US dollar •2004 •1 year	•Annual intervention implementation for 100 residents (direct costs only) •Mean hospitalisation cost for injurious fall (from the literature)	•USD 2 per resident per day (annual cost for 100 resident facility USD 73,000)	•Assuming 35 injurious falls per 100 residents per year, annual cost savings for 100 resident facility if 12% fewer injurious falls USD 429, USD 232, 953 if 50% fewer injurious falls	
Houghton 2014 (Sach 2015) •Residents from care homes with average age > 65, registered with GP in local area and registered with Care Quality Commission for at least 6 months •Analysis 5.1 •Cost analysis, detailed micro-costing	• Multiprofessional medication review (N = 826)	•NHS and care homes •Pound sterling •2012 •1 year	•Intervention costs: personnel and resources, Staff costs for time spent on reviews, travel time and costs •Medication costs •Healthcare resource use •Hospitalisations	•GBP 104.80 (SD 50.91) per resident		

(*Continued*)

•Haines 2013 (analysis of Haines 2011) •Acute and Rehabilitation hospital inpatients age ≥ 60, Brisbane and Perth, Australia, mean age 75 (SD 11) years •Analysis 20.1, Analysis 20.2 •Cost-effectiveness analysis	•Multimedia patient education programme with physiotherapist follow up (total N = 1,206)	•Health service provider •Australian dollar •2008 •Period of hospitalisation	•Acute care costs •Rehabilitation costs •Direct falls related costs: radiological investigations, medical costs, nursing costs, medication costs, on-call payment costs, suture procedure costs, orthoses costs, and other tests costs		Intervention/ control group costs post consent per participant (mean (SD) AUD) Subgroup cognitively intact: Intervention cost (complete programme) •Acute care AUD 10,774 (18,344) •Rehabilitation AUD 11,197 (18,906) •Direct falls related costs AUD 1 (7) Control cost •Acute care AUD 8,481 (12,856) •Rehabilitation AUD 10,964 (19,972) •Direct falls related costs intact AUD 8 (47) Subgroup cognitively impaired: Intervention cost (complete programme) •Acute care AUD 11,128 (28,570) •Rehabilitation AUD 21,740 (37,130) •Direct falls related costs AUD 187 (1,602) Control group	For subgroup who were cognitively intact: • AUD 294 per fall prevented •AUD 526 per faller prevented

(*Continued*)

					•Acute care AUD 5,140 (8,142) •Rehabilitation AUD 26,050 (36,776) •Direct falls related costs AUD 15 (85)	
•Meyer 2009 •Nursing home residents in Hamburg, Germany, mean age 86 (SD 6) years •Analysis 8.1, Analysis 8.2, Analysis 8.3 •Cost description	•Administer standardised risk assessment tool (Downton Index) monthly (N = 574, 29 nursing homes) vs usual care (N = 551, 29 nursing homes)	•Nursing care facility •Euro •2006 •1 year	•Nurse time for training and assessing using the Downton Index	•Not reported (total during the study EUR 10, 500 (USD 16, 170, GBP 8160)		
•Mulrow 1994 •Residents from 9 nursing homes in San Antonio, Texas, USA, dependent in ≥ 2 activities of daily living, mean age 80 (SD 8) years •Analysis 3.1, Analysis 3.2, Analysis 2.1, Analysis 2.2 •Cost analysis	•One-on-one physical therapy sessions (N = 97) vs friendly visits (N = 97) 3 x week for 4 months	•Not stated •US dollar •Not stated •4 months	•Intervention delivery (wages, travel expenses, equipment, overheads) •Nursing home, hospitalisation, physician and other health professional visits, emergency department visits, procedures, and medication charges	•USD 1220 (95% CI 412 to 1832) for physical therapy programme, USD 189 (95% CI 80 to 298) control group	•Healthcare charges (81% nursing home, 15% hospitalisation) USD 11, 398 (95% CI 10,929 to 11, 849) per participant (NS)	
•Schnelle 2003 •Residents of 4 nursing homes, incontinence of urine, US, mean age 88 (SD 8) years •Analysis 11.1, Analysis 11.2,	•Low-intensity functionally orientated exercise and incontinence care 5 days a week every 2 hours between 8: 00 am and 4:00	•Not stated •US dollar •1997/98 •8 months	•Diagnostic tests, treatment related to each acute condition (dermatological, genitourinary, gastrointestinal, res-		•USD 24.42 per resident per week to evaluate and treat the selected conditions intervention group, USD 38. 36 control group	

(*Continued*)

Analysis 11.3 •Cost analysis	pm for 8 months (N = 92) vs usual care (N = 98)		piratory and cardiovascular systems; falls; pain; psychiatric and nutritional disturbances)		(NS)	
•Van de Ven 2014 •Dementia special care units, diagnosed with dementia, ≥1 neuropsychiatric symptom, mean age 84.7 years •Analysis 8.1 •Cost analysis	•Dementia care mapping (DCM), 4 months intervention delivered twice during the study (N = 154) vs usual care (N=164)	•Healthcare perspective •US dollar •Not stated •18 months	•Intervention costs: DCM basic and advanced training, mapping exercise, inter-rater reliability test, observation, preparing the DCM reports, feedback sessions •Hospital costs: outpatient, inpatient, emergency department & ambulance •Psychotropic drugs •Nursing home healthcare professional costs	•Intervention cost per resident per day USD 0.63 (SD 0.23)	•Healthcare consumption and drug use per resident per day at 18 months (mean(SD): intervention group USD 4.25 (0.59) vs usual care USD 4.4 (0.57)	
•Wald 2011 •Medical inpatients at University of Colorado Hospital, USA, aged ≥ 70 years, mean age 81 (SD 7) years •Analysis 19.1 •Cost analysis	•Hospitalist run acute care service for elderly people (N = 122) vs usual hospital inpatient care (N = 95)	•Not stated •US dollar •2007 •6 months	•"Hospital charges"		•Mean "hospital charges" USD 24,617 (SD 15,828) intervention vs USD 21,488 (SD 13,407) usual care, P = 0.12	
•Zermansky 2006 •Residents of 65 nursing care facilities in Leeds, UK taking ≥ 1 medicines, mean age 85 (interquartile range	•Clinical medication review by pharmacist (N = 331) vs usual general practitioner care (N = 330)	•Not stated •Pound sterling •2003 •6 months	• Pharmaceutical use		•Mean medication cost per patient per 28 days medication review group GBP 42.24 (SD 38.33) vs GBP 42.95	

(*Continued*)

80 to 91) years •Analysis 5.1, Analysis 5.2 •Cost analysis					(SD 41.01) control group, mean difference GBP -0.70 (95% CI -7.28 to 5.71)	

CMAI: Cohen-Mansfield agitation inventory
NS: difference between groups not statistically significant
QALY: quality adjusted life year
SD: standard deviation

Appendix 11. Additional detail for other identified systematic reviews including meta-analyses

Additional detailed discussion of comparisons of the current review with other identified systematic reviews is provided.

Exercise

Lee 2017 included 21 studies of exercise in care facilities, 15 with exercise as a single intervention, six with exercise combined with one or more interventions. Data were pooled from studies comparing exercise with other interventions, usual care or placebo. In the current review, comparisons of alternate exercise programmes were not pooled with trials of exercise in comparison with usual care. Lee 2017 reported that pooled data from all trials showed a decrease in the rate of falls (RaR 0.81, 95% CI 0.68 to 0.97) but not risk of falling (RR 0.93, 95% CI to 0.86 to 1.01). When exercise was combined with other falls interventions (which were considered as multifactorial interventions in our review) the effect on the rate of falls was greater (RR 0.61, 95% CI 0.52-0.72) and there was a reduction in the risk of falling (RR 0.85, 95% CI 0.77 to 0.95). Post-hoc analysis in Lee 2017 indicated that gait, balance, and functional training with mechanical devices (two studies, Shimada 2004; Sihvonen 2004) reduced the rate of falls. The current review has pooled gait, balance, and functional training with mechanical devices in Sihvonen 2004 with the functional walking arm of Faber 2006, Kerse 2008 and the Sakamoto 2006 one-leg standing arm as interventions of gait, balance, and functional training compared to usual care and found no change in the rate of falls. A post-hoc analysis of balance and strength training in Lee 2017 that shows a reduction in the rate of falls also considers different studies within this category to the current review.

Vitamin D supplementation

Le Blanc 2015, in a systematic review examining trials conducted in both institutionalised or community settings, found that vitamin D significantly reduced the number of falls per person (5 trials, RR 0.66, 95%CI 0.50 to 0.88) but did not significantly reduce the risk of falling (5 trials, RR 0.84, 95%CI 0.69 to 1.02, I^2 = 70%), consistent with the findings in care facilities in this review. The authors found subgroup analyses based on institutionalisation, baseline 25-hydroxyvitamin D level, vitamin D dosage study duration and age did not explain the heterogeneity in the risk of falling outcome. Heterogeneity was reduced to zero when two studies treating with a combination of vitamin D and calcium were excluded; vitamin D treatment alone decreased the risk of falling (3 studies, RR 0.65, 95%CI 0.51 to 0.81, I^2 = 0%). The two included studies conducted in institutionalised settings are included in this review. The other trials included patients of an older age (>70 years), with mobility problems or multiple co-morbidities. Pooled analysis of four trials and one nested case-control study did not find a significant effect on the risk of any fracture (RR 0.98 95%CI, 0.82 to 1.16, I^2 = 32%) or hip fracture (4 trials; RR, 0.96 95%CI, 0.72 to 1.29, I^2 = 46%).
Bolland 2014, pooled outcomes from six randomised trials conducted in care facilities or hospitals and found no significant reduction in falls with vitamin D supplementation, with or without calcium supplementation (RR 0.96, 95%CI 0.88 to 1.05). These authors conducted a sequential analysis of trials in any setting and considered that a risk reduction of less than 15% was not clinically relevant for an individual, but also considered a threshold of 10% as a sensitivity analysis. It was proposed that smaller treatment benefits are unlikely to be considered attractive to an individual. It was concluded that supplementation with vitamin D does not reduce risk of falling by 15% or more and that future trials are unlikely to alter this conclusion. One study included as institutional in the Bolland 2014 review was excluded from this review as 51% of participants were residing in the community (Graafmans 1996); all other studies were included in this review. The Bolland 2014 review focused on analysis of falls risk but also acknowledges that it is useful to consider

the rate of falls from a public health perspective due to plausible effects on multiple fallers. The authors conducted a secondary analysis of rate of falls of studies conducted in any setting, and did not consider pooling to be appropriate due to high heterogeneity (I^2 = 92%). This Cochrane Review has focused on studies conducted in care facilities or hospitals and found that whilst vitamin D supplementation did not reduce the risk of falling, it did reduce the rate of falls in care facilities. Our analysis included data on the rate of falls from the same four studies pooled for the risk of falling and whilst there was heterogeneity for the pooled rate of falls outcome (I^2 = 62%), it was lower than observed in Bolland 2014 for studies overall. Consideration of the acceptability of the intervention should be explored in a cost-effectiveness analysis and/or discrete choice experiments to gain insight into consumer preferences.

Appendix 12. Contribution of authors for the first version of this review

Contribution of authors for the first version of this review

Ian Cameron and Lesley Gillespie initiated splitting the previous review, entitled *'Interventions for preventing falls in elderly people'*, into separate reviews for older people living in the community and for older people in nursing care facilities and hospitals. The protocol was adapted by Geoffrey Murray from the previous review with guidance from Lesley Gillespie and Ian Cameron. All authors then met to finalise the protocol before preparation by Geoffrey Murray. Geoffrey Murray was primarily responsible for locating studies, and both he and Ian Cameron decided independently and then by consensus which studies met inclusion criteria. All seven authors assessed quality and extracted data from included studies. Keith Hill adjudicated differences in quality assessments and data in most studies and Geoffrey Murray adjudicated the others. Geoffrey Murray prepared the drafts and did the primary data entry and analysis into RevMan. Lesley Gillespie and Clare Robertson provided guidance with this process. Clare Robertson prepared the generic inverse data for entry into RevMan. All authors commented on re-analyses and revisions at all stages. Ian Cameron is the guarantor of the review.

Contribution of authors for the 2012 update of this review

Ian Cameron, the guarantor of the review, conceived and designed the review and for this update carried out 'Risk of bias' assessment and data extraction, assisted with categorisation of trial interventions using the ProFaNE taxonomy, and commented on drafts of the review. Lesley Gillespie conceived the review and for this update co-ordinated the review, modified the search strategies, carried out the searches,
screened search results and obtained papers, screened retrieved papers against inclusion criteria, carried out 'Risk of bias' assessment and data extraction, entered data into RevMan, and wrote the review. Clare Robertson carried out 'Risk of bias' assessment and data extraction for all newly included trials, managed data and carried out statistical calculations, wrote the economic evaluation section and Appendix 9, and wrote the review. Geoff Murray conceived and designed the review, and for this update screened retrieved papers against inclusion criteria, updated the Characteristics of included studies table, Appendix 3, Appendix 4 and Appendix 5, assisted with categorisation of trial interventions using the ProFaNE taxonomy, and commented on drafts of the review. Keith Hill carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review. Robert Cumming carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review. Ngaire Kerse carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review.

WHAT'S NEW

Last assessed as up-to-date: 3 August 2017.

Date	Event	Description
7 September 2018	Amended	NIHR acknowledgement added

HISTORY

Protocol first published: Issue 3, 2005

Review first published: Issue 1, 2010

Date	Event	Description
10 August 2018	New search has been performed	For this update, the following changes were made. • Search updated to August 2017. • 35 new trials added. • The classification of social environment interventions has been reconsidered. Stenvall 2007 has been reclassified as a social environment intervention (previously multifactorial). Koh 2009 and Van Gaal 2011b are still classified within the social environment ProFaNE category but considered as organisational service model change rather than staff training. • Trials on medication review in care facilities reclassified according to medication target, rather than according to the type of health professional performing the review. • Additional subgroup analysis by level of care conducted for multifactorial interventions in hospitals. • Background section revised and citations updated. • Risk of bias conducted for additional items for previously included trials according to current Cochrane guidelines. • Overall quality of evidence for main comparisons assessed according to GRADE. • A new cost-effectiveness analysis of Haines 2011 (Haines 2013) has been added. • Exercise interventions are reported according to the ProFaNE exercise category and the comparator arm of the trial.
10 August 2018	New citation required and conclusions have changed	New evidence, the reclassification of some intervention categories and the implementation of new methods, including assessment of the quality of the evidence using GRADE, has resulted in some changed conclusions Changes made to authorship, including addition of new authors
27 February 2013	Feedback has been incorporated	Changes relate to two pieces of feedback, received 19 December 2013 and 12 February 2013. Two Summary [of feedback] and Reply entries were added to the Feedback section. There were no changes to the review in relation to the second piece of feedback. Changes

(*Continued*)

		in relation to the first piece included: 1. Appendix 6 was revised and Appendices 7 and 8 were deleted. 2. A new Appendix 7, containing raw data, was added. 3. Sections of the review (principally, the 'Description of studies') were revised to reflect these changes
9 November 2012	New search has been performed	For this update, published in Issue 12, 2012, the following changes were made: 1. Search updated to March 2012 2. Twenty additional trials (35,270 participants) included in this update 3. One previously included trial recruiting people post stroke (Barreca 2004) excluded, as no longer within the scope of this version of the review 4. Kerse 2008 reclassified as an exercise intervention (formerly multifactorial) 5. Additional trials testing multifactorial interventions with results for subgroups with and without cognitive impairment 6. Evidence relating to additional interventions, these include: patient education in hospital (Ang 2011; Haines 2011), dementia care mapping (Chenoweth 2009), motion sensors (Clifton 2009), decision-support software (Dykes 2010; Lapane 2011), multivitamin supplementation (Grieger 2009), low-low beds (Haines 2010), multisensory stimulation (Klages 2011), guideline implementation (Koh 2009; Van Gaal 2011a; Van Gaal 2011b), a fall risk assessment tool (Meyer 2009), increased sunlight exposure (Sambrook 2012), lavender oil stimulation (Sakamoto 2012), an acute care service for elderly people (Wald 2011) 7. One newly included trial included a cost-effectiveness analysis (Chenoweth 2009) 8. Background section revised and citations updated 9. 'Risk of bias' item relating to 'Allocation concealment' split into two: 'Sequence generation' and 'Allocation concealment' and applied to all included studies 10. Subgroup analyses revised
9 November 2012	New citation required and conclusions have changed	1. In response to the external referee's comments, the title of this review has been changed to reflect the fact that facilities which do not include nursing care are also included in this review. 2. Change in conclusion for multifactorial interventions in care facilities from no evidence of effect to a suggestion of possible benefits. Evidence from one trial

(*Continued*)

		for the effectiveness of an educational session targeting identified risk factors in acute hospital setting
30 November 2009	Amended	Correction of two minor errors
23 September 2009	Amended	The published review 'Interventions for preventing falls in elderly people' (Gillespie 2003) is not being updated. Due to its size and complexity it was split into two reviews: 'Interventions for preventing falls in older people living in the community' and 'Interventions for preventing falls in older people in nursing care facilities and hospitals'
1 April 2009	Amended	Converted to new review format

CONTRIBUTIONS OF AUTHORS

ID Cameron, the guarantor for this review, conceived and designed the review and for this update contributed to assessment of retrieved studies against inclusion criteria, carried out 'Risk of bias' assessment, data extraction and assessment of GRADE quality of the evidence, assisted with categorisation of trial interventions using the ProFaNE taxonomy, and commented on drafts of the review.

SM Dyer for this update co-ordinated the review, carried out trial registry searches, screened search results and obtained papers, screened retrieved papers against inclusion criteria, carried out 'Risk of bias' assessment, data extraction and assessment of GRADE quality of the evidence, managed data and carried out statistical calculations, entered data into Review Manager, and drafted the review.

CE Panagoda screened search results and obtained papers, screened retrieved papers against inclusion criteria, carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review.

GR Murray carried out 'Risk of bias' assessment, data extraction and assessment of GRADE quality of the evidence, assisted with categorisation of trial interventions using the ProFaNE taxonomy, and commented on drafts of the review.

KD Hill carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review.

RG Cumming carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review.

N Kerse carried out 'Risk of bias' assessment and data extraction, and commented on drafts of the review.

See Appendix 12 for 'Contribution of authors' for the previous version of this review.

DECLARATIONS OF INTEREST

Four review authors were investigators for seven included studies: ID Cameron (Cumming 2008; Sambrook 2012) and RG Cumming (Barker 2016; Cumming 2008; Sambrook 2012); KD Hill (Barker 2016; Haines 2004; Haines 2011); N Kerse (Kerse 2004; Kerse 2008). Authors did not assess risk of bias in their own trials.

SOURCES OF SUPPORT

Internal sources

- John Walsh Centre for Rehabilitation Research, Kolling Institute, The University of Sydney, Australia.

Salary, administration, computing, and library services (IDC, RGC)

- Illawarra Shoalhaven Local Health Network, Warrawong, Australia.

Computing and library services (GM)

- Curtin University, Perth, Australia.

Salary, administration, computing, and library services (KDH)

- University of Auckland, New Zealand.

Salary, administration, computing and library services (NK)

External sources

- National Health and Medical Research Council, Practitioner Fellowship, Australia.

Salary contribution (IDC)

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Latest update

'Risk of bias' assessment

In this version of the review, we now exclusively assess risk of bias of each included study based on the recommended tool described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b). We also assessed bias in the recall of falls due to less reliable methods of ascertainment (Hannan 2010).

Assessing the quality of the evidence and 'Summary of findings' tables

We now use the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the quality of the body of evidence. We prepared 'Summary of findings' tables for each of the main categories of interventions, for the listed outcomes. The risk of bias has been assessed according to the Cochrane tool for assessing risk of bias, plus two items relating to method of ascertaining falls and baseline imbalance.

Data synthesis

Where the reported trial outcomes did not include falls during the intervention period, we did not pool these data with those of other trials.

Subgroup analysis and investigation of heterogeneity

In addition to subgroup analyses by intervention types according to the Prevention of Falls Network Europe (ProFaNE) fall-prevention taxonomy (Lamb 2007; Lamb 2011), we conducted sensitivity analyses of exercise trials excluding those with 20 participants or less in each arm of the trial. We also conducted a sensitivity analysis of medication review excluding one trial with three participants with more than 30 falls in the intervention arm of the trial. In the previous version of this review, subgroup analyses were conducted according to level of cognition and level of care in care facilities. In this update, we have added subgroup analysis by level of care (setting) in hospitals. We have conducted a sensitivity analysis for the rate of falls analysis for exercise versus usual care in care facilities to test the exclusion of one trial with zero falls recorded in the intervention arm of the trial.

Reconsideration of categorisation of some interventions according to ProFaNE

Upon further consideration, we have re-categorised some interventions across different ProFaNE categories that fall within the social environment classification. Stenvall 2007 has been reclassified as a social environment intervention (previously multifactorial). Koh 2009 and Van Gaal 2011b remain classified within the social environment ProFaNE category but are considered as organisational service model change rather than staff training as these interventions are primarily to introduce new guidelines and staff training was secondary.

Update in 2012

Criteria for considering studies for this review

Trials including only participants after stroke were excluded as a protocol for a Cochrane Review on interventions for preventing falls in people after stroke has been published (Verheyden 2010).

Separation of analyses by setting

We reported the results for care facilities and hospitals separately as the primary analyses because this is likely to be more useful to the users of this review. Interventions will be organised differently in these two types of settings and there may be different effectiveness of similar interventions between the two settings.

'Risk of bias' assessment

The protocol was completed and submitted for publication prior to the general release of RevMan 5 and the supporting version of the *Cochrane Handbook for Systematic Reviews of Interventions* (version 5.0) in February 2008. In the protocol, we stated that we would assess methodological quality using the 11-item tool used in Gillespie 2003.
For this version of the review, we used three criteria from the Cochrane tool for assessing risk of bias: 'Random sequence generation', 'Allocation concealment', and 'Blinding of outcome assessment', and eight items from the 11-item tool (*see* Appendix 2). The items relating to allocation concealment and blinding of outcome assessors have not been used (now redundant). Also, the item relating to appropriateness of duration of clinical surveillance was not used due to very poor agreement between assessors during preparation of the first version of this review.

Other changes

Interventions were classified using the Prevention of Falls Network Europe (ProFaNE) fall-prevention taxonomy (Lamb 2007; Lamb 2011). Subgroup analyses were conducted to explore heterogeneity where appropriate.

INDEX TERMS

Medical Subject Headings (MeSH)

*Hospitals [statistics & numerical data]; *Nursing Homes [statistics & numerical data]; Accidental Falls [*prevention & control; statistics & numerical data]; Calcium, Dietary [administration & dosage]; Exercise; Randomized Controlled Trials as Topic; Safety Management; Vitamin D [administration & dosage]; Vitamins [administration & dosage]

MeSH check words

Aged; Aged, 80 and over; Female; Humans; Male

Cochrane Library

Cochrane Database of Systematic Reviews

Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews (Review)

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH.
Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews.
Cochrane Database of Systematic Reviews 2017, Issue 4. Art. No.: CD011279.
DOI: 10.1002/14651858.CD011279.pub3.

www.cochranelibrary.com

TABLE OF CONTENTS

[Overview of Reviews]

Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews

Louise J Geneen[1], R Andrew Moore[2], Clare Clarke[3], Denis Martin[4], Lesley A Colvin[5], Blair H Smith[1]

[1]Division of Population Health Sciences, University of Dundee, Dundee, UK. [2]Pain Research and Nuffield Department of Clinical Neurosciences (Nuffield Division of Anaesthetics), University of Oxford, Oxford, UK. [3]Division of Population Health Sciences, University of Dundee, Dundee, UK. [4]Institute of Health and Social Care, Teesside University, Middlesbrough, UK. [5]Anaesthesia & Pain Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK

Contact address: Louise J Geneen, Division of Population Health Sciences, University of Dundee, Dundee, UK. louise.geneen@dph.ox.ac.uk.

Editorial group: Cochrane Pain, Palliative and Supportive Care Group.
Publication status and date: Edited (no change to conclusions), published in Issue 4, 2017.

Citation: Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. *Cochrane Database of Systematic Reviews* 2017, Issue 4. Art. No.: CD011279. DOI: 10.1002/14651858.CD011279.pub3.

ABSTRACT

Background

Chronic pain is defined as pain lasting beyond normal tissue healing time, generally taken to be 12 weeks. It contributes to disability, anxiety, depression, sleep disturbances, poor quality of life, and healthcare costs. Chronic pain has a weighted mean prevalence in adults of 20%.

For many years, the treatment choice for chronic pain included recommendations for rest and inactivity. However, exercise may have specific benefits in reducing the severity of chronic pain, as well as more general benefits associated with improved overall physical and mental health, and physical functioning.

Physical activity and exercise programmes are increasingly being promoted and offered in various healthcare systems, and for a variety of chronic pain conditions. It is therefore important at this stage to establish the efficacy and safety of these programmes, and furthermore to address the critical factors that determine their success or failure.

Objectives

To provide an overview of Cochrane Reviews of adults with chronic pain to determine (1) the effectiveness of different physical activity and exercise interventions in reducing pain severity and its impact on function, quality of life, and healthcare use; and (2) the evidence for any adverse effects or harm associated with physical activity and exercise interventions.

Methods

We searched the *Cochrane Database of Systematic Reviews* (CDSR) on the Cochrane Library (CDSR 2016, Issue 1) for systematic reviews of randomised controlled trials (RCTs), after which we tracked any included reviews for updates, and tracked protocols in case of full review publication until an arbitrary cut-off date of 21 March 2016 (CDSR 2016, Issue 3). We assessed the methodological quality of the reviews using the AMSTAR tool, and also planned to analyse data for each painful condition based on quality of the evidence.

We extracted data for (1) self-reported pain severity, (2) physical function (objectively or subjectively measured), (3) psychological function, (4) quality of life, (5) adherence to the prescribed intervention, (6) healthcare use/attendance, (7) adverse events, and (8) death.

Due to the limited data available, we were unable to directly compare and analyse interventions, and have instead reported the evidence qualitatively.

Main results

We included 21 reviews with 381 included studies and 37,143 participants. Of these, 264 studies (19,642 participants) examined exercise versus no exercise/minimal intervention in adults with chronic pain and were used in the qualitative analysis.

Pain conditions included rheumatoid arthritis, osteoarthritis, fibromyalgia, low back pain, intermittent claudication, dysmenorrhoea, mechanical neck disorder, spinal cord injury, postpolio syndrome, and patellofemoral pain. None of the reviews assessed 'chronic pain' or 'chronic widespread pain' as a general term or specific condition. Interventions included aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi.

Reviews were well performed and reported (based on AMSTAR), and included studies had acceptable risk of bias (with inadequate reporting of attrition and reporting biases). However the quality of evidence was low due to participant numbers (most included studies had fewer than 50 participants in total), length of intervention and follow-up (rarely assessed beyond three to six months). We pooled the results from relevant reviews where appropriate, though results should be interpreted with caution due to the low quality evidence.

Pain severity: several reviews noted favourable results from exercise: only three reviews that reported pain severity found no statistically significant changes in usual or mean pain from any intervention. However, results were inconsistent across interventions and follow-up, as exercise did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point.

Physical function: was the most commonly reported outcome measure. Physical function was significantly improved as a result of the intervention in 14 reviews, though even these statistically significant results had only small-to-moderate effect sizes (only one review reported large effect sizes).

Psychological function and quality of life: had variable results: results were either favourable to exercise (generally small and moderate effect size, with two reviews reporting significant, large effect sizes for quality of life), or showed no difference between groups. There were no negative effects.

Adherence to the prescribed intervention: could not be assessed in any review. However, risk of withdrawal/dropout was slightly higher in the exercising group (82.8/1000 participants versus 81/1000 participants), though the group difference was non-significant.

Healthcare use/attendance: was not reported in any review.

Adverse events, potential harm, and death: only 25% of included studies (across 18 reviews) actively reported adverse events. Based on the available evidence, most adverse events were increased soreness or muscle pain, which reportedly subsided after a few weeks of the intervention. Only one review reported death separately to other adverse events: the intervention was protective against death (based on the available evidence), though did not reach statistical significance.

Authors' conclusions

The quality of the evidence examining physical activity and exercise for chronic pain is low. This is largely due to small sample sizes and potentially underpowered studies. A number of studies had adequately long interventions, but planned follow-up was limited to less than one year in all but six reviews.

There were some favourable effects in reduction in pain severity and improved physical function, though these were mostly of small-to-moderate effect, and were not consistent across the reviews. There were variable effects for psychological function and quality of life.

The available evidence suggests physical activity and exercise is an intervention with few adverse events that may improve pain severity and physical function, and consequent quality of life. However, further research is required and should focus on increasing participant numbers, including participants with a broader spectrum of pain severity, and lengthening both the intervention itself, and the follow-up period.

PLAIN LANGUAGE SUMMARY

Physical activity and exercise for chronic pain in adults - an overview of Cochrane Reviews

Background

Chronic (long-term) pain is pain that has lasted beyond the body's usual healing time. It is often described as pain that has lasted for at least three months. Chronic pain causes many problems, beyond the pain itself, including fatigue, anxiety, depression, and a poor quality of life.

In the past, people with chronic pain were told to rest. However, general advice now is to keep active - whether to affect the pain directly or to combat the other problems associated with it. Therefore, research studies have attempted to examine the effect of physical activity in people with chronic pain.

This overview aimed to bring together and analyse any reviews published by Cochrane that looked at physical activity and exercise studies in any chronic pain condition, including arthritis, back and neck pain, and menstrual (period) pain.

Key results and quality of the evidence

In January 2016, we identified 21 Cochrane Reviews which covered 10 different diagnoses (osteoarthritis (a joint disease), rheumatoid arthritis (joint pain and swelling), fibromyalgia (widespread pain condition), low back pain, intermittent claudication (cramping pain in the legs), dysmenorrhoea (period pain), mechanical neck disorders (neck pain), spinal cord injury, postpolio syndrome (a condition occurring in people who have had polio), patellofemoral pain (pain at the front of the knee)). The physical activity or exercise programme used in the trials ranged in frequency, intensity, and type, including land- and water-based activities, those focusing on building strength, endurance, flexibility and range of motion, and muscle activation exercises.

The quality of the evidence was low. This was mostly due to the small numbers of people with chronic pain who participated in each reviewed study. Ideally, a study should have hundreds of people assigned to each group, whereas most of the studies included in the review process here had fewer than 50 people in total.

There was evidence that physical activity reduced the severity of pain, improved physical function, and had a variable effect on both psychological function and quality of life. However, these results were not found in all studies. The inconsistency could be due to the quality of the studies or because of the mix of different types of physical activity tested in the studies. Additionally, participants had predominantly mild-to-moderate pain, not moderate-to-severe pain.

Conclusions

According to the available evidence (only 25% of included studies reported on possible harm or injury from the intervention), physical activity did not cause harm. Muscle soreness that sometimes occurs with starting a new exercise subsided as the participants adapted to the new activities. This is important as it shows physical activity in general is acceptable and unlikely to cause harm in people with chronic pain, many of whom may have previously feared it would increase their pain further.

Future studies should focus on increasing participant numbers, including a wider range of severity of pain (more people with more severe pain), and lengthening both the intervention (exercise programme) itself, and the follow-up period. This pain is chronic in nature, and so a long-term intervention, with longer periods of recovery or follow-up, may be more effective.

BACKGROUND

Description of the condition

Chronic pain has been defined as pain lasting beyond normal tissue healing time, generally taken to be 12 weeks (International Association for the Study of Chronic Pain; Merskey 2011). It contributes to disability, anxiety and depression, sleep disturbances, poor quality of life, and healthcare costs (Leadley 2014; Moore 2014a; Park 2012).

Chronic pain has a weighted mean prevalence in adults of 20% (Breivik 2006; Moore 2014a), which increases as the population ages (32% of adults aged 25 to 34 years, 62% of adults over 75

years; Abdulla 2013; Elliott 1999). This is a greater proportion than people with asthma (To 2012) or diabetes (IDF 2012) in the same population (van Hecke 2013a). The World Health Organization (WHO) recognises chronic pain as a public health problem throughout the world, with one systematic review assessing the growing evidence that the prevalence of chronic pain in the general population is high internationally (34% in low-income countries and 30% in high-income countries; Elzahaf 2012). Chronic painful conditions comprise four of the 10 highest ranking conditions for years lived with disability in 2013 (Vos 2015), and are responsible for considerable loss of quality of life and employment, and increased healthcare costs (Moore 2014b). Despite this, the term 'chronic pain' was only added as a MeSH term in MEDLINE in January 2012 (National Library of Medicine), highlighting the relatively small proportion of specific research dedicated to this population.

Certain factors can contribute to an increased risk of chronic pain (female gender, older age, lower socioeconomic status, geographical and cultural background, and genetics; Smith 2007; van Hecke 2013b). Other factors associated with chronic pain conditions are modifiable, such as smoking status, alcohol intake, nutrition, obesity, comorbidities, employment status and occupational factors, and physical activity level (Smith 2007; van Hecke 2013a).

A review of current issues in the treatment of chronic pain strongly suggests that health professionals traditionally focus on biomedical views of pain, utilising pharmacology first and foremost, and sometimes not addressing potential non-pharmacological approaches such as physical activity and changing attitudes towards chronic pain (Schofield 2011). Guidance often suggests that lifestyle advice is important: for example, the National Institute for Health and Care Excellence (NICE) osteoarthritis guidelines state that "exercise should be a core treatment ... irrespective of age, comorbidity, pain severity and disability. Exercise should include: local muscle strengthening [and] general aerobic fitness" (NICE 2014). Non-pharmacological treatments have been developed, investigated, and implemented, with Cochrane Reviews and protocols evaluating the available evidence for psychological, physical, and other non-medical interventions (e.g. cognitive behavioural and behavioural therapy, Eccleston 2014; Williams 2012; TENS, Nnoaham 2008; low-impact/intensity movement/exercise therapy, Wieland 2013; dietary, Straube 2015; and patient education, Engers 2008; Gross 2009). While evidence for the effectiveness of these interventions is of variable quantity and quality, the 2013 Scottish Intercollegiate Guideline Network (SIGN) guidelines on the management of chronic pain made strong recommendations on the use of exercise, based on evidence drawn from randomised controlled trials (RCTs), stating: "exercise and exercise therapies, regardless of their form, are recommended in the management of patients with chronic pain" (SIGN 2013).

Description of the interventions

Physical activity has been defined by the WHO as "any bodily movement produced by skeletal muscles that requires energy expenditure, including activities undertaken while working, playing, carrying out household chores, travelling, and engaging in recreational pursuits" (WHO 2015). WHO also states that "exercise ... is a sub-category of physical activity that is planned, structured, repetitive, and aims to improve or maintain one or more components of physical fitness" (WHO 2015).

Physical activity for health can take many different forms: it can be structured exercise, such as in classes, gym-based, or a DVD or programme performed at home; or unstructured and involve adding just a few small activities each day (activities of daily living). Physical activity and exercise can also vary in intensity, duration, and type: aerobic (such as walking) or more focused on increasing flexibility, strength, or balance. Physical activity and exercise can also be taught (or led) by another individual such as an exercise professional, or initiated and maintained through the person's own initiative and motivation.

Both physical activity and exercise can be performed on land or in the water, and can range from whole-body to localised (body site-specific) training. Most forms of exercise can also be modified to be performed where there is restricted movement (e.g. in a chair, a bed, or another assistive device).

How the intervention might work

Physical activity and exercise can be adapted for an individual, and is something people can do to help themselves. It is likely to be associated with minimal adverse effects, such as interactions with medication and potential for abuse in adults with chronic pain, when compared to pharmaceutical and surgical interventions. It is therefore an attractive option to help manage an individual's pain if the systematic reviews show benefit. However, current evidence suggests that simply giving an individual advice to exercise is insufficient to bring about significant change (SIGN 2013), and a badly prescribed intervention that does not consider the individual's conditions and present state of health and fitness, such as one that does not incorporate pacing or gradual progression, may bring about adverse events such as pain 'flare-ups', or lead to cardiac or respiratory events (American College of Sports Medicine 2007). This suggests that supervised or structured interventions may be more fruitful, though this is currently unconfirmed.

Since the 1980s, primary care physician advice for treating pain has changed, moving away from "rest", to minimising or eliminating bedrest and instead remaining active (back pain, Waddell 1987). Exercise may have specific benefits in reducing the severity of chronic pain, as well as more general benefits associated with improved overall physical and mental health, and physical functioning of people with chronic pain, as depression (Finan 2013), deconditioning (Bousema 2007), and obesity are commonly observed in these people (headache/migraine, Bigal 2012; fibromyalgia, Ursini 2011). For example, studies have revealed that a sin-

gle bout of exercise increases the production of endogenous opioids, leading to transient anti-nociception in both animals and humans, and repeated exercise produces long-lasting anti-nociception in otherwise untreated animals (Stagg 2011). Aerobic exercise is also strongly linked to weight loss (Messier 2013), which in turn has implications for the management of chronic pain as the pressure on joints is reduced. Alternatively, resistance exercise, or other forms of strength training, can improve the person's capacity to support bone and cartilage through improved musculature supporting movement around a joint, with potential to relieve stiffness (Mayer 2008) and bringing about some pain relief. Resistance training through repetitive full range-of-motion exercise around the lumbar spine (in chronic low back pain) may affect disc metabolism itself, with the possibility that the exercise programme could improve metabolic exchange in the lumbar discs and aid in repair (Mooney 2006). Training to improve balance and flexibility also has benefits as it reduces the risk of falls, and the potential for further pain or injury (Harvard 2013).

Why it is important to do this overview

If physical activity and exercise interventions are shown to effectively and safely reduce pain intensity or frequency (or both), they are likely to be a preferable alternative or adjunct therapy to pharmacological/surgical treatments for chronic pain. The interventions could promote personal involvement of individuals in the management of their pain, thus increasing self-efficacy and the ability to self-manage. In turn this could lead to an increase in overall quality of life and a consequent reduction in healthcare use. In addition, exercise is of great importance for cardiovascular (Vigorito 2014) and bone health (Sakuma 2012). Reduced physical function and consequent lack of mobility in people with chronic pain is associated with increased all-cause and cardiovascular mortality (Nüesch 2011), with other studies linking severe chronic pain to general increased all-cause mortality (Moore 2014a; Torrance 2010).

Physical activity and exercise programmes are increasingly being promoted and offered in various healthcare systems (American College of Sports Medicine (ACSM) 'Exercise is Medicine' global pledge at the Inaugural World Congress 2010) and for a variety of chronic pain conditions, including arthritis (Fransen 2014; Silva 2010), fibromyalgia (Busch 2013), and dysmenorrhoea (Brown 2010). At this stage it is important to establish the efficacy and safety of these programmes, and furthermore to address the critical factors that determine their success or failure.

It is therefore important to identify whether (and how) exercise interventions can be effectively and safely applied in people with chronic pain.

With a number of systematic reviews published by Cochrane evaluating the effectiveness of exercise in various painful conditions, it is timely and important to bring together all relevant published information to evaluate the current evidence, and identify the availability and quality of evidence-based exercise interventions. This overview will determine the extent to which the published systematic reviews have accurately assessed the evidence for exercise in chronic pain conditions/syndromes, which will help to direct future guidelines and identify current research gaps.

OBJECTIVES

To provide an overview of Cochrane Reviews of adults with chronic pain to determine (1) the effectiveness of different physical activity and exercise interventions in reducing pain severity and its impact on function, quality of life, and healthcare use; and (2) the evidence for any adverse effects or harm associated with physical activity and exercise interventions.

METHODS

Criteria for considering reviews for inclusion

We included only systematic reviews of RCTs of physical activity and exercise in participants with chronic pain, and published in the *Cochrane Database of Systematic Reviews*. The included reviews had to fulfil the following criteria:

Participants

Adults (aged 18 years and over) reporting chronic non-cancer pain, including persistent (e.g. chronic back pain, fibromyalgia) and intermittent (e.g. migraine, dysmenorrhoea) pain, for at least three months (12 weeks) in any body site.

Intervention

Reviews of RCTs assessing physical activity or exercise as the intervention (any reviews where that assessed physical activity or exercise as a stand-alone intervention). This included physical activity interventions that could be initially taught by an exercise professional, or involve periodical/ongoing supervision.

Exclusions

Interventions not deemed physical activity or exercise using the WHO definition, such as manipulation, mobilisation, or passive movement. Any multi-modal interventions were excluded if physical activity/exercise could not be assessed for effect (the effect of exercise must have been measured distinctly).

Comparison

Usual care, waiting list control, placebo/sham treatment, other treatment, or a combination of treatments (as long as the effect of exercise could be measured distinctly).

Primary outcome

- self-reported pain (severity).

This could be presented and analysed as change on a continuous scale, the proportion of participants who 'responded', or, ideally, in a dichotomised format as the proportion of participants in each group who achieved a predetermined threshold of improvement (e.g. outcome in individual participants of at least 50% pain intensity reduction, or no worse than mild pain, at the end of the trial, with at least 30% pain intensity reduction as a secondary outcome, or recovery; Moore 2013).

Secondary outcomes

- Physical function (objectively or subjectively measured).
- Psychological function.
- Quality of life.
- Adherence to the prescribed intervention.
- Healthcare use/attendance.
- Adverse events (not death).
- Death.

Reviews may not always report specifically on activity or exercise for chronic pain in adults. We anticipated two possible circumstances which might have arisen.

- A review included some interventions of interest or reported only some outcomes of interest. In this case we extracted the interventions and outcomes of interest, but we did not include interventions or outcomes outside the scope of this overview.
- Reviews occasionally included papers that included children and adults together, but the results for adults were not reported or analysed separately in the included papers or the review. In this case we made a judgement as to whether the review could be included based on the proportion of adults. Our intention was to include only those reviews where more than 80% of participants were adults.

Search methods for identification of reviews

We searched the *Cochrane Database of Systematic Reviews* (CDSR), 2016, Issue 1, on the Cochrane Library for relevant reviews using the search strategy: (*pain or migraine or headache*) and (*exercise or activity or physical*). We did not seek non-Cochrane reviews.

Data collection and analysis

Two overview authors (LG, CC) independently carried out searches and selected reviews for inclusion. Disagreements were resolved through discussion, and a third overview author (RAM) acted as arbitrator where necessary.

Two overview authors (independently carried out assessment of methodological quality (LG, CC), and extracted data (LG, RAM). Any disagreements were resolved through discussion, or involving a third overview author if necessary (DM).

One overview author (LG) tracked results of the search for the most up to date version of each review and protocol that fulfilled the inclusion criteria.

Selection of reviews

Included reviews assessed RCTs of the effects of exercise for pain management in adults (as defined by individual reviews), compared with any of the listed comparators, and included:

- a clearly defined clinical question;
- details of inclusion and exclusion criteria;
- details of databases searched and relevant search strategies;
- participant-reported pain severity (primary outcome measure);
- summary results for at least one other desired outcome.

Data extraction and management

Two overview authors (LG, RAM) independently extracted data from the included review using a standardised data extraction form and checked for agreement prior to entry into Microsoft Excel for Windows. We did not extract data from reports included in the reviews again, neither did we undertake any re-analysis of data from reviews. Data were not entered for analysis into Cochrane's statistical software due to the lack of relevant and comparable data (RevMan 2014).

We collected the following information (where available) from the reviews:

- number of included studies and participants;
- intervention (exercise or activity type) and dose (frequency/intensity);
- comparator;
- condition treated;
- time of assessment;
- duration of follow-up;
- relevant outcomes.

Where possible we extracted risk ratio (RR), number needed to treat for an additional beneficial outcome (NNTB), mean difference (MD), and standardised mean difference (SMD), and other relevant statistical data for the primary and secondary outcomes. This included:

- obtaining 50% pain relief (participant-reported);

- obtaining any other measure of 'improvement' (participant-reported);
- adverse events;
- death;
- withdrawals.

Assessment of methodological quality of included reviews

Quality of included reviews

Two overview authors (LG, CC) independently assessed each included review to see if it satisfied the criteria specified in the 'assessment of multiple systematic reviews' (AMSTAR) measurement tool (Shea 2007), for rigorous methodological quality. Arbitration by a third overview author (DM) was necessary for some fields. High quality reviews were required to fulfil each of the established AMSTAR criteria (further criteria to fulfil each field is listed in Table 1).

For each review we also planned to assess the likelihood of publication bias by calculating the number of participants in studies with zero effect (relative benefit of one) that would be needed to give an NNTB too high to be clinically relevant (Moore 2008). In this case we would have considered an NNTB of 10 or greater for the outcome of participant-reported pain relief of 30% or greater to be the cut-off for clinical relevance. This method is used as statistical tests for the presence of publication bias have been shown to be unhelpful (Thornton 2000). However, assessment of publication bias was not possible due to the lack of specificity of the populations included within the reviews, and so we were unable to extract comparable data.

Quality of evidence in included reviews

We planned to use two main indicators for the quality of evidence: all included reviews must have used only primary studies that were both randomised and double-blind, so minimising the risk of bias from these items; and all included reviews must have included only people with at least moderate pain intensity at baseline (visual analogue scale greater than 30/100, categorical rating scale greater than 1/3, and numerical rating scale greater than 3/10, Collins 1997), providing a sensitive assay of intervention efficacy. Subsequently, we planned to analyse data for each painful condition in three tiers, according to outcome and freedom from known sources of bias.

- The first tier used data meeting current best standards, where studies reported the outcome of at least 50% pain intensity reduction from baseline (where 50% was the cut-off for a dichotomous (yes/no) outcome: was a 50% reduction in pain observed?), or its equivalent, without using last observation carried forward (LOCF) or other imputation method for dropouts, reported an intention-to-treat (ITT) analysis, lasted eight or more weeks, had a parallel-group design, and had at least 200 participants (preferably at least 400) in the comparison (Moore 2010). These top-tier results were usually reported first.
- The second tier used any available data, but where one or more of these conditions were not met, for example reporting at least 30% pain intensity reduction, using LOCF or a completer analysis, lasting four to eight weeks, and where the numbers of participants were at least 200.
- A third tier of evidence related to small amounts of data (fewer than 200 participants), or short studies of less than four weeks, or where there was obvious major heterogeneity between studies, or where there were other shortcomings in allocation concealment, considerable attrition, and incomplete outcome data. For this third tier of evidence, no data synthesis was reasonable, and may have been misleading, but an indication of beneficial effects might be possible.

This overview examined the quality of all included reviews according to current best standards for reporting in pain. These included the attempt and ability of the reviews to identify studies/interventions with the maximum evidence of effectiveness, and minimum risk of bias, including the reporting of the following.

- Outcomes in trials of the proportion of participants obtaining at least 50% pain intensity reduction, or no worse than mild pain, at the end of the trial (with at least 30% pain intensity reduction as a secondary outcome). We did not consider the use of mean changes in pain scores as high quality because responses to pain interventions are not Gaussian, and few people have the mean response.
- Duration of included studies of eight weeks or longer.
- Imputation method of baseline observation carried forward (BOCF), LOCF, or worst observation carried forward (WOCF) if adverse event withdrawals were similar in active and control groups.
- At least 200 participants per treatment group in included studies, with at least two trials, as a minimum criterion for trustworthiness of any analysis. Pooled analysis of small studies may be considered good quality if at least 400 participants were involved, but we regarded these as being potentially subject to bias.

We extracted the 'Risk of bias' as assessed by the original review authors from included reviews. Counts of low risk of bias were extracted from relevant studies in the included reviews and tabulated under the following headings to evaluate the proportion of studies achieving a low risk of bias for each:

- random sequence generation (selection bias);
- allocation concealment (selection bias);
- blinding of participants and personnel (performance bias);
- blinding of outcome assessment (detection bias);
- incomplete outcome data (attrition bias);
- selective reporting (reporting bias);
- sample size;

- any other biases.

Data synthesis

Additional quantitative analyses were not required, since we only considered results from properly conducted (Cochrane) reviews. The aim was to concentrate on specific outcomes such as the proportion of participants with at least 50% pain relief, all-cause or adverse event discontinuations, or serious adverse events, and to explore how these can be compared across different treatments for the same condition. We planned to compare only like with like (where possible); for example in study duration, which can be an additional source of bias if insufficient in length (Moore 2010).
However due to the limited data available, we were unable to directly compare and analyse interventions, and have instead reported the evidence qualitatively only. We had also planned to employ subgroup analyses assessing age, condition, and intervention type/intensity, though this was not feasible using the available data from included reviews. For this reason we have also been unable to include a 'Summary of findings' table as planned and stated in the protocol.
Importantly, we have tried to highlight issues of low trial quality, inadequate size, and whether trials were truly valid for the particular condition in making between-therapy comparisons.
We approached each review with four main questions/focus, and extracted data accordingly.

- Did they report exercise versus non-exercise studies?
- Did the review or studies included in the review (or both) have low risk of bias?
- Did they have our main outcome?
- What were the actual intervention/s included in the review?

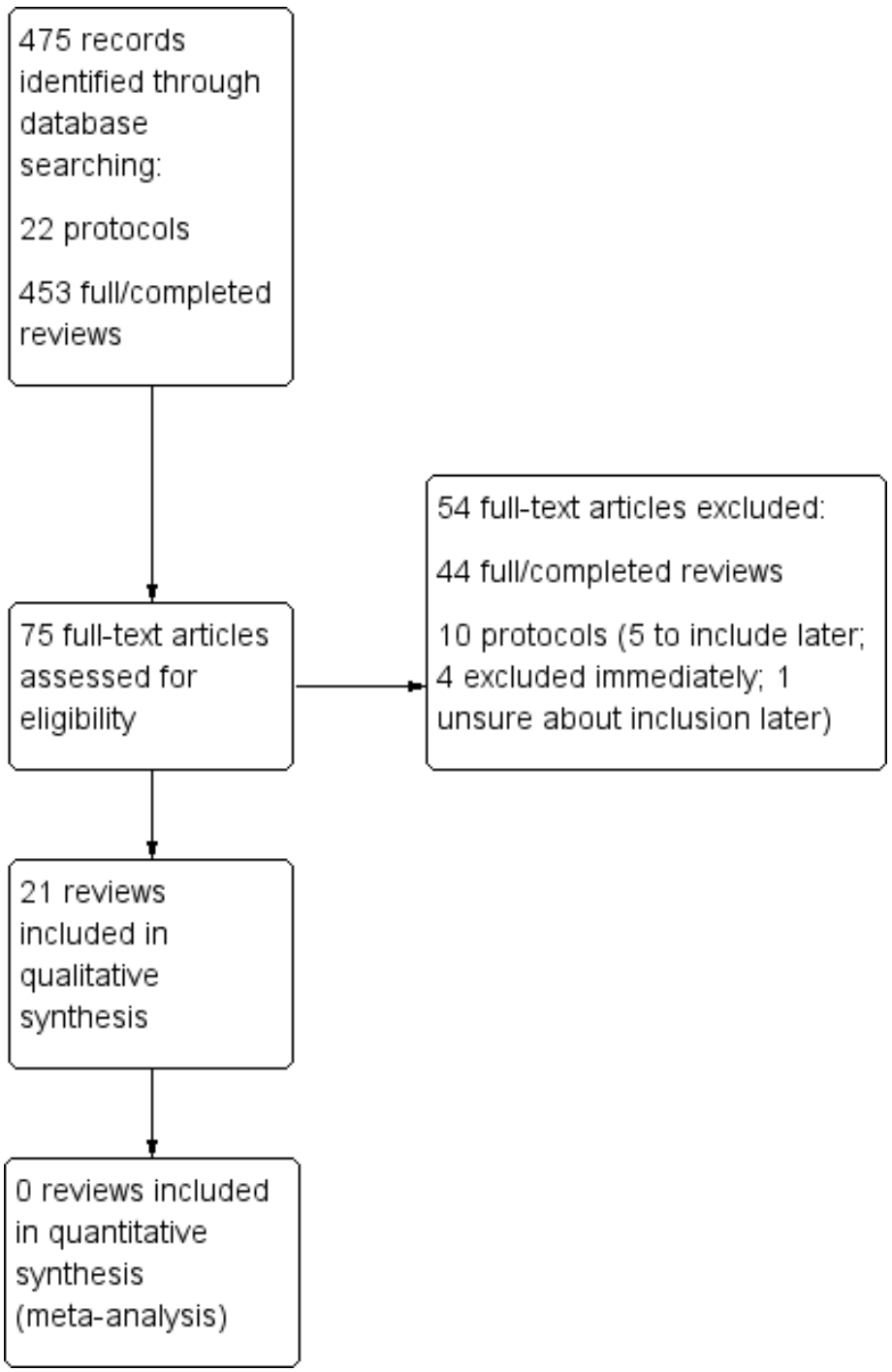
RESULTS

We included 21 reviews with 381 included studies, totalling 37,143 participants. Of these, 264 studies (19,642 participants) examined exercise versus no exercise/minimal intervention in adults with chronic pain (the focus of this overview) and so were used in the qualitative analysis.

Description of included reviews

The search strategy was performed in the Cochrane Library only, and revealed 475 potentially relevant titles, of which 75 were assessed as full papers.
The search was undertaken on 31 January 2016 (CDSR 2016, Issue 1), after which any included reviews were tracked for updates, and protocols were followed in case of full review publication until 21 March 2016 (CDSR 2016, Issue 3).
All extracted data and methodological quality assessment were taken from the most recent published version of the full review.
Ultimately, of the 75 titles requiring further assessment, 10 were reviews at protocol stage only (five of which have potential to be included once published as a full review, one which was unclear, and four that were excluded based on information within the protocol). Hence, we excluded 54 titles (10 protocols and 44 full reviews; Figure 1), reasons for which are listed in Table 2.

Figure 1. Study flow diagram.



Detailed information about the included reviews is available in Table 3. Trial and participant number, age, and gender distribution is reported in Table 4.

Specificity of chronic pain condition of included reviews

Following abstract and full paper assessment, 21 reviews fulfilled the inclusion criteria: four in rheumatoid arthritis (Cramp 2013; Han 2004; Hurkmans 2009; Silva 2010), four in osteoarthritis (Bartels 2007; Fransen 2014; Fransen 2015; Regnaux 2015), three in fibromyalgia (Bidonde 2014; Busch 2007; Busch 2013), three in low back pain (Hayden 2005; Saragiotto 2016; Yamato 2015), two in intermittent claudication (Lane 2014; Lauret 2014), one in dysmenorrhoea (Brown 2010), one in mechanical neck disorder (Gross 2015a), one in spinal cord injury (Boldt 2014), one in postpolio syndrome (Koopman 2015), and one in patellofemoral pain (van der Heijden 2015). None of the included reviews assessed 'chronic pain' or 'chronic widespread pain' as a general term or specific condition.

The 21 included reviews were published by five different Cochrane Review groups: 11 from the Cochrane Musculoskeletal Group (Bartels 2007; Bidonde 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Han 2004; Hurkmans 2009; Regnaux 2015; Silva 2010); four from the Cochrane Neck and Back Group previously the Cochrane Back Group) (Gross 2015a; Hayden 2005; Saragiotto 2016; Yamato 2015); two from the Cochrane Peripheral Vascular Diseases Group (Lane 2014; Lauret 2014); one from the Cochrane Menstrual Disorders and Subfertility Group (Brown 2010); one from the Cochrane Injuries Group (Boldt 2014); one from the Cochrane Neuromuscular Group (Koopman 2015); and one from the Cochrane Bone, Joint and Muscle Trauma Group (van der Heijden 2015).

Protocols that may be included in updates of this overview focus on osteoarthritis (Østerås 2013 from the Cochrane Musculoskeletal Group), migraine (Brønfort 2015 from the Cochrane Pain, Palliative and Supportive Care Group), chronic low back pain (Hayden 2012 from the Cochrane Back Group), ankylosing spondylitis (Regnaux 2014 from the Cochrane Musculoskeletal Group), and temporomandibular disorders (Craane 2006 from the Cochrane Oral Health Group).

Exercise and physical activity interventions implemented in the included reviews

Interventions assessed included: any specified style of land-based exercise or physical activity such as one designed to improve strength, range of movement, aerobic capacity, or a combination of these (Boldt 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hurkmans 2009; Koopman 2015; Regnaux 2015; van der Heijden 2015); a single style of land-based exercise only (tai chi only: Han 2004, walking only: Lauret 2014, walking or jogging only: Brown 2010; Lane 2014, balance training only: Silva 2010, motor control exercise only: Saragiotto 2016, Pilates method only: Yamato 2015); any pool-based or aquatic therapy (Bartels 2007; Bidonde 2014; Cramp 2013), or "any exercise therapy" (Hayden 2005).

Aquatic exercise

Any exercise performed in water. This can include swimming, though many studies will be referring to exercises performed vertically in the water (not horizontally), either using the water to support the body through the exercise, or as resistance against the body.

Range of motion and flexibility exercise

Can be performed in water or on land. The intention is to increase the range of motion around a joint through progressive stretching and mobilising of the muscles around and crossing the joint. For the purposes of this overview, we only included active movement where the movement was brought about by the participant, and not passively moved by an external force such as a therapist.

Aerobic exercise

Can be performed in water or on land. Exercise usually performed continuously to raise the heart rate and breathing rate for a prolonged period. Examples include walking, jogging, running, cycling, and swimming. Often presented as a percentage of the participant's heart rate max (HRmax) - the highest heart rate reached when performing at their absolute maximum. Similarly it may be presented as a percentage of VO_2max or VO_2peak (a proportion of the maximum amount of oxygen the muscle can take up per minute), or as an absolute value (mL/kg/minute).

Strength/resistance exercise

Can be performed in water or on land. Exercise performed against a progressive resistance with the intention of improving muscle strength, muscle endurance, muscle power, or a combination of these. Resistance can come from fixed or free weights, elastic bands, body weight (against gravity), and water resistance. It may also involve static or isometric strength (holding a position or weight without moving against it). Often presented as a percentage of the participant's one repetition maximum (1-RM) - the maximum weight they can lift/move if they only have to do it once.

Motor control exercise

Can be performed in water or on land. Exercise to bring about activation of the deep trunk muscles, targeting the restoration of control and co-ordination of these 'core muscles' (Saragiotto 2016).

Balance (proprioceptive) training

Can be performed in water or on land (water may be used initially for support). Exercise emphasises the maintenance of balance during visual and perturbation challenges with eyes open or closed, range of motion, and maintaining stability over reduced areas of support and unstable surface (Silva 2010), that is improving balance in increasingly unstable situations.

Tai chi

An ancient Chinese discipline developed from martial arts, involving a continuous series of very controlled (and usually slow) movements designed to improve physical and mental wellbeing.

Yoga

Arising out of Hindu philosophy. Exercise includes breath control, simple meditation, and the adoption of specific bodily postures. It is widely practised for health, relaxation, and control (physically and mentally). Incorporates stretching and flexibility training with isometric strength training (holding certain poses, with no movement against a resistance).

Pilates

Developed by Joseph Pilates in the 20th Century, it is a system of exercises (often using special apparatus) designed to improve physical strength, flexibility, and posture, and enhance mental awareness.

Duration and dose (frequency/intensity) of the exercise and physical activity interventions

A detailed breakdown of each review can be seen in Table 5.

Duration of intervention

Interventions assessed by the included reviews varied in length from a single session (Fransen 2015) to 30 months (Fransen 2015). Only five reviews enforced a minimum intervention period to reduce risk of bias, and were able to attribute any effects to the intervention (Brown 2010; Busch 2013; Gross 2015a; Hurkmans 2009; Silva 2010).

Frequency

There was large variation in the exercise or physical activity intervention being implemented, ranging from just once a week (Bidonde 2014; Busch 2007; Fransen 2014; Fransen 2015; Han 2004; Saragiotto 2016), to twice a day (Boldt 2014), and some performing a short series of exercises (two-minute duration) every 15 minutes during the day (Gross 2015a). However, when reported, most included studies in the reviews implemented the programme twice a week (or stated at least twice a week).

Intensity

Few studies quantified the intensity of each session. Baseline intensity was often accepted as low/moderate, with the aim to progress over the intervention period to 70% to 85% of HRmax or heart rate reserve (HRR) for aerobic interventions (Brown 2010; Cramp 2013; Hurkmans 2009), 70% to 80% of an individual's 1-RM, or 50% to 70% maximum voluntary contraction (Koopman 2015) in strength/resistance training programmes (Busch 2013; Hurkmans 2009). In other reviews, intensity was described more loosely as "variable" or "low intensity (very light) to maximum effort (vigorous)" (Bidonde 2014; Fransen 2014; Lane 2014; Regnaux 2015), "low intensity" (Fransen 2014; Gross 2015a; Han 2004; Silva 2010), or "moderate or moderate-to-high" (Cramp 2013; Fransen 2015).

Duration (per session)

Individual sessions varied in length from two minutes (Gross 2015a), to 90 minutes (Busch 2013; Cramp 2013; Han 2004) or 120 minutes (Boldt 2014), but mostly situated around 45 to 60 minutes. However, it is important to note that the shorter sessions were often performed more regularly than longer sessions. With more information it would have been possible to calculate total volume of exercise or physical activity (session duration × frequency per week × number of weeks), for a more accurate and detailed analysis.

Intervention specificity for chronic pain in the included reviews

The focus of this overview was exercise versus no-exercise interventions with the intention of answering the original question: is exercise beneficial, detrimental, or ineffective for people with chronic pain when compared to inactivity? Two of the 21 reviews did not include/locate any studies that examined simply exercise versus no exercise (Lauret 2014; Silva 2010). However, many of the included reviews compared varying exercise modality, duration, intensity, and frequency. The "no-exercise" intervention referred to the control group where there was a minimal intervention (such as sham exercise or education) or wait-list control/no treatment (see Table 3 for more information on control group activity).

Time points reported

Four of the 19 reviews that reported data, reported results at a single time point only ('post-intervention': Bidonde 2014; Busch 2007; Cramp 2013; Han 2004). Reviews also analysed outcome measures immediately post-intervention and at one or more follow-up points. Each review defined short-, intermediate-, and long-term follow-up according to their own assessment, so when the time period was not mentioned explicitly, we grouped the reviews according to the review authors' own classification only, and where a time period (weeks, month, years) was explicitly listed but not defined by the authors, we grouped them as short-term (follow-up as under six months), intermediate-term (six to 12 months), and long-term (longer than 12 months): short-term: Busch 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Lane 2014; Regnaux 2015; Saragiotto 2016; intermediate-term: Bartels 2007; Fransen 2015; Gross 2015a; Hayden 2005; Lane 2014; Regnaux 2015; Saragiotto 2016; long-term: Gross 2015a; Hayden 2005; Regnaux 2015; Saragiotto 2016. Five reviews did not report "post-intervention" but at short-term, mid/intermediate-term, and long-term postrandomisation (short, mid, and long term: Boldt 2014; short and intermediate term: Koopman 2015; Yamato 2015; short and long-term: Hurkmans 2009; van der Heijden 2015). One review assessed participants in an ongoing fashion "over three menstrual cycles" (Brown 2010).

Long-term follow-up

Of the seven reviews claiming to report "long term" follow-up, one classed long-term as longer than six weeks (intermediate term as one to six weeks' follow-up) (Boldt 2014). The remaining six reviews defined long-term follow up as over 12 months (one year) post-intervention (Gross 2015a; Hayden 2005; Hurkmans 2009; Regnaux 2015; Saragiotto 2016; van der Heijden 2015).

Methodological quality of included reviews

AMSTAR quality assessment of included reviews

No review achieved a perfect score of 11/11, though five achieved 10/11 (Boldt 2014; Busch 2013; Hayden 2005; Koopman 2015; Regnaux 2015) and eight scored 9/11 (Cramp 2013; Gross 2015a; Hurkmans 2009; Lane 2014; Lauret 2014; Saragiotto 2016; van der Heijden 2015; Yamato 2015). The lowest score was 6/11 (Silva 2010) though five categories were not applicable (n/a) due to there being no included studies. Quality assessment results for each individual review are presented in Table 6.

All reviews except one (Bidonde 2014) fulfilled the basic criteria (questions one to three of Table 1); to follow an 'a priori' design as Cochrane implements a system of protocol publication before undertaking the full reviews, where it also specifies dual study selection and data extraction from a comprehensive literature search. One review did not fulfil the 'a priori' design as this was an update and separation from a broader review series, and so the criteria had not been explicitly listed prior to publication for this specific title (Bidonde 2014).

Criteria which scored badly using the AMSTAR tool were characteristics of included studies (question six of Table 1), reporting of publication bias (question 10 of Table 1), and conflict of interest declarations (question 11 of Table 1).

- Included study characteristics were limited, often reporting the "inclusion criteria" used to recruit participants in the study instead of the characteristics of actual included participants, and excluding information such as participants' age, gender split, ethnicity, and disease status.

- Assessment of publication bias was omitted entirely in five reviews (Bartels 2007; Fransen 2014; Fransen 2015; Han 2004; Hurkmans 2009), and when it was assessed, it was reported using only a simple statement (with no test values, analyses used, or diagrams to demonstrate the result; Busch 2007; Koopman 2015). Two reviews mentioned in the methods as planned analyses, though was not mentioned again (Brown 2010; van der Heijden 2015), and a third review mentioned it in the methods, but appeared to use it interchangeably with reporting bias causing great confusion (Bidonde 2014).

- Conflicts of interest were sufficiently reported in only three out of 21 of the included reviews (Hayden 2005; Koopman 2015; Silva 2010). In the remaining reviews, a cursory statement was commonly made regarding the review authors' conflicts of interests, however, fulfilling the AMSTAR criteria also requires a statement to be made regarding any conflict of interest for any of the included studies.

Risk of bias in included reviews

The original review authors assessed risk of bias (see Table 7). The table shows the number of studies assessed as low risk of bias only, and excluded those that were assessed as unclear or high risk of bias.

Selection bias (randomisation and allocation concealment)

Selection bias had the largest proportion of included studies with low risk of bias (63% and 42% of studies adequately undertaking and reporting the methods used).

Performance and detection bias (blinding participants, personnel, outcome assessors)

With any exercise or physical activity intervention it is very difficult to blind both participants and personnel to the allocation, though some studies included in reviews attempted to by offering sham exercise.

Due to the difficulty of blinding participants to their group allocation, review authors assessed the risk of bias in different ways,

which may cause confusion: whereas the majority declared this lack of possible blinding to be high risk of bias or unclear, two reviews labelled such cases as low risk of bias in order not to exclude these studies unnecessarily from their analysis (Lane 2014; Lauret 2014). Without these two reviews, only a small percentage (7.8% or 18/229) of the included studies would have scored low risk of performance bias (blinding of participants and personnel), but by including them (all 35 studies from those two reviews assessed as low risk of bias) the overall proportion of studies assessed as having low risk of bias was closer to 20% (53/264).

Attrition (incomplete outcome data, withdrawals/dropouts)

About 55% (144/264) of the studies included in these reviews showed low risk of bias.

Reporting bias (selective reporting)

Reporting bias was classed as low risk in only 46% of included studies. However, it is important to note this was not due to the remainder having high risk of bias, but instead 'unclear', as trial protocols were not always published or accessible to the review authors to accurately assess/interpret.

Study/sample/group size

Sample size was not always included within the risk of bias assessment. It was therefore extracted directly from each review's table of included study characteristics by a single overview author (LG), and assessed as being low risk of bias when there was a minimum of 50 participants per arm, or 100 in total. Numbers were then separated for the proportion of studies with greater than 100 participants per arm (or 200 in total), and 200 participants per arm (or 400 in total), as this could then be considered higher tiered evidence.

Only 26 out of 264 included studies (10%) across the 21 reviews reported over 100 participants in total (or 50 per arm), a further 6% (15/264) included over 200 participants per arm. The remaining 223 studies (84%) had fewer than 50 participants per arm (or sample size was not reported), often not reaching 50 in total.

Other bias

The format for reporting bias has changed, and therefore some earlier reviews (that are yet to be updated) did not assess bias using the same format. Others reported additional criteria as 'other bias' including the similarity of baseline characteristics, and similarity of timing points.

Interpretation of results/conclusions by original review authors

For conclusions made by the original review authors, see Table 8. We assessed whether these conclusions/interpretations of the results accurately reflected the information provided within the review, and if any further information should have been included. This final assessment of the review is an important stage in determining any author bias within the review process, as many readers, funders, and policy makers will focus on the author conclusions without a full appraisal of the actual presented data.

Eleven of the 21 reviews reported appropriate conclusions based on the data available in the context of the quality of evidence (Bidonde 2014; Boldt 2014; Busch 2007; Busch 2013; Fransen 2015; Gross 2015a; Koopman 2015; Regnaux 2015; Saragiotto 2016; Silva 2010; Yamato 2015); five reviews had appropriate conclusions, did not mention quality of the evidence in the conclusion, but did discuss it in detail earlier in the review (Bartels 2007; Cramp 2013; Han 2004; Hayden 2005; Lauret 2014); two reviews had appropriate conclusions but had only limited discussion of quality or did not adequately consider the quality of the evidence in the interpretation of the results (Hurkmans 2009; Lane 2014); and three reviews needed further comment as the strength of the conclusions were not appropriate based on the available data (Brown 2010; Fransen 2014), or we were unable to agree with their interpretation due to difficulty in extracting the data (van der Heijden 2015).

Effect of interventions

We have interpreted results using data reported in the reviews, and did not return to the original studies. Where data have been reported as MDs or as an absolute or relative change score we have used the appropriate scales (where possible) to determine whether this was clinically significant. When data have only been presented as SMD, with or without 95% confidence intervals (CI), with or without level of significance (P value), we have cautiously used the interpretation by Cohen 1988 who defined effect size using the SMD as small (SMD 0.2 to 0.5), moderate (SMD 0.5 to 0.8), or large (SMD greater than 0.8).

For the purposes of clarity, we have used the term 'intervention' to refer to the exercise or physical activity intervention, and 'control' to refer to the included comparison group which did not involve any exercise or physical activity element.

Primary outcome

Self-reported pain (severity)

Part of the inclusion criteria for this overview was for pain severity to be listed as an outcome measure.

Two of the 21 reviews did not include/identify any studies that examined intervention versus control (Lauret 2014; Silva 2010). Of the remaining reviews that did report studies examining intervention versus control (no physical activity or exercise, or minimal intervention), two did not report pain as an absolute or relative score of severity, intensity, or change as a result of the intervention (Brown 2010; Han 2004), and one review assessed pain-free time and distance during exercise (they did not assess pain using a mean/usual pain scale; Lane 2014). We could not extract relevant data for one review as they compared two different exercise interventions and a control but did not report the data compared to the control (Regnaux 2015).

The remaining 15 reviews reported a mean or usual pain score for exercise (intervention) and no-exercise (control) groups (Bartels 2007; Bidonde 2014; Boldt 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Hurkmans 2009; Koopman 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015).

Reported baseline pain score

Of the 15 reviews that were able to assess pain (Table 9), only three reviews reported actual baseline pain scores (Bidonde 2014; Boldt 2014; Hayden 2005). Three reviews reported change data (Bartels 2007; Busch 2007; Busch 2013), but we were able to use control group baseline and earliest control group scores as assumed or approximate baseline measures for the intervention groups in nine reviews (Bartels 2007; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Koopman 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015). Overall, only three reviews that assessed pain did not provide baseline or control group scores for comparison (Busch 2013; Cramp 2013; Hurkmans 2009).

Intervention group at baseline	Control group at baseline	Control group at earliest follow-up
Median pain score 70.9/100 (based on 7 studies, n = 382; Bidonde 2014)	WOMAC 9.1/20 (2 studies, n = 380) VAS ~ 55/100 (3 studies, n = 117) HAQ 1.05/3 (1 study, n = 249) (Bartels 2007)	Mean pain score ~ 29/100 (9 studies, n = 549; Fransen 2014)
11.05 to 22.6 on a 0 to 150 WUSPI score (1 study, n = 35; Boldt 2014)	VAS 35/100 to 61/100 (4 studies, n = 204; Busch 2007)	44/100 (44 studies, n = 3537; Fransen 2015)
Mean pain score 46/100 (95% CI 41 to 50) (8 studies, n = 370; Hayden 2005)	-	40/100 to 60/100 (2 studies, n = 147; Gross 2015a)
-	-	44/100 SD 24 (1 study, n = 55; Koopman 2015)
-	-	range 25/100 to 56/100 (4 studies, n = 291; Saragiotto 2016)
-	-	2.1/10 to 6.0/10 (2 studies, n = 41; van der Heijden 2015)
-	-	range 18/100 to 52/100 (6 studies, n = 148; Yamato 2015)
Range: 46 to 70.9 on a 0 to 100 scale 16 studies, n = 787	**Range: 35 to 55 on a 0 to 100 scale 10 studies, n = 950**	**Range: 18 to 60 on a 0 to 100 scale 68 studies, n = 4768**

CI: confidence interval; HAQ: Health Assessment Questionnaire; n: number of participants; SD: standard deviation; VAS: visual analogue score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; WUSPI: Wheelchair User's Shoulder Pain Index

HAQ: mean of different category scores, 0 or 1 (mild to moderate disability), up to 2 or 3 (severe to very severe disability); WOMAC pain score: 5 items summed to 0 (no pain) to 20 (worst pain ever); WUSPI: 15 items of 0 to 10 VAS scores, summed to form total

CI: confidence i
ment Questionn
SD: standard de
score; WOMAC
ter Universities

(*Continued*)

of 0 (no pain) to 150 (worst pain ever)	Wheelchair User HAQ: mean of 1 (mild to mod (severe to very se score: 5 items s (worst pain ever) VAS scores, sum to 150 (worst pa

This suggests the majority of participants reviewed had mild-to-moderate pain (only one review reported a mean of severe pain (aquatic exercise for fibromyalgia, Bidonde 2014) at the commencement of each intervention (less than 30/100 mild pain, 30/100 to 60/100 moderate pain, more than 60/100 severe pain; Collins 1997), though labelling the majority as having only mild-to-moderate pain should be interpreted with caution due to the lack of specific data available - the baseline data of the intervention group would have been preferable to the proxies we have had to use.

Quality judgement/ tiered quality (first, second, third tier evidence)

Our assessment criteria stated that we would accept the information as graded evidence when reported as the number of participants achieving a 50% (first tier evidence) or 30% (second tier evidence) reduction in pain, but none of the included reviews reported results in this way, and so instead we used the reported absolute and relative change values.

None of the included reviews fulfilled the requirements for first tier evidence (at least 50% pain reduction from baseline, study duration longer than eight weeks, and more than 200 participants per arm).

Second tier evidence (at least 30% pain reduction from baseline, study duration between four and eight weeks, and more than 200 participants in total or 100 participants per arm) was also lacking in these reviews; three reviews found at least 30% reduction in pain from baseline (Busch 2007; Busch 2013; van der Heijden 2015), one of which also used long enough exercise programmes (eight to 21 weeks' intervention, Busch 2013) but totalled only 81 participants across two studies. The other two reviews did not fulfil the study duration criteria (interventions from 2.5 weeks, Busch 2007; and three weeks, van der Heijden 2015) or study size criteria.

Consequently results from relevant reviews have been pooled (all tier three quality) where appropriate, though results should be interpreted with caution due to the low quality evidence.

Treatment effect

Data that could be extracted for pain can be seen in Table 9 for all reviews. Only three reviews found no statistically significant changes in usual or mean pain from any intervention (Cramp 2013; Hurkmans 2009; Koopman 2015 (assumed due to lack of presented data)). The remaining reviews reported a statistically significant effect of the intervention at one or more time points, in at least one subgroup.

Three reviews found at least 30% pain reduction from baseline (post-intervention - strength training: Busch 2007; Busch 2013, at short-term follow-up: van der Heijden 2015). Additionally, seven reviews reported clinically significant results (minimally important difference: reduction in pain from baseline of at least 10 points on a 0 to 100 scale or an absolute improvement of at least 10% to 20%, Dworkin 2008) as a result of the exercise intervention (1.3/10 from aerobic training, Busch 2007; 12/100 (95% CI 10 to 15), Fransen 2015,; 14.9/100 (95% CI 7.39 to 22.40), Gross 2015a; 10.2/100 (95% CI 1.31 to 19.09), Hayden 2005; 2.5/10 (95% CI 1.52 to 3.48), Boldt 2014; 10.01/100 (95% CI 4.35 to 15.67), Saragiotto 2016; 14.05/100 (95% CI 9.19 to 18.91), Yamato 2015). Three reviews found statistically significant improvements as a result of the intervention, but they did not reach clinical significance (post-intervention, $P = 0.02$, Bartels 2007; "small to moderate" benefit post-intervention and at six-month follow-up, $P < 0.001$, Fransen 2014; "moderate effect" of 7% (95% CI 3 to 11) benefit post-intervention, Bidonde 2014).

Overall, results were inconsistent across interventions and follow-up (see Table 9), as exercise did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point.

Secondary outcomes

Physical function (objectively or subjectively measured)

Measures of physical function were the primary outcome measure in eight out of 21 reviews (Busch 2013; Han 2004; Hayden 2005;

Hurkmans 2009; Koopman 2015; Lane 2014; Lauret 2014; Silva 2010), and a reported (non-primary) outcome measure in nine more reviews (Bartels 2007; Bidonde 2014; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Regnaux 2015; Saragiotto 2016; van der Heijden 2015, plus some which assessed disability; Cramp 2013; Saragiotto 2016; Yamato 2015). Only Boldt 2014 and Brown 2010 did not list physical function (or disability, or activity limitation) as a potential outcome measure.

Treatment effect

Data that could be extracted for physical function are shown in Table 10. Two reviews which reported physical function had no data to extract (Lauret 2014; Silva 2010), and for one review we were unable to extract the relevant data (Regnaux 2015). Two reviews found no significant difference in physical function between the intervention and control groups (Han 2004; Hurkmans 2009, both rheumatoid arthritis, 8 studies, n = 240). The remaining 14 reviews showed that the intervention produced a statistically significant benefit over the control at a minimum of one reported time point (Bartels 2007; Bidonde 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Koopman 2015; Lane 2014; Saragiotto 2016; van der Heijden 2015; Yamato 2015; 129 studies, n greater than 9559 (exact number unknown due to some participant numbers not being reported)).

Many of these statistically significant results were of small or moderate effect size (as reported by the review authors, or using the definition by Cohen 1988 if unreported; small effect size: Bartels 2007; Bidonde 2014; Fransen 2014; Fransen 2015; Gross 2015a; Koopman 2015; Saragiotto 2016; Yamato 2015, moderate effect size: Busch 2007; Fransen 2015; Yamato 2015).

Only one review reported statistical significance and large effect size (both short-term and long-term follow-up: SMD 1.10 (95% CI 0.58 to 1.63) and 1.62 (95% CI 0.31 to 2.94), van der Heijden 2015). However, the original review authors highlighted the low to very low quality of the evidence as many studies had high or unclear risk of bias across multiple domains (van der Heijden 2015).

Psychological function

Only five out of 21 reviews assessed psychological function as mental health (Bartels 2007; Bidonde 2014; Busch 2013), anxiety (Cramp 2013), and depression (Boldt 2014; Busch 2013; Cramp 2013).

Treatment effect

Data that could be extracted for psychological function can be seen in Table 11. There were significant effects in favour of the intervention for mental health (Bartels 2007) and depression (Busch 2013) scores, and "variable effect" for depression (Cramp 2013). However, there was also no effect or no differences between control and intervention groups reported for mental health (Bidonde 2014; Busch 2013), anxiety (Cramp 2013), and depression (Boldt 2014).

Quality of life

A version of quality of life assessment was reported in nine reviews. Six were termed quality of life or health-related quality of life (HRQoL) (Bartels 2007; Boldt 2014; Fransen 2014; Fransen 2015; Gross 2015a; Lauret 2014).

Other reviews assessed global perceived effect (Gross 2015a), global wellbeing (Busch 2007), global assessment (Hayden 2005), global impression of recovery (Saragiotto 2016; Yamato 2015), health assessment questionnaire (Silva 2010), multi-dimensional function (Bidonde 2014; Busch 2013), and work status (Hayden 2005). These have been reported separately to quality of life (Table 12).

Treatment effect

Data that could be extracted for quality of life can be seen in Table 12. Four reviews found no significant difference between intervention and control groups in health-related quality of life post-intervention (9 studies, n = 556) (HRQoL: Boldt 2014; Fransen 2014; Gross 2015a, global assessment: Bidonde 2014; Gross 2015a)), three reviews did not or were unable to report any data (HRQoL: Lauret 2014, global assessment: Hayden 2005, other assessment: Silva 2010), and seven reviews found a significant improvement as a result of the intervention (34 studies, n = 2700) (HRQoL: Bartels 2007, Fransen 2015, global assessment: Busch 2007; Saragiotto 2016; Yamato 2015, other assessment: Bidonde 2014; Busch 2013).

Two reviews assessing strength/resistance training interventions found significantly large effect sizes (SMD greater than 0.8, as defined by Cohen 1988) in favour of the intervention (global wellbeing measure, SMD 1.43 (95% CI 0.76 to 2.10), Busch 2007; Fibromyalgia Impact Questionnaire, SMD 1.27 (95% CI 0.72 to 1.83), Busch 2013). Other statistically significant changes reported in the included reviews were of small-to-moderate effect size (SMD 0.2 to 0.8, Cohen 1988).

Adherence to the prescribed intervention

Only one review reported adherence to the intervention as an outcome measure (Regnaux 2015), but the authors were unable to perform an analysis on attendance as most studies did not clearly report attendance or compliance (Regnaux 2015). However, five reviews assessed withdrawals or dropouts (Bidonde 2014; Fransen 2014; Han 2004; Regnaux 2015; Saragiotto 2016), one reported all-cause attrition (Busch 2013), and another reported the discontinuation rate (Silva 2010).

Data that could be extracted for adherence, withdrawals, and attrition can be seen in Table 13. Pooling all available data for withdrawals/dropout/attrition gave an RR of 1.02 (95% CI 0.94 to 1.12) in favour of the control group (6 reviews, 30 studies, n = 2256, control withdrawal 81/1000, intervention withdrawal 82.8/1000).

One clinically controlled trial (CCT) in one review reported statistically significant improvement in enjoyment of exercise/rest (P = 0.0002) and self-reported benefit from exercise/rest (P = 0.006) at both post-intervention (end of therapy, 10 weeks) and follow-up (four months later) (n = 95, Han 2004).

Healthcare use/attendance

None of the reviews reported healthcare use/attendance.

Adverse events (not death)

Eighteen out of 21 reviews reported adverse effects (three reviews did not report adverse events as an outcome measure due to lack of studies or other undisclosed reasons; Brown 2010; Lauret 2014; Silva 2010). Two reviews only assessed a specific adverse event ("amputation" Lane 2014; "motor unit survival" Koopman 2015), one review observed "safety - pain and radiological damage" (Hurkmans 2009), and another referred to any "side-effects" (Han 2004).

Data that could be extracted for adverse events (not death) can be seen in Table 14. The total number of reported adverse events (not death) was 137 events across 39 studies out of 61 studies that had adverse events as an outcome measure (over one-third of all trials that reported them found no adverse events related to the intervention): six reviews reported no adverse events from the included trials (Bartels 2007; Busch 2013; Cramp 2013; Hurkmans 2009; Koopman 2015; Yamato 2015) though the authors questioned whether this was due to lack of reporting by the trial authors, or whether there were no adverse events.

Adverse events were largely reported as a total number per trial, though one review separately reported results for the intervention group versus the control group (Saragiotto 2016), and two others reported adverse events for the intervention group only (Boldt 2014; Regnaux 2015). Only one review calculated an RR for the adverse events, showing a reduced risk for amputation in the intervention group (two amputations in the usual care/control group: RR 0.20, 95% CI 0.01 to 4.15, based on one study in one review, Lane 2014).

Death

Only one out of 21 reviews reported death separately to other adverse events (Lane 2014). Based on five studies within the review, death had an RR of 0.71 (95% CI 0.28 to 1.78) in favour of exercise as being protective, though was not statistically significant (P = 0.47).

DISCUSSION

Specificity of the condition: despite the heterogeneous nature of chronic pain, in this overview we have combined several painful conditions covering a number of conditions and diagnoses. Regardless of aetiology, the impact of chronic pain is broadly similar across many conditions.

Summary of main results

Pain severity: there were favourable results in a number of reviews as a result of exercise: only three reviews found no statistically significant changes in usual or mean pain from any intervention. However, results were inconsistent across interventions and follow-up, as the intervention did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point. The exercise or physical activity interventions did not have a negative effect on the outcome (did not worsen the pain). A factor in the lack of statistical and clinically significant result may be the baseline pain severity of participants. The majority of the included population had an assumed mild-to-moderate pain severity score (assumed only due to lack of exact group data at baseline). This is often the desired outcome (post-intervention) of many drug therapies for pain, and it may therefore be difficult to show a clinically significant improvement in these people.

Physical function: physical function/disability was the most commonly reported outcome measure, and was the primary measure in eight out of the 21 reviews. Physical function was significantly (statistically) improved as a result of the intervention in 14 reviews, though even these statistically significant results had only small-to-moderate effect sizes in all but one review.

Psychological function and quality of life: there were variable results for psychological function and quality of life: results were either favourable to exercise (two reviews reporting significantly large effect sizes for quality of life), or showed no difference between groups. There were no negative effects.

Adherence to the prescribed intervention: could not be assessed in any included review. However, risk of withdrawal/dropout was slightly higher in the exercising group (82.8/1000 participants versus 81/1000 participants), though the group difference was not significant.

Healthcare use/attendance: not reported in any included review.

Adverse events, potential harm, and death: importantly, exercise caused no actual harm, with most adverse events being increased soreness or muscle pain, which reportedly subsided after several weeks of the intervention. One review reported a non-significant reduction in risk of death as a result of the intervention.

Overall completeness and applicability of evidence

Of the 21 included reviews, seven could be considered out of date as they were most recently assessed as up-to-date prior to

2010 such that any recent controlled trials assessing pain severity have not been included in this overview (Cochrane recommends updating reviews every two years) (Bartels 2007; Brown 2010; Busch 2007; Han 2004; Hayden 2005; Hurkmans 2009; Silva 2010). We included these reviews in the overview, but they may not be as relevant now due to the elapsed time since they were updated. One protocol that had potential to be included was published in 2006 with no full review available yet (Craane 2006).

Available data suggest that participants in the included reviews and studies would generally be characterised as having mild-moderate pain (moderate greater than 30/100 or 3/10) with only one review reporting moderate-severe pain (severe greater than 60/100 or 6/10). Therefore whether the evidence of change or no change seen here as a result of each intervention is applicable to people further along on the pain spectrum (with higher pain scores/worse pain) is debatable. However, it can be argued that those people are more likely to be assigned medical or surgical interventions than physical activity and exercise alone (where available), and as a group they may be less able to engage in exercise, and may therefore be more difficult to recruit into exercise-only studies. Having said this, the labelling of participants as having mild-moderate pain was a cautious one within this overview due to the lack of specific data available at baseline assessment; only three reviews included baseline pain scores in the intervention group, and two further reviews provided control group baseline scores.

There are still gaps in the available literature, and therefore also within this overview. None of the included reviews examined generalised or widespread chronic pain as a global condition, each instead examined specific conditions that included chronic pain as a symptom or result of the ongoing condition (rheumatoid arthritis, osteoarthritis, fibromyalgia, low back pain, intermittent claudication, dysmenorrhoea, mechanical neck disorder, spinal cord injury, postpolio syndrome, and patellofemoral pain). The pain in these cases can occur secondary to other symptoms such as fatigue, muscle stiffness, difficulty sleeping, and depression, all of which could separately (and more effectively) be influenced by the intervention. Additionally, only 25% of included studies actively reported adverse events. This may affect the completeness of the evidence as conclusions have been drawn based on the available data. The included reviews did not discuss the possible impact of this non-reporting by the original trials, and this may lead to underestimating possible adverse events from an intervention, or overestimating its safety.

The exercise interventions examined in the included reviews were broad; including aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi. Many of these interventions can be accessed in the community by the general public and people with chronic pain, either individually or in classes (yoga, Pilates, tai chi). Other exercise intervention programmes, such as the motor control exercise and proprioceptive (balance) training, requires at least initial supervision by a therapist to teach the correct techniques and provide feedback for progression.

Quality of the evidence

In assessing the quality of the evidence, we employed the AMSTAR tool to examine the reviews, extracted data on risk of bias to examine the available primary evidence, and evaluated the authors' conclusions to ensure that they were appropriate based on the available data.

The AMSTAR tool is useful in assessing the reporting of a systematic review, though it does not inform us of the actual undertaking or conduct of the review process. All 21 included reviews scored well across the AMSTAR assessment, though this is likely due to the stringent reporting guidelines implemented by Cochrane prior to publication. However, it may be necessary or advisable for the Cochrane guidelines to be further expanded and detailed with regards to reporting study characteristics, publication bias, and conflicts of interest, as these areas often did not meet the requirements laid out in the AMSTAR criteria (Table 1).

Data extracted from the reviews regarding their assessment of bias (risk of bias) showed moderate level scores at best across all included studies within the included reviews. Other than issues surrounding blinding (which are problematic in exercise intervention studies due to the nature of the intervention), the trials did not consistently and adequately report potential attrition and reporting biases, with less than half of studies within these reviews at low risk of bias.

However, the most prominent issue with regards to bias in these exercise and physical activity intervention studies is the sample size used. This subcategory is not used as standard in the assessment of bias in Cochrane Reviews, despite the increasing volume of research available suggesting that small studies of fewer than 100 participants per arm (Moore 2010; Nüesch 2010) are at increased risk of succumbing to the random effects in estimating both direction and magnitude of treatment effects (Moore 1998; Turner 2013) due to greater heterogeneity within and between small studies (IntHout 2015).

Studies within the included reviews here were very small (often fewer than 50 participants in total). For greater quality and a more reliable effect, at least 100 participants per arm should be analysed for a study to potentially be classed as tier two evidence (200 per arm for tier one); small studies are known to overestimate the treatment effect by up to 32% in comparison with larger studies (Deschartes 2013).

Assessing studies for risk of bias based on study size (total number or per arm) should be included in any review or meta-analysis in future, to adequately assess the influence of small trials on the estimated treatment effect (Nüesch 2010). Inclusion in the standard assessment process may in turn influence the design and undertaking of future research trials to increase the sample size, and produce more consistent clinically and statistically accurate results.

Of the 21 included reviews, 12 used a pain measure as their primary outcome (Bartels 2007; Boldt 2014; Brown 2010; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Regnaux 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015), and the remaining nine reviews included the measure as a secondary outcome only. Other outcomes were shared, including physical and psychological function, and quality of life. Likewise, each review team will have included studies that did not use their chosen outcome measures as the primary measure, and that were therefore powered according to a different primary outcome. On collating the evidence, some studies may appear underpowered for the outcome(s) of interest to us (Turner 2013), yet were adequately powered for the studies' primary measure. To increase the power of the results of this overview, and the intermediary reviews we have included, intervention studies that focus on painful conditions should include pain intensity as the primary outcome, or at least as a prominent secondary outcome; alternatively review authors should seek to include only those studies that were adequately powered for pain intensity as a primary outcome measure.

Intervention length ranged from a single session to regular sessions over a period of 30 months, though the majority were between eight and 12 weeks. Durations of this length are common among exercise and physical activity intervention studies to allow for physiological adaptation and familiarisation. In contrast, the follow-up period was often inadequate, as many reviews reported only a single follow-up point (immediately post-intervention), or repeated measures over the short-term (less than six months): only six of the 21 reviews planned to assess participants over the long term (over 12 months: Gross 2015a; Hayden 2005; Hurkmans 2009; Regnaux 2015; Saragiotto 2016; van der Heijden 2015). With chronic conditions, it would be advisable to include longer follow-up periods (beyond 12 months post-randomisation) as long-term solutions may be more relevant to their control or pain management. It is also possible that initial adaptation and potential benefits as a result of an exercise intervention may take longer to manifest in comparison to a 'healthy' person due to the possible limitations in exercise intensity and progression (a training threshold) beyond which any additional physical training may be detrimental to the underlying pathophysiological mechanisms (Daenen 2015) or simply be additional physical stress with no additional physical benefit (Benton 2011).

We grouped outcome measurement points in this overview into short term (less than six months), intermediate term (six to 12 months), and long term (longer than 12 months). The broad time window for 'short term' outcomes (less than six months) is a potential source of heterogeneity as the early period is the one where time of measurement is most likely to result in variable outcomes. These initial problems could be overcome by use of standard reporting periods in exercise intervention studies (suggested four-weekly within the 'short term' period to assess both neural adaptation and other physiological changes). This would allow review authors to use the data gathered closest to the time point they are assessing, for more accurate analyses. Additionally, by extending the follow-up period beyond one year (long-term follow-up), heterogeneity may be reduced further.

Reviews generally did not enforce a minimum exercise requirement for inclusion in their review. Additionally, not all exercise sessions were supervised or baseline fitness/physical ability was assessed subjectively, and consequently it was not reported whether the intervention was fulfilled as described, or whether the dose was enough to elicit a physiological response. Studies often rely on the self-report of participants as to the actual physical activity and exercise being undertaken, which can lead to a greater risk of bias, and reduced study quality as it is questionable as to whether the effect can be truly attributed to the intervention. This was examined in a previous review, where it was concluded that non-subjective physical assessment should be performed where possible (Perruchoud 2014), though these still have challenges regarding implementation.

In summary, the quality of the evidence was low (third tier): within this overview we found no tier one or tier two evidence. This is largely due to the small sample sizes and potentially underpowered studies. A number of studies within the reviews had adequately long interventions, but planned follow-up was limited to less than one year (12 months) in all but six reviews.

Interpretation of the available data, and conclusions drawn by the review authors, were appropriate, although the conclusions were sometimes stronger than warranted by the available data. Occasionally results were not discussed with regards to the quality of the evidence or risk of bias: it is important to discuss the findings in the context of the quality of the evidence, with complete transparency, as this may affect future research, and implications for patients, funders, and policy makers.

Potential biases in the overview process

While we have attempted to include all relevant reviews in the overview process, we do concede that by only searching the Cochrane Library, and including only current Cochrane Reviews we may have missed some key literature. However previous publications have referred to the higher quality grading (high AMSTAR score) in Cochrane Reviews due to the basic criteria necessary for publication at any stage (protocol or full review) suggesting they may be the most reliable source of evidence (O'Connell 2013).

Agreements and disagreements with other studies or reviews

This is a summary overview of current Cochrane Reviews, we are not aware of any overviews or reviews summarising non-Cochrane reviews.

AUTHORS' CONCLUSIONS

There is limited evidence of improvement in pain severity as a result of exercise. There is some evidence of improved physical function and a variable effect on both psychological function and quality of life. However, results are inconsistent and the evidence is low quality (tier three). Promisingly however, none of the physical and activity interventions assessed appeared to cause harm to the participants.

Implications for practice

For clinicians and people with chronic pain

The evidence in this overview suggests that the broad spectrum of physical activity and exercise interventions assessed here (aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi) are potentially beneficial, though the evidence for benefit is low quality and inconsistent. The most commonly reported adverse events were increased soreness or muscle pain, which subsided after several weeks of the intervention.

Physical activity and exercise may improve pain severity as well as physical function and quality of life.

For policy makers

The evidence showed variable results, though in some reviews there was a clinical and statistical benefit in pain relief and physical function (based on low quality evidence). The evidence suggests that physical activity or exercise is an acceptable intervention in people with chronic pain, with minimal negative adverse effects. However based on this low quality evidence, we cannot provide direction to the content of an exercise programme should clinicians decide to implement one.

Implications for research

There is a clear need for further research into exercise and physical activity for chronic pain in adults.

General implications

- Future research should report baseline values for outcome measures in both intervention and control groups, together with detailed relevant information about the participants. Knowing the baseline value is relevant to interpreting any change observed as a result of the intervention, and understanding the broader value of the intervention.
- Where possible, pain results should be reported as the number of people achieving 50%, 30%, and 10% pain relief, and the number who did not meet that point (dichotomous outcome). These are clinically important cut-offs in pain intervention research, and reporting in this way allows readers to observe the clinical effect more effectively.
- Reporting should include median and range as well as mean and standard deviation (SD) of results. This will allow readers to review the effects of any outliers that may have skewed the data, which often goes unnoticed in the reporting of mean and SD alone.
- The importance of clear intervention reporting is underestimated: often studies report both intervention and control programmes simply, where other researchers and clinicians alike are unable to replicate the trial or intervention. Recommendations for reporting are based on the Consolidated Standards of Reporting Trials (CONSORT) statement (www.consort-statement.org/), but this alone does not detail the extent of necessary intervention and control programmes reporting. The template for intervention description and replication (TIDieR) approach (Hoffman 2014) is intended as an extension to CONSORT item 5 ("The interventions for each group with sufficient details to allow replication, including how and when they were actually administered") and is a checklist for detailing the programmes using: why (rationale), what (materials and procedures), who, how, where, when, and how much.

Design

- One previous review highlighted the increased bias often present in questionnaires and other self-report measures of physical activity in people with chronic pain, and as a result made the recommendation to use objective measures instead, such as accelerometers, or the use of direct and indirect calorimetry, where possible (Perruchoud 2014), though these still have challenges regarding implementation. This would allow direct and exact comparison and analyses of actual energy expenditure and treatment effect.

Population/participants/sample

- There needs to be a focus on participants with generalised and/or widespread chronic pain, instead of (or as well as) condition-specific populations.
- Studies should include people with higher pain severity (greater than 50/100 on a 100-point visual analogue scale) at baseline. People with mild-moderate pain should still be included, but it would be advisable to separate the results for analysis, ensuring the study is adequately powered to allow this subgroup analysis in advance. This way we could determine if exercise has benefit overall, or affects one group more than another, and tailor exercise programmes according to the individual needs.
- It has been previously suggested that for 20% to 25% of participants undertaking an exercise programme there is little to

no favourable response (Timmons 2014), while a small percentage (5% to 10%) have adverse events (Bouchard 2012). It is therefore vitally important that much larger sample sizes are used: ideally *more than 200 participants per arm*, though even this number in total would increase the quality of the evidence in the first instance. In this way we may be able to learn to identify individuals who will benefit, and those who will require further intervention.

Interventions

- Different forms of exercise should be researched in detail. For the purposes of this overview, we combined all physical activity and exercise interventions under one banner to determine if there was any effect. However a number of reviews separately analysed resistance (strength) training, aerobic (endurance), and combination programmes. It is important to continue to examine different modalities, but currently there is not enough high quality evidence to exclude or prioritise one specific mode (resistance, endurance, stability) or medium (land/water based), or the proportion of a combination programme to be assigned to each, as all may have individual benefits for people with chronic pain.
- Intensity of exercise, duration of individual sessions, and frequency should be investigated. It is this dose alongside duration (of the entire intervention) and adherence that may determine the actual efficacy.
- More reviews and trials should attempt to minimise intervention heterogeneity by implementing minimum and maximum requirements. Only this way will the research community be able to determine more accurately the direction and magnitude of effect of a specific programme or intervention. Many of these important restrictions can be implemented as subgroup analyses, though if this is the case it is important to have adequate study numbers (ideally 200 participants per arm or subgroup).
- Due to the chronicity and long-term nature of the condition, physiological and psychological changes may take longer to manifest. It is widely accepted that there is a delay in muscular hypertrophy as a result of exercise, and initial gains within the first few weeks of any training programme will be as a result of neural factors (Enoka 1997); this is also in line with the grading of evidence (tier two evidence or higher requires a minimum of a four-week intervention). This suggests that longer interventions may be necessary (eight weeks for tier one evidence), though assessing participants at regular intervals, including at four weeks, would be beneficial to examine the effect of the neural adaptation alone.

Measurement (end-points)

- Randomised controlled trials with long-term follow-up are needed. Chronic pain is defined by its chronic nature, and therefore long-term follow-up of results is equally important as the initial short-term effect (if not more so): outcomes should be assessed beyond one year after randomisation. In turn this will inform the direct effect of the intervention, as well as the proportion of the population who maintains the programme of exercise employed in the intervention, or something else under the guise of physical activity as a result of participation.
- The broad time window for 'short term' outcomes (less than six months) is a potential source of heterogeneity as the early period is the one where time of measurement is most likely to result in variable outcomes. These initial problems could be overcome by use of standard reporting periods in exercise intervention studies (suggested four-weekly assessment within the 'short term' period to assess both neural adaptation and other physiological changes). This would allow review authors to use the data recorded closest to the time point they are assessing, for more accurate and comparable analyses.
- Outcome measures used by researchers should be standardised across trials and studies. Recommendations for selecting the most appropriate and important outcome measures to those who live with chronic pain have previously been published (Initiatives on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) Consensus Recommendations, Dworkin 2005; Turk 2003).

Other

- It would be of interest in future research to determine the reasons for non-participation in regular physical activity or non-compliance to a prescribed exercise intervention in people with chronic pain, and how to overcome these barriers.
- Future Cochrane Reviews could include: exercise for chronic pain or chronic widespread pain (and not specific conditions such as osteoarthritis, fibromyalgia, etc.), and exercise for neuropathic pain. These areas have not been covered by Cochrane with an exercise or physical activity intervention.

ACKNOWLEDGEMENTS

Thanks to Dr Amanda Williams, Professor Chris Eccleston, and Anna Erskine at the Cochrane Pain, Palliative and Supportive Care Review Group (PaPaS), for their support and rigour in the development of the concept of this overview, title proposal, and protocol. Thanks also to peer reviewers Dr Neil O'Connell and Dr Clare Stevinson, and consumer referees Jennifer Patrick and Dr Shailen Mehta.

Cochrane Review Group funding acknowledgement: the National Institute for Health Research (NIHR) is the largest single funder of the Cochrane PaPaS Group.

Disclaimer: the views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, National Health Service (NHS), or the Department of Health.

External funding support from the National Institute of Academic Anaesthesia (BJA RCoA NIAA) and the MRC Lifelong Health and Wellbeing Grant to complete the overview as part of two larger projects.

REFERENCES

References to included reviews

Bartels EM, Lund H, Hagen KB, Dagfinrud H, Christensen R, Danneskiold-Samsøe B. Aquatic exercise for the treatment of knee and hip osteoarthritis. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD005523.pub2]

Bidonde J, Busch AJ, Webber SC, Schachter CL, Danyliw A, Overend TJ, et al. Aquatic exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011336]

Boldt I, Eriks-Hoogland I, Brinkhof MWG, de Bie R, Joggi D, von Elm E. Non-pharmacological interventions for chronic pain in people with spinal cord injury. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD009177.pub2]

Brown J, Brown S. Exercise for dysmenorrhoea. *Cochrane Database of Systematic Reviews* 2010, Issue 2. [DOI: 10.1002/14651858.CD004142.pub2]

Busch AJ, Barber KA, Overend TJ, Peloso PMJ, Schachter CL. Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD003786.pub2]

Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010884]

Cramp F, Hewlett S, Almeida C, Kirwan JR, Choy EHS, Chalder T, et al. Non-pharmacological interventions for fatigue in rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD008322.pub2]

Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2008, Issue 4. [DOI: 10.1002/14651858.CD004376.pub2]

Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: 10.1002/14651858.CD007912.pub2]

Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD004376.pub3]

Gross A, Kay TM, Paquin JP, Blanchette S, Lalonde P, Christie T, et al. Exercises for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD004250.pub5]

Han A, Judd M, Welch V, Wu T, Tugwell P, Wells GA. Tai chi for treating rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2004, Issue 3. [DOI: 10.1002/14651858.CD004849]

Hayden J, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database of Systematic Reviews* 2005, Issue 3. [DOI: 10.1002/14651858.CD000335.pub2]

Hurkmans E, van der Giesen FJ, Vliet Vlieland TPM, Schoones J, Van den Ende ECHM. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2009, Issue 4. [DOI: 10.1002/14651858.CD006853.pub2]

Kay TM, Gross A, Goldsmith CH, Rutherford S, Voth S, Hoving JL, et al. Exercises for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 2012, Issue 8. [DOI: 10.1002/14651858.CD004250.pub4]

Koopman FS, Beelen A, Gilhus NE, de Visser M, Nollet F. Treatment for postpolio syndrome. *Cochrane Database of Systematic Reviews* 2015, Issue 5. [DOI: 10.1002/14651858.CD007818.pub3]

Lane R, Ellis B, Watson L, Leng GC. Exercise for intermittent claudication. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD000990.pub3]

Lauret GJ, Fakhry F, Fokkenrood HJP, Hunink MGM, Teijink JAW, Spronk S. Modes of exercise training for intermittent claudication. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD009638.pub2]

Regnaux JP, Lefevre-Colau MM, Trinquart L, Nguyen C, Boutron I, Brosseau L, et al. High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis. *Cochrane Database of Systematic Reviews* 2015, Issue 10. [DOI: 10.1002/14651858.CD010203.pub2]

Saragiotto BT, Maher CG, Yamato TP, Costa LOP, Menezes Costa LC, Ostelo RWJG, et al. Motor control exercise for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2016, Issue 1. [DOI: 10.1002/

14651858.CD012004]
Silva KNG, Mizusaki Imoto A, Almeida GJM, Atallah ÁN, Peccin MS, Fernandes Moça Trevisani V. Balance training (proprioceptive training) for patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2010, Issue 5. [DOI: 10.1002/14651858.CD007648.pub2]
van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SMA, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD010387.pub2]
Yamato TP, Maher CG, Saragiotto BT, Hancock MJ, Ostelo RWJG, Cabral CMN, et al. Pilates for low back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 7. [DOI: 10.1002/14651858.CD010265.pub2]

References to excluded reviews

Aggarwal VR, Lovell K, Peters S, Javidi H, Joughin A, Goldthorpe J. Psychosocial interventions for the management of chronic orofacial pain. *Cochrane Database of Systematic Reviews* 2011, Issue 11. [DOI: 10.1002/14651858.CD008456.pub2]
Bierma-Zeinstra SMA, Brinks A, Verhagen AP, Van Rijn RM, Koes BW, Verhaar JAN. Interventions for lateral hip pain (tendinopathy or bursitis). *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.CD008924]
Brønfort G, Haas M, Evans RL, Goldsmith CH, Assendelft WJJ, Bouter LM. Non-invasive physical treatments for chronic/recurrent headache. *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD001878.pub3]
Brønfort G, Evans RL, Goldsmith CH, Haas M, Leininger B, Levin M, et al. Spinal rehabilitative exercise and manual treatment for the prevention of migraine attacks in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 8. [DOI: 10.1002/14651858.CD011848]
Choi BKL, Verbeek JH, Tam WWS, Jiang JY. Exercises for prevention of recurrences of low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD006555.pub2]
Craane B, De Laat A, Dijkstra PU, Stappaerts K, Stegenga B. Physical therapy for the management of patients with temporomandibular disorders and related pain. *Cochrane Database of Systematic Reviews* 2006, Issue 1. [DOI: 10.1002/14651858.CD005621]
Dagfinrud H, Hagen KB, Kvien TK. Physiotherapy interventions for ankylosing spondylitis. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD002822.pub3]
Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. *Cochrane Database of Systematic Reviews* 2010, Issue 6. [DOI: 10.1002/14651858.CD007612.pub2]
Dal Bello-Haas V, Florence JM. Therapeutic exercise for people with amyotrophic lateral sclerosis or motor neuron disease. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD005229.pub3]
de Souza RF, Lovato da Silva CH, Nasser M, Fedorowicz Z, Al-Muharraqi MA. Interventions for the management of temporomandibular joint osteoarthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD007261.pub2]
Fokkenrood HJP, Bendermacher BLW, Lauret GJ, Willigendael EM, Prins MH, Teijink JAW. Supervised exercise therapy versus non-supervised exercise therapy for intermittent claudication. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD005263.pub3]
Franke H, Fryer G, Ostelo RWJG, Kamper SJ. Muscle energy technique for non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: 10.1002/14651858.CD009852.pub2]
Green S, Buchbinder R, Hetrick SE. Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD004258]
Gross A, Aker P, Goldsmith CH, Peloso PMJ. Physical medicine modalities for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 1998, Issue 2. [DOI: 10.1002/14651858.CD000961]
Gross A, Forget M, St George K, Fraser MMH, Graham N, Perry L, et al. Patient education for neck pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. [DOI: 10.1002/14651858.CD005106.pub4]
Gross A, Langevin P, Burnie SJ, Bédard-Brochu MS, Empey B, Dugas E, et al. Manipulation and mobilisation for neck pain contrasted against an inactive control or another active treatment. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD004249.pub4]
Hayden JA, Cartwright J, van Tulder MW, Malmivaara A. Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD009790]
Heintjes EM, Berger M, Bierma-Zeinstra SMA, Bernsen RMD, Verhaar JAN, Koes BW. Exercise therapy for patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2003, Issue 4. [DOI: 10.1002/14651858.CD003472]
Henschke N, Ostelo RWJG, van Tulder MW, Vlaeyen JWS, Morley S, Assendelft WJJ, et al. Behavioural treatment for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 7. [DOI: 10.1002/14651858.CD002014.pub3]
Heymans MW, van Tulder MW, Esmail R, Bombardier C, Koes BW. Back schools for non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD000261.pub2]
Hilde G, Hagen KB, Jamtvedt G, Winnem M. Advice to stay active as a single treatment for low-back pain and sciatica. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD003632.pub2]
Hoving JL, Lacaille D, Urquhart DM, Hannu TJ, Sluiter JK, Frings-Dresen MHW. Non-pharmacological interventions

for preventing job loss in workers with inflammatory arthritis. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD010208.pub2]
Hurley M, Dickson K, Walsh N, Hauari H, Grant R, Cumming J, et al. Exercise interventions and patient beliefs for people with chronic hip and knee pain: a mixed methods review. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010842]
IJzelenberg W, Rubinstein SM, Hayden J, Koes BW, van Tulder MW. Exercise therapy for acute non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2011, Issue 10. [DOI: 10.1002/14651858.CD009365]
Jones G, Crotty M, Brooks P. Interventions for treating psoriatic arthritis. *Cochrane Database of Systematic Reviews* 2000, Issue 3. [DOI: 10.1002/14651858.CD000212]
Jordan JL, Holden MA, Mason EEJ, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD005956.pub2]
Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RWJG, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. *Cochrane Database of Systematic Reviews* 2014, Issue 9. [DOI: 10.1002/14651858.CD000963.pub3]
Karjalainen KA, Malmivaara A, van Tulder MW, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary rehabilitation for fibromyalgia and musculoskeletal pain in working age adults. *Cochrane Database of Systematic Reviews* 1999, Issue 3. [DOI: 10.1002/14651858.CD001984]
Karjalainen KA, Malmivaara A, van Tulder MW, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD002194]
Larun L, Brurberg KG, Odgaard-Jensen J, Price JR. Exercise therapy for chronic fatigue syndrome. *Cochrane Database of Systematic Reviews* 2016, Issue 2. [DOI: 10.1002/14651858.CD003200.pub4]
Liddle SD, Pennick V. Interventions for preventing and treating low-back and pelvic pain during pregnancy. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD001139.pub4]
Liu H, Yao K, Zhang J, Li L, Wu T, Brox JI, et al. Sling exercise therapy for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 9. [DOI: 10.1002/14651858.CD010689]
Miller J, Gross A, Kay TM, Graham N, Burnie SJ, Goldsmith CH, et al. Manual therapy with exercise for neck pain. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD011225]
Moi JHY, Sriranganathan MK, Edwards CJ, Buchbinder R. Lifestyle interventions for chronic gout. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD010039.pub2]
O'Brien K, Nixon S, Glazier R, Tynan AM. Progressive resistive exercise interventions for adults living with HIV/AIDS. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD004248.pub2]
O'Connell NE, Wand BM, McAuley J, Marston L, Moseley GL. Interventions for treating pain and disability in adults with complex regional pain syndrome- an overview of systematic reviews. *Cochrane Database of Systematic Reviews* 2013, Issue 4. [DOI: 10.1002/14651858.CD009416.pub2]
Page MJ, O'Connor D, Pitt V, Massy-Westropp N. Exercise and mobilisation interventions for carpal tunnel syndrome. *Cochrane Database of Systematic Reviews* 2012, Issue 6. [DOI: 10.1002/14651858.CD009899]
Page MJ, Green S, Kramer S, Johnston RV, McBain B, Chau M, et al. Manual therapy and exercise for adhesive capsulitis (frozen shoulder). *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD011275]
Peters S, Page MJ, Coppieters MW, Ross M, Johnston V. Rehabilitation following carpal tunnel release. *Cochrane Database of Systematic Reviews* 2013, Issue 6. [DOI: 10.1002/14651858.CD004158.pub2]
Preston NJ, Seers K, Mortimer PS. Physical therapies for reducing and controlling lymphoedema of the limbs. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD003141.pub2]
Proctor M, Murphy PA, Pattison HM, Suckling JA, Farquhar C. Behavioural interventions for dysmenorrhoea. *Cochrane Database of Systematic Reviews* 2007, Issue 3. [DOI: 10.1002/14651858.CD002248.pub3]
Radner H, Ramiro S, Buchbinder R, Landewé RBM, van der Heijde D, Aletaha D. Pain management for inflammatory arthritis (rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis and other spondyloarthritis) and gastrointestinal or liver comorbidity. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.CD008951.pub2]
Regnaux JP, Trinquart L, Boutron I, Nguyen C, Brosseau L, Ravaud P. High-intensity versus low-intensity physical activity or exercise in patients with hip or knee osteoarthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD010203]
Regnaux JP, Lefevre-Colau MM, Palazzo C, Poiraudeau S, Ravaud P, Boutron I. Exercise therapy for ankylosing spondylitis. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011321]
Richards BL, Whittle SL, Buchbinder R. Neuromodulators for pain management in rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.CD008921.pub2]
Riemsma RP, Kirwan JR, Taal E, Rasker HJJ. Patient education for adults with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD003688]
Schaafsma FG, Whelan K, van der Beek AJ, van der Es-Lambeek LC, Ojajärvi A, Verbeek JH. Physical conditioning as part of a return to work strategy to reduce sickness absence for workers with back pain. *Cochrane Database*

of Systematic Reviews 2013, Issue 8. [DOI: 10.1002/14651858.CD001822.pub3]
Steultjens EEMJ, Dekker JJ, Bouter LM, Schaardenburg DD, Kuyk MAMAH, Van den Ende ECHM. Occupational therapy for rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2004, Issue 1. [DOI: 10.1002/14651858.CD003114.pub2]
Stones W, Cheong YC, Howard FM, Singh S. Interventions for treating chronic pelvic pain in women. *Cochrane Database of Systematic Reviews* 2005, Issue 2. [DOI: 10.1002/14651858.CD000387]
Takken T, Van Brussel M, Engelbert RH, van der Net JJ, Kuis W, Helders PPJM. Exercise therapy in juvenile idiopathic arthritis. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD005954.pub2]
van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SMA, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010387]
van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD011142.pub2]
White CM, Pritchard J, Turner-Stokes L. Exercise for people with peripheral neuropathy. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD003904.pub2]
Zammit GV, Menz HB, Munteanu SE, Landorf KB, Gilheany MF. Interventions for treating osteoarthritis of the big toe joint. *Cochrane Database of Systematic Reviews* 2010, Issue 9. [DOI: 10.1002/14651858.CD007809.pub2]
Østerås N, Kjeken I, Smedslund G, Moe RH, Slatkowsky-Christensen B, Uhlig T, et al. Exercises for hand osteoarthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010388]

Additional references

Abdulla 2013
Abdulla A, Adams N, Bone M, Elliott AM, Gaffin J, Jones D, et al. British Geriatric Society. Guidance on the management of pain in older people. *Age and Ageing* 2013;**42**(Suppl 1):i1–57. [DOI: 10.1093/ageing/afs200; PUBMED: 23420266]

American College of Sports Medicine 2007
American College of Sports Medicine, American Heart Association. Exercise and acute cardiovascular events: placing the risks into perspective. *Medicine and Science in Sport and Exercise* 2007;**39**(5):886–97. [DOI: 10.1249/mss.0b013e3180574e0e; PUBMED: 17468590]

Benton 2011
Benton MJ, Kasper MJ, Raab SA, Waggener GT, Swan PD. Short-term effects of resistance training frequency on body composition and strength in middle-aged women. *Journal of Strength and Conditioning Research/National Strength & Conditioning Association* 2011;**25**(11):3142–9. [DOI: 10.1519/JSC.0b013e31820f505f]

Bigal 2012
Bigal ME, Rapoport AM. Obesity and chronic daily headache. *Current Pain and Headache Reports* 2012;**16**(1):101–9. [DOI: 10.1007/s11916-011-0232-0; PUBMED: 22076673]

Bouchard 2012
Bouchard C, Blair SN, Church TS, Earnest CP, Hagberg JM, Häkkinen K, et al. Adverse metabolic response to regular exercise: is it a rare or common occurrence?. *PLoS One* 2012;**7**(5):e37887. [DOI: 10.1371/journal.pone.0037887; PUBMED: 22666405]

Bousema 2007
Bousema EJ, Verbunt JA, Seelen HA, Vlaeyen JW, Knottnerus JA. Disuse and physical deconditioning in the first year after the onset of back pain. *Pain* 2007;**130**(3):279–86. [PUBMED: 17467902]

Breivik 2006
Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *European Journal of Pain* 2006;**10**(4):287–333. [PUBMED: 16095934]

Cohen 1988
Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edition. Hillsdale, NJ: L. Erlbaum Associates, 1988. [ISBN 0805802835]

Collins 1997
Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? . *Pain* 1997;**72**(1-2):95–7. [DOI: 10.1016/S0304-3959(97)00005-5; PUBMED: 9272792]

Daenen 2015
Daenen L, Varkey E, Kellmann M, Nijs J. Exercise, not to exercise, or how to exercise in patients with chronic pain? Applying science to practice. *Clinical Journal of Pain* 2015;**31**(2):108–14. [DOI: 10.1097/AJP.0000000000000099; PUBMED: 24662498]

Deschartes 2013
Deschartes A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ* 2013;**346**:f2304. [DOI: 10.1136/bmj.f2304]

Dworkin 2005
Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;**113**:9–19. [PUBMED: 15621359]

Dworkin 2008
Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Consensus Statement: Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *Journal of Pain* 2008;**9**(2):105–21. [DOI: 10.1016/j.jpain.2007.09.005; PUBMED: 18055266]

Eccleston 2014
Eccleston C, Fisher E, Craig L, Duggan GB, Rosser BA, Keogh E. Psychological therapies (Internet-delivered) for the management of chronic pain in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD010152.pub2]

Elliott 1999
Elliott AM, Smith BH, Penny KI, Smith WC, Chambers WA. The epidemiology of chronic pain in the community. *Lancet* 1999;**354**:1248–52. [PUBMED: 10520633]

Elzahaf 2012
Elzahaf RA, Tashani OA, Unsworth BA, Johnson MI. The prevalence of chronic pain with an analysis of countries with a Human Development Index less than 0.9: a systematic review without meta-analysis. *Current Medical Research & Opinion* 2012;**28**(7):1221–9. [DOI: 10.1185/03007995.2012.703132]

Engers 2008
Engers A, Jellema P, Wensing M, van der Windt DA, Grol R, van Tulder MW. Individual patient education for low back pain. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD004057.pub3; PUBMED: 18254037]

Enoka 1997
Enoka RM. Neural adaptations with chronic physical activity. *Journal of Biomechanics* 1997;**30**(5):447–55. [PUBMED: 9109556]

Exercise is Medicine
American College of Sports Medicine (ACSM). Exercise is Medicine. http://www.exerciseismedicine.org/ (accessed 27 June 2014).

Finan 2013
Finan PH, Smith MT. The comorbidity of insomnia, chronic pain, and depression: dopamine as a putative mechanism. *Sleep Medicine Reviews* 2013;**17**(3):173–83. [DOI: 10.1016/j.smrv.2012.03.003; PUBMED: 22748562]

Gross 2009
Gross A, Forget M, St George K, Fraser MM, Graham N, Perry L, et al. Patient education for neck pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. [DOI: 10.1002/14651858.CD005106.pub4; PUBMED: 22419306]

Harvard 2013
Better balance: mental and physical fitness are both essential: daily physical activity, targeted balance exercise, and steps to maintain your brain all help prevent falls. *Harvard Men's Health Watch* 2013;**17**(11):5. [PUBMED: 24000423]

Hoffman 2014
Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**. [DOI: 10.1136/bmj.g1687; PUBMED: 24609605]

IDF 2012
Guariguata L. By the numbers: new estimates from the IDF Diabetes Atlas Update for 2012. *Diabetes Research and Clinical Practice* 2012;**98**(3):524–5. [DOI: 10.1016/j.diabres.2012.11.006; PUBMED: 23217268]

IntHout 2015
IntHout J, Ioannidis JPA, Borm GF, Goeman JJ. Small studies are more heterogeneous than large ones: a meta-meta-analysis. *Journal of Clinical Epidemiology* 2015;**68**(8):860–9. [DOI: 10.1016/j.jclinepi.2015.03.017; PUBMED: 25959635]

Leadley 2014
Leadley RM, Armstrong N, Reid KJ, Allen A, Misso KV, Kleijnen J. Healthy aging in relation to chronic pain and quality of life in Europe. *Pain Practice* 2014;**14**(6):547–58. [DOI: 10.1111/papr.12125; PUBMED: 24138082]

Mayer 2008
Mayer J, Mooney V, Dagenais S. Evidence-informed management of low back pain with lumbar extensor strengthening exercises. *The Spine Journal* 2008;**8**(1):96–113. [DOI: 10.1016/j.spinee.2007.09.008; PUBMED: 18164458]

Merskey 2011
Merskey H, Bogduk N. Part III: Pain terms: a current list with definitions and notes on usage. *Classification of Chronic Pain*. 2nd Edition. Seattle: IASP Task Force on Taxonomy, IASP Press, 2011 [update from 1994].

Messier 2013
Messier SP, Mihalko SL, Legault C, Miller GD, Nicklas BJ, DeVita P, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. *JAMA* 2013;**310**(12):1263–73. [DOI: 10.1001/jama.2013.277669; PUBMED: 24065013]

Mooney 2006
Mooney V, Verna J, Morris C. Clinical management of chronic, disabling low back syndromes. In: Morris C editor(s). *Rehabilitation of the Spine: a Practitioner's Manual*. New York: McGraw-Hill, 2006.

Moore 1998
Moore RA, Gavaghanb D, Tramer MR, Collins SL, McQuay HJ. Size is everything - large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. *Pain* 1998;**78**:209–16. [PUBMED: 9870574]

Moore 2008
Moore RA, Barden J, Derry S, McQuay HJ. Managing potential publication bias. In: McQuay HJ, Kalso E, Moore RA editor(s). *Systematic Reviews in Pain Research: Methodology Refined*. Seattle: IASP Press, 2008:15–23. [ISBN: 978–0–931092–69–5]

Moore 2010
Moore RA, Moore OA, Derry S, Peloso PM, Gammaitono AR, Wang H. Responder analysis for pain relief and numbers needed to treat in a meta-analysis of etoricoxib

osteoarthritis trials: bridging a gap between clinical trials and clinical practice. *Annals of Rheumatic Diseases* 2010;**69** (2):374–9. [DOI: 10.1136/ard.2009.107805; PUBMED: 19364730]

Moore 2013
Moore RA, Straube S, Aldington D. Pain measures and cut-offs - 'no worse than mild pain' as a simple, universal outcome. *Anaesthesia* 2013;**68**(4):400–12. [DOI: 10.1111/anae.12148; PUBMED: 23347230]

Moore 2014a
Moore RA, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. *Pain Practice* 2014;**14**(1):79–94. [PUBMED: 23464879]

Moore 2014b
Moore RA, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. Pain Practice 2014; Vol. 14, issue 1:79–94. [DOI: 10.1111/papr.12050]

National Library of Medicine
National Library of Medicine. Medical Subject Headings. search "chronic pain" at www.nlm.nih.gov/mesh/MBrowser.html (accessed 22 August 2014). [MeSH unique ID: D059350]

NICE 2014
National Institute for Health and Care Excellence (NICE). Osteoarthritis. Care and management in adults (CG177), February 2014. www.nice.org.uk/guidance/cg177 (accessed 4 December 2015). [ISBN: 978–1–4731–0426–6]

Nnoaham 2008
Nnoaham KE, Kumbang J. Transcutaneous electrical nerve stimulation (TENS) for chronic pain. *Cochrane Database of Systematic Reviews* 2008, Issue 3. [DOI: 10.1002/14651858.CD003222.pub2; PUBMED: 18646088]

Nüesch 2010
Nüesch E, Trelle S, Reichenbach S, Rutjes AW, Tschannen B, Altman DG, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;**341**:c3515. [PUBMED: 20639294]

Nüesch 2011
Nüesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Jüni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011;**342**:d1165. [10.1136/bmj.d1165]

Park 2012
Park J, Hughes AK. Non-pharmacological approaches to the management of chronic pain in community-dwelling older adults: a review of empirical evidence. *Journal of the American Geriatric Society* 2012;**60**(3):555–68. [DOI: 10.1111/j.1532-5415.2011.03846.x; PUBMED: 22288789]

Perruchoud 2014
Perruchoud C, Buchser E, Johanek LM, Aminian K, Paraschiv-Ionescu A, Taylor RS. Assessment of physical activity of patients with chronic pain. *Neuromodulation* 2014;**17**:42–7. [DOI: 10.1111/ner.12036]

RevMan 2014 [Computer program]
The Nordic Cochrane Centre, Cochrane. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, Cochrane, 2014.

Sakuma 2012
Sakuma M, Endo N. Space flight/bedrest immobilization and bone. Exercise training for osteoporosis. *Clinical Calcium* 2012;**22**(12):1903–7. [DOI: CliCa121219031907; PUBMED: 23187084]

Schofield 2011
Schofield P, Clarke A, Jones D, Martin D, McNamee P, Smith BH. Chronic pain in later life: a review of current issues and challenges. *Aging Health* 2011;**7**(4):551–6. [DOI: 10.2217/ahe.11.41]

Shea 2007
Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007;**7**:10. [DOI: 10.1186/1471-2288-7-10; PUBMED: 17302989]

SIGN 2013
Scottish Intercollegiate Guidelines Network (SIGN). *Management of chronic pain. SIGN publication no. 136.* 1st Edition. Edinburgh, UK: SIGN, 2013. [ISBN: 978–1–909103–17–7]

Smith 2007
Smith BH, Macfarlane GJ, Torrance N. Epidemiology of chronic pain, from the laboratory to the bus stop: time to add understanding of biological mechanisms to the study of risk factors in population-based research?. *Pain* 2007;**127**(1-2):5–10. [DOI: 10.1016/j.pain.2006.11.001; PUBMED: 17140732]

Stagg 2011
Stagg NJ, Mata HP, Ibrahim MM, Henriksen EJ, Porreca F, Vanderah TW, et al. Regular exercise reverses sensory hypersensitivity in a rat neuropathic pain model: role of endogenous opioids. *Anaesthesiology* 2011;**114**(4):940–8. [DOI: 10.1097/ALN.0b013e318210f880; PUBMED: 21386701]

Straube 2015
Straube S, Derry S, Moore RA, McQuay HJ. Vitamin D for the treatment of chronic painful conditions in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 5. [DOI: 10.1002/14651858.CD007771.pub3; PUBMED: 25946084]

Thornton 2000
Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *Journal of Clinical Epidemiology* 2000;**53**(2):207–16. [DOI: 10.1016/S0895-4356(99)00161-4]

Timmons 2014
Timmon J, Phillips SM. Do we have a valid public health strategy for exercise prescription?. *Sport and Exercise Scientist* 2014;**39**:26–7.

To 2012
To T, Stanojevic S, Moores G, Gershon AS, Bateman ED, Cruz AA, et al. Global asthma prevalence in adults: findings from the cross-sectional world health study. *BMC Public Health* 2012;**12**:204. [DOI: 10.1186/1471-2458-12-204; PUBMED: 22429515]

Torrance 2010
Torrance N, Elliott AM, Lee AJ, Smith BH. Severe chronic pain is associated with increased 10 year mortality. A cohort record linkage study. *European Journal of Pain* 2010;**14**(4): 380–6. [DOI: 10.1016/j.ejpain.2009.07.006]

Turk 2003
Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003; **106**:337–45. [PUBMED: 14659516]

Turner 2013
Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane Reviews. *PLoS One* 2013;**8**(3):e59202. [DOI: 10.1371/journal.pone.0059202; PUBMED: 23544056]

Ursini 2011
Ursini F, Naty S, Grembiale RD. Fibromyalgia and obesity: the hidden link. *Rheumatology International* 2011;**31**(11): 1403–8. [DOI: 10.1007/s00296-011-1885-z; PUBMED: 21476098]

van Hecke 2013a
van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology - where do lifestyle factors fit in?. British Journal of Pain 2013; Vol. 7, issue 4:209–17. [DOI: 10.1177/2049463713493264]

van Hecke 2013b
van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology and its clinical relevance. *British Journal of Anaesthesia* 2013;**111**(1):13–8. [DOI: 10.1093/bja/aet123; PUBMED: 23794640]

Vigorito 2014
Vigorito C, Giallauria F. Effects of exercise on cardiovascular performance in the elderly. *Frontiers in Physiology* 2014;**5**: 51. [DOI: 10.3389/fphys.2014.00051]

Vos 2015
Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;**386**(9995):743–800. [DOI: 10.1016/S0140-6736(15)60692-4; PUBMED: 26063472]

Waddell 1987
Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine* 1987;**12**(7):632–44. [PUBMED: 2961080]

WHO 2015
World Health Organization. Physical activity. Fact sheet 385. www.who.int/mediacentre/factsheets/fs385/en/ (accessed 4 December 2015).

Wieland 2013
Wieland LS, Skoetz N, Manheimer E, Pilkington K, Vempati R, Berman BM. Yoga treatment for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 7. [DOI: 10.1002/14651858.CD010671]

Williams 2012
Williams ACDC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD007407.pub3]

References to other published versions of this review

Geneen 2014
Geneen L, Smith B, Clarke C, Martin D, Colvin LA, Moore RA. Physical activity and exercise for chronic pain in adults: an overview of Cochrane reviews. *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD011279]

* *Indicates the major publication for the study*

ADDITIONAL TABLES

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews

Criteria	Specific requirements (possible answers: yes, no, cannot answer, not applicable)
1. Was an 'a priori' design used?	The research question and inclusion criteria should be established before the conduct of the review *Note: need to refer to a protocol, ethics approval, or predetermined/a priori published research objectives to score a "yes."*

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews (*Continued*)

2. Was there duplicate study selection and data extraction?	There should be at least 2 independent data extractors and a consensus procedure for disagreements should be in place *Note: 2 people do study selection, 2 people do data extraction, consensus process or 1 person checks the other person's work.*
3. Was a comprehensive literature search performed?	At least 2 electronic sources should be searched. The report must include years and databases used (e.g. CENTRAL, MEDLINE, and Embase). Keywords or MeSH terms (or both) must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialised registers, or experts in the particular field of study, and by reviewing the references in the studies found *Note: if at least 2 sources + 1 supplementary strategy used, select "yes" (Cochrane register/ CENTRAL counts as 2 sources; a grey literature search counts as supplementary).*
4. Was the status of the publication (i.e. grey literature) used as inclusion criteria?	The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language, etc *Note: if review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished literature.*
5. Was a list of studies (included and excluded) provided?	A list of included and excluded studies should be provided. *Note: acceptable if the excluded studies were referenced. If there was an electronic link to the list but the link is no longer active, select "no. "*
6. Were the characteristics of the included studies provided?	In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions, and outcomes. The ranges of characteristics in all the studies analysed, e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported *Note: acceptable if not in table format as long as they are described as above.*
7. Was the scientific quality of the included studies assessed and documented?	'A priori' methods of assessment should be provided (e.g. for effectiveness studies if the author(s) chose to include only randomised, double-blind, placebo-controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant *Note: can include use of a quality scoring tool or checklist, e.g. Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some type of result for EACH study ("low" or "high" is acceptable, as long as it is clear which studies scored "low" and which*

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews (*Continued*)

	scored "high;" a summary score/range for all studies is not acceptable).
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations *Note: might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.*
9. Were the methods used to combine findings of studies appropriate?	For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi^2 test for homogeneity, I^2 statistic). If heterogeneity exists, a random-effects model should be used or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?), or both *Note: indicate "yes" if they mention or describe heterogeneity, i.e. if they explain that they cannot pool because of heterogeneity/variability between interventions.*
10. Was the likelihood of publication bias assessed?	An assessment of publication bias should include a combination of graphical aids (e.g. funnel plot, other available tests) or statistical tests (e.g. Egger regression test), or both *Note: if no test values or funnel plot included, score "no." Score "yes" if they mention that publication bias could not be assessed because there were fewer than 10 included studies.*
11. Was the conflict of interest stated?	Potential sources of support should be clearly acknowledged in both the systematic review and the included studies *Note: to get a "yes," must indicate source of funding or support for the systematic review AND for each of the included studies.*

Table 2. Reasons for exclusion

Review	Reason for exclusion from overview
Aggarwal 2011	Not exercise/physical activity
Brønfort 2015	Protocol stage only - possibly include when published as full review
Bierma-Zeinstra 2011	Protocol stage only - exclude when published as full review
Brønfort 2014	Withdrawn from the Cochrane Library
Choi 2010	Not chronic using definition of > 3 months
Craane 2006	Protocol stage only - possibly include when published as full review

Table 2. Reasons for exclusion (*Continued*)

Dagfinrud 2008	Physiotherapy - required therapist to perform intervention
Dahm 2010	Acute pain, not chronic. Intervention was advice
Dal Bello-Haas 2013	Malignant condition
de Souza 2012	Drug- and surgery-based interventions
Fokkenrood 2013	Did not include RCTs (excluded studies with control groups)
Franke 2015	Not exercise/physical activity
Green 2003	Physiotherapy - required therapist to perform intervention
Gross 1998	Withdrawn from the Cochrane Library
Gross 2012	Not exercise/physical activity
Gross 2015b	Not exercise/physical activity
Hayden 2012	Protocol stage only - possibly include when published as full review
Heintjes 2003	Withdrawn from the Cochrane Library January 2015
Henschke 2010	Not exercise/physical activity
Heymans 2004	Exercise could not be assessed as stand-alone intervention
Hilde 2006	Withdrawn from the Cochrane Library
Hoving 2014	No exercise intervention, and no pain outcome measure
Hurley 2013	Protocol stage only - exclude when published as full review
IJzelenberg 2011	Protocol stage only - exclude when published as full review
Jones 2000	Drug-based interventions
Jordan 2010	Intervention to improve adherence to exercise, not exercise itself
Kamper 2014	Exercise could not be assessed as stand-alone intervention
Karjalainen 1999	Exercise could not be assessed as stand-alone intervention
Karjalainen 2003	Exercise could not be assessed as stand-alone intervention

Table 2. Reasons for exclusion (*Continued*)

Larun 2016	Chronic fatigue, not chronic pain
Liddle 2015	Pain in pregnancy only, not chronic pain
Liu 2013	Protocol stage only - unsure about inclusion when published as full review
Miller 2014	Protocol stage only - exclude when published as full review
Moi 2013	Exercise could not be assessed as stand-alone intervention
O'Brien 2004	No pain outcome measure
O'Connell 2013	Overview of reviews, not systematic review
Østerås 2013	Protocol stage only - possibly include when published as full review
Page 2012	No pain outcome measure
Page 2014	Manual therapy - required therapist to perform intervention
Peters 2013	Exercise could not be assessed as stand-alone intervention
Preston 2004	No pain outcome measure
Proctor 2007	Exercise could not be assessed as stand-alone intervention
Radner 2012	Drug-based interventions
Regnaux 2014	Protocol stage only - possibly include when published as full review
Richards 2012	Not exercise/physical activity
Riemsma 2003	Not exercise/physical activity
Schaafsma 2013	No pain outcome measure
Steultjens 2004	Occupational therapy - exercise could not be assessed as stand-alone intervention
Stones 2005	Exercise cannot be assessed as stand-alone intervention
Takken 2008	Aged < 18 years - not adults
van Dessel 2014	Not chronic pain and no specific pain outcome measure
White 2004	No pain outcome measure
Williams 2012	Not exercise/physical activity

Table 2. Reasons for exclusion (*Continued*)

Zammit 2010	Surgery or required therapist to perform intervention

RCT: randomised controlled trial.

Table 3. Characteristics of included reviews

Review and Cochrane Review Group	Assessed as up to date	Chronic pain condition	Duration of pain/ diagnosis	Intervention description	Control description	Outcomes with data reported	Time points reported
Bartels 2007 Cochrane Musculoskeletal Group	Aug 2007	Hip or knee OA	Not reported	All types of exercises developed in the therapeutic/heated indoor pool (ROM, dynamics, aerobics, etc.) were permitted	No treatment or other treatment.	Function, quality of life, mental health, pain, adverse events	Post-intervention (immediate), 6-month follow-up
Bidonde 2014 Cochrane Musculoskeletal Group	Oct 2013	Fibromyalgia	12 yr (range 6 to 24)	Aquatic exercise training intervention defined as "exercise conducted in a vertical standing position."	Treatment as usual, physical activity as usual, wait list control, placebo or sham, education-only, water immersion-only, and attention only	Multi-dimensional function (wellness), self-reported physical function (wellness), pain (symptoms), stiffness (symptoms), muscle strength (physical fitness), submaximal cardiorespiratory function (physical fitness), withdrawals (safety and acceptability), adverse effects (safety and acceptability)	Post-intervention (4 to 32 wk)

Table 3. Characteristics of included reviews (*Continued*)

Boldt 2014 Cochrane Injuries Group	Mar 2011	Spinal cord injury	Mean 66 months, and 1 to 24 yr when reported	"Exercise": stretching and strengthening exercises aimed at mobilising painful shoulder joint	Wait list control or no intervention.	Pain, depression, quality of life, adverse effects	Short term (within 24 hours of last intervention, i.e. post-intervention) and intermediate term (1 to 6 wk post-intervention) and long term (> 6 wk post-intervention)
Brown 2010 Cochrane Menstrual Disorders and Subfertility Group	Aug 2009	Primary dysmenorrhoea in the majority (≥ 50%) of cycles	Ongoing/not appropriate	12-wk walk or jog training programme at an intensity of 70% to 85% of the HR range. Training for 3 days/wk and duration of aerobic phase was 30 minutes with 15-minute warm-up and cool-down periods	Asked not to exercise during the experimental period.	Pain: menstrual disorders questionnaire (MDQ) score	Ongoing - over 3 menstrual cycles
Busch 2007 Cochrane Musculoskeletal Group	Aug 2007	Fibromyalgia	Not reported	Exercise-only interventions included aerobic-only training, strength-only training, flexibility-only training, or mixed exercise-only interventions	"Untreated."	Pain, global wellbeing, objectively measured physical function	Post-intervention (strength exercise 21 wk, aerobic exercise 6 to 23 wk)
Busch 2013 Cochrane Musculoskeletal Group	Mar 2013	Fibromyalgia	mean range from 4 yrs (SD 3.1) to 12 yrs (SD 4)	Defined resistance training as exercise performed against a progressive	Untreated control conditions (treatment as usual, activity as usual, wait	Multidimensional function, self-reported physical function,	Post-intervention, follow-up (12 wk) in 1 study only

Table 3. Characteristics of included reviews (*Continued*)

				resistance on a minimum of 2 days/wk (on non-consecutive days) with the intention of improving muscle strength, muscle endurance, muscle power, or a combination of these	list control, and placebo), other types of exercise or physical activity interventions (e.g. aerobic, flexibility), and other resistance training interventions (head-to-head comparisons)	pain, tenderness, muscle strength, adverse effects, all-cause attrition	
Cramp 2013 Cochrane Musculoskeletal Group	Oct 2012	Rheumatoid arthritis	Not reported	Included pool-based therapy (twice/wk, moderate intensity, music-paced), yoga (6 wk, twice/wk, 1.5-hour sessions), dynamic strength training (home-based after inpatient programme, all main muscle groups using dumbbells and elastic bands), stationary cycling (70% HRmax, 5 minute excluding: 1-minute of rest, increased duration), low-impact aerobics (class	"Could have been placebo, an alternative intervention (pharmacological or non-pharmacological) or usual care."	Fatigue, pain, anxiety, depression, disability, tender and swollen joints, adverse events	Post-intervention (only a single time point analysed)

Table 3. Characteristics of included reviews (*Continued*)

				at fitness centre and video at home, individual HR targets), tai chi (1-hour group sessions)			
Fransen 2014 Cochrane Musculoskeletal Group	May 2013	Hip OA	Not reported	Any land-based therapeutic exercise regimens aiming to relieve the symptoms of hip OA, regardless of content, duration, frequency, or intensity. This included any exercise designed to improve muscle strength, range of joint movement or aerobic capacity (or combinations of the three). Programmes could be designed and supervised by physiotherapists or other professionals, or provided as a home programme with minimal monitoring	Wait-list control, usual care, GP education.	Self-reported pain, physical function, quality of life, withdrawal or dropouts, adverse events	post-intervention (immediate in 9/10 studies) follow-up 3 to 6 months
Fransen 2015 Cochrane Musculoskeletal Group	May 2013	Knee OA	Often not reported: some less than 1yr, oth-	"land-based therapeutic exercise." Along	No exercise: active (any no-exercise intervention) or no	Knee pain, self-reported physical function, quality of	Immediately at the end of treatment (post-

Table 3. Characteristics of included reviews (*Continued*)

			ers over 10yr	with delivery mode and content, treatment 'dosage' (duration, frequency, intensity) varied widely between studies	treatment (including waiting list)	life	treatment), 2 to 6 months after cessation of monitored study treatment and longer than six months after cessation of monitored study treatment
Gross 2015a Cochrane Back Group	May 2014	Mechanical neck disorders	"Chronic" (not subacute or acute)	Cervical stretch/ROM exercises + cervical/ scapulothoracic strengthening + static/ dynamic cervical/shoulder stabilisation	Wait list control.	Pain intensity, function, quality of life, global perceived effect, adverse effects	Immediately post-treatment (≤ 1 day), short-term follow-up (1 day to 3 months), intermediate-term follow-up (3 months up to, but not including, 1 yr), and long-term follow-up (≥ 1 yr)
Han 2004 Cochrane Musculoskeletal Group	Apr 2004	Rheumatoid arthritis	Not reported	Only trials of exercise programmes with tai chi instruction or incorporating principles of tai chi philosophy	Not reported.	Function, tender and swollen joints, ROM, strength, enjoyment, withdrawals, adverse effects	Post-intervention (8 to 10 wk)
Hayden 2005 Cochrane Back Group	Sep 2004	Non-specific low back pain	Chronic, i.e. longer than 12 wk: 5.6 yr (95% CI 3.4 to 7.8)	Exercise therapy defined as "a series of specific movements with the aim of training or developing the	No exercise: no treatment or placebo treatment, other conservative therapy, or another ex-	Pain, functional ability, work status, global assessment, adverse events	Earliest, 6 wk, 6 months, 12 months

Table 3. Characteristics of included reviews (*Continued*)

				body by a routine practice or as physical training to promote good physical health;" only 54% adequately described the exercise intervention	ercise group		
Hurkmans 2009 Cochrane Musculoskeletal Group	Jun 2009	Rheumatoid arthritis	5 to 14 yr	Dynamic exercise programmes - aerobic capacity and muscle strength training; short-term muscle strength training (high quality); short-term dynamic exercise to improve aerobic capacity (not high methodological quality); exercise frequency of at least 20 minutes twice a week. Duration of exercise programme at least 6 wk (duration < 3 months was considered short-term; duration > 3 months was considered long-term)	Not reported	Functional ability, aerobic capacity, muscle strength, safety (pain and radiological damage)	Follow-up (12 wk and 24 months)

Table 3. Characteristics of included reviews (*Continued*)

				. Exercise programme performed under supervision Aerobic exercise intensity at least 55% of the maximum HR; or intensity starting at 40% to 50% of the maximum oxygen uptake reserve or HR maximum reserve. Furthermore, the intensity was increased up to 85% during the intervention. Progressively strengthening exercise loads starting at 30% to 50% and increasing to 80% of maximum (defined as the percentage of either 1 repetition maximum, 1 MVC, maximum speed, or as maximal subjective exertion)			
Koopman 2015 Cochrane Neuromuscular Group	Jul 2014	Postpolio syndrome (PPS)	Not reported	Exercise therapy (e.g. aerobic exercise, muscle strengthening exercise, respiratory muscle	Placebo, usual care or no treatment.	Self-perceived activity limitations, muscle strength, muscle endurance, fatigue, pain, adverse events	3 and 6 months

Table 3. Characteristics of included reviews (*Continued*)

				training, warm climate training, hydro training)		(minor and serious)	
Lane 2014 Cochrane Peripheral Vascular Diseases Group	Sep-2013	intermittent claudication	not reported	Any exercise programme used in the treatment of intermittent claudication was included, such as walking, skipping and running. Inclusion of trials was not affected by the duration, frequency or intensity of the exercise programme but these issues were taken into account in the meta-analysis	Exercise was compared to six different modes of treatment, the most common being usual care or placebo. Two early trials compared exercise with placebo tablets but in more recent studies usual care was used as the control comparator. Exercise was compared with the following drug therapies: antiplatelet agents pentoxifylline, iloprost, and vitamin E. One study compared exercise with pneumatic foot and calf compression	maximal walking time, pain-free walking time, pain-free walking distance, maximum walking distance, ankle brachial index (ABI), peak exercise calf blood flow, mortality, amputation	Post-intervention, 3-month follow up, six-month follow up
Lauret 2014 Cochrane Peripheral Vascular Diseases Group	Jul 2013	Intermittent claudication	Not reported	Supervised walking programme needed to be supervised at least twice a week for a	Alternative exercise.	Maximum walking distance (METs), pain-free walking distance (METs), health-related	n/a

Table 3. Characteristics of included reviews (*Continued*)

				consecutive 6 wk of training		quality of life and functional impairment	
Regnaux 2015 Cochrane Musculoskeletal Group	Jun 2014	Hip or knee OA	> 6 months	High-intensity physical activity or exercise programme.	Low-intensity physical activity or exercise programme and control (no-exercise) group in 1 study.	Pain, physical function, quality of life, adverse effects (related to intervention), severe adverse events or withdrawal (due to intervention)	Post-intervention, intermediate term (6 to 12 months), long-term (over 12 months) follow-up
Saragiotto 2016 Cochrane Back and Neck Group	Apr 2015	Low back pain	> 12 wk	MCE: activation of the deep trunk muscles, targeting the restoration of control and co-ordination of these muscles	Placebo, no treatment, another active treatment, or when MCE was added as a supplement to other interventions. When MCE was used in addition to other treatments, it had to represent at least 50% of the total treatment programme to be included	Pain intensity and disability, function, quality of life, global impression of recovery, return to work, adverse events and recurrence	Post-intervention, short term (4 to 10 wk), intermediate term (3 to 6 months), long term (12 to 36 months)
Silva 2010 Cochrane Musculoskeletal Group	Jun 2009	Rheumatoid arthritis	No studies found	Balance training (proprioceptive training).	No intervention or other intervention.	ACR-50, pain, disease activity score (DAS), Health Assessment Questionnaire (HAQ for function), gait, adverse effects, discontinuation rate	n/a

Table 3. Characteristics of included reviews (*Continued*)

van der Heijden 2015 Cochrane Bone, Joint and Muscle Trauma Group	May 2014	Adolescents and adults with patellofemoral pain	3 wk to 8 months (as minimum requirement) ; reported pain 4 wk to 9 yr	Exercise therapy for patellofemoral pain syndrome; exercises could be performed at home or under supervision of a therapist - various descriptions in the included trials, including knee exercises, hip and knee exercises, home exercises, supervised exercises, closed kinetic chain, open kinetic chain	No treatment, placebo, or waiting list controls. This also included 'exercise therapy + another intervention (e.g. taping) versus the other intervention alone (e.g. taping).'	Pain during activity, usual pain, functional ability, recovery	4- to 12-wk follow-up (short term) and 16 wk to 12 months (long term)
Yamato 2015 Cochrane Back Group	Mar 2014	Low back pain	Acute, subacute, chronic (i.e. no minimum)	Explicitly stated as based on Pilates principles, or the therapists who provided the interventions had previous training in Pilates exercises or the therapists were described as certified Pilates instructors	No intervention, placebo, or other interventions.	Pain intensity, disability, global impression of recovery, quality of life, return to work, adverse effects	Short term (4 to 8 wk), intermediate term (3 to 6 months)

ACR: American College of Rheumatology; GP: general practitioner; HR: heart rate; MCE: motor control exercise; MET: metabolic equivalents; n/a: not applicable; OA: osteoarthritis; ROM: range of motion; wk: week; yr: year.

Table 4. Further characteristics of included reviews

Review	Number of trials included	Total number of participants	Gender distribution	Participants ages
Bartels 2007	6 (4 exercise vs no exercise)	800 (674 exercise vs no exercise)	50% to 86% Female	Means ranged from 66 to 71 yr
Bidonde 2014	16 (9 exercise vs no exercise)	881 (519 exercise vs no exercise)	513 female, 6 male	Means ranged from 46.3 to 48.3 yr
Boldt 2014	16 (3 exercise vs no exercise)	616 (149 exercise vs no exercise)	115 male, 41 female across 3 studies	Range 19 to 65 yr and mean 35 to 45 yr
Brown 2010	1	36	100% female	Not reported
Busch 2007	34 (in meta-analysis - strength training vs control: 2; aerobic training vs control: 4)	2276 total (in meta-analysis - strength: 47, aerobic: 269)	96.4% female when reported (in 2197 participants)	Range reported as 27.5 to 60.2 yr
Busch 2013	5 studies as 7 publications (exercise vs control: 3 publications, 2 studies)	219 with fibromyalgia (exercise vs control: 81)	100% female	Not reported
Cramp 2013	24 (only 6 using physical activity interventions)	2882 (physical activity interventions: 371)	"A higher percentage of females"… when reported	"Mainly within the fifth decade"
Fransen 2014	10	> 549	75% to 80% female when reported	58 to 70 yr (means) when reported
Fransen 2015	54	5362	When reported 55% to 100% female	When reported mean age 60 to 70 yr
Gross 2015a	27 (16 chronic pain)	2485	Not reported	Not reported
Han 2004	4 (3 RCTs). Pain not reported in any included study	206 total; pain not reported in any included study	Not reported	Range 38 to 72 yr
Hayden 2005	61 (43 chronic low back pain)	6390 (3907 chronic low back pain)	Chronic: 46% male (95% CI 39 to 52)	Chronic: 42 yr (95% CI 40 to 44)
Hurkmans 2009	8 RCTs (5 exercise vs no-exercise)	575	"Mainly female"	52 yr

Table 4. Further characteristics of included reviews (*Continued*)

Koopman 2015	13 (2 exercise vs no exercise)	675 (68 exercise vs no exercise) - 1 study used 3 arms (no treatment in cold, exercise in cold, exercise in warm; we have excluded the warm exercise arm as cannot compare directly to the control)	~ 25% male	Mean 58 and 65 yr
Lane 2014	30	1822 total	Not reported	Mean > 65 yr
Lauret 2014	5 (0 for exercise vs no exercise)	184 (0 for exercise vs no exercise)	n/a	n/a
Regnaux 2015	6 (1 for exercise vs no exercise) only 1 study that had a no exercise control	656 (102 for exercise vs no exercise)	79 female	62.6 yr
Saragiotto 2016	29 (7 for exercise vs no exercise/minimal intervention)	2431 (671 for exercise vs no exercise)	"Mixed"	Median 40.9 yr (IQR 11.2) (range 20.8 to 54.8)
Silva 2010	None	None	n/a	n/a
van der Heijden 2015	31 (10 for exercise vs control)	1690	0% to 100% female; equally distributed across range	Mean 25 to 50 yr
Yamato 2015	10 (6 exercise vs minimal intervention (control))	478 (265 exercise vs control)	2 trials were all female, the others included both genders	Mean 38 yr (range 22 to 50)

CI: confidence interval; GP: general practitioner; IQR: interquartile range; OA: osteoarthritis; RCT: randomised controlled trial; ROM: range of motion; wk: week; yr: year.

Table 5. Dose and duration of exercise interventions in included reviews

Review	Duration	Frequency (sessions per day/wk/month)	Intensity	Duration (per session)	Other description
Bartels 2007	Not reported	Not reported	"Muscle maintenance" and "range of motion"	Not reported	No minimum requirement for inclusion. Actual intervention only reported by 2

Table 5. Dose and duration of exercise interventions in included reviews *(Continued)*

					of 6 included studies
Bidonde 2014	17 wk (range 4 to 32)	1 to 4/wk	Very light (< 57% HRmax) to vigorous (95% HRmax), self-selected, and not specified	45 minutes (range 30 to 70)	No minimum requirement for inclusion. None of the studies met the ACSM exercise guidelines specified for aerobic or strength training. Only 1 study met the ACSM guidelines for flexibility training
Boldt 2014	12 wk to 9 months	2/day to 2/wk	Not reported	Reported for 1 study only (90 to 120 minutes)	No minimum requirement for inclusion. Stretching and strengthening exercises aimed at mobilising painful shoulder joint
Brown 2010	≥ 12 wk	3/wk	70% to 85% HRR	1 hour	No minimum requirement for inclusion.
Busch 2007	3 wk to 6 months	1 to 5/wk	Not reported	Not reported	No minimum requirement for inclusion. Assessed as whether they "met ACSM recommendations."
Busch 2013	8 to 21 wk (median 16 wk)	≥ 2/wk	> 4/10 RPE rating progressing to 70% to 80% 1RM	40 to 90 minutes	Assessed as whether they "met ACSM recommendations."
Cramp 2013	6 wk (when reported)	2/wk	"Low impact", "moderate", and 70% HRmax	1 to 1.5 hours, when reported	No minimum requirement for inclusion.
Fransen 2014	6 to 12 wk (median 8)	1 to 3/wk	"Low intensity" to "max effort"	30 to 60 minutes	No minimum requirement for inclusion. Intensity only re-

Table 5. Dose and duration of exercise interventions in included reviews (*Continued*)

					ported in 2 of 10 studies.
Fransen 2015	single session to 30 months	1 to 5/wk	"Moderate to moderately high intensity"	15 to 60 minutes	No minimum requirement for inclusion. Varied in dose and duration.
Gross 2015a	2 wk to 3 months	5/wk to every 15 minutes/day	Low intensity	2 to 20 minutes	-
Han 2004	8 to 10 wk (when reported)	1 to 7/wk (median 1/wk)	Tai chi = low intensity	1 to 1.5 hours	No minimum requirement for inclusion.
Hayden 2005	Not reported	Not reported	Not reported	Not reported	No minimum requirement for inclusion. Could not extract actual data.
Hurkmans 2009	≥ 6 wk	2/wk	Aerobic: ≥ 55% HRmax increasing to 85% HRmax strength: start 30% 1RM increasing to 80% 1RM	20 minutes	-
Koopman 2015	4 to 12 wk	Daily to 3/wk	Reported in 1 study: 50% to 70% MVC	45 minutes	No minimum requirement for inclusion. 1 study: supervised progressive resistance training consisting of 3 sets of 8 isometric contractions of the thumb muscles 1 study: combination of individual and group therapy with daily treatment in a swimming pool (45 minutes), physiotherapy, individually adapted training programme

Table 5. Dose and duration of exercise interventions in included reviews *(Continued)*

Lane 2014	3 to 12 months	≥ 2/wk	"Variable"	~ 60 minutes	No minimum requirement for inclusion.
Lauret 2014	≥ 6 wk	≥ 2/wk	Not reported	Not reported	No minimum requirement for inclusion. Must be supervised.
Regnaux 2015	8 wk	3/wk	Compared high vs low intensity vs control	30 to 50 minutes	Every 2 wk 1RM was retested and increased by 5% as tolerated in each group Supervision: an experienced therapist. 3 arms (n=34 per arm): high intensity, low intensity, control (no exercise)
Saragiotto 2016	20 days to 12 wk (median 8 wk (IQR 2.0))	1 to 5/wk (median 12 sessions (IQR 6.0))	Not reported	20 to 90 minutes (median 45 (IQR 30) minutes)	MCE is usually delivered in 1:1 supervised treatment sessions, and sometimes involves ultrasound imaging, the use of pressure biofeedback units or palpation to provide feedback on the activation of trunk muscles
Silva 2010	≥ 6 wk	2/wk	Balance training only	≥ 30 minutes	No studies found.
van der Heijden 2015	3 to 16 wk	2/wk to daily	Not reported	Not reported	No minimum requirement for inclusion. Assessed by duration (< or > 3 months), frequency (sev- eral times, or once a week), medium (land or water), etc

Table 5. Dose and duration of exercise interventions in included reviews *(Continued)*

Yamato 2015	10 to 90 days (mostly 8 wk)	2/wk (mean session number 15.3, range 6 to 30)	Not reported	1 hour	No minimum requirement for inclusion. Must be supervised (for the Pilates technique).

1RM: one repetition maximum; ACSM: American College of Sport Medicine; HRmax: maximum heart rate; HRR: heart rate reserve, IQR: interquartile range; MCE: motor control exercise; MVC: maximum voluntary contraction; RPE: rating of perceived exertion; wk: week.

Table 6. Methodological quality of included reviews using the AMSTAR tool

Review	Criteria											Total "Y"	Total "N"	Total "n/a"
	1	2	3	4	5	6	7	8	9	10	11			
Bartels 2007	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Bidonde 2014	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N	8	3	-
Boldt 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Brown 2010	Y	Y	Y	N	Y	Y	Y	Y	n/a	N	N	7	3	1
Busch 2007	Y	Y	Y	N	Y	Y	Y	Y	Y	N	N	8	3	-
Busch 2013	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Cramp 2013	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Fransen 2014	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Fransen 2015	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-

Table 6. Methodological quality of included reviews using the AMSTAR tool (*Continued*)

Gross 2015a	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Han 2004	Y	Y	Y	Y	Y	N	Y	Y	N	N	N	7	4	-
Hayden 2005	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	10	2	-
Hurkmans 2009	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9	2	-
Koopman 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	10	1	-
Lane 2014	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	9	2	-
Lauret 2014	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Regnaux 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Saragiotto 2016	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Silva 2010	Y	Y	Y	Y	Y	n/a	n/a	n/a	n/a	n/a	Y	6	0	5
van der Heijden 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9	2	-
Yamato 2015	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Total "Y"	**20**	**21**	**21**	**19**	**21**	**10**	**20**	**20**	**17**	**10**	**3**	-	-	-
Total "N"	**1**	**-**	**-**	**2**	**-**	**10**	**-**	**-**	**2**	**10**	**18**	-	-	-

Table 6. Methodological quality of included reviews using the AMSTAR tool (*Continued*)

Total "n/a"	-	-	-	-	-	1	1	1	2	1	-	-	-	-

N: no; n/a: not applicable; Y: yes; out of maximum summative score of 11.
Following arbitration, the authors removed the response "cannot answer" due to no responses as such.

Table 7. Risk of bias - studies assessed as low risk of bias

Review	Number of studies in assessment	Selection bias		Performance bias	Detection bias	Attrition bias	Reporting bias	Other bias	
		Random sequence generation (studies)	Allocation concealment (studies)	Blinding of participants and personnel (studies)	Blinding of outcome assessment (studies)	Incomplete outcome data (studies)	Selective reporting (studies)	Sample size	Other biases (studies)
Bartels 2007	6	Not reported	3	Not reported	2	3	Not reported	2, n > 100 per arm	-
Bidonde 2014	9	5	3	2	8	8	5	1, n > 50 per arm	7
Boldt 2014	3	1	1	0	1	2	3	0	1
Brown 2010	1	0	0	0	0	1	1	1, n > 50 per arm	-
Busch 2007	34	17	10	8	20	Unclear	32	5, n > 50 per arm	-
Busch 2013	5	4	2	1	2	5	3	0, n > 50 per arm	-
Cramp 2013	7	5	2	0	Not reported	6	4		1
Fransen 2014	10	8	7	0	0	7	4	1, n > 50 per arm	7
Fransen 2015	54	40	22	3	4	29	10	5, total n > 200	
Gross 2015a	16	8	8	1	0	11	0	0	11

Table 7. Risk of bias - studies assessed as low risk of bias *(Continued)*

Han 2004	4	2	0	0	0	0	Not reported	0	
Hayden 2005	43	27	22	Not reported	12	29	Not reported	10, total n > 100 + 5, total n > 200	-
Hurkmans 2009	8	8	1	-	4	5	-	1, total n > 200	1
Koopman 2015	2	1	0	0	0	0	0	0	1
Lane 2014	30	16	14	30	7	19	29	3, total n > 100	
Lauret 2014	5	4	2	5	3	4	5	1, total n > 100	4
Regnaux 2015	1	1	0	0	1	0	0	1, total n > 100	1
Saragiotto 2016	7	5	4	1	1	2	7	1, total n > 100 + 1, total n > 200	7
Silva 2010	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
van der Heijden 2015	10	8	6	0	0	6	9	2, total n > 100	10
Yamato 2015	9	5	5	2	7	7	9	0	9
Studies with low risk of bias (number)	**264**	**165**	**112**	**53**	**72**	**144**	**121**	**total n > 100: 26 total n > 200: 15 total n > 400: 0**	**71**
Studies with low risk of bias (per-	**-**	**63%**	**42%**	**20%**	**27%**	**55%**	**46%**	**total n > 100: 10% total n >**	**27%**

Table 7. Risk of bias - studies assessed as low risk of bias *(Continued)*

centage)								**200: 6% total n > 400: 0%**	

n: number of participants, n/a: not applicable.

Table 8. Interpretation of results by original review authors

Review	Review authors' conclusions	Overview authors' assessment of conclusions
Bartels 2007	"Aquatic exercise has some short-term beneficial effects on the condition of OA patients with hip or knee OA or both. The controlled and randomised studies in this area are still too few to give further recommendations on how to use this therapy... No long-term effects have been found."	Appropriate conclusions based on available data. No mention of quality/risk of bias in conclusions, though found to be high quality in results section
Bidonde 2014	"Low to moderate quality evidence relative to control suggests that aquatic training is beneficial for improving wellness, symptoms, and fitness in adults with fibromyalgia. Very low to low quality evidence suggests that there are benefits of aquatic and land-based exercise, except in muscle strength (very low quality evidence favoring land). No serious adverse effects were reported."	Appropriate conclusions based on available data.
Boldt 2014	"Evidence is insufficient to suggest that non-pharmacological treatments are effective in reducing chronic pain in people living with SCI. The benefits and harms of commonly used non-pharmacological pain treatments should be investigated in randomised controlled trials with adequate sample size and study methodology"	Appropriate conclusions based on available data.
Brown 2010	"There is a lack of available evidence to support the use of exercise in the alleviation of symptoms associated with dysmenorrhoea. The limited evidence implies that there are no adverse effects associated with exercise."	Review authors should not have commented on lack of adverse events as this was not reported in the included study. The comment on lack of adverse events contravened present Cochrane guidance
Busch 2007	"There is moderate quality evidence that short-term aerobic training (at the intensity recommended for increases in cardiorespiratory fitness) produces important benefits in people with FM in global outcome measures, physical function, and possibly pain and tender points. There is limited evidence that strength training improves a number of outcomes including pain, global wellbeing, physical function,	Appropriate conclusions based on available data.

Table 8. Interpretation of results by original review authors (*Continued*)

	tender points and depression. There is insufficient evidence regarding the effects of flexibility exercise. Adherence to many of the aerobic exercise interventions described in the included studies was poor."	
Busch 2013	"We have found evidence in outcomes representing wellness, symptoms, and physical fitness favoring resistance training over usual treatment and over flexibility exercise, and favoring aerobic training over resistance training. Despite large effect sizes for many outcomes, the evidence has been decreased to low quality based on small sample sizes, small number of randomized clinical trials (RCTs), and the problems with description of study methods in some of the included studies."	Appropriate conclusions based on available data.
Cramp 2013	"There is some evidence that physical activity interventions ... may help to reduce fatigue in RA. However, the optimal parameters and components of these interventions are not yet established."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion despite low/unclear quality score in results and discussion sections No conclusions about effect on pain (insufficient data).
Fransen 2014	"There is currently high-level evidence that land-based exercise will reduce hip pain, and improve physical function, among people with symptomatic hip osteoarthritis."	Evidence was good quality though sample sizes were often small (i.e. it is debatable if this was high level evidence as claimed by authors). Agree that results demonstrate small but significant benefit from intervention
Fransen 2015	"High-quality evidence suggests that land-based therapeutic exercise provides benefit in terms of reduced knee pain and quality of life and moderate-quality evidence of improved physical function among people with knee OA… Despite the lack of blinding we did not downgrade the quality of evidence for risk of performance or detection bias."	Appropriate conclusions based on available data. May have been generous with quality assessment but this was stated in conclusions for transparency
Gross 2015a	"…there is still no high quality evidence and uncertainty about the effectiveness of exercise for neck pain… Moderate quality evidence supports the use specific strengthening exercises as a part of routine practice … Moderate quality evidence supports the use of strengthening exercises, combined with endurance or stretching exercises may also yield similar beneficial results. However, low quality evidence notes when only stretching or only endurance type exercises … there may be minimal beneficial effects for both neck pain and function."	Appropriate conclusions based on available data.

Table 8. Interpretation of results by original review authors (*Continued*)

Han 2004	"Tai chi appears to have no detrimental effects on the disease activity of RA in terms of swollen/tender joints and activities of daily living…tai chi appears to be safe, since only 1 participant out of 121 withdrew due to adverse effects and withdrawals were greater in the control groups than the tai chi groups."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias in conclusion despite very low quality score in results section
Hayden 2005	"Evidence from randomized controlled trials demonstrates that exercise therapy is effective at reducing pain and functional limitations in the treatment of chronic low-back pain, though cautious interpretation is required due to limitations in this literature."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion despite low quality score in results and discussion sections
Hurkmans 2009	"Short-term, land-based dynamic exercise programs have a positive effect on aerobic capacity (aerobic capacity training whether or not combined with muscle strength training) and muscle strength (aerobic capacity training combined with muscle strength training) immediately after the intervention, but not after a follow-up period. Short-term, water-based dynamic exercise programs have a positive effect on functional ability and aerobic capacity directly after the intervention but it is unknown whether these effects are maintained after follow-up. Long-term, land-based dynamic exercise programs (aerobic capacity and muscle strength training) have a positive effect on functional ability, aerobic capacity, and muscle strength immediately after the intervention but it is unknown whether these effects are maintained after follow-up... Based on the evidence, aerobic capacity training combined with muscle strength training is recommended for routine practice in patients with RA."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion No conclusions regarding pain severity.
Koopman 2015	"Data from two single trials suggested that muscle strengthening of thumb muscles (very low-quality evidence) ... are safe and beneficial for improving muscle strength ... with unknown effects on activity limitations." "We found evidence varying from very low quality to high quality that ... rehabilitation in a warm or cold climate are not beneficial in PPS." "Due to a lack of good-quality data and randomised studies, it was impossible to draw definitive conclusions about the effectiveness of interventions in people with PPS."	Appropriate conclusions based on available data.
Lane 2014	"… Exercise therapy should play an important part in the care of selected patients with intermittent claudication, to improve walking times and distances. Ef-	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of stud-

Table 8. Interpretation of results by original review authors (*Continued*)

	fects were demonstrated following three months of supervised exercise although some programmes lasted over one year."	ies in conclusion No conclusions regarding pain severity.
Lauret 2014	"There was no clear evidence of differences between supervised walking exercise and alternative exercise modes in improving the maximum and pain-free walking distance of patients with intermittent claudication…. The results indicate that alternative exercise modes may be useful when supervised walking exercise is not an option for the patient."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion (in discussion)
Regnaux 2015	"We found very low- to low-quality evidence for no important clinical benefit of high-intensity compared to low-intensity exercise programs in improving pain and physical function in the short term.... The included studies did not provide any justification for the levels of intensity of exercise programs. No authors reported evidence for the minimal and maximal intensity that could be delivered."	Appropriate conclusions based on available data. This overview has only used one study of the six included as it alone included a control group, for which we could not extract data as the control comparison was not used in the analysis by the review authors
Saragiotto 2016	"There is very low to moderate quality evidence that MCE has a clinically important effect compared with a minimal intervention for chronic low back pain... As MCE appears to be a safe form of exercise and none of the other types of exercise stands out, the choice of exercise for chronic low back pain should depend on patient or therapist preferences, therapist training, costs and safety."	Appropriate conclusions based on available data.
Silva 2010	"We were not able to provide any evidence to support the application of balance exercises (proprioceptive training) alone in patients with RA."	Appropriate conclusions based on available data (no included studies)
van der Heijden 2015	"This review has found very low quality but consistent evidence that exercise therapy for patellofemoral pain syndrome (PFPS) may result in clinically important reduction in pain and improvement in functional ability."	No subgroup analysis to differentiate between acute, subacute, and chronic pain made it difficult to extract appropriate data for this review
Yamato 2015	"No definite conclusions or recommendations can be made as we did not find any high quality evidence for any of the treatment comparisons, outcomes or follow-up periods investigated. However, there is low to moderate quality evidence that Pilates is more effective than minimal intervention in the short and intermediate term as the benefits were consistent for pain intensity and disability, with most of the effect sizes being considered medium."	Appropriate conclusions based on available data. There was no subgroup analysis to differentiate between acute, subacute, and chronic pain made it difficult to extract appropriate data for this review (one included study had subacute back pain (> 6 weeks), all others were chronic back pain (> 12 weeks)) but results are presented altogether as chronic pain

FM: fibromyalgia; MCE: motor control exercise; OA: osteoarthritis; PPS: postpolio syndrome; RA: rheumatoid arthritis; SCI: spinal cord injury.

Table 9. Pain severity

Review	Number of trials (and participants) assessing 'pain severity'	Baseline pain score	Post-intervention reported result or change data (or if only one data point reported in review)	Follow-up	Overall comment/ statement
Bartels 2007 (osteoarthritis)	Hip + knee OA: Post-intervention: 4 (638) Follow-up: 1 (310) Hip only: follow-up: 1 (17) Knee only: post-intervention: 1 (46)	Control baseline: Hip + knee OA WOMAC 0 to 20 (2 studies): 9.10 (SD 3.14) VAS 0 to 100 (1 study): 55.3 (SD 24.6) HAQ 0 to 3 (1 study): 1.05 (SD 0.61) Hip only VAS 0 to 100 (1 study): 56 (SD 21.89) Knee only VAS 0 to 10 (1 study): 5.6 (SD 1.4)	Hip + knee OA A minor effect of a 3% absolute reduction (0.6 fewer points on WOMAC 0 to 20 scale) and 6.6% relative reduction SMD 0.19 (95% CI 0.04 to 0.35) (P = 0.02) Knee only SMD 0.86 (95% CI 0.25 to 1.47) (P = 0.005) Absolute difference 12% (1.2 fewer points on a 0 to 10 scale) Relative change 22% improvement	Hip + knee OA Follow-up at 6 months: SMD 0.11 (95% CI -0.12 to 0.33) (ns) No difference Hip only SMD 1.00 (95% CI -0.04 to 2.04) (P = 0.06, ns)	Statistically significant post-intervention in hip + knee OA group, but not clinically significant Knee-only OA had moderate to large effect size (statistically significant) immediately post-intervention
Bidonde 2014 (fibromyalgia)	Post-intervention: 7 (382)	Weighted mean score at baseline (all participants): 69.59 median value for pain was 70.9 in studies comparing aquatic training to control	On 100-point scale: MD -6.59 (95% CI -10.71 to -2.48) SMD -0.53 (95% CI -0.76 to -0.31) Absolute difference -7% (95% CI -11 to -3) NNTB 5 (95% CI 3 to 8)	3 studies at 12, 48, or 52 weeks' post-intervention could not be combined. 2 studies showed SMD favouring intervention at follow-up.	"We found a moderate effect favouring the aquatic exercise training for pain" …"similar improvements in pain in the low pain groups (SMD -0.60, 95% CI -0.98 to -0.23) and in the high pain groups (SMD -0.57, 95% CI -1.11 to -0.03)." Among the

Table 9. Pain severity (*Continued*)

					major wellness outcomes, none of the outcomes met the threshold for clinically relevant differences (15%)
Boldt 2014 (spinal cord injury)	Post-intervention: 3 (149)	WUSPI score 22.6 (exercise group) to 11.05 (control group) in 1 group at baseline Not reported for 2 studies	WUSPI change score: Exercise group: -7.7 (SD 19.01) Control group: 12.8 (SD 12.74) SF-36 (pain experience): -1.9 (95% CI -3.4 to -0.4) favoured exercise (P = 0.01) VAS (0 to 10): MD -2.8 (95% CI -3.77 to -1.83) favoured exercise (P < 0.00001)	1 study at 4 weeks: VAS (0 to 10): -2.50 (95% CI -3.48 to -1.52) (P < 0.00001) WUSPI: -26.40 (95% CI -37.62 to -15.18favoured exercise (P < 0.00001)	"All three studies were fraught with high overall risk of bias. In particular, the comparison with 'no treatment' or waiting lists as control interventions likely leads to an overestimation of the effectiveness of the exercise programmes provided in these studies. Consequently, no conclusion on their effectiveness can be drawn."
Busch 2007 (fibromyalgia)	Strength training: 1 (21) Aerobic training: 3 (183)	Control baseline: Aerobic: 6.1/10 (VAS) (SD 1.97) Strength: 35/100 (VAS) (SD 19)	Aerobic training: SMD 0.65 (95% CI -0.09 to 1.39) (ns) Weighted absolute change 13% (1.3 cm lower on 10-cm scale) Relative change 21% Strength training: SMD 3.00 (95% CI 1.68 to 4.32) (ns) Weighted absolute change 49% (49 points lower on 100-point scale) Relative change 140%, NNTB 2	n/a	">30% improvement was seen in the strength training group as compared to an untreated control group in pain." Aerobic training led to an improvement of 1.3/10.

Table 9. Pain severity (*Continued*)

Busch 2013 (fibromyalgia)	Post-intervention: 2 (81) Follow-up at 8 weeks, 16 weeks, 28 weeks: 1 (60)	Not reported - change data only	Change score on VAS (in cm): MD -3.30 (95% CI -6.35 to -0.26) (P = 0.03) SMD -1.89 (95% CI -3.86 to 0.07) Relative % change 44.6% (95% CI 3.5 to 85.9) favoured exercise NNTB 2 (95% CI 1 to 34)	8 weeks: MD -0.68 (95% CI -1.62 to 0.26) (ns) 16 weeks: MD -1.79 (95% CI -2.70 to -0.88) (P < 0.001) 28 weeks: MD -0.85 (95% CI -1.77 to 0.07) (P = 0.07, ns) Overall (n = 180): MD -1.12 (95% CI -1.65 to -0.58) (P < 0.0001)	> 30% improvement post-intervention.
Cramp 2013 (rheumatoid arthritis)	4 (not reported)	Not reported	In narrative only - Harkcom 1985: statistics not reported separately for pain data, but reported as improvement over time; Hakkinen 2003: "stat significant improvement in 24 months"; Evans 2012 and Wang 2008: no statistically significant effects	Not reported	"Improvement over time" with "significant improvement in 24 months." No actual data available.
Fransen 2014 (OA)	End of treatment: 9 (549) 3 to 6 months: 5 (391)	Not reported; land based exercise vs no exercise: mean pain in control group ~ 29/100 (based on 9 studies' control values)	End of treatment: SMD -0.38 (95% CI -0.55 to -0.20) "small to moderate" favoured exercise (P < 0.0001)	3 to 6 months: SMD -0.38 (95% CI -0.58 to -0.18) "small to moderate" favoured exercise (P = 0.0002)	"Small to moderate" statistically significant improvement, but only mild pain at baseline
Fransen 2015 (OA)	End of treatment: 44 (3537) Follow-up (2 to 6 months): 12 (1468) Follow-up (> 6 months): 8 (1272)	Not reported; land-based exercise vs no exercise: mean pain in control group 44/100 (based on 1 study control values)	Land-based exercise vs no exercise: Mean pain in intervention groups was 0.49 SDs lower (95% CI 0.39 to 0.59 lower). This translates to an absolute mean reduction of 12 points	2 to 6 months: SMD -0.24 (95% CI -0.35 to -0.14) favoured exercise (P < 0.00001) > 6 months: SMD -0.52 (95% CI -1.01 to -0.03) favoured exercise (P = 0.04)	Absolute improvement of 12/100 post-intervention (statistically significant)

Table 9. Pain severity (*Continued*)

			(95% CI 10 to 15) compared with control group on a 0 to 100 scale SMD -0.49 (95% CI -0.39 to -0.59) (P < 0.00001) Absolute reduction 12% (95% CI 10% to 15%) Relative change 27% (95% CI 21% to 32%) NNTB 4 (95% CI 3 to 5)		
Gross 2015a (mechanical neck disorders)	12-week treatment: 2 (147) 24 week (or 12-week treatment + 12-week follow-up) : 2 (140)	Not reported, but control scores at end of treatment 40 to 60/100 (moderate pain)	12 weeks: pooled MD -14.90 (95% CI -22.40 to -7.39) favoured exercise (P = 0.0001)	24 weeks: pooled MD -10.94 (95% CI -18.81 to -3.08) favoured exercise (P = 0.0064)	2 trials showed a moderate (statistically significant) reduction in pain post-intervention (14.9/100)
Hayden 2005 (low back pain)	Earliest follow-up: 8 (370) Follow-up (time since randomisation) Short term (6 weeks): 6 (268) Intermediate term (6 months): 5 (249) Long term (12 months): 2 (126)	"Chronic group" at baseline: mean 46/100 (95% CI 41 to 50) (moderate pain)	Earliest: MD -10.20 (95% CI -19.09 to -1.31) (P = 0.02)	Short term: MD -8.58 (95% CI -18.46 to 1.29) (P = 0.09, ns) Intermediate term: MD -12.48 (95% CI -22.69 to -2.27) (P = 0.02) Long term: MD -3.93 (95% CI -9.89 to 2.02) (P = 0.2, ns)	Reduction of ~ 10/100 at earliest measurement point.
Hurkmans 2009 (rheumatoid arthritis)	4 studies (total 188 participants) in different categories (results not combined)	Not reported	Short-term (12 weeks): Short-term land-based (aerobic and strength training) SMD -0.53 (95% CI -1.09 to 0.04) Short-term land-based (aerobic only) SMD -0.27 (95% CI -0.79 to 0.26) Short-term water-based SMD 0.06	Long-term (24 months) land-based (aerobic and strength training) SMD 0.35 (95% CI -0.46 to 1.16)	No significant difference between control and intervention.

Table 9. Pain severity (*Continued*)

			(95% CI -0.43 to 0.54)		
Koopman 2015 (postpolio syndrome)	1 (55)	Not reported, but control scores at end of treatment mean 44 (SD 24) on a 0 to 100 scale (moderate pain)	3 months post-intervention: VAS (0 to 100): MD 11.00 (95% CI -0.98 to 22.98) (P = 0.072)	n/a	No significant effect/no difference between groups.
Regnaux 2015 (OA)	Only 1 study that had a no-exercise control: 1 (68) - excluded data for control (no exercise) from analysis (n = 34)	Not reported	Post-intervention: WOMAC (0 to 20) Change data presented for high- vs low-intensity groups only, not compared to control	n/a	Actual individual study data was extracted (where possible) instead of pooled MD or SMD due to comparison this overview wishes to make (exercise vs no-exercise only) Could not extract exercise vs control data.
Saragiotto 2016 (low back pain)	Short term (< 3 months): 4 (291) Intermediate term (3 to 12 months): 4 (348) Long term (> 12 months): 3 (279)	Not reported, but control scores at follow-up range 25 to 56/100 (mild-moderate pain)	Short term: MD -10.01 (95% CI -15.67 to -4.35) favoured exercise (P < 0.001)	Intermediate term: MD -12.61 (95% CI -20.53 to -4.69) favoured exercise (P = 0.002) Long term: MD -12.97 (95% CI -18.51 to -7.42) favoured exercise (P < 0.001)	Medium effect size favouring exercise at all follow-up assessments (moderate quality evidence at short- and long-term, low quality evidence at intermediate term) Clinically important effect.
van der Heijden 2015 (patellofemoral pain syndrome)	3 studies with pain > 3 months (135 participants), 2 studies used in analysis (41 participants) Long-term follow-up: 1 (94)	Not reported, but control scores at follow-up range 2.1 to 6.0/10 (mild-moderate pain)	Short-term (4 to 8 weeks): MD for usual pain in the exercise group was 0.93 (95% CI 1.60 to 0.25) SDs lower SMD -0.93 (95% CI -1.60 to -0.25) (P = 0.008)	"Long term" (16 weeks) VAS (0 to 10): MD -4.42 (95% CI -7.75 to -0.89) favoured exercise (P = 0.01)	Reduction in pain of 4/10 at 16 weeks' follow-up.
Yamato 2015 (low back pain)	Short term: 6 (265) Intermediate term:	Not reported, but control scores at ear-	Short-term follow-up (< 3 months):	Intermediate term (3 to 12 months):	"Low quality evidence (downgraded

Table 9. Pain severity (*Continued*)

	2 (148)	liest follow-up range 18 to 52/100 (mild-moderate pain)	MD -14.05 (95% CI -18.91 to -9.19) (P < 0.001)	MD -10.54, (95% CI -18.54 to -2.62) (P = 0.009)	due to imprecision and risk of bias) that Pilates reduces pain compared with minimal intervention at short-term follow-up, with a medium effect size.. intermediate-term follow-up, two trials, provided moderate quality evidence (downgraded due to imprecision) that Pilates reduces pain compared with minimal intervention, with a medium effect size"

CI: confidence interval; HAQ: Health Assessment Questionnaire; MD: mean difference; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; ns: not significant; OA: osteoarthritis; SD: standard deviation; SF-36: 36-item Short Form; SMD: standardised mean difference; VAS: visual analogue score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; WUSPI; Wheelchair User Shoulder Pain Index.

Table 10. Physical function

Review	Outcome measure	Number of trials (and participants) used in analysis	Post-intervention result (or if only 1 result reported)	Short-term follow-up (or if only 1 follow-up point reported)	Intermediate-term follow-up	Long-term follow-up	Overall comment/ statement
Bartels 2007 (OA)	Self-reported function (WOMAC and HAQ) and walking ability, and DRI	Post-intervention Hip + knee function: 4 (648) walking ability: 2 (355) Hip only function: 1 (28) Follow-up function hip + knee: 1 (306) hip only: 1 (17)	Function (hip + knee): SMD 0.26 (95% CI 0.11 to 0.42) favoured exercise (P < 0.001) Walking (hip + knee): SMD 0.18 (95% CI -0.03 to 0.39) favoured exercise (P = 0.08, ns)	Hip only Disability, SMD 1.00 (95% CI -0.04 to 2.04) favoured exercise (P = 0.06, ns)	Hip + knee (6 months) Function, SMD 0.10 (95% CI -0.12 to 0.33) (ns)	n/a	Function was significantly improved in people with hip + knee OA immediately post-intervention only - small effect size only

Table 10. Physical function (*Continued*)

			Function (hip only): SMD 0.76 (95% CI -0.02 to 1.53) favours exercise (P = 0.06, ns)				
Bidonde 2014 (fibromyalgia)	Self-reported physical function (0 to 100 scale)	5 (285)	MD -4.35 (95% CI -7.77 to -0.94) SMD -0.44 (95% CI -0.76 to -0.11) Absolute difference -4 (95% CI -8 to -1) NNTB 6 (95% CI 3 to 22)	n/a	n/a	n/a	Small difference (improvement) in aquatic exercise group. Among the major wellness outcomes, none of the outcomes met the threshold for clinically relevant differences (15%)
Busch 2007 (fibromyalgia)	Physical function	Aerobic: 4 (253) Strength: 2 (47)	Aerobic: SMD 0.66 (95% CI 0.41 to 0.92) favoured exercise (P < 0.0001) Strength: SMD 0.52 (95% CI -0.07 to 1.10) favoured exercise (P = 0.08, ns)	n/a	n/a	n/a	Function was significantly improved from aerobic exercise training, strength training neared significance Moderate effect size.
Busch 2013 (fibromyalgia)	HAQ and SF-36 for function	3 (107)	Change score MD -6.29 (95% CI -10.45 to -2.13) favoured exercise (P < 0.01)	n/a	n/a	n/a	Significantly favourable effect of exercise.
Cramp 2013 (rheumatoid arthritis)	Disability	4 (not reported)	n/a	n/a	n/a	n/a	"Studies investigating hydrotherapy and tai chi demon-

Table 10. Physical function (*Continued*)

							strated statistically significant improvements in the intervention arm compared to the control arm between baseline and follow-up. The studies investigating strength training and Ivengar yoga did not demonstrate a statistically significant difference between study arms."
Fransen 2014 (OA)	Physical function	Post-intervention: 9 (521) Follow-up (3 to 6 months): 5 (365)	SMD -0.30 (95% CI -0.54 to -0.05) "significant benefit" favoured exercise (P = 0.02) The demonstrated effect size for exercise was equivalent to an improvement of physical function of 7 points (95% CI 1 to 12) on a 0 to 100 scale compared with a control group	SMD -0.37 (95% CI -0.57 to -0.16) favoured exercise (P < 0.001)	n/a	n/a	Statistically significant, but small effect size only.
Fransen 2015 (OA)	Physical function	Post-intervention: 44 (3913) Follow-up (2 to 6 months):	SMD -0.52 (95% CI -0.64 to -0.39) favoured exercise (P < 0.	SMD -0.15 (95% CI -0.26 to -0.04) favoured exercise (P = 0.	SMD -0.57 (95% CI -1.05 to -0.10) favoured exercise (P = 0.02)	n/a	Significant effect from exercise at every follow-up point.

Table 10. Physical function (*Continued*)

		10 (1279) Follow-up (> 6 months): 8 (1266)	0001); an improvement of 10 points (95% CI 8 to 13) on a 0- to 100-point scale	008)			Moderate effect size at short- and long-term follow-up, but only small effect at intermediate-term follow-up
Gross 2015a (mechanical neck disorders)	Physical function	12 wk: 2 (147) 24 wk: 2 (140)	12 wk treatment: pooled SMD -0.50 (95% CI -1.04 to 0.03) favoured exercise (P = 0.07, ns)	24 wk treatment (or 12 wk' treatment + 12 wk follow-up) : pooled SMD -0.40 (95% CI -0.74 to -0.06) favoured exercise (P = 0.02)	n/a	n/a	2 trials showed a moderate (statistical) improvement in function
Han 2004 (rheumatoid arthritis)	Functional assessment and 50-feet walk test	Function: 2 (52) Walk test: 2 (48)	Function: MD 0.01 (95% CI -2.94 to 2.97) (ns) Walk test: MD 0.35 seconds (95% CI -1.14 to 1.84) (ns)	n/a	n/a	n/a	No significant effect.
Hayden 2005 (low back pain)	Function	Earliest: 7 (337) Short term: 6 (268) Intermediate term: 4 (216) Long term: 2 (126)	Earliest: MD -2.98 (95% CI -6.48 to 0.53) favoured exercise (P = 0.09, ns)	Short term: MD -3.03 (95% CI -6.35 to 0.53) favoured exercise (P = 0.07, ns)	Intermediate term: MD -3.84 (95% CI -7.06 to -0.61) favoured exercise (P = 0.02)	Long term: MD -4.22 (95% CI -7.99 to -0.46) favoured exercise (P = 0.03)	Favoured exercise from the earliest measure, but only reached statistical significance at intermediate and long term after randomisation
Hurkmans 2009 (rheumatoid arthritis)	Functional ability	Land-based aerobic: 2 (66) Land-based aerobic + strength: 2 (74)	n/a	Short-term training (12 wk) Land-based aerobic only training	n/a	n/a	No significant difference between control and intervention groups

Table 10. Physical function (*Continued*)

				SMD 0.03 (95% CI -0.46 to 0.51) (ns) Land-based aerobic and strength training SMD -0.4 (95% CI -0.86 to 0.06) (ns)			
Koopman 2015 (postpolio syndrome)	Muscle strength; and activity limitation (Sunnaas ADL-index range 0 to 36; Rivermead Mobility Index (RMI) range 0 to 15)	Strength: 1 (10) Activity limitation: 1 (53)	Isometric muscle strength (postintervention): MD 39.00% (95% CI 6.12 to 71.88) Activity limitation: 3 months' postintervention: ADL-index: MD -2.70 (95% CI -4.53 to -0.87) Rivermead Mobility Index (RMI): MD -1.50 (95% CI -2.93 to -0.07)	Activity limitation: 6-months postintervention: ADL-index: MD -2.90 (95% CI -4.73 to -1.07) RMI: MD -1.80 (95% CI -3.19 to -0.41)	n/a	n/a	Activity limitation: favoured intervention at both assessment points "The baseline imbalance in favour of the usual care group probably biased these results."
Lane 2014 (intermittent claudication)	Maximal walking time and maximal walking distance	Postintervention Walking time: 12 (577) Walking distance: 9 (480) 3-month follow-up Walking time: 5 (174) Walking distance: 3 (116) 6-month follow-up Walking time:	Time: MD 4.51 minutes (95% CI 3.11 to 5.92) favoured exercise (P < 0.00001) Distance: 108.99 m (95% CI 38.20 to 179.78) favoured exercise (P = 0.003)	Time: MD 6.05 minutes (95% CI 5.47 to 6.62) favoured exercise (P < 0.00001) Distance: MD 104.46 m (95% CI -64.33 to 273.24) favoured exercise (ns)	Time: MD 3.20 minutes (2.04 to 4.36) favoured exercise (P < 0.0001) Distance: MD 138.36 m (95% CI 22.39 to 254.34) favoured exercise (P = 0.02)	n/a	Objectively measured walking time and distance showed significant improvement

Table 10. Physical function (*Continued*)

		4 (295) Walking distance: 3 (156)					
Lauret 2014 (intermittent claudication)	Maximal walking time (mins) and maximal walking distance (metres)	No relevant studies	n/a	n/a	n/a	n/a	No relevant studies.
Regnaux 2015 (OA)	WOMAC (0 to 68) disability scale, and muscle strength	1 (68) - excluded control (no-exercise data: n = 34)	n/a	n/a	n/a	n/a	Could not extract exercise vs control data - data presented for high vs low intensity groups only, not compared to control
Saragiotto 2016 (low back pain)	Disability (Oswestry Disability Index, Roland Morris Disability Questionnaire)	Short-term follow-up (< 3 months): 5 (332) Intermediate term (3 to 12 months): 4 (348) Long term (> 12 months): 3 (279)	-	MD -8.63 (95% CI -14.78 to -2.47) (P < 0.01)	MD -5.47 (95% CI -9.17 to -1.77) (P = 0.004)	MD -5.96 (95% CI -9.81 to -2.11) (P = 0.002)	Small effect sizes, favoured exercise. Short term: CI included a clinically important effect.
Silva 2010 (rheumatoid arthritis)	HAQ function	No studies found	n/a	n/a	n/a	n/a	No studies found.
van der Heijden 2015 (patellofemoral pain syndrome)	Functional ability	Short-term follow-up: 7 (483) Long-term follow-up: 3 (274)	n/a	Short-term (4 to 8 wk): SMD 1.10 (95% CI 0.58 to 1.63) favoured exercise (P < 0.0001)	n/a	SMD 1.62 (95% CI 0.31 to 2.94) favoured exercise (P = 0.02)	Significant effect of exercise. Very large effect size at short- and long-term follow-up.
Yamato 2015 (low back pain)	Disability (all measures con-	Short-term (< 3 months) fol-	n/a	MD -7.95 (95% CI -13.	MD -11.17 (95% CI -18.	n/a	"Low quality ev-

Table 10. Physical function (*Continued*)

	verted to 0 to 100 scale)	low-up: 5 (248) -Interme-diate-term (3 to 12 months) follow-up: 2 (146)		23 to -2.67) (P = 0.003)	41 to -3.92) (P = 0.0025)		idence (down-graded due to imprecision and inconsis-tency) that Pi-lates improves disability at short-term follow-up compared with minimal inter-vention, with a small effect size ... interme-diate-term fol-low-up, two trials pro-vided moder-ate quality ev-idence (down-graded due to im-precision) of a significant ef-fect in favour of Pilates, with a medium ef-fect size"

ADL: activities of daily living; CI: confidence interval; DRI: Disability Rating Index; HAQ: Health Assessment Questionnaire; MD: mean difference; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; ns: not significant; OA: osteoarthritis; SF-36: 36-item Short Form; SMD: standardised mean difference; wk: week; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index,

Table 11. Psychological function

Review	Outcome measure	Number of trials (and participants) reporting psychological function	Outcome result (postintervention or if only one measurement point)	Follow-up	Additional statement/comment
Mental health					**Mental health**
Bartels 2007	-	4 studies	SMD 0.16 (95% CI 0.01 to 0.032)	No significant difference at 6 months, 1	Very small effect size postintervention.

Table 11. Psychological function (*Continued*)

			favoured aquatic exercise	study	
Busch 2013	SF-36 - Mental health scale	1 study	-	n/a	No group differences.
Bidonde 2014	SF-36 - mental Health scale SF-12 - Mental Health scale	4 studies, n = 243	MD -3.03 (95% CI -8.06 to 2.01)	n/a	No effect.
Anxiety					**Anxiety**
Cramp 2013	Brief Symptom Inventory	1 study	"No significant effect"	n/a	-
Depression					**Depression**
Boldt 2014	CES-D	1 study, n = 34	MD -6.0 (95% CI -15.87 to 3.87) (P = 0.23)	n/a	No effect.
Busch 2013	HADS - Depression Beck Depression Index	1 study, n = 21	MD -3.70 (95% CI -6.37 to -1.03) Relative difference 57%	n/a	Significant effect, favoured resistance training.
Cramp 2013	CES-D	Not reported	"Variable effect" reported in text only	n/a	-

CES-D: Centre for Epidemiological Studies-Depression; CI: confidence interval; HADS: Hospital Anxiety and Depression Scale; MD: mean difference; n: number of participants; n/a: not applicable; SF-12: 12-item Short Form; SF-36: 36-item Short Form; SMD: standardised mean difference.

Table 12. Quality of life

Review	Outcome measure	Number of trials (and participants) reporting Quality of Life (QoL)	Outcome result	Additional statement/ comment	
(Health-related) Quality of Life					**(Health-related)**
Bartels 2007	QoL: SF-12 (Physical), PQoL, EuroQoL	Hip + knee OA (post-intervention): 3 studies, n = 599 Hip only OA (post-intervention): 1 study, n = 28	Hip + knee (post-intervention): SMD 0.32 (95% CI 0.03 to 0.61) (P = 0.028) Hip only (post-intervention): SMD 0.76 (95% CI	Significantly favoured aquatic exercise post-intervention in hip + knee OA Small effect size only (when statistically	

Table 12. Quality of life (*Continued*)

		Hip only OA (follow-up): 1 study, n = 17	-0.02 to 1.53) (ns) Hip only (follow-up): SMD 1.00 (95% CI -0.04 to 2.04) (ns)	significant).
Boldt 2014	PQoL (perceived quality of life) SQoL (subjective quality of life)	Post-intervention: 1 study, n = 34, PQoL; 1 study, n = 80, SQoL Follow-up (intermediate term): 1 study, n = 80, SQoL	Post-intervention: PQoL MD 10.8 (95% CI -4.2 to 25.8) (P = 0.16) SQoL MD 0.3 (95% CI -0.22 to 0.82) (P = 0.25) Follow-up: SQoL MD 0.5 (95% CI -0.03 to 1.03) (P = 0.07)	No difference between groups.
Fransen 2014	QoL	Post-intervention: 3 studies, n = 183	SMD 0.07 (95% CI -0.23 to 0.36) (ns)	No difference between groups.
Fransen 2015	QoL: self-report questionnaire, scale 0 to 100 (100 is maximum QoL)	Post-intervention: 13 studies, n = 1073	SMD 0.28 (95% CI 0.15 to 0.40) (P < 0.0001) Absolute difference 4% (95% CI 2% to 5%) relative difference 9% (95% CI 5% to 13%)	Statistically significant, but equates to an absolute improvement of 4 points (95% CI 2 to 5) on a 0 to 100 scale Small effect size only.
Gross 2015a	QoL: SF-36 (Physical Function subscale)	Post-intervention: 2 studies, n = 143	12-wk intervention: MD -2.22 (95% CI -5.17 to 0.72) (ns) 24-wk intervention: MD 0.06 (95% CI -4.06 to 4.17) (ns)	No significant difference between groups.
Lauret 2014	HRQoL	No relevant studies	n/a	n/a
Global assessment				**Global assessme**
Busch 2007	Global wellbeing	Strength: 2 studies, n = 47 Aerobic: 4 studies, n = 269	Strength: SMD 1.43 (95% CI 0.76 to 2.10) Aerobic: SMD 0.49 (95% CI 0.23 to 0.75)	Favoured exercise - higher score showed better QoL, Strength: very large effect size. Aerobic: small-to-moderate effect size only.
Bidonde 2014	Participant-rated global (10-cm VAS)	1 study, n = 46	MD -0.87 (95% CI -1.74 to 0.00)	No effect.
Gross 2015a	Global perceived effect	1 study, n = 70	"No significant difference"	No significant difference.
Hayden 2005	Global assessment	7 studies, n = 16	Not reported	n/a

Table 12. Quality of life (*Continued*)

Saragiotto 2016	Global impression of recovery	1 study, n = 154	Short term, MD 1.30 (95% CI 0.30 to 2.30) (P = 0.01) Intermediate term, MD 1.20 (95% CI 0.31 to 2.09) (P = 0.008) Long term, MD 1.50 (95% CI 0.61 to 2.39) (P < 0.001)	Medium effect size.
Yamato 2015	Global impression of recovery	1 study, n = 86	Short term (< 3 months): MD 1.50 (95% CI 0.70 to 2.30) Intermediate term (3 to 12 months): MD 0.70 (95% CI -0.11 to 1.51)	"Low quality evidence (downgraded due to imprecision and inconsistency), we found a significant short-term effect, with a small effect size, but not for intermediate/mid-term follow up."
Other method of assessment				**Other method o**
Bidonde 2014	Multi-dimensional function- FIQ	7 studies, n = 367	MD -5.97 (95% CI -9.06 to -2.88) SMD -0.55 (95% CI -0.83 to -0.27) Absolute difference -6 (95% CI -9 to -3) NNTB 5 (95% CI 3 to 9)	Favoured aquatic exercise - lower score showed reduced impact of pain on life "Moderate difference."
Busch 2013	Multi-dimensional function - FIQ	1 study, n = 60	SMD -1.27 (95% CI -1.83 to -0.72) Absolute difference -16.75 FIQ units (95% CI -23.31 to -10.19)	Favoured exercise - lower score showed reduced impact of pain on life Very large effect size.
Hayden 2005	Work status	9 studies, n = 21	Not reported	n/a
Silva 2010	Health Assessment Questionnaire (HAQ)	No included studies	n/a	n/a

FIQ: Fibromyalgia Impact Questionnaire; HRQoL: health-related quality of life; MD: mean difference; n: number of participants; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; OA: osteoarthritis; PQoL: perceived quality of life; QoL: quality of life; SF-36: 36-item Short Form; SMD: standardised mean difference; SQoL: subjective quality of life; VAS: visual analogue scale.

Table 13. Adherence/withdrawals

Review	Number of trials (and participants) reporting withdrawals	Number withdrawn (per 1000) - intervention group	Number withdrawn (per 1000) - control group	RR or OR
Bidonde 2014 (fibromyalgia)	8 studies, n = 472	151 (imputed from reported 38/252)	129 (imputed from reported 30/232)	RR 1.13 (95% CI 0.73 to 1.77) (P = 0.45)
Busch 2013 (fibromyalgia)	3 studies, n = 107	134 (95% CI 30 to 439)	39	RR 3.50 (95% CI 0.79 to 15.49)
Fransen 2014 (osteoarthritis)	7 studies, n = 715	59 (95% CI 30 to 114)	34	OR 1.77 (95% CI 0.86 to 3.65)
Han 2004 (rheumatoid arthritis)	4 studies, n = 189	109 (imputed from reported 11/101)	284 (imputed from reported 25/88)	RR 0.37 (95% CI 0.19 to 0.72)
Regnaux 2015 (osteoarthritis)	1 study, n = 102	44 (imputed from reported 3/68 (4%) ; all from high-intensity group)	0	Calculated RR 3.55 (95% CI 0.19 to 66.8)
Saragiotto 2016 (low back pain)	7 studies, n = 671	0	0	-
Silva 2010 (rheumatoid arthritis)	No included studies	n/a	n/a	n/a
Total	**30 studies, n = 2256**	**82.8/1000**	**81/1000**	**Calculated RR 1.02 (95% CI 0.94 to 1.12) Calculated OR 1.05 (95% CI 0.88 to 1.25)**

CI: confidence interval; n: number of participants; n/a: not applicable; OR: odds ratio; RR: risk ratio.

Table 14. Adverse events (not death)

Review	Total number of trials (and participants) in review reporting exercise vs control in chronic pain population	Number of trials (and participants) reporting adverse events	Number of adverse events	Overall statement
Bartels 2007	4 (674)	2 (148)	0	Adverse events were recorded (and reported), but none occurred

Table 14. Adverse events (not death) (*Continued*)

Bidonde 2014	9 (519)	0	0	Review stated that no included studies actively reported on adverse events (some reported withdrawal)
Boldt 2014	3 (149)	2 (115)	5 events over 2 studies	"Neck, shoulder and elbow injuries in five participants in the intervention group."
Busch 2007	34 (2276)	6 (strength training: 115, aerobic: 1264)	Strength training: 3 Aerobic training: 5	-
Busch 2013	3 (81)	2 (86 exercising participants)	0	Adverse events were recorded (and reported), but none occurred
Cramp 2013	6 (371)	3	0	Adverse events were recorded (and reported), but none occurred
Fransen 2014	10 (> 549)	5	7 events over 3 studies	-
Fransen 2015	54 (5362)	11	42 events over 8 studies	-
Gross 2015a	16 (2485)	11	41 events over 6 studies	-
Han 2004	3 (206)	2	1 event in 1 study	In narrative: "approximately one-third of the patients complained of soreness in the knee, shoulder or lower back during the first 3 weeks… pain eventually subsided for all patients… only exception was one patient, who complained of knee pain."
Hayden 2005	43 (3907)	10	23 events over 10 studies	"Negative reported: 16 events over 7 trials."
Hurkmans 2009	5 (575)	2	0	Adverse events were recorded (and reported), but none occurred

Table 14. Adverse events (not death) (*Continued*)

Koopman 2015	2 (68)	1 (10)	0	Adverse events were recorded (and reported), but none occurred "The study investigated deleterious effects of this training on motor unit survival through motor unit number estimates (MUNE). Results showed that the MUNE did not change at the end of the training."
Lane 2014	30 (1822)	1 (88 exercising participants)	2 events in control group in 1 study	RR 0.20 (95% CI 0.01 to 4.15) in favour of exercise group.
Regnaux 2015	1 (102)	1 (68 exercising participants over 2 groups: low and high resistance)	3 events in 1 study	"3 participants in high resistance group discontinued the exercise intervention due to severe knee pain."
Saragiotto 2016	7 (671)	1 (154)	5 events in 1 study	"Five patients (three from the MCE [motor control exercise] group and two from the minimal intervention group) had mild adverse effects during the study (all temporary exacerbations of pain)."
van der Heijden 2015	10 (1690)	0	0	Of the relevant studies, none actively reported on adverse events
Yamato 2015	6 (265)	1 (86)	0	Adverse events were recorded (and reported), but none occurred
Total	**246 studies (> 21,772)**	**61 studies (> 2134 participants)**	**137 events over 39 studies**	**61/246 (25%) of studies have reported on adverse events; of which 39/61 (64%) did have adverse events occur as a result of the intervention or control.**

n: number of participants; RR: risk ratio.

WHAT'S NEW

Last assessed as up-to-date: 31 January 2016.

Date	Event	Description
18 April 2017	New citation required but conclusions have not changed	Conclusions not changed; retrospective open access.
10 April 2017	Amended	See Published notes.

CONTRIBUTIONS OF AUTHORS

LG conceived the idea of the overview, and wrote the protocol and full overview.

LG, BHS, LC, and RAM developed the concept and details of the overview (participants, intervention, comparison, outcomes).

LG and CC carried out searches and selected reviews for inclusion (RAM and DM acted as arbitrators).

LG and CC carried out assessment of methodological quality using the AMSTAR tool (DM acted as arbitrator).

LG and RAM extracted data and interpreted initial findings.

LG, RAM, LC, and BHS formulated the focus of the discussion and made suggestions for future study and review authors.

All authors were involved in the interpretation of results, and in approving the final review.

LG and BHS will be responsible for future updates.

DECLARATIONS OF INTEREST

LG: none known.

RAM has received grant support from Grünenthal relating to individual patient level analyses of trial data regarding tapentadol in osteoarthritis and back pain (2015). He has received honoraria for attending boards with Menarini concerning methods of analgesic trial design (2014), with Novartis (2014) about the design of network meta-analyses, and RB on understanding pharmacokinetics of drug uptake (2015). He has received honoraria from Omega Pharma (2016) and Futura Pharma (2016) for providing advice on trial and data analysis methods.

CC: none known.

DM: none known.

LC: none known. For transparency, LC has received honoraria for speaking at educational meetings to healthcare professionals on a range of chronic pain topics (Pfizer (October 2015), Astellas (June 2014, March 2015)); editor on the *British Journal of Anaesthesia* (receives an honorarium plus a contribution toward related departmental expenses (October 2010 - to date)). LC is a medical clinician attending patients in the NHS Lothian Pain Service.

BHS: none known. BHS is a medical clinician attending patients in the NHS Tayside Pain Service.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- MRC, Lifelong Health and Wellbeing Grant, ID 91029, UK.
- National Institute of Academic Anaesthesia (BJA RCoA NIAA), UK.

NOTES

This overview review was re-published in April 2017 with retrospective open access.

Clinical Interventions in Aging

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Effect of exercise on cognitive function in chronic disease patients: a meta-analysis and systematic review of randomized controlled trials

This article was published in the following Dove Press journal:
Clinical Interventions in Aging
11 May 2017
Number of times this article has been viewed

Hong Cai*
Guichen Li*
Shanshan Hua
Yufei Liu
Li Chen

School of Nursing, Jilin University, Changchun, Jilin, People's Republic of China

*These authors contributed equally to this work

Background: The purpose of this study was to conduct a meta-analysis and systematic review to assess the effect of exercise on cognitive function in people with chronic diseases.

Methods: PubMed, Web of Science, Embase, the Cochrane Library, CINAHL, PsycINFO, and three Chinese databases were electronically searched for papers that were published until September 2016. This meta-analysis and systematic review included randomized controlled trials that evaluated the effect of exercise on cognitive function compared with control group for people with chronic diseases.

Results: Totally, 35 studies met the inclusion criteria, with 3,113 participants. The main analysis revealed a positive overall random effect of exercise intervention on cognitive function in patients with chronic diseases. The secondary analysis revealed that aerobic exercise interventions and aerobic included exercise interventions had a positive effect on cognition in patients with chronic diseases. The intervention offering low frequency had a positive effect on cognitive function in patients with chronic diseases. Finally, we found that interventions offered at both low exercise intensity and moderate exercise intensity had a positive effect on cognitive function in patients with chronic diseases. The secondary analysis also revealed that exercise interventions were beneficial in Alzheimer's disease patients when grouped by disease type.

Conclusion: This meta-analysis and systematic review suggests that exercise interventions positively influence cognitive function in patients with chronic diseases. Beneficial effect was independent of the type of disease, type of exercise, frequency, and the intensity of the exercise intervention.

Keywords: exercise, cognitive function, physical activity

Abbreviations

ACE-R, Addenbrooke's Cognitive Examination–Revised; ADAS-Cog, Alzheimer's Disease Assessment Scale–cognitive subscale; AIDS, acquired immune deficiency syndrome; BCPT-cog, Breast Cancer Prevention Trial symptom checklist–cognitive problems scale; BDI-cog, Beck Depression Inventory–cognitive subscale; BDNF, brain-derived neurotrophic factor; BOPI-cog, Breast Cancer Prevention Trial Symptom Checklist-Cognitive Problems Scale; CDT, clock drawing test; CI, confidence interval; CINAHL, Cumulative Index to Nursing and Health Literature; CNKI, China National Knowledge Infrastructure; EORTC-cog, European Organization for Research and Treatment of Cancer–cognitive; ERFC, Rapid Evaluation of Cognitive Function; FAB, Frontal Assessment Battery; FACS, functional assessment of communication skills; HR_{max}, maximum heart rate; MCI, mild cognitive impairment; IL, interleukin; MD, mean difference; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive

Correspondence: Li Chen
School of Nursing, Jilin University, No 965 Xinjiang Street, Changchun, Jilin 130021, People's Republic of China
Tel/fax +86 431 8561 9366
Email chen_care@126.com

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/CIA.S135700

Assessment; PFS-cog, Piper Fatigue Scale-cognitive; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLCT, six-letter cancellation test; SMD, standardized mean difference; TCD, total cognitive domain; TNF-α, tumor necrosis factor-alpha.

Introduction

Chronic diseases are long-term conditions with slow disease progression and without an effective cure,[1] and 38 million people die from chronic diseases each year. In addition, 16 million of these deaths occur before the age of 70 years. Chronic diseases may lead to alteration in brain structure and function and are associated with cognitive change.[2–5] Some of these changes may be related to neurodegenerative diseases (such as Alzheimer's disease and other types of dementia), increased dementia incidence, and cognitive decline. Strategies are needed to reduce disease-related cognitive impairment in chronic disease patients.

Exercise, the aim of which is to improve or maintain physical fitness, is a subset of physical activity that is planned, structured, and flexible, in addition to promoting aerobic endurance.[6] Exercise is essential in maintaining physical function and physiological health. The results of animal studies have identified that engagement in physical activity may enhance neurotrophic factor levels,[7,8] neurogenesis,[9,10] and vascularization[11] and may even reduce aggregation of pathogenic proteins,[12,13] mediate neuroinflammation,[13] and inhibit neuronal dysfunction.[14] Exercise also appears to be associated with the maintenance of brain health and cognitive performance in cognitively normal older adults. Most experimental studies have identified increased lifetime physical activity to be associated with reduced risk of suffering from dementia in cognitively normal older persons.[15–17] The results of a meta-analysis of prospective studies on physical activity and the risk of cognitive decline, which included 15 prospective studies (12 cohorts) with 33,816 nondemented subjects followed for 1–12 years, showed that a total of 3,210 patients demonstrated cognitive decline during the follow-up period. The results of the cumulative analysis indicated that subjects who performed high levels of physical activity were significantly less likely to demonstrate cognitive decline during the follow-up period.[18]

The results of a meta-analysis of the effect of physical activity on cognitive function in patients with dementia suggested that physical activity interventions positively influenced cognitive function in patients with dementia.[19] The results of another meta-analysis of aerobic exercise implied that this practice promotes cognitive function in older adults with mild cognitive impairment, finding that aerobic exercise was associated with an improvement in global cognitive ability.[20] However, a comprehensive evaluation of the effect of exercise interventions on cognitive function in chronic disease patients has not been conducted. We therefore conducted a meta-analysis and systematic review of randomized controlled trials investigating the effect of exercise intervention on cognitive outcomes in chronic disease patients.

Methods

Protocol and registration

The meta-analysis was conducted and reported in accordance with the PRISMA guidelines[21] to ensure comprehensive and transparent reporting of our methods and results.

Search strategy

PubMed, Web of Science, Embase, the Cochrane Library, CINAHL, PsycINFO, and three Chinese databases (CNKI, WanFang Data, and VIP) were electronically searched for papers that were published until September 2016. The search strategy included various combinations of the terms "Cognition", "Cognitive function", and "MMSE", with exercise intervention terms such as "Exercise" or "Muscle Stretching Exercises" or "Resistance Training" or "Running" or "Swimming" or "Walking" or "Cycling" or "Physical activity" or "Aerobic" or "Yoga" or "Tai Chi" or "Qigong". Randomized controlled trials were specifically targeted using the following search terms: "Randomized controlled trial", "Controlled clinical trial", or "Randomized" or "Randomly" or "Trial" or "Group". The search was limited to human studies.

Eligibility criteria

Types of studies

Only randomized controlled trials were included in this review. No publication date restrictions were imposed on the initial search.

Types of participants

The participants were adults (≥18 years) who had been diagnosed with a chronic disease (eg, arthritis, asthma, cancer, COPD, diabetes, heart disease, or AIDS).[1] Participants with mental problems were excluded.

Types of interventions

The inclusion criteria were as follows: 1) the intervention group underwent exercise intervention. When a study included two or more intervention groups that were found to be eligible criteria, we included all in the meta-analysis; 2) the control

group did not undergo any type of exercise intervention. However, studies in which exercise training was part of an intervention with multiple components (eg, combined with a drug intervention) were excluded.

Types of outcomes

The studies were required to report global cognitive function as the outcome measure. Any studies reporting only the results for a specific cognition scale, including scales assessing memory, attention, language, verbal fluency, visuospatial ability, or executive ability, were excluded.

Study selection

The study selection process is outlined in Figure 1. The eligibility assessment was performed by two independent reviewers in a standardized manner. All papers identified using the search strategy were assessed for eligibility, as indicated based on the previously defined inclusion criteria, by reviewing their titles and/or abstracts. If insufficient information was available to evaluate the inclusion or exclusion of an article, then a full-text version was obtained. Full-text versions of all the relevant studies were obtained and reviewed by two independent reviewers to ensure that the studies met the inclusion criteria. Disagreements were resolved by discussion with a third reviewer. When insufficient information or data were available in the included articles, the authors were contacted to obtain additional information if possible.

Quality assessment

Two reviewers independently assessed the quality of all included studies, using the Downs and Black Quality Index. The scales are designed to assess the methodological quality of randomized studies of health care interventions[22] and include reporting, external validity, bias, confounding, and power, and their maximum scores are 11, 3, 7, 6, and 5, respectively. The maximum possible total score is 32. Quality was then rated on a four-category scale: poor (<18), moderate (18–23), good (24–29), and excellent (≥30).

Data extraction and statistical analysis

Data were extracted from the included articles using a data extraction form (Table 1). Sample characteristics were collected, including the sample size, intervention and control group sizes, diagnoses, baseline MMSE scores, and age. Details on exercise interventions were collected, including intervention category, frequency, duration, HR_{max}, and exercise intensity. The effects of the exercise training interventions, including cognitive function measures and the study results, were extracted. One investigator performed the data extraction, which was checked by a second investigator.

The statistical analyses were performed using version 5.3 of the RevMan meta-analysis software. The intervention effect sizes for continuous variables were measured by determining the SMDs between the intervention and control groups with regard to the change observed between the baseline and follow-up cognitive scores and their corresponding 95% CI. According to the recommendations of the *Cochrane Handbook for Systematic Reviews of Interventions*, published by Cochrane Collaboration and Wiley, the selection of fixed- or random-effects model is based on the underlying effect of the intervention.[23] Due to the expected heterogeneity across studies (eg, different intervention types and cognitive outcome measurements), we performed random-effects meta-analysis. Additionally, when the heterogeneity identified across studies was high, a subgroup analysis was performed to identify potential causes of heterogeneity, including exercise types, types of disease, exercise frequency, and intensity. Heterogeneity

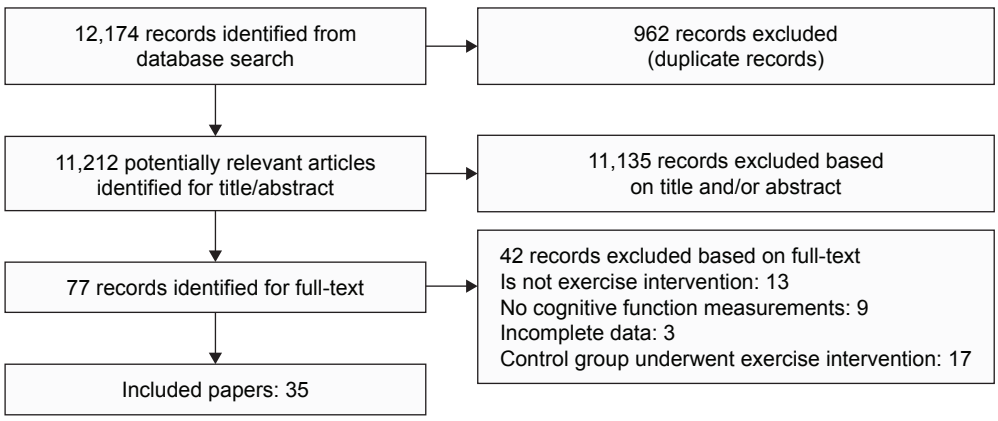


Figure 1 Flow diagram of literature search.

Table 1 Patient characteristics of included studies

First author (year)	Sample characteristics					Intervention characteristics					Outcome	
	Total N	Intervention/ control, n	Diagnosis	Baseline MMSE	Age, years	Intervention category	Frequency, min/wk	Duration, weeks	HR_{max}	Exercise intensity	Selected cognitive measurement	Conclusion
Arcoverde (2013)[24]	20	10/10	Alzheimer's disease	21.15	78.75	Aerobic	60	12	60%	Moderate	MMSE	Yes
Oh (2012)[42]	81	44/37	Cancer	Unknown	62	Aerobic	180	10	Unknown	Low	EORTC-cog	Yes
Tsai (2013)[44]	55	28/27	Osteoarthritic knee	25.45	78.91	Aerobic	60–120	20	Unknown	Low	MMSE	No
Moore (2015)[41]	40	20/20	Stroke	28.5	69	Resistance	135–180	19	70%–80%	High	ACE-R	Yes
Derry (2015)[29]	200	100/100	Cancer	Unknown	51.6	Aerobic	90	12	Unknown	Low	BCPT-CPS	Yes
Varela (2011)[51]	32	17/15	MCI	20.77	78.3	Aerobic	90	12	40%	Low	MMSE	No
Varela (2011)[51]	30	15/15	MCI	21.29	77.9	Aerobic	90	12	60%	Moderate	MMSE	No
Hildreth (2015)[33]	53	26/27	MCI	28.7	64.5	Aerobic	180	18	50%–60%	Moderate	ADAS-cog	No
Lindsay (2014)[40]	168	47/121	Heart failure	Unknown	68.98	Aerobic	180	12	Unknown	Moderate	MMSE	No
Vadiraja (2009)[49]	88	44/44	Cancer	Unknown	47.5	Aerobic	180	6	Unknown	Low	EORTC-cog	Yes
Schmidt (2015)[45]	98	49/49	Cancer	Unknown	52.7	Resistance	120	12	Unknown	High	EORTC-cog	No
Gowans (2001)[31]	31	15/16	Fibromyalgia	Unknown	47.94	Aerobic	90	23	60%–75%	Moderate	BDI-cog	Yes
Kim (2005)[35]	35	18/17	Cancer	Unknown	33.58	Aerobic	210	6	Unknown	Low	PFS-cog	Yes
Ohman (2016)[43]	140	70/70	Alzheimer's disease	17.75	77.9	Combined	120	48	Unknown	Low	CDT	No
Ohman (2016)[43]	140	70/70	Alzheimer's disease	18.1	78.2	Combined	120	48	Unknown	Moderate	CDT	No
Cancela (2016)[25]	189	73/116	Alzheimer's disease	15.03	82.02	Aerobic	105	60	Unknown	Low	MMSE	Yes
Chattha (2008)[26]	120	59/61	Climacteric syndrome	Unknown	48.5	Aerobic	300	8	Unknown	Low	SLCT	Yes
Cheng (2014)[27]	74	39/35	Alzheimer's disease	18.79	81.38	Aerobic	180	12	Unknown	Low	MMSE	Yes
Christofoletti (2008)[54]	30	17/20	Alzheimer's disease	13.7	76.7	Resistance	180	24	Unknown	Moderate	MMSE	No
Cott (2002)[28]	56	30/26	Alzheimer's disease	5.8	82.5	Aerobic	150	16	Unknown	Low	FACS	No
Hashimoto (2015)[32]	38	19/19	Parkinson's disease	28.25	68.77	Aerobic	60	12	50%–70%	Moderate	FAB	Yes
Hashimoto (2015)[32]	40	21/19	Parkinson's disease	28.41	65.86	Resistance	60	12	50%–70%	Moderate	FAB	Yes
Holthoff (2015)[34]	30	15/15	Alzheimer's disease	20.6	72.4	Combined	90	12	Unknown	Moderate	MMSE	Yes
Lam (2015)[37]	278	147/131	MCI	25.7	75.45	Aerobic	180	48	Unknown	Low	MMSE	No
Lam (2012)[57]	261	171/218	MCI	24.5	77.82	Aerobic	90	48	Unknown	Low	MMSE	No
Lautenschlager (2011)[38]	170	85/85	MCI	Unknown	68.65	Combined	150	24	Unknown	Moderate	ADAS-cog	Yes
Lu (2016)[39]	46	23/23	MCI	26.82	69.73	Resistance	180	12	Unknown	Moderate	ADAS-cog	Yes
Miu (2008)[30]	85	36/49	Alzheimer's disease	18.9	76.66	Aerobic	120	12	Unknown	Moderate	MMSE	No
Nascimento (2014)[58]	37	20/17	MCI	Unknown	67.7	Aerobic	180	16	Unknown	Moderate	MoCA	Yes
Kwak (2008)[56]	30	15/15	Alzheimer's disease	14.0	81.0	Combined	40	48	40%–60%	Moderate	MMSE	Yes
Kemoun (2010)[55]	38	20/18	Alzheimer's disease	12.8	81.8	Aerobic	180	15	Unknown	Moderate	ERFC	Yes
Kim (2011)[36]	44	32/12	Metabolic syndrome	26.18	68.19	Aerobic	120	24	50%–80%	Moderate to high	TCD	No
Stevens (2016)[46]	54	24/30	Alzheimer's disease	Unknown	80.2	Aerobic	90	12	Unknown	Moderate	CDT	Yes
Suzuki (2013)[47]	100	50/50	MCI	26.55	75.3	Combined	180	24	60%	Moderate	MMSE	No
Suzuki (2012)[48]	50	25/25	MCI	26.7	76.05	Combined	180	48	60%	Moderate	MMSE	Yes
Van (2004)[50]	25	15/10	Alzheimer's disease	Unknown	81.65	Combined	210	12	Unknown	Moderate	MMSE	Yes
Vreugdenhil (2011)[52]	40	20/20	Alzheimer's disease	22.00	74.1	Combined	210	17.3	Unknown	Moderate	MMSE	Yes
Wei (2014)[53]	60	30/30	MCI	24.67	66	Aerobic	120	24	Unknown	Moderate	MMSE	Yes

Abbreviations: ACE-R, Addenbrooke's Cognitive Examination-Revised; ADAS-cog, Alzheimer's Disease Assessment Scale-cognitive subscale; BCPT-CPS, Breast Cancer Prevention Trial Symptom Checklist-Cognitive Problems Scale; BDI-cog, Beck Depression Inventory-cognitive; CDT, clock drawing test; EORTC-cog, European Organization for Research and Treatment of Cancer-cognitive; ERFC, Rapid Evaluation of Cognitive Function; FAB, Frontal Assessment Battery; FACS, functional assessment of communication skills; HR_{max}, maximum heart rate; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; PFS-cog, Piper Fatigue Scale-cognitive; SLCT, six-letter cancellation test; TCD, total cognitive domain.

was assessed using Higgins I^2 values. The significance level was set at $P<0.05$.

Results

Study search and selection

We identified 12,174 studies based on the database searches; 962 articles were excluded because of duplicate records. Based on title and abstract, we excluded 11,135 studies. The full-text papers of 77 studies were reviewed and 42 studies were excluded. These studies were excluded because they did not assess exercise interventions, reported incomplete data, their control group underwent an exercise intervention, or they did not report cognitive function as an outcome measurement. Finally, 35 studies with 3,113 participants were included in the final analysis (Figure 1).[24–58]

Description of studies

The characteristics of the included articles are stated in Table 1. In the included studies, the sample size ranged from 20[24] to 278[37] participants. Thirteen studies examined the effect of exercise on cognitive function in patients with Alzheimer's disease.[24,25,27,28,30,34,43,46,50,52,54–56] Nine studies examined the effect of exercise on cognitive function in patients with mild cognitive impairment.[33,37–39,47,48,53,57,58] Five studies examined the effect of exercise on cognitive function in patients with cancer.[29,35,42,45,49] One study examined the effect of exercise on cognitive function in patients with Parkinson's disease.[32] One study examined the effect of exercise on cognitive function in patients with heart failure.[40] One study examined the effect of exercise on cognitive function in patients with stroke.[41] One study examined the effect of exercise on cognitive function in patients with metabolic syndrome.[36] One study examined the effect of exercise on cognitive function in patients with osteoarthritic knee.[44] And one study examined the effect of exercise on cognitive function in patients with fibromyalgia.[31] One study examined the effect of exercise on cognitive function in patients with climacteric syndrome.[26] The mean baseline MMSE ranged from 5.8[28] to 28.7,[33] excluding 11 studies that did not report baseline MMSE. The mean age ranged from 47.5 years[49] to 82.5 years.[28] The interventions were then divided into three exercise modes: aerobic, resistance, and a combination of aerobic and resistance, according to the American College of Sports Medicine.[59] The frequency of the exercise intervention varied from 40 min[56] to 300 min[26] per week. The duration of the total training period varied from 6 weeks[35,49] to 60 weeks.[25] Fitness level was divided into three modes: low, moderate, and high, according to the American College of Sports Medicine.[60]

Seventeen studies used the MMSE,[24,25,27,30,34,37,40,44,47,48,50–54,56,57] three studies used the ADAS-cog,[33,38,39] three studies used the EORTC questionnaire,[42,45,49] two studies used the CDT,[43,46] one study used the FAB,[32] one study used the ACE-R,[41] one study used the BOPI-cog,[29] one study used the BDI-cog,[31] one study used the PFS-cog,[35] one study used the FACS,[28] one study used the SLCT,[26] one study used the MoCA,[58] one study used the ERFC,[55] and one study used the TCD assessment tool.[36] These different tools were applied to evaluate the same cognitive domain within a study or between studies.

Among the included studies, three studies compared two intervention types with a control group (Table 1).[32,43,51]

Quality assessment

Thirty-five studies were included in the quality assessment. The assessment of bias in each domain across the included studies is shown in Table 2. The quality of the majority of the studies was moderate, with a mean score of 22.05. Four studies were rated as of poor quality, 18 studies were deemed to be of moderate quality, and 13 studies were deemed to be of good quality.

Main analysis: effects of exercise intervention on cognitive function

Thirty-five studies with 3,113 participants evaluated the effect of exercise on cognitive function in patients with chronic diseases.[24–58] The main analysis revealed a positive overall random effect of the exercise interventions on cognitive function in patients with chronic diseases (Table 3).

Secondary analyses

Types of disease

Thirteen studies containing 958 participants examined the effect of exercise on cognitive function in patients with Alzheimer's disease.[24,25,27,28,30,34,43,46,50,52,54–56] We found positive overall random effect of exercise intervention on cognitive function in studies evaluating Alzheimer's disease patients. Nine studies containing 1,117 participants examined the effect of exercise on cognitive function in patients with MCI.[33,37–39,47,48,51,53,57,58] In these studies, the difference observed in postintervention cognitive function did not differ between the exercise group and the control group. Five studies containing 502 participants examined the effect of exercise on cognitive function in patients with cancer.[29,35,42,45,49] In these studies, the difference observed in postintervention cognitive function did not differ between the exercise group and the control group in cancer patients.

Table 2 Quality of included studies

First author (year)	Reporting (11 points)	External validity (3 points)	Bias (7 points)	Confounding (6 points)	Power (5 points)	Total (32 points)	Quality as per cutoff described
Arcoverde (2013)[24]	10	3	6	3	5	27	Good
Oh (2012)[42]	10	3	5	3	5	26	Good
Tsai (2013)[44]	8	3	5	3	0	19	Moderate
Moore (2015)[41]	9	3	6	1	5	24	Good
Derry (2015)[29]	8	2	5	2	5	17	Poor
Varela (2011)[51]	8	3	5	3	0	19	Moderate
Hildreth (2015)[33]	8	2	5	4	0	16	Poor
Lindsay (2014)[40]	8	3	5	4	0	20	Moderate
Vadiraja (2009)[49]	10	2	5	4	0	21	Moderate
Schimidt (2015)[45]	9	3	5	4	0	21	Moderate
Gowans (2001)[31]	9	3	5	4	5	26	Good
Kim (2005)[35]	9	2	5	3	0	19	Moderate
Ohman (2016)[43]	9	2	5	5	0	21	Moderate
Cancela (2016)[25]	9	3	5	4	5	27	Good
Chattha (2008)[26]	9	3	5	3	5	25	Good
Cheng (2014)[27]	9	3	5	4	5	26	Good
Cott (2002)[28]	9	2	5	3	0	19	Moderate
Christofoletti (2008)[54]	7	2	5	2	0	16	Poor
Hashimoto (2015)[32]	9	2	5	1	5	22	Moderate
Holthoff (2015)[34]	9	2	5	2	5	23	Moderate
Lam (2015)[37]	10	2	6	4	0	22	Moderate
Lam (2012)[57]	8	3	5	3	0	19	Moderate
Lautenschlager (2011)[38]	9	2	5	3	5	24	Good
Lu (2016)[39]	9	2	5	3	5	24	Good
Miu (2008)[30]	10	2	6	4	5	27	Good
Nascimento (2014)[58]	8	2	5	2	0	17	Poor
Kwak (2008)[56]	9	1	5	4	5	24	Moderate
Kemoun (2010)[55]	7	3	5	2	5	22	Moderate
Kim (2011)[36]	9	3	6	3	5	26	Good
Stevens (2016)[46]	8	2	5	3	0	18	Moderate
Suzuki (2013)[47]	9	2	5	4	0	20	Moderate
Suzuki (2012)[48]	9	2	5	4	5	25	Good
Van (2004)[50]	8	2	5	3	5	23	Moderate
Vreugdenhil (2011)[52]	9	3	5	4	5	26	Good
Wei (2014)[53]	7	2	5	2	5	21	Moderate

Three studies assessed the effect of exercise intervention on cognitive function in patients with osteoarthritic knee,[44] heart failure,[40] and metabolic syndrome.[36] In these studies, the difference observed in postintervention cognitive function did not differ between the exercise and control groups. Four studies evaluated the effect of exercise on cognitive function in patients with fibromyalgia,[31] stroke,[41] Parkinson's disease,[32] and climacteric syndrome,[26] and the results of these studies indicated that exercise had a positive effect on cognitive function.

Types of exercise intervention

Twenty-three studies containing 2,120 participants examined the effect of aerobic exercise on cognitive function in patients with chronic disease.[24–33,35–37,40,42,44,46,49,51,53,55,57,58] Five studies containing 261 participants examined the effect of resistance exercise on cognitive function in patients with chronic disease.[32,39,41,45,54] Eight studies containing 725 participants examined the effect of combined exercise on cognitive function in patients with chronic disease.[34,38,43,47,48,50,52,56] We identified an overall positive random effect for aerobic exercise interventions but not for resistance exercise interventions and combined exercise interventions. Twenty-seven studies containing 2,845 participants examined the effect of interventions including aerobic exercise (both combined exercise and aerobic exercise intervention) on cognitive function in patient with chronic disease.[24–38,40,42–44,46–53,55–58] We identified positive overall random effect for the included aerobic exercise interventions.

Frequency of exercise intervention

According to the World Health Organization recommendations, a weekly schedule of 150 min exercise was used to distinguish between high- and low-frequency interventions.[61]

Table 3 Meta-analysis of effect of exercise on cognitive function

Parameters	Included studies	N	*P*-value	*I*	*Z*	SMD (95% CI)
Main outcome	35	3,113	0.0007	74%	3.41	0.26 (0.11, 0.41)
Types of disease						
Alzheimer's disease	13	958	0.004	77%	2.88	0.42 (0.14, 0.71)
Mild cognitive impairment	9	1,117	0.21	77%	1.25	0.17 (–0.10, 0.44)
Cancer	5	502	0.66	86%	0.44	0.11 (–0.39, 0.61)
Types of exercise intervention						
Aerobic	23	2,120	0.0008	70%	3.34	0.29 (0.12, 0.47)
Resistance	5	261	0.69	73%	0.40	0.10 (–0.39, 0.59)
Combined	8	725	0.14	83%	1.46	0.29 (–0.10, 0.67)
Aerobic included	31	2,845	0.0005	75%	3.47	0.28 (0.12, 0.45)
Frequency of exercise intervention						
High frequency	18	1,494	0.07	83%	1.82	0.25 (–0.02, 0.52)
High frequency (excluded resistance)	16	1,450	0.02	82%	2.24	0.31 (0.04, 0.57)
Low frequency	16	1,547	0.0008	40%	2.67	0.19 (0.05, 0.33)
Low frequency (excluded resistance)	14	1,409	0.15	60%	1.43	0.13 (–0.05, 0.32)
Intensity of exercise intervention						
Low	13	1,609	0.03	77%	2.12	0.24 (0.02, 0.47)
Moderate	21	1,322	0.03	72%	2.13	0.24 (0.02, 0.46)
Moderate (excluded resistance)	19	1,199	0.008	76%	2.66	0.34 (0.09, 0.59)
High	2	138	0.30	0%	1.04	0.18 (–0.16, 0.51)

Abbreviations: CI, confidence interval; SMD, standardized mean difference.

Eighteen studies containing 1,494 participants examined the effect of high-frequency exercise intervention on cognitive function in patients with chronic disease.[26–28,33,35,37–40,42,47–50,52,54,55,58] Sixteen studies containing 1,547 participants examined the effect of low-frequency exercise intervention on cognitive function in patients with chronic disease.[24,25,29–32,34,36,43–46,51,53,56,57] We identified a positive overall random effect for low-frequency interventions but not for high-frequency interventions. Further investigations revealed that after the exclusion of resistance exercise interventions, the effect of low-frequency exercise interventions was not significant. Further investigations revealed that after the exclusion of resistance interventions, the effect of the high-frequency exercise interventions was significant. Moore et al[41] examined the effect of an intervention including 135–180 min of weekly exercise on cognitive function in patients with stroke and found that the exercise intervention did not have an effect on cognitive function.

Intensity of exercise intervention

Thirteen studies containing 1,609 participants examined the effect of low-intensity exercise intervention on cognitive function in patients with chronic diseases.[25–29,35,37,42–44,49,51,57] Twenty-one studies containing 1,322 participants examined the effect of moderate-intensity exercise intervention on cognitive function in patients with chronic diseases.[24,30–33,38–40,43,46–48,50–56,58] Two studies containing 138 participants examined the effect of high-intensity exercise intervention on cognitive function in patients with chronic diseases.[41,45] We found positive random effects for low-intensity and moderate-intensity exercise interventions but not for high-intensity exercise interventions. Further investigations revealed that after the exclusion of resistance interventions, the effect of moderate-intensity exercise intervention on the cognitive function was significant.

Discussion

In this meta-analysis of randomized controlled trials, we identified a positive overall effect of exercise interventions on cognitive function in patients with chronic diseases. Aerobic exercise interventions were found to have a positive effect on cognitive function in patients with chronic disease. In addition, the effect of exercise on cognitive function was independent of the presence of Alzheimer's disease. Furthermore, we found that low-frequency exercise interventions had a positive effect on cognitive function in chronic disease patients. Finally, we observed positive effects of low-intensity and moderate-intensity exercise intervention on cognitive function in chronic disease patients.

In this meta-analysis and systematic review, we found that the exercise interventions were beneficial in the current sample of chronic disease patients. Exercise has been reported to cause physiological state changes that disrupt brain homeostasis.[41] The brain has been found to modify its resource allocation in response to these changes. Studies have suggested that maintenance of physical activity may

be associated with increased neural resources in some brain regions and reduced neural resources in other brain regions.[47,58,62] Exercise affects cognitive function by causing a significant reduction in the peripheral concentrations of IL-6 and TNF-α, as well as a significant increase in peripheral levels of BDNF in individuals with chronic diseases.[58] Exercise also leads to structural changes in the brain, such as increases in dendritic length and branching and hippocampal neurogenesis,[62] as well as maintains the atrophy levels of the whole brain cortex.[47]

Cognition is a complex term that includes various domains. Some studies have proposed relationships between specific exercise regimens and specific cognitive domains in chronic disease patients. Aerobic exercise has been reported to contribute to further beneficial effects on the memory domain.[32,36,48] The results of an animal study investigating the effects of 12 weeks of voluntary running on the restoration of place recognition memory in 20-month-old rats emphasized the unique synaptic effects of exercise on the aged brain and their specific relevance to the hippocampal-based system for place recognition memory.[63] Dancing involves paying attention to music and signals while envisaging the next movement, and these feature may help patients to perform better in the verbal fluency category.[50] Lu et al[39] found that variations in position changes and movement configurations during dumbbell-training sessions were associated with changes in the spatiotemporal orientation, selective attention, and executive control of participants.

The results of our study showed that aerobic exercise interventions had a positive effect on cognitive function. This result was consistent with the recommendations of the World Health Organization for a weekly minimum of 150 min of moderate-intensity aerobic or 75 min of vigorous-intensity aerobic activity with additional muscle-strengthening exercises.[61] Two meta-analyses of the effect of aerobic exercise on cognitive function found that aerobic exercise improved cognitive function.[19,20] These results were similar to the results of our study. Aerobic exercises improve the maximum oxygen uptake and increase and redistribute cerebral blood flow, enhance antioxidant action via repair enzymes and proinflammatory cytokines, as well as increase beta-amyloid degradation, levels of neurotrophic factors, neurogenesis, and angiogenesis.[24,58] In this meta-analysis, we did not find resistance exercises to have an effect on cognitive function in chronic disease patients, which may be due to the difficulties related to controlling for some methodological and sampling biases and the short follow-up periods.[54]

In this meta-analysis and systematic review, we found that exercise interventions were beneficial for cognitive function in Alzheimer's disease patients. The results of this study were similar to those of previous studies that reported that exercise has a positive effect on cognitive function in Alzheimer disease patients.[19,64,65] These trends may indicate that the practice of regular physical exercise might contribute to slower declines in cognitive function. In our study, the exercise interventions were not found to have a positive effect on cognitive function in patients with MCI. There may be insufficient evidence for an effect of exercise intervention on MCI patients. A meta-analysis and systematic review of the effect of aerobic exercise on cognitive function in older adults with MCI showed that aerobic exercise significantly improved global cognitive ability (MMSE scores: MD =0.98, 95% CI: 0.5–1.45; $P<0.0001$).[20] The cited meta-analysis and systematic review evaluated interventions encompassing the practice of any aerobic exercises regardless of the style (eg, yoga, Tai Chi, or treadmill) for at least 4 weeks, with >1 exercise session per week. In addition, the outcomes assessed included global cognitive ability and any specific domains of cognition assessed in the aforementioned meta-analysis. Our meta-analysis and systematic review was not restricted by type of exercise, and the outcome of interest was global cognitive function. In our study, the exercise intervention was not found to have a positive effect on cognitive function in patients with cancer. The studies that assessed the effects of exercise interventions on cognitive function in cancer patients all provided data for short-term interventions that did not exceed 12 weeks.

In this meta-analysis and systematic review, we found that low-frequency exercise intervention was beneficial in chronic disease patients. The studies providing low-frequency exercise intervention all exceeded 12 weeks' duration, and two studies even exceeded 48 weeks' duration.[25,43] The beneficial effect of low-frequency exercise on cognitive function may be associated with a good performance of the functional capacity.[24] Additionally, all the included studies of low-frequency exercise interventions were of good methodological quality.

The American College of Sports Medicine has suggested that moderate-intensity physical exercise may lead to significant changes in brain health and cognitive performance, with potential effects on a broad range of cognitive domains.[66] In this meta-analysis and systematic review, we found that both low-intensity and moderate-intensity exercise interventions appeared to be effective in improving cognitive function in chronic disease patients. Moderate-intensity exercise might

be an effective alternative to reduce the level of systemic inflammation and decrease cognitive decline.[58] Additionally, most of the studies evaluating moderate-intensity exercise were of good methodological quality. The included studies indicated that low-intensity exercise such as Tai Chi[27,42,44] or yoga[29] exerted positive effects on cognitive function. Patients involved in low-exercise intensity interventions demonstrated more notable changes in physical functioning, contributing to the positive effects observed in psychological well-being.[42] Lower levels of distress and fatigue may have contributed to the beneficial effects of low-intensity exercise interventions on cognitive function.[29]

Limitations

This study had some limitations. First, the weekly duration of exercise intervention used in the included studies varied from 40 min to 300 min per week, and the overall duration of the exercise intervention used in the included studies varied from 6 weeks to 60 weeks. These differences may have affected the relationship observed between the specific types of exercise and improvements in the evaluated cognitive domains. Second, it was not possible to blind participants to the exercise intervention. Therefore, performance bias may have been unavoidable.

Conclusion

The findings of this meta-analysis support the efficacy of exercise interventions in improving cognitive function in individuals affected by chronic disease. Beneficial effects were observed independent of the type of clinical disease, type of exercise, frequency, and intensity of the exercise intervention.

Acknowledgment

This work was supported by the Youth Program of Health and Family Planning Commission of Jilin Province (grant number 2015Q022).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Wang C, Collet JP, Lau J. The effect of Tai Chi on health outcomes in patients with chronic conditions: a systematic review. *Arch Intern Med*. 2004;164(5):493–501.
2. Graff-Radford NR. Can aerobic exercise protect against dementia? *Alzheimers Res Ther*. 2011;3(1):1–6.
3. Kirk-Sanchez NJ, McGough EL. Physical exercise and cognitive performance in the elderly: current perspectives. *Clin Interv Aging*. 2014;9:51–62.
4. Nelson ME, Rejeski WJ, Blair SN, et al; American College of Sports Medicine; American Heart Association. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Circulation*. 2007;116(9):1094–1105.
5. Hess LM, Huang HQ, Hanlon AL, et al. Cognitive function during and six months following chemotherapy for front-line treatment of ovarian, primary peritoneal or fallopian tube cancer: an NRG oncology/gynecologic oncology group study. *Gynecol Oncol*. 2015;139(3):541–545.
6. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep*. 1985;100(2):126–131.
7. Radak Z, Hart N, Sarga L, et al. Exercise plays a preventive role against Alzheimer's disease. *J Alzheimers Dis*. 2010;20(3):777–783.
8. Adlard PA, Perreau VM, Pop V, Cotman CW. Voluntary exercise decreases amyloid load in a transgenic model of Alzheimer's disease. *J Neurosci*. 2005;25(17):4217–4221.
9. Sahay A, Scobie KN, Hill AS, et al. Increasing adult hippocampal neurogenesis is sufficient to improve pattern separation. *Nature*. 2011;472(7344):U466–U539.
10. Kronenberg G, Bick-Sander A, Bunk E, Wolf C, Ehninger D, Kempermann G. Physical exercise prevents age-related decline in precursor cell activity in the mouse dentate gyrus. *Neurobiol Aging*. 2006;27(10):1505–1513.
11. Ding YH, Luan XD, Li J, et al. Exercise-induced overexpression of angiogenic factors and reduction of ischemia/reperfusion injury in stroke. *Curr Neurovasc Res*. 2004;1(5):411–420.
12. Leem YH, Lim HJ, Shim SB, Cho JY, Kim BS, Han PL. Repression of tau hyperphosphorylation by chronic endurance exercise in aged transgenic mouse model of tauopathies. *J Neurosci Res*. 2009;87(11):2561–2570.
13. Belarbi K, Burnouf S, Fernandez-Gomez FJ, et al. Beneficial effects of exercise in a transgenic mouse model of Alzheimer's disease-like tau pathology. *Neurobiol Dis*. 2011;43(2):486–494.
14. Um HS, Kang EB, Koo JH, et al. Treadmill exercise represses neuronal cell death in an aged transgenic mouse model of Alzheimer's disease. *Neurosci Res*. 2011;69(2):161–173.
15. Buchman AS, Boyle PA, Yu L, Shah RC, Wilson RS, Bennett DA. Total daily physical activity and the risk of AD and cognitive decline in older adults. *Neurology*. 2012;78(17):1323–1329.
16. Hamer M, Chida Y. Physical activity and risk of neurodegenerative disease: a systematic review of prospective evidence. *Psychol Med*. 2009;39(1):3–11.
17. Scarmeas N, Luchsinger JA, Schupf N, et al. Physical activity, diet, and risk of Alzheimer disease. *JAMA*. 2009;302(6):627–637.
18. Sofi F, Valecchi D, Bacci D, et al. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. *J Intern Med*. 2011;269(1):107–117.
19. Groot C, Hooghiemstra AM, Raijmakers PG, et al. The effect of physical activity on cognitive function in patients with dementia: a meta-analysis of randomized control trials. *Ageing Res Rev*. 2016;25:13–23.
20. Zheng G, Xia R, Zhou W, Tao J, Chen L. Aerobic exercise ameliorates cognitive function in older adults with mild cognitive impairment: a systematic review and meta-analysis of randomised controlled trials. *Br J Sports Med*. 2016;50:1443–1450.
21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62(10):1006–1012.
22. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of healthcare interventions. *J Epidemiol Community Health*. 1998;52(6):7.
23. Higgins JPT. *Cochrane Handbook for Systematic Reviews of Interventions Version*. London: The Cochrane Collaboration; 2011.
24. Arcoverde C, Deslandes A, Moraes H, et al. Treadmill training as an augmentation treatment for Alzheimer's disease: a pilot randomized controlled study. *Arq Neuropsiquiatr*. 2014;72(3):190–196.

25. Cancela JM, Ayan C, Varela S, Seijo M. Effects of a long-term aerobic exercise intervention on institutionalized patients with dementia. *Med Sci Sports Sci*. 2016;19(4):293–298.
26. Chattha R, Nagarathna R, Padmalatha V, Nagendra HR. Effect of yoga on cognitive functions in climacteric syndrome: a randomised control study. *BJOG*. 2008;115(8):991–1000.
27. Cheng ST, Chow PK, Song YQ, et al. Mental and physical activities delay cognitive decline in older persons with dementia. *Am J Geriatr Psychiatry*. 2014;22(1):63–74.
28. Cott CA, Dawson P, Sidani T, Wells T. The effects of a walking/talking program on communication, ambulation, and functional status in residents with Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2002;16(2):81–87.
29. Derry HM, Jaremka LM, Bennett JM, et al. Yoga and self-reported cognitive problems in breast cancer survivors: a randomized controlled trial. *Psychooncology*. 2015;24(8):958–966.
30. DKY Miu SS, Mak YF. A randomised controlled trial on the effect of exercise on physical, cognitive and affective function in dementia subjects. *Asian J Gerontol Geriatr*. 2008;3:8–16.
31. Gowans SE, deHueck A, Voss S, Silaj A, Abbey SE, Reynolds WJ. Effect of a randomized, controlled trial of exercise on mood and physical function in individuals with fibromyalgia. *Arthritis Rheum*. 2001; 45(6):519–529.
32. Hashimoto H, Takabatake S, Miyaguchi H, Nakanishi H, Naitou Y. Effects of dance on motor functions, cognitive functions, and mental symptoms of Parkinson's disease: a quasi-randomized pilot trial. *Complement Ther Med*. 2015;23(2):210–219.
33. Hildreth KL, Van Pelt RE, Moreau KL, et al. Effects of pioglitazone or exercise in older adults with mild cognitive impairment and insulin resistance: a pilot study. *Dement Geriatr Cogn Dis Extra*. 2015;5(1): 51–63.
34. Holthoff VA, Marschner K, Scharf M, et al. Effects of physical activity training in patients with Alzheimer's dementia: results of a pilot RCT study. *PLoS One*. 2015;10(4):e0121478.
35. Kim SD, Kim HS. Effects of a relaxation breathing exercise on fatigue in haemopoietic stem cell transplantation patients. *J Clin Nurs*. 2005;14(1):51–55.
36. Kim SH, Kim M, Ahn YB, et al. Effect of dance exercise on cognitive function in elderly patients with metabolic syndrome: a pilot study. *J Sports Sci Med*. 2011;10(4):671–678.
37. Lam LCW, Chan WC, Leung T, Fung AWT, Leung EMF. Would older adults with mild cognitive impairment adhere to and benefit from a structured lifestyle activity intervention to enhance cognition? A cluster randomized controlled trial. *PLoS One*. 2015;10(3):e0118173.
38. Lautenschlager NT, Cox KL, Flicker L, et al. Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: a randomized trial. *JAMA*. 2009;301(3):1027–1037.
39. Lu JJ, Sun MY, Liang LC, Feng Y, Pan XY, Liu Y. Effects of momentum-based dumbbell training on cognitive function in older adults with mild cognitive impairment: a pilot randomized controlled trial. *Clin Interv Aging*. 2016;11:9–16.
40. Miller LA. *The longitudinal effects of cardiac rehabilitation on cognition in older adults with heart failure* [dissertation]. Kent: Kent State University; 2014.
41. Moore SA, Hallsworth K, Jakovljevic DG, et al. Effects of community exercise therapy on metabolic, brain, physical, and cognitive function following stroke: a randomized controlled pilot trial. *Neurorehabil Neural Repair*. 2015;29(7):623–635.
42. Oh B, Butow PN, Mullan BA, et al. Effect of medical Qigong on cognitive function, quality of life, and a biomarker of inflammation in cancer patients: a randomized controlled trial. *Support Care Cancer*. 2012;20(6):1235–1242.
43. Ohman H, Savikko N, Strandberg TE, et al. Effects of exercise on cognition: the Finnish Alzheimer disease exercise trial: a randomized, controlled trial. *J Am Geriatr Soc*. 2016;64(4):731–738.
44. Tsai PF, Chang JY, Beck C, Kuo YF, Keefe FJ. A pilot cluster-randomized trial of a 20-Week Tai Chi Program in elders with cognitive impairment and osteoarthritic knee: effects on pain and other health outcomes. *J Pain Symptom Manage*. 2013;45(4):660–669.
45. Schmidt ME, Wiskemann J, Armbrust P, Schneeweiss A, Ulrich CM, Steindorf K. Effects of resistance exercise on fatigue and quality of life in breast cancer patients undergoing adjuvant chemotherapy: a randomized controlled trial. *Int J Cancer*. 2015;137(2):471–480.
46. Stevens J, Killeen M. A randomised controlled trial testing the impact of exercise on cognitive symptoms and disability of residents with dementia. *Contemp Nurse*. 2006;21(1):32–40.
47. Suzuki T, Shimada H, Makizako H, et al. A randomized controlled trial of multicomponent exercise in older adults with mild cognitive impairment. *PLoS One*. 2013;8(4):e61483.
48. Suzuki T, Shimada H, Makizako H, et al. Effects of multicomponent exercise on cognitive function in older adults with amnestic mild cognitive impairment: a randomized controlled trial. *BMC Neurol*. 2012;12:128.
49. Vadiraja HS, Rao MR, Nagarathna R, et al. Effects of yoga program on quality of life and affect in early breast cancer patients undergoing adjuvant radiotherapy: a randomized controlled trial. *Complement Ther Med*. 2009;17(5–6):274–280.
50. Van de Winckel A, Feys H, De Weerdt W, Dom R. Cognitive and behavioural effects of music-based exercises in patients with dementia. *Clin Rehabil*. 2004;18(3):253–260.
51. Varela S, Ayan C, Cancela JM, Martin V. Effects of two different intensities of aerobic exercise on elderly people with mild cognitive impairment: a randomized pilot study. *Clin Rehabil*. 2012;26(5):442–450.
52. Vreugdenhil A, Cannell J, Davies A, Razay G. A community-based exercise programme to improve functional ability in people with Alzheimer's disease: a randomized controlled trial. *Scand J Caring Sci*. 2012;26(1):12–19.
53. Wei XH, Ji LL. Effect of handball training on cognitive ability in elderly with mild cognitive impairment. *Neurosci Lett*. 2014;566:98–101.
54. Christofoletti G, Oliani MM, Gobbi S, Stella F, Bucken Gobbi LT, Renato Canineu P. A controlled clinical trial on the effects of motor intervention on balance and cognition in institutionalized elderly patients with dementia. *Clin Rehabil*. 2008;22(7):618–626.
55. Kemoun G, Thibaud M, Roumagne N, et al. Effects of a physical training programme on cognitive function and walking efficiency in elderly persons with dementia. *Dement Geriatr Cogn Disord*. 2010;29(2): 109–114.
56. Kwak YS, Um SY, Son TG, Kim DJ. Effect of regular exercise on senile dementia patients. *Int J Sports Med*. 2008;29(6):471–474.
57. Lam LCW, Chau RCM, Wong BML, et al. A 1-year randomized controlled trial comparing mind body exercise (Tai Chi) with stretching and toning exercise on cognitive function in older Chinese adults at risk of cognitive decline. *J Am Med Dir Assoc*. 2012;13(6): 568.e15–568.e20.
58. Nascimento CMC, Pereira JR, de Andrade LP, et al. Physical exercise in MCI elderly promotes reduction of pro-inflammatory cytokines and improvements on cognition and BDNF peripheral levels. *Curr Alzheimer Res*. 2014;11(8):799–805.
59. American College of Sports Medicine, Chodzko-Zajko WJ, Proctor DN, et al. American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc*. 2009; 41(7):1510–1530.
60. Thompson PD, Arena R, Riebe D, Pescatello LS. ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription, ninth Edition. *Curr Sports Med Rep*. 2013;12(4):215–217.
61. Organization WH. *Global Recommendations on Physical Activity for Health*. Geneva: World Health Organization; 2015.
62. Swain RA, Harris AB, Wiener EC, et al. Prolonged exercise induces angiogenesis and increases cerebral blood volume in primary motor cortex of the rat. *Neuroscience*. 2003;117(4):1037–1046.
63. Siette J, Westbrook RF, Cotman C, et al. Age-specific effects of voluntary exercise on memory and the older brain. *Biol Psychiatry*. 2013; 73(5):435–442.
64. Farina N, Rusted J, Tabet N. The effect of exercise interventions on cognitive outcome in Alzheimer's disease: a systematic review. *Int Psychogeriatr*. 2014;26(1):9–18.

65. Heyn P, Abreu BC, Ottenbacher KJ. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Arch Phys Med Rehabil*. 2004;85(10):1694–1704.

66. American College of Sports Medicine Position Stand. Exercise and physical activity for older adults. *Phys Sportsmed*. 1999;27(11):115.

frontiers in Neuroscience

REVIEW
published: 23 August 2016
doi: 10.3389/fnins.2016.00372

Emerging Relationships between Exercise, Sensory Nerves, and Neuropathic Pain

Michael A. Cooper[1], *Patricia M. Kluding*[2] *and Douglas E. Wright*[1]*

[1] Department of Anatomy and Cell Biology, University of Kansas Medical Center, Kansas City, KS, USA, [2] Department of Physical Therapy and Rehabilitation Science, University of Kansas Medical Center, Kansas City, KS, USA

The utilization of physical activity as a therapeutic tool is rapidly growing in the medical community and the role exercise may offer in the alleviation of painful disease states is an emerging research area. The development of neuropathic pain is a complex mechanism, which clinicians and researchers are continually working to better understand. The limited therapies available for alleviation of these pain states are still focused on pain abatement and as opposed to treating underlying mechanisms. The continued research into exercise and pain may address these underlying mechanisms, but the mechanisms which exercise acts through are still poorly understood. The objective of this review is to provide an overview of how the peripheral nervous system responds to exercise, the relationship of inflammation and exercise, and experimental and clinical use of exercise to treat pain. Although pain is associated with many conditions, this review highlights pain associated with diabetes as well as experimental studies on nerve damages-associated pain. Because of the global effects of exercise across multiple organ systems, exercise intervention can address multiple problems across the entire nervous system through a single intervention. This is a double-edged sword however, as the global interactions of exercise also require in depth investigations to include and identify the many changes that can occur after physical activity. A continued investment into research is necessary to advance the adoption of physical activity as a beneficial remedy for neuropathic pain. The following highlights our current understanding of how exercise alters pain, the varied pain models used to explore exercise intervention, and the molecular pathways leading to the physiological and pathological changes following exercise intervention.

Keywords: exercise, pain management, neuropathy, inflammation, neurotrophins, dorsal root ganglion

OPEN ACCESS

Edited by:
Soroku Yagihashi,
Hirosaki University, Japan

Reviewed by:
Isaura Tavares,
Faculty of Medicine of Porto, Portugal
Nigel Calcutt,
University of California, San Diego, USA
Mark Yorek,
University of Iowa, USA

***Correspondence:**
Douglas Wright
dwright@kumc.ed

Specialty section:
This article was submitted to Diabetes, a section of the journal Frontiers in Neuroscience

Received: *20 June 2016*
Accepted: *02 August 2016*
Published: *23 August 2016*

Citation:
Cooper MA, Kluding PM and Wright DE (2016) Emerging Relationships between Exercise, Sensory Nerves, and Neuropathic Pain. Front. Neurosci. 10:372. doi: 10.3389/fnins.2016.00372

INTRODUCTION

Twenty five million Americans are encumbered by acute pain and over 50 million suffer from varying chronic pain syndromes, leading to a medical cost of over $635 billion a year (Gaskin and Richard, 2012). This enormous health plight highlights the need to find novel interventions to reduce the burden of chronic pain. Generally speaking, chronic pain undergoes a progressive movement from peripheral tissues, such as the hands and feet, to the central nervous system which often leads to even more debilitating and chronic effects as the disease progresses (Tesfaye et al., 2013; Jones et al., 2016). The perception of pain is a very broad and complex mechanism to study, having multiple origins including nerve damage, metabolic disease, and numerous others. Each

form of pain may be unique not only in its development but also in the treatments necessary to provide relief. Unfortunately, current therapies available for the treatment of these pain states are still associated with pain abatement and do not address underlying mechanisms driving the development of varying forms and levels of sensory discomfort (Schreiber et al., 2015).

Physical activity offers a wide array of benefits and is well documented to help in a myriad of diseases, however the mechanisms by which exercise exerts its benefits are poorly understood. The complexities of understanding how global cross organ communication and changes induce molecular changes to provide benefits in disease makes exercise research often hard to perform on a basic level. However, the clear benefits of exercise provide a strong rational to continue to study this complex intervention.

Nociceptive and neuropathic pain syndromes both receive physiological and behavioral benefits from exercise intervention, even though they are thought to have separate physiological characteristics. Nociceptive pain results from an expected noxious stimulus, while neuropathic pain occurs in the absence of a stimulus, or with a normally innocuous stimulus. The neuronal pathway of nociceptive pain starts with a noxious stimulus detected by a peripheral sensory peripheral terminal, of an Aδ- or C-fiber. The electrical signal is then propagated up through spinal and thalamic pathways to terminate in an appropriate somatotopic region of the cortex (Serpell, 2006). In the case of neuropathic pain, adaptations occur in Schwann cells, satellite cells, the peripheral immune system, spinal microglia, and astrocytes that lead to the development of a painful syndrome when one would not normally exist (Scholz and Woolf, 2007). Important areas to examine in these pain pathways are interneuronal interactions and the molecular and cellular changes that are initiated within them. This is an important aspect of any therapeutic target for pain due to the activity-dependent neuronal plasticity that occurs in the nervous system (Zhuo et al., 2011).

In response to new information about neuronal activity-dependent plasticity, a new and rapidly growing area within both pain research and neural physiology has begun to examine the effects of exercise on peripheral and central nervous system components. However, the scarcity of well-controlled basic research in this area hampers the utilization of exercise as a therapy for neuropathic and other chronic pain syndromes. While exercise intervention is growing quickly as a clinical therapeutic tool for many diseases, its use to reduce pain states is still relatively new and the research available leaves an incomplete picture of the molecular pathways affected. Continued research therefore is vital to gain a better understanding of how exercise benefits the management of various pain syndromes and for the implementation of this therapeutic technique on a broader scale by physicians.

Abbreviations: EPR, exercise pressor reflex; HR, heart rate; TRPV1, transient receptor potential vanilloid 1; ASIC3, acid sensing ion channel 3; DRG, dorsal root ganglion; BDNF, brain-derived neurotrophic factor; NGF, nerve growth factor; NT-3, neurotrophin 3; SNAP1, synapsin I; GAP43, growth associated protein 43; TrkA, tropomyosin receptor kinase A; NGF, nerve growth factor; NMDA, N-methyl-D-aspartate; CNS, central nervous system; CGRP, calcitonin gene-related peptide; TNF- α, tumor necrosis factor alpha; IL1-β, interleukin 1 beta; IL-10, interleukin 10; IL-6, interleukin 6; IL1- α, interleukin 1 alpha; IL-2, interleukin 2; Il-4, interleukin 4; IFN-γ, interferon gamma; TGF-β, transforming growth factor beta; IL-1RA, IL-1 receptor agonist; HSP72, heat shock protein-72; DPN, diabetic peripheral neuropathy; TRPM8, transient receptor potential cation channel subfamily M member 8; NT-3, neurotrophin 3; IENFD, intraepidermal nerve fiber density; VAS, visual analog scale; QOL, quality of life; EIH, exercise-induced hypoalgesia; TSP, temporal summation of pain; ALA, alpha-lipoic acid treatment.

SENSORY PATHWAYS SENSITIVE TO EXERCISE

A well-established effect of exercise is its activation of afferent sensory nerves from active muscles to the spinal cord. Activity in sensory fibers of working muscles is increased throughout exercise and provides important feedback on the cardiovascular and respiratory systems during physical activity (Mitchell, 1985). One example of afferent nerve activity affected by exercise is the exercise pressor reflex (EPR), which is responsible for the control of blood pressure and heart rate (HR) changes during physical activity through sympathetic nerve activation (O'Leary et al., 1999; Amann et al., 2011). This reflex is partially mediated by the transient receptor potential vanilloid 1 (TRPV1) receptor, the sensory receptor respondent to capsaicin that is stimulated from temperature and pH level changes (Smith et al., 2010). Similar to TRPV1, the acid sensing ion channel 3 (ASIC3) found on sensory nerve terminals in active skeletal muscle is involved in the regulation of arterial pressure through the EPR (Tsuchimochi et al., 2011). However, EPR is additionally modulated by sodium channel (Na_v) function; these channels in turn, may be modulated through reactive oxygen species levels in the dorsal root ganglion (DRG) (Wang et al., 2011). The EPR sensory pathway crosses over with known painful sensory pathways through the TRPV1, ASIC3 and sodium channel alterations; this cross talk can provide possible avenues by which exercises known benefits may also connect to painful sensory states.

Studies now demonstrate molecular and cellular changes in DRG sensory neurons can be induced by exercise. The benefits of exercise can appear quickly as seen in the improvement in regeneration after nerve injury following as little as 3 or 7 days of exercise (Molteni et al., 2004). This effect may be related to increased production of a number of molecular signals, including brain-derived neurotrophic factor (BDNF), nerve growth factor (NGF), neurotrophin 3 (NT-3), synapsin I (SNAP1), and growth associated protein 43 (GAP43) in sensory ganglia, thereby stimulating axonal growth (Molteni et al., 2004; López-Álvarez et al., 2015). Consistent with this idea, exercised animals display clear alterations in molecular mediators in their DRGs, in large DRG neurons undergo changes in mRNA expression that are associated with neuronal plasticity and apoptosis in response to prolonged exercise, including higher BDNF, NT3, SNAP1, and GAP43 mRNA levels compared to sedentary animals (Keeler et al., 2012). Exercise of animals that have been given a high fat diet suggests that exercise can reverse alterations in neurotrophin changes that are associated with a high fat diet, insulin resistance

and pain (Groover et al., 2013). For example, exercise of diabetic mice induces significant increases in GDNF in the spinal cord and sciatic nerve, along with axonal transport in the sciatic nerve (Wright, unpublished observations).

Our own studies suggest that important phenotypic changes can occur in peripheral terminals of epidermal axons in response to exercise (Groover et al., 2013). A high fat diet increases the number of epidermal axons that express tropomyosin receptor kinase A (TrkA), the high affinity receptor for nerve growth factor (NGF). This phenotypic change in peripheral axons corresponds to an increase in pain thresholds of the mice. Importantly, however, continuous exercise reverses this phenotypic change and normalizes pain thresholds (Groover et al., 2013). Finally, Schwann cell proliferation is increased following exercise and may play an important role in the increase in axonal regeneration necessary for appropriate response to peripheral nerve injury The benefits seen with peripheral nerve regeneration are significant enough to achieve improved values in both functional and morphological markers of nerve and motor function post exercise (Bobinski et al., 2011). These studies bolster the idea that axonal regeneration responds positively to exercise.

Exercise's benefits are not only limited to the periphery, as they also display a substantial value to the central nervous system. The numerous benefits of exercise on both the peripheral and central sensory nervous system are highlighted in **Figure 1**. Centrally, the brain imparts bi-directional control of pain processing and pain modulation that alters the transmission and perception of pain and sensation (Denk et al., 2014). The effects of physical activity on this system are grossly understudied and this important central modulation of pain and sensation would benefit by continued examination of the metabolic, inflammatory, and ionic changes within the CNS.

Although only a few studies have been published, regular physical exercise has been reported to prevent the development of chronic muscle pain and exercise induced muscle pain, possibly by reducing phosphorylation of the NR1 subunit of the N-methyl-D-aspartate (NMDA) receptor in the brainstem, modulating nociception and individual experiences (Sluka et al., 2013). Utilizing either aerobic and resistance exercise, an increase in circulating nitrate levels is seen in both the plasma and cerebrospinal fluid, this observation as well as a loss of analgesic benefit of exercise by nitric oxide inhibitors suggest the nitric oxide/cyclic GMP pathway may provide an antinociceptive benefit during physical activity (Galdino et al., 2010a, 2015a,b).

Additionally in the brain, exercise increases the endogenous opioid content in brainstem regions important in pain modulation, suggesting that exercise-induced reversal of neuropathic pain may include an up-regulation of endogenous opioids (Stagg et al., 2011). This may be a key analgesic mechanism as patients with chronic pain display a reduced endogenous pain inhibition system and creating an imbalance between pain modulation systems (Denk et al., 2014). This highlights another benefit of exercise in which it can increase endogenous analgesic systems known to be critically important in modulating pain. However, the endogenous opioid system however has been disputed in its role in modulating internal antinociceptive effects during physical activity (Galdino et al., 2010b, 2014a). This group has instead suggested the endogenous endocannabinoid system is playing a prominent role in the antinociceptive benefits of exercise (Galdino et al., 2014b,c). There is a definitive need to further explore these endogenous systems that are sensitive to exercise and play a prominent role in antinociception.

The primary benefits of physical activity may have an additive effect when paired with pharmacological interventions. One such study reported that the osteoporosis drug risedronate combined with treadmill running had the most efficacious effects on improving bone mineral density and decreasing sensory nerve calcitonin gene-related peptide (CGRP) expression in osteoporotic rats, as compared to rats receiving only the drug (Orita et al., 2010). In conclusion, while exercise is often thought of as a preventative intervention, there also appears to be a clear benefit after injury. The benefits of exercise affect the nervous system at multiple levels and multiple sites associated with sensory function. Emerging evidence is also revealing the molecular pathways that seem sensitive to, including axonal growth, altered neurotrophin levels, and phenotypic changes in both the periphery and central components of the nervous system.

THE RELATIONSHIP OF INFLAMMATION AND EXERCISE

The immune and nervous systems interact substantially in chronic pain states via immune cells, glia and neurons that coordinate immune responses and the excitation of the pain pathway. Many of these interactions include the synthesis and release of inflammatory mediators and neurotransmitters (Ren and Dubner, 2010). When injured, damaged tissue will signal mast cell degranulation and pro-inflammatory cytokine release such as tumor necrosis factor alpha (TNF- α) and interleukin 1 beta (IL1-β). Additional actions include increased NGF signaling via TrkA that enhances substance P release and pain signal propagation in the spinal cord. In response to peripheral tissue injury, immune cells also synthesize and secrete anti-inflammatory cytokines (IL-10 and IL-6), pro-resolution lipid mediators and opioid peptides to suppress the pain from pro-inflammation cytokines (Rittner et al., 2008; Xu et al., 2010). Many of these pro-inflammatory signals are present acutely after exercise, however, chronic examination of these markers post-exercise often show a robust anti-inflammatory signaling cascade in response to these acute pro-inflammatory markers (Woods et al., 2012). However, it is important to note that in non-healthy patients, there is variability in the acute and chronic inflammatory effects. Overall, however, the chronic effects of exercise on inflammation are still viewed as beneficial to reduce inflammatory signaling in disease (Ploeger et al., 2009). For this reason, exercise has been utilized as an intervention that can activate this natural anti-inflammatory mechanism that causes cells to secrete anti-inflammatory cytokines that suppress the pain induced by pro-inflammatory cytokines (Paley and Johnson, 2015).

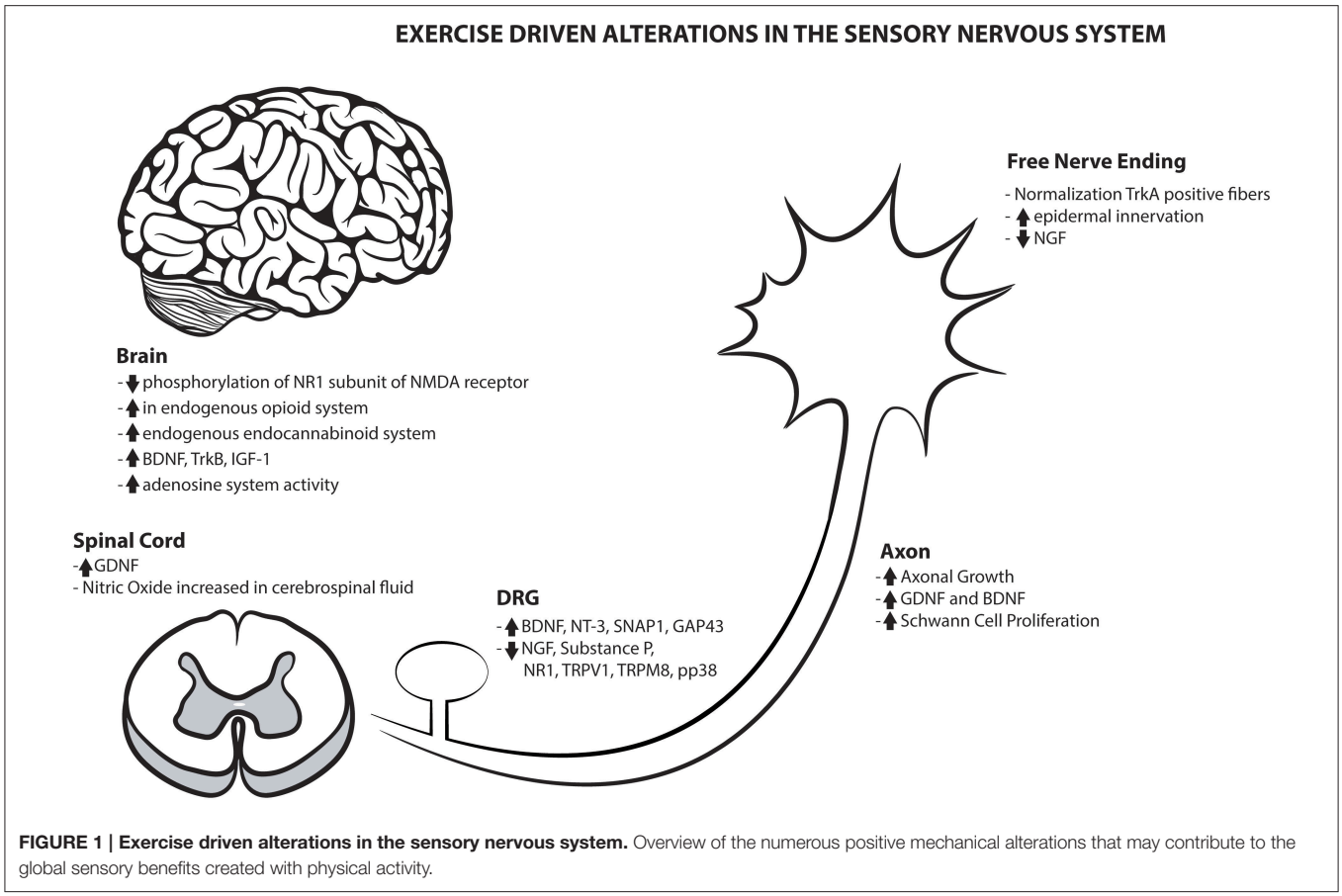


FIGURE 1 | Exercise driven alterations in the sensory nervous system. Overview of the numerous positive mechanical alterations that may contribute to the global sensory benefits created with physical activity.

Numerous cytokines [interleukin 1 alpha (IL1- α), IL1-β, interleukin 2 (IL-2), interleukin 4 (Il-4), interleukin 6 (IL-6), interleukin 10 (IL-10), interferon gamma (IFN-γ), transforming growth factor beta (TGF-β), TNF-α] become active in the spinal cord and DRG during conditions of neuropathic pain (Hopkins and Rothwell, 1995; Ledeboer et al., 2007; Mika et al., 2008; Racz et al., 2008; Wei et al., 2013). Pro-inflammatory cytokines have been implicated in neuropathic and inflammatory nociceptive conditions in a wide array of research (Mika et al., 2013). The most prominently studied inflammation marker, TNF-α, has been implicated as having a key role in both the peripheral and central mechanisms of sensitization to painful stimuli (Leung and Cahill, 2010). Due to their recurring presence during painful stimuli, inflammatory cytokines have been investigated as a sensible target for the explanation of the reduction in allodynia and nociceptive symptoms observed in neuropathic models that utilize endurance and resistance exercise as a rehabilitative technique (Zdziarski et al., 2015).

A prominent benefit of both endurance and resistance exercise programs is their reduction of pro-inflammatory cytokines and their increase in anti-inflammatory markers as displayed in **Figure 2** (Gleeson et al., 2011). The pro-inflammatory acute affects of exercise are proposed to cause a subsequent spike in anti-inflammatory cytokines that are long-lasting after completion of the exercise bout. Regular exercise has been found to decrease inflammatory markers in both young and older humans (Mattusch et al., 2000; Tsukui et al., 2000; Geffken et al., 2001; Colbert et al., 2004). During and after exercise, skeletal muscle increases levels of IL-6, which appears to be responsible for the rise in levels of anti-inflammatory cytokines such as IL-10 and IL-1 receptor agonist (IL-1RA) (Pedersen, 2009). IL-6 has been described as a myokine, a cytokine that is released from muscle fibers during contraction while exerting its effects on other organs (Petersen and Pedersen, 2005). When given as an intravenous infusion, IL-6 provide anti-inflammatory effects similar to a bout of exercise and suppressed pro-inflammatory cytokines such as TNF- α suggesting that IL-6 levels are the cause of anti-inflammatory benefits seen from exercise (Starkie et al., 2003).

Anti-inflammatory markers respond to the rise in IL-6 induced by exercise and have compounding effects that can cause a decrease in allodynia. Increases in IL-10 are able to decrease the expression of pro-inflammatory cytokines and, in-turn, increase the ability of T cells to provide inflammatory responses (Maynard and Weaver, 2008). In mice that exercise via running wheels, T cells were increased in number and associated with a reduction in pro-inflammatory cytokines and an increase in anti-inflammatory cytokines (Wang et al., 2012).

Research showing a decrease in allodynia and hyperalgesia due to exercise intervention suggests that the alterations in

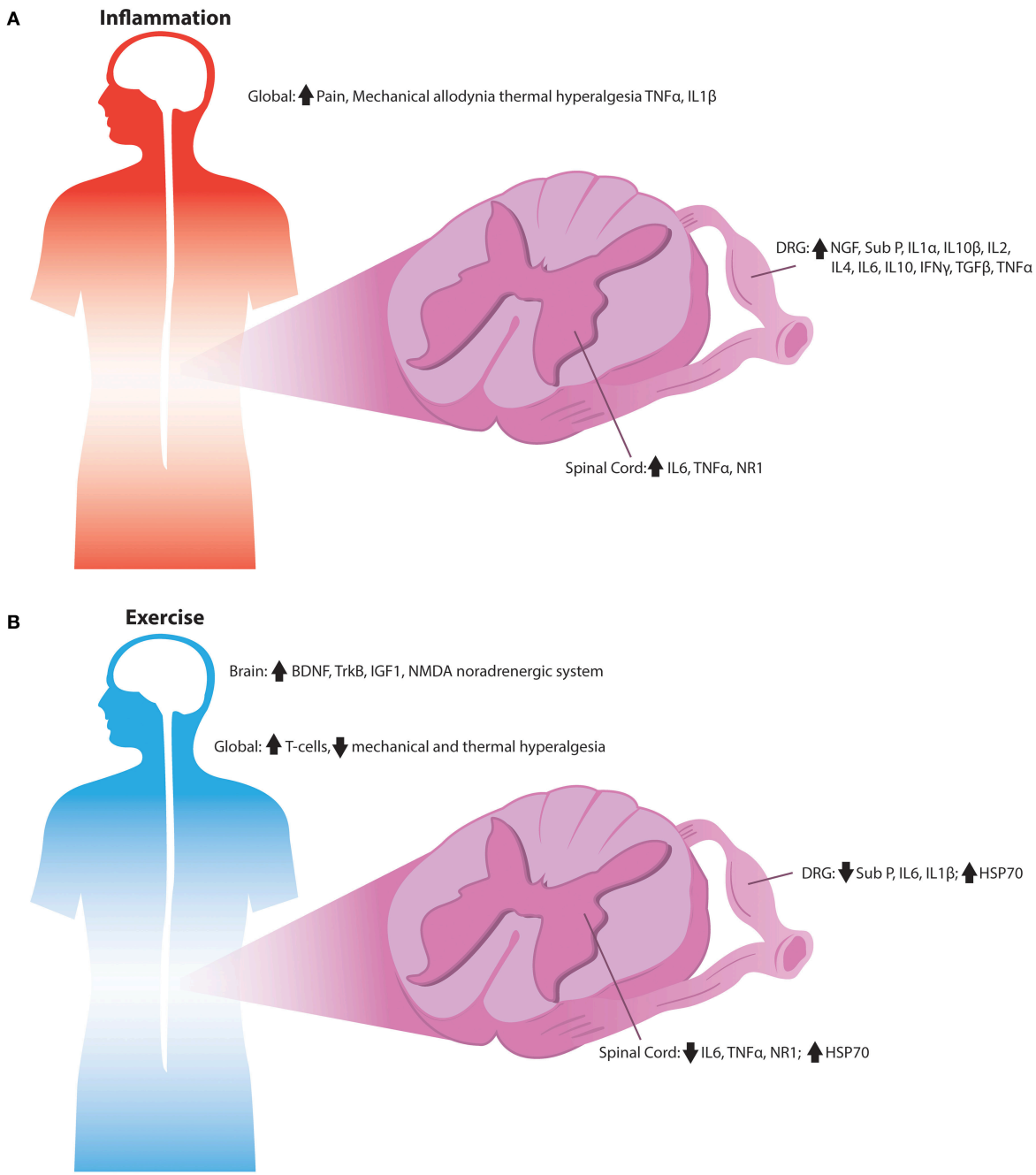


FIGURE 2 | Signaling alterations associated with inflammation by exercise in the sensory nervous system. (A) Overview of the inflammatory alterations in the sensory nervous system of rodents. Inflammatory cytokines are increased throughout the DRG, spinal cord, and peripheral tissues during numerous pain states, which is associated with the development of mechanical and thermal hypersensitivity. **(B)** Anti-inflammatory signaling observed in the sensory nerves of exercised rodents. Exercise's anti-inflammatory signals may reduce pro-inflammatory cytokines; while increasing heat shock proteins and T-cells leading to reduced mechanical and thermal hypersensitivity normally associated with inflammation.

painful sensations are strongly influenced by the increase in anti-inflammatory and decrease in pro-inflammatory cytokines. For example, forced treadmill running reduces substance P, IL-6, TNF- α, NR1, and IL-1β levels after the development of mechanical allodynia induced by skin/muscle incision and retraction (Chen et al., 2013b, 2014). Swimming and treadmill exercise decrease mechanical allodynia, cold allodynia, and heat hyperalgesia while also decreasing TNF- α and IL1-β production

(Chen et al., 2012; Yoon et al., 2015). Diabetes associated neuropathic pain is markedly reduced by progressive exercise training, possibly mediated by an increase of heat shock protein-72 (HSP72) without increases in TNF- α and IL-6 (Chen et al., 2013a). HSP72 is suggested to have a role in the inhibition of TNF- α and IL-6 as well as many other cytokines (Moseley, 1998). A greater expression of HSP72 and a reduction in diabetes-associated neuropathic pain, including thermal hyperalgesia and mechanical allodynia is seen after exercise (Chen et al., 2013a). HSP72 has an essential role in blocking inflammation and insulin resistance associated with a high-fat diet that can lead to type 2 diabetes (Chung et al., 2008). Extended swimming reduces mechanical allodynia and thermal hyperalgesia in rats with formalin and nerve injury-induced inflammatory pain (Kuphal et al., 2007). In conclusion, the interactions between release of pro-inflammatory cytokines and exercise's anti-inflammatory role via the up-regulation and release of anti-inflammatory myokines provides a mechanism that includes multiple sites and actions by which exercise can benefit overactive pain neurons.

EXPERIMENTAL STUDIES OF NEUROPATHIC PAIN AND EXERCISE

Many of these studies discussed above have centered on the ability of exercise to alleviate neuropathic pain associated with diabetes as it is the most prevalent form of neuropathic pain being investigated in hopes of improving diabetic neuropathy (Zilliox and Russell, 2011). **Table 1** highlights studies in rodents in which exercise was used as an intervention. A delay in onset of diabetes-associated neuropathy with continuous exercise may be associated to changes in calcium channel function in the DRG allowing for an alteration in nociceptive signaling from the periphery (Shankarappa et al., 2011). Additional benefits such as increased motor nerve conduction velocities occur in diabetic patients, modifying or delaying the natural course of diabetic peripheral neuropathy (DPN) (Balducci et al., 2006). The mechanism by which exercise is alleviating neuropathic pain is still far from known however, and may very well be the combination of a multitude of changes driven by physical activity. This may especially be true in instances where exercise is seen to provide alleviation in differing forms of neuropathy as well as different types of allodynia.

In a type I DPN rodent model, running is able to rescue many different forms of allodynia including mechanical, cold, and heat hyperalgesia. Associated with these forms of allodynia are key molecular markers such as TRPV1 (heat) and transient receptor potential cation channel subfamily M member 8 (TRPM8) (cold), which were also positively altered by forced running (Yoon et al., 2015). Voluntary aerobic exercise has a recovery effect on nociceptive symptoms and behavior developed from early stages of diabetes such as pre-diabetes induced by a high-fat diet (Groover et al., 2013). However, researchers and clinicians must always be careful not to exceed the level of exercise at which activity is no longer a therapeutic tool, as there has been limited evidence that exercise can increase negative outcome variables in some instances, as discussed in the Clinical section of the current review. An over zealous training protocol can lead to the loss of many benefits seen with activity due to the bodies self protection through the downward activation of glial cells in both motor and sensory neurons (Pereira et al., 2015).

A survey of the literature suggests that diabetes is the most researched disease relative to exercise and neuropathic pain; however, physical activity is also a useful therapeutic tool for pain derived from other sources. Paclitaxel-induced neuropathy frequently occurs in patients undergoing chemotherapy and induces a loss of sensation and sensory fiber loss in the skin. Treadmill exercise reduces the symptoms of sensory loss and increases epidermal nerve fiber density in paclitaxel-treated mice (Park et al., 2015). Additional results include the ability of exercise to decrease abnormal levels of detyrosinated tubulin in paclitaxel-treated nerves, highlighting important anti-neurotoxic effects of exercise (Park et al., 2015). Other studies have shown that moderate intensity exercise reduces hyperalgesia and increases in the neurotrophin, neurotrophin 3 (NT-3), which acts in an analgesic fashion in a number of different pain conditions (Sharma et al., 2010). In studies of spinal cord injury associated pain, treadmill training improves sensory function, ameliorated allodynia, and restores normal sensation after within 5 weeks of the spinal cord injury (Hutchinson et al., 2004). Physical activity also decreases the presence of phagocytic and reactive glial cells following spinal cord injury, suggesting that the positive impact of exercise is limited not only to pain scores, but may also lead to improved functional scores and improved neuronal tissue health (Sandrow-Feinberg et al., 2009). This finding suggests that rhythmic, weight-bearing exercise may be an effective intervention to counter spinal cord injury induced allodynia. Finally, studies of pain associated with sciatica report that physical exercise has a negative influence on nociception. Although this model is not a widely utilized model of pain, in this case, exercise producing more hyperalgesia in rats with sciatica than in a control, non-exercised group (Bertolini et al., 2011). Thus, studies demonstrating negative actions of exercise demand that caution be used to address pain with exercise, as all forms of pain do not respond in the same fashion to exercise.

Varying modes and intensities of exercise have been tested to treat neuropathic pain, almost all of which have a positive effect (Hutchinson et al., 2004; Balducci et al., 2006; Kuphal et al., 2007; Sharma et al., 2010; Shankarappa et al., 2011; Stagg et al., 2011; Sluka et al., 2013). However, not all forms and types of exercise provide the same type or degree of benefit, particularly related to the intensity of exercise (Seo et al., 2009). For instance, treadmill running will increase neurite outgrowth with low intensity, but not high intensity exercise levels. However, studies have not rigorously investigated how varying modes of exercise impact a single model and a single sensory dysfunction. This is likely because researchers focus on a single exercise modality throughout their research study for consistency and control among experimental studies.

Regardless of mode, almost all prominent exercise methods demonstrate beneficial effects as seen with the number of studies previously discussed utilizing traditional aerobic and resistance exercise methods, however even less common forms such as swimming provide a benefit to the nervous system.

TABLE 1 | Summary of various rodent studies addressing sensory dysfunction associated with pain.

Species	Mode of exercise	Pain model	Benefit	References
Rat	Forced running	Skin/muscle incision	↓ Substance P, TNF-α, IL-1β	Chen et al., 2013b, 2014
		Chronic muscle pain	↓ NR1 phosphorylation	Sluka et al., 2013
		Spinal cord injury and acidic saline	↓ Mechanical allodynia	Hutchinson et al., 2004; Sharma et al., 2010
		Sciatic nerve constriction	↓ Heat hyperalgesia and cold allodynia	Chen et al., 2012
		Sciatic nerve crush	↑- Schwann cell proliferation	Seo et al., 2009
		Lumbar spinal nerve ligation	↑- Endogenous opioids	Stagg et al., 2011
		Sciatic nerve cut	↓ NGF and BDNF	López-Álvarez et al., 2015
		Sciatic nerve cut	Normalized NKCC1 regulation	López-Álvarez et al., 2015
		Osteoporosis	↓ CGRP fibers in bone	Orita et al., 2010
		Paclitaxel-induced neuropathy	↑- Epidermal axon innervation	Park et al., 2015
		Streptozotocin	↑- HSP72	Chen et al., 2013a
		Streptozotocin	↓ TRPM8, TRPV1, and pp38	Yoon et al., 2015
		Acute antinociception	Activated endogenous cannabinoid system	Galdino et al., 2014b
		Acute antinociception	Activation of nitrous Oxide/cGMP pathway	Galdino et al., 2010a, 2015a
Rat	Swimming	Nerve constriction and inflammation	↓ Mechanical allodynia and heat hyperalgesia	Kuphal et al., 2007; Chen et al., 2012
		Streptozotocin	↓ TNF-alpha and IL-1β	Yoon et al., 2015
		CRPS type I	↑- Adenosine	Martins et al., 2013
Rat	Resistance exercise	Acute antinociception	Activated endogenous cannabinoid system	Galdino et al., 2014a
		Acute antinociception	Activated nitrous oxide/cGMP/KATP pathway	Galdino et al., 2015b
Mouse	Running wheel	High fat diet/pre-diabetes	↓ Mechanical allodynia	Groover et al., 2013
		High fat diet/pre-diabetes	↓ TrkA positive fibers	Groover et al., 2013
		High fat diet/pre-diabetes	↓ NGF, ↑- BDNF	Groover et al., 2013
		Nerve crush	↑- BDNF, NT3, GAP43, and SNAP1 (mRNA)	Molteni et al., 2004
Mouse	Treadmill	Sciatic nerve crush	↑- Nerve regeneration	Bobinski et al., 2011

A range of species, modes of exercises, pain model, and primary outcomes are provided.

Swimming provides positive results as a therapy for induced nerve injury in rats, reducing both mechanical allodynia and thermal hyperalgesia (Shen et al., 2013). In addition, swimming reduces pain hypersensitivity in a number of experimental models, including formalin and nerve injury-induced animal models of persistent pain, decreasing nerve injury, induced cold allodynia, thermal hyperalgesia in rats, and decreased nerve injury-induced hyperalgesia in mice (Kuphal et al., 2007). The mechanism by which swimming exercise reduces mechanical allodynia may involve endogenous adenosine and adenosine A_1 receptors (Martins et al., 2013). It is reported that agonists to the adenosine A1 receptor reduce mechanical allodynia in a neuropathic pain model of diabetes, suggesting another possible mechanism in which exercise may reduce pain (Katz et al., 2015). These positive results from swimming offer an extremely attractive exercise modality for patients with neuropathic pain due to the reduced load on pain-affected extremities and problems with coordination many patients, especially in elderly patients. For this reason, additional research in swimming regimens is needed to understand benefits for neuropathic pain, as its utilization could be the best avenue for relief for many patients.

Just as exercise intensity and mode may be key factors in the benefits of physical activity, time of exercise onset and duration may prove to be important as well. Intense short-burst exercise significantly reduced mechanical allodynia in a chronic constriction injury model of neuropathic pain, resulting in better recovery of sensorimotor function (Cobianchi et al., 2010). The relationship of time between the onset of injury and the start of exercise is not clear, however, there are multiple studies that have reported positive results with exercise starting within 1 week of injury. Initiation of treadmill running 3 days after an induced injury had an immediate and long-lasting reduction in pain that was independent of the duration of exercise (Cobianchi et al., 2013). Exercise training beginning 5 days after injury was sufficient to prevent the development of neuropathic pain (Detloff et al., 2014). Also, exercise initiation 7 days after spinal nerve ligation was able to reduce thermal and tactile hypersensitivity (Stagg et al., 2011). These studies suggest that there may be no need to wait for a certain amount of time to pass after injury before the introduction of exercise as a

therapeutic aide; however, understanding of how a given injury may affect motor control should be taken into consideration as improper exercise technique can minimize benefits or even increase negative effects of an injury.

CLINICAL USE OF EXERCISE

The use of physical exercise as a therapeutic treatment to specifically address pain is a relatively new and developing field. The majority of research on exercise for peripheral pain syndromes in human subjects is associated with diabetic or pre-diabetic neuropathic pain. The few studies, which have utilized physical activity, are displayed in **Table 2**. Historically, clinicians may have been reluctant to encourage exercise in patients with diabetic neuropathy due to the risk of possible adverse outcomes such as foot ulcers in insensate feet or increased pain. People with fibromyalgia have expressed exercise as a pain-inducing stimulus, and report an increase in negative symptoms due to exercise, however, cumulatively exercise has been shown to improve patients quality of life (Nijs et al., 2012; Daenen et al., 2015). Additionally, people with painful diabetic neuropathy (PDN) have reported higher ratings of perceived exertion and muscle pain during exercise and no improvements in thermal pain ratings following exercise (Knauf and Koltyn, 2014). Importantly, however, it has been reported by numerous groups that exercise can be performed safely in patients with type 2 diabetic neuropathies and exercise intervention produces a marked improvement in certain nerve functions (Fisher et al., 2007; Kluding et al., 2015).

Aerobic exercise is an often-studied modality in clinical programs for people with diabetes. Diabetics have experienced a benefit in both motor and sensory neuropathy measures. Aerobic exercise reduces the development of diabetic neuropathy (Balducci et al., 2006), as well as increasing the intraepidermal nerve fiber density (IENFD) and visual analog scale (VAS) pain measure in people with diabetes (Smith et al., 2006; Fisher et al., 2007; Kluding et al., 2012; Singleton et al., 2014, 2015). Similarly, exercise induced improvement in metabolic syndrome patients saw an increase in cutaneous IENFD even though these patients were non-diabetic (Singleton et al., 2015). Physical exercise when paired with diet counseling has resulted in partial cutaneous re-innervation in pre-diabetic individuals, highlighting that exercise may have on early symptoms and possible prevention of neuropathic symptoms (Smith et al., 2006).

With sensory changes, it is important to remember that anatomical changes are not the only factor to examine, functional changes are just as important for clinical implications. Exercise training's benefit through the reinforcement of existent sensorimotor pathways rather than promoting generation of new pathways may be a significant reason to examine functional changes as outcome measures (de Leon et al., 1998). For instance, patients with motor and sensory neuropathy see a gain in strength with exercise training, but only a marginal functional increase (Allet et al., 2010; Song et al., 2011; Mueller et al., 2013; Dixit et al., 2014). These sensorimotor benefits such as improved balance, mobility, and a decrease in peripheral neuropathy, can combine to significant whole measure outcome increases such as quality of life (QOL) (Streckmann et al., 2014). Even with the knowledge that exercise is beneficial in multiple diseases, providing benefits through a multitude of mechanisms, there exists a great limitation in the breadth of knowledge as to how exercise truly exerts its benefit.

One extremely well documented result of exercise on the sensorimotor pathways is the observation of exercise-induced hypoalgesia (EIH) resulting in a myriad of populations and testing conditions (Koltyn and Arbogast, 1998; Koltyn, 2002; Koltyn and Umeda, 2006; Kodesh and Weissman-Fogel, 2014; Koltyn et al., 2014; Vaegter et al., 2014, 2015b). This induction of hypoalgesia occurs independent of exercise mode with benefits occurring with both interval and traditional aerobic exercise, as well as with resistance exercise (Koltyn and Arbogast, 1998; Kodesh and Weissman-Fogel, 2014; Vaegter et al., 2015b). In line with the development of hypoalgesia, numerous groups have reported a reduction in temporal summation of pain (TSP) following physical activity (Koltyn et al., 2013; Naugle and Riley, 2014; Vaegter et al., 2015a). TSP is commonly used to reflect the amount of CNS involved nociception and is often hypothesized as being sensitive to alteration in acute and chronic pain states, suggesting that exercise may be providing a benefit in both the peripheral and central nervous system pathways important for pain.

A unique aspect of clinical trials in human patients is the ability to perform voluntary resistance based exercise as well as aerobic exercise. A study associated with metabolic features of diabetes that combined aerobic and resistance training did not observe an increase in detrimental affects when compared to a program that utilized only one form of exercise (Sigal et al., 2007). This study saw an improvement in glycemic control through measurement of hemoglobin A1C values in elderly adults that completed a moderate intensity weight program. Combining resistance training with balance training and vibration as opposed to aerobic training saw an improvement in balance, muscle strength, and hemoglobin A1C levels when compared to balance and vibration alone (Lee et al., 2013). Resistance training combined with high intensity training significantly improved muscle strength, blood pressure regulation in long-standing, insulin-treated type 2 diabetics with diabetic neuropathy (Praet et al., 2008). Therefore, individuals with type 2 diabetes looking to improve glycemic control through physical activity should be encouraged to perform both aerobic and resistance training. This point needs to be addressed in future clinical studies associated with pain.

Moderate aerobic exercise helps to preserve peripheral nerve function and help to combat health behaviors associated with DPN in type 2 diabetes (Dixit et al., 2014). In adults with and without diabetic neuropathy, aerobic exercise benefited gait changes, reaction times, and balance measures, although it did not reduce the rate of falls in these groups (Morrison et al., 2014). Combination therapy of backward walking and alpha-lipoic acid treatment (ALA) to reduce and prevent free radical damage through antioxidant action was more effective than just ALA alone when examining plantar pressure in patients with DPN, suggesting that just as seen with mammalian models, exercise

TABLE 2 | Summary of human studies addressing sensory dysfunction associated with pain.

Mode of exercise	Pain model	Benefit	References
Aerobic exercise	Diabetic neuropathy	↑- Motor conduction velocity	Balducci et al., 2006
	DPN and metabolic syndrome	↑-Epidermal innervation	Kluding et al., 2012; Singleton et al., 2015
	DPN	↓ Pain ratings	Kluding et al., 2012
Vibration Platform	DPN	↓ Pain levels and improved gait	Hong et al., 2013
	DPN	↓ Neuropathic pain scale	Kessler and Hong, 2013
	DPN	↓ Pain ratings	Kessler and Hong, 2013

Modes of exercises, pain model, and primary outcomes are provided.

may have an additive affect when paired with pharmacological agents (Zhang et al., 2014). A 2013 case study showed that a 4-week vibration treatment designed to simulate movement on the feet of patients with diabetic peripheral small fiber neuropathy significantly improved pain levels and gait (Hong et al., 2013). Another study reported that whole body vibration significantly reduced acute pain in the visual analog pain scale and chronic reduction in neuropathic pain scales (Kessler and Hong, 2013). This proposes that movement alone, or simulation of movement, may be able to provide a benefit for painful symptoms. However, there still remains a significant gap in our knowledge about the molecular pathways altered by exercise.

CONCLUSIONS

The use of exercise as a therapeutic tool is a rapidly growing field in biomedical research. However, there is a dire need for increased research into understanding the role of exercise in sensory nerve disorders. The lack of understanding in the pathways affected by exercise and the molecular changes that lead to the benefits seen with exercise is a hindrance to the medical community working to utilize this tool for their patients. It is our expectation that certain types of pain may benefit from exercise, though different mechanisms driving the development of pain can vary. For instance the benefits of exercise on diabetic pain may be influenced by concurrent correction of metabolic abnormalities, while nerve damage associate pain may be associated with local, acute alterations in gene expression and inflammation. Overall however, the prevailing literature suggests that for the vast majority of nerve related disorders; exercise offers a benefit and can be an attractive therapeutic aide for clinicians. However, the clinical use of exercise requires the investment of the patient and their willingness to expend the effort which exercise requires. To motivate patients to exercise, they must perceive there will be approximately two times greater improvement of symptoms than without exercise (Anderson et al., 2015). Clinicians and researchers therefore must continue to examine and highlight the myriad of benefits which exercise provides. Future research should continue to examine the use of exercise in a clinical setting, looking to answer what changes occur in different neural compartments that underlie reductions in pain. Finally, the use of exercise in human subjects in a larger array of diseases will help expose the clinical benefits of exercise for a larger portion of the health care community.

AUTHOR CONTRIBUTIONS

MC: Primary author, performed literature review which was the basis of the submitted review. PK: Was consulted with over the clinical research included in the review as well as the shaping of the entire article. DW: Corresponding author, primary editor of the review as well as advisor of all work performed by MC.

ACKNOWLEDGMENTS

This work was supported by the NIH grants P20GM1033418 and RO1NS43314 (DW) as well as core facility support from NICHD P30 HD002528 (Kansas IDDRC).

REFERENCES

Allet, L., Armand, S., Aminian, K., Pataky, Z., Golay, A., de Bie, R. A., et al. (2010). An exercise intervention to improve diabetic patients' gait in a real-life environment. *Gait Posture* 32, 185–190. doi: 10.1016/j.gaitpost.2010.04.013

Amann, M., Runnels, S., Morgan, D. E., Trinity, J. D., Fjeldstad, A. S., Wray, D. W., et al. (2011). On the contribution of group III and IV muscle afferents to the circulatory response to rhythmic exercise in humans. *J. Physiol.* 589(Pt 15), 3855–3866. doi: 10.1113/jphysiol.2011.209353

Anderson, R. J., Hurley, R. W., Staud, R., and Robinson, M. E. (2015). Cognitive-motivational influences on health behavior change in adults with chronic pain. *Pain Med*. doi: 10.1111/pme.12929. [Epub ahead of print].

Balducci, S., Iacobellis, G., Parisi, L., Di Biase, N., Calandriello, E., Leonetti, F., et al. (2006). Exercise training can modify the natural history of diabetic peripheral neuropathy. *J. Diabetes Complicat.* 20, 216–223. doi: 10.1016/j.jdiacomp.2005.07.005

Bertolini, G. R., do Nascimento, C. M., Artifon, E. L., Cunha, D. M., and Meireles, A. (2011). Swimming training and functional assessment of sciatic nociception in rats. *Rev. Bras. Reumatol.* 51, 249–259.

Bobinski, F., Martins, D. F., Bratti, T., Mazzardo-Martins, L., Winkelmann-Duarte, E. C., Guglielmo, L. G., et al. (2011). Neuroprotective and neuroregenerative effects of low-intensity aerobic exercise on sciatic nerve crush injury in mice. *Neuroscience* 194, 337–348. doi: 10.1016/j.neuroscience.2011.07.075

Chen, Y. W., Hsieh, P. L., Chen, Y. C., Hung, C. H., and Cheng, J. T. (2013a). Physical exercise induces excess hsp72 expression and delays the development

of hyperalgesia and allodynia in painful diabetic neuropathy rats. *Anesth. Analg.* 116, 482–490. doi: 10.1213/ANE.0b013e318274e4a0

Chen, Y. W., Lin, M. F., Chen, Y. C., Hung, C. H., Tzeng, J. I., and Wang, J. J. (2013b). Exercise training attenuates postoperative pain and expression of cytokines and N-methyl-D-aspartate receptor subunit 1 in rats. *Reg. Anesth. Pain Med.* 38, 282–288. doi: 10.1097/AAP.0b013e31828df3f9

Chen, Y. W., Li, Y. T., Chen, Y. C., Li, Z. Y., and Hung, C. H. (2012). Exercise training attenuates neuropathic pain and cytokine expression after chronic constriction injury of rat sciatic nerve. *Anesth. Analg.* 114, 1330–1337. doi: 10.1213/ANE.0b013e31824c4ed4

Chen, Y. W., Tzeng, J. I., Lin, M. F., Hung, C. H., and Wang, J. J. (2014). Forced treadmill running suppresses postincisional pain and inhibits upregulation of substance P and cytokines in rat dorsal root ganglion. *J. Pain.* 15, 827–834. doi: 10.1016/j.jpain.2014.04.010

Chung, J., Nguyen, A. K., Henstridge, D. C., Holmes, A. G., Chan, M. H., Mesa, J. L., et al. (2008). HSP72 protects against obesity-induced insulin resistance. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1739–1744. doi: 10.1073/pnas.0705799105

Cobianchi, S., Casals-Diaz, L., Jaramillo, J., and Navarro, X. (2013). Differential effects of activity dependent treatments on axonal regeneration and neuropathic pain after peripheral nerve injury. *Exp. Neurol.* 240, 157–167. doi: 10.1016/j.expneurol.2012.11.023

Cobianchi, S., Marinelli, S., Florenzano, F., Pavone, F., and Luvisetto, S. (2010). Short- but not long-lasting treadmill running reduces allodynia and improves functional recovery after peripheral nerve injury. *Neuroscience* 168, 273–287. doi: 10.1016/j.neuroscience.2010.03.035

Colbert, L. H., Visser, M., Simonsick, E. M., Tracy, R. P., Newman, A. B., Kritchevsky, S. B., et al. (2004). Physical activity, exercise, and inflammatory markers in older adults: findings from the Health, Aging and Body Composition Study. *J. Am. Geriatr. Soc.* 52, 1098–1104. doi: 10.1111/j.1532-5415.2004.52307.x

Daenen, L., Varkey, E., Kellmann, M., and Nijs, J. (2015). Exercise, not to exercise, or how to exercise in patients with chronic pain? Applying science to practice. *Clin. J. Pain* 31, 108–114. doi: 10.1097/AJP.0000000000000099

de Leon, R. D., Hodgson, J. A., Roy, R. R., and Edgerton, V. R. (1998). Locomotor capacity attributable to step training versus spontaneous recovery after spinalization in adult cats. *J. Neurophysiol.* 79, 1329–1340.

Denk, F., McMahon, S. B., and Tracey, I. (2014). Pain vulnerability: a neurobiological perspective. *Nat. Neurosci.* 17, 192–200. doi: 10.1038/nn.3628

Detloff, M. R., Smith, E. J., Quiros Molina, D., Ganzer, P. D., and Houlé, J. D. (2014). Acute exercise prevents the development of neuropathic pain and the sprouting of non-peptidergic (GDNF- and artemin-responsive) c-fibers after spinal cord injury. *Exp. Neurol.* 255, 38–48. doi: 10.1016/j.expneurol.2014.02.013

Dixit, S., Maiya, A., and Shastry, B. (2014). Effect of aerobic exercise on quality of life in population with diabetic peripheral neuropathy in type 2 diabetes: a single blind, randomized controlled trial. *Qual. Life Res.* 23, 1629–1640. doi: 10.1007/s11136-013-0602-7

Fisher, M. A., Langbein, W. E., Collins, E. G., Williams, K., and Corzine, L. (2007). Physiological improvement with moderate exercise in type II diabetic neuropathy. *Electromyogr. Clin. Neurophysiol.* 47, 23–28.

Galdino, G., Romero, T., Andrade, I., Duarte, I., and Perez, A. (2014a). Opioid receptors are not involved in the increase of the nociceptive threshold induced by aerobic exercise. *Neurosciences (Riyadh).* 19, 33–37.

Galdino, G., Romero, T. R., Silva, J. F., Aguiar, D. C., de Paula, A. M., Cruz, J. S., et al. (2014c). The endocannabinoid system mediates aerobic exercise-induced antinociception in rats. *Neuropharmacology* 77, 313–324. doi: 10.1016/j.neuropharm.2013.09.022

Galdino, G., Romero, T., Silva, J. F., Aguiar, D., Paula, A. M., Cruz, J., et al. (2014b). Acute resistance exercise induces antinociception by activation of the endocannabinoid system in rats. *Anesth. Analg.* 119, 702–715. doi: 10.1213/ANE.0000000000000340

Galdino, G. S., Cortes, S. F., Duarte, I. D., and Perez, A. C. (2010a). Involvement of the nitric oxide/(C)GMP/K(ATP) pathway in antinociception induced by exercise in rats. *Life Sci.* 86, 505–509. doi: 10.1016/j.lfs.2010.02.004

Galdino, G. S., Duarte, I. D., and Perez, A. C. (2010b). Participation of endogenous opioids in the antinociception induced by resistance exercise in rats. *Braz. J. Med. Biol. Res.* 43, 906–909. doi: 10.1590/S0100-879X2010007500086

Galdino, G. S., Duarte, I. D., and Perez, A. C. (2015a). Central release of nitric oxide mediates antinociception induced by aerobic exercise. *Braz. J. Med. Biol. Res.* 48, 790–797. doi: 10.1590/1414-431X20144160

Galdino, G. S., Xavier, C. H., Almeida, R., Silva, G., Fontes, M. A., Menezes, G., et al. (2015b). The Nitric oxide/CGMP/KATP pathway mediates systemic and central antinociception induced by resistance exercise in rats. *Int. J. Neurosci.* 125, 765–773. doi: 10.3109/00207454.2014.970256

Gaskin, D. J., and Richard, P. (2012). The economic costs of pain in the United States. *J. Pain* 13, 715–724. doi: 10.1016/j.jpain.2012.03.009

Geffken, D. F., Cushman, M., Burke, G. L., Polak, J. F., Sakkinen, P. A., and Tracy, R. P. (2001). Association between physical activity and markers of inflammation in a healthy elderly population. *Am. J. Epidemiol.* 153, 242–250. doi: 10.1093/aje/153.3.242

Gleeson, M., Bishop, N. C., Stensel, D. J., Lindley, M. R., Mastana, S. S., and Nimmo, M. A. (2011). The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease. *Nat. Rev. Immunol.* 11, 607–615. doi: 10.1038/nri3041

Groover, A. L., Ryals, J. M., Guilford, B. L., Wilson, N. M., Christianson, J. A., and Wright, D. E. (2013). Exercise-mediated improvements in painful neuropathy associated with prediabetes in mice. *Pain* 154, 2658–2667. doi: 10.1016/j.pain.2013.07.052

Hong, J., Barnes, M., and Kessler, N. (2013). Case study: use of vibration therapy in the treatment of diabetic peripheral small fiber neuropathy. *J. Bodyw. Mov. Ther.* 17, 235–238. doi: 10.1016/j.jbmt.2012.08.007

Hopkins, S. J., and Rothwell, N. J. (1995). Cytokines and the nervous system. I: expression and recognition. *Trends Neurosci.* 18, 83–88. doi: 10.1016/0166-2236(95)80029-2

Hutchinson, K. J., Gómez-Pinilla, F., Crowe, M. J., Ying, Z., and Basso, D. M. (2004). Three exercise paradigms differentially improve sensory recovery after spinal cord contusion in rats. *Brain* 127(Pt 6), 1403–1414. doi: 10.1093/brain/awh160

Jones, R. C. 3rd., Lawson, E., and Backonja, M. (2016). Managing neuropathic pain. *Med. Clin. North Am.* 100, 151–167. doi: 10.1016/j.mcna.2015.08.009

Katz, N. K., Ryals, J. M., and Wright, D. E. (2015). Central or peripheral delivery of an adenosine A1 receptor agonist improves mechanical allodynia in a mouse model of painful diabetic neuropathy. *Neuroscience* 285, 312–323. doi: 10.1016/j.neuroscience.2014.10.065

Keeler, B. E., Liu, G., Siegfried, R. N., Zhukareva, V., Murray, M., and Houlé, J. D. (2012). Acute and prolonged hindlimb exercise elicits different gene expression in motoneurons than sensory neurons after spinal cord injury. *Brain Res.* 1438, 8–21. doi: 10.1016/j.brainres.2011.12.015

Kessler, N. J., and Hong, J. (2013). Whole body vibration therapy for painful diabetic peripheral neuropathy: a pilot study. *J. Bodyw. Mov. Ther.* 17, 518–522. doi: 10.1016/j.jbmt.2013.03.001

Kluding, P. M., Pasnoor, M., Singh, R., D'Silva, L. J., Yoo, M., Billinger, S. A., et al. (2015). Safety of aerobic exercise in people with diabetic peripheral neuropathy: single-group clinical trial. *Phys. Ther.* 95, 223–234. doi: 10.2522/ptj.20140108

Kluding, P. M., Pasnoor, M., Singh, R., Jernigan, S., Farmer, K., Rucker, J., et al. (2012). The effect of exercise on neuropathic symptoms, nerve function, and cutaneous innervation in people with diabetic peripheral neuropathy. *J. Diabetes Complicat.* 26, 424–429. doi: 10.1016/j.jdiacomp.2012.05.007

Knauf, M. T., and Koltyn, K. F. (2014). Exercise-induced modulation of pain in adults with and without painful diabetic neuropathy. *J. Pain* 15, 656–663. doi: 10.1016/j.jpain.2014.02.008

Kodesh, E., and Weissman-Fogel, I. (2014). Exercise-induced hypoalgesia - interval versus continuous mode. *Appl. Physiol. Nutr. Metab.* 39, 829–834. doi: 10.1139/apnm-2013-0481

Koltyn, K. F. (2002). Exercise-induced hypoalgesia and intensity of exercise. *Sports Med.* 32, 477–487. doi: 10.2165/00007256-200232080-00001

Koltyn, K. F., and Arbogast, R. W. (1998). Perception of pain after resistance exercise. *Br. J. Sports Med.* 32, 20–24. doi: 10.1136/bjsm.32.1.20

Koltyn, K. F., Brellenthin, A. G., Cook, D. B., Sehgal, N., and Hillard, C. (2014). Mechanisms of exercise-induced hypoalgesia. *J. Pain* 15, 1294–1304. doi: 10.1016/j.jpain.2014.09.006

Koltyn, K. F., Knauf, M. T., and Brellenthin, A. G. (2013). Temporal summation of heat pain modulated by isometric exercise. *Eur. J. Pain* 17, 1005–1011. doi: 10.1002/j.1532-2149.2012.00264.x

Koltyn, K. F., and Umeda, M. (2006). Exercise, hypoalgesia and blood pressure. *Sports Med.* 36, 207–214. doi: 10.2165/00007256-200636030-00003

Kuphal, K. E., Fibuch, E. E., and Taylor, B. K. (2007). Extended swimming exercise reduces inflammatory and peripheral neuropathic pain in rodents. *J. Pain* 8, 989–997. doi: 10.1016/j.jpain.2007.08.001

Ledeboer, A., Jekich, B. M., Sloane, E. M., Mahoney, J. H., Langer, S. J., Milligan, E. D., et al. (2007). Intrathecal interleukin-10 gene therapy attenuates paclitaxel-induced mechanical allodynia and proinflammatory cytokine expression in dorsal root ganglia in rats. *Brain Behav. Immun.* 21, 686–698. doi: 10.1016/j.bbi.2006.10.012

Lee, K., Lee, S., and Song, C. (2013). Whole-body vibration training improves balance, muscle strength and glycosylated hemoglobin in elderly patients with diabetic neuropathy. *Tohoku J. Exp. Med.* 231, 305–314. doi: 10.1620/tjem.231.305

Leung, L., and Cahill, C. M. (2010). TNF-alpha and neuropathic pain–a review. *J. Neuroinflammation* 7, 27. doi: 10.1186/1742-2094-7-27

López-Álvarez, V. M., Modol, L., Navarro, X., and Cobianchi, S. (2015). Early increasing-intensity treadmill exercise reduces neuropathic pain by preventing nociceptor collateral sprouting and disruption of chloride cotransporters homeostasis after peripheral nerve injury. *Pain* 156, 1812–1825. doi: 10.1097/j.pain.0000000000000268

Martins, D. F., Mazzardo-Martins, L., Soldi, F., Stramosk, J., Piovezan, A. P., and Santos, A. R. (2013). High-intensity swimming exercise reduces neuropathic pain in an animal model of complex regional pain syndrome type I: evidence for a role of the adenosinergic system. *Neuroscience* 234, 69–76. doi: 10.1016/j.neuroscience.2012.12.042

Mattusch, F., Dufaux, B., Heine, O., Mertens, I., and Rost, R. (2000). Reduction of the plasma concentration of C-reactive protein following nine months of endurance training. *Int. J. Sports Med.* 21, 21–24. doi: 10.1055/s-2000-8852

Maynard, C. L., and Weaver, C. T. (2008). Diversity in the contribution of interleukin-10 to T-cell-mediated immune regulation. *Immunol. Rev.* 226, 219–233. doi: 10.1111/j.1600-065X.2008.00711.x

Mika, J., Korostynski, M., Kaminska, D., Wawrzczak-Bargiela, A., Osikowicz, M., Makuch, W., et al. (2008). Interleukin-1 alpha has antiallodynic and antihyperalgesic activities in a rat neuropathic pain model. *Pain* 138, 587–597. doi: 10.1016/j.pain.2008.02.015

Mika, J., Zychowska, M., Popiolek-Barczyk, K., Rojewska, E., and Przewlocka, B. (2013). Importance of glial activation in neuropathic pain. *Eur. J. Pharmacol.* 716, 106–119. doi: 10.1016/j.ejphar.2013.01.072

Mitchell, J. H. (1985). Cardiovascular control during exercise: central and reflex neural mechanisms. *Am. J. Cardiol.* 55, 34D–41D. doi: 10.1016/0002-9149(85)91053-7

Molteni, R., Zheng, J. Q., Ying, Z., Gómez-Pinilla, F., and Twiss, J. L. (2004). Voluntary exercise increases axonal regeneration from sensory neurons. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8473–8478. doi: 10.1073/pnas.0401443101

Morrison, S., Colberg, S. R., Parson, H. K., and Vinik, A. I. (2014). Exercise improves gait, reaction time and postural stability in older adults with type 2 diabetes and neuropathy. *J. Diabetes Complicat.* 28, 715–722. doi: 10.1016/j.jdiacomp.2014.04.007

Moseley, P. L. (1998). Heat shock proteins and the inflammatory response. *Ann. N. Y. Acad. Sci.* 856, 206–213. doi: 10.1111/j.1749-6632.1998.tb08327.x

Mueller, M. J., Tuttle, L. J., Lemaster, J. W., Strube, M. J., McGill, J. B., Hastings, M. K., et al. (2013). Weight-bearing versus nonweight-bearing exercise for persons with diabetes and peripheral neuropathy: a randomized controlled trial. *Arch. Phys. Med. Rehabil.* 94, 829–838. doi: 10.1016/j.apmr.2012.12.015

Naugle, K. M., and Riley, J. L. 3rd (2014). Self-reported physical activity predicts pain inhibitory and facilitatory function. *Med. Sci. Sports Exerc.* 46, 622–629. doi: 10.1249/MSS.0b013e3182a69cf1

Nijs, J., Kosek, E., Van Oosterwijck, J., and Meeus, M. (2012). Dysfunctional endogenous analgesia during exercise in patients with chronic pain: to exercise or not to exercise? *Pain Physician* 15(3 Suppl), ES205–ES213.

O'Leary, D. S., Augustyniak, R. A., Ansorge, E. J., and Collins, H. L. (1999). Muscle metaboreflex improves O2 delivery to ischemic active skeletal muscle. *Am. J. Physiol.* 276(4 Pt 2), H1399–H1403.

Orita, S., Ohtori, S., Koshi, T., Yamashita, M., Yamauchi, K., Inoue, G., et al. (2010). The effects of risedronate and exercise on osteoporotic lumbar rat vertebrae and their sensory innervation. *Spine* 35, 1974–1982. doi: 10.1097/BRS.0b013e3181d5959e

Paley, C. A., and Johnson, M. I. (2015). Physical activity to reduce systemic inflammation associated with chronic pain and obesity: a narrative review. *Clin. J. Pain*. 32, 365–370. doi: 10.1097/AJP.0000000000000258

Park, J. S., Kim, S., and Hoke, A. (2015). An exercise regimen prevents development paclitaxel induced peripheral neuropathy in a mouse model. *J. Peripher. Nerv. Syst.* 20, 7–14. doi: 10.1111/jns.12109

Pedersen, B. K. (2009). Edward F. Adolph distinguished lecture: muscle as an endocrine organ: IL-6 and other myokines. *J. Appl. Physiol.* 107, 1006–1014. doi: 10.1152/japplphysiol.00734.2009

Pereira, B. C., Lucas, G., da Rocha, A. L., Pauli, J. R., Ropelle, E. R., Cintra, D., et al. (2015). Eccentric exercise leads to glial activation but not apoptosis in mice spinal cords. *Int. J. Sports Med.* 36, 378–385. doi: 10.1055/s-0034-1395589

Petersen, A. M., and Pedersen, B. K. (2005). The anti-inflammatory effect of exercise. *J. Appl. Physiol.* 98, 1154–1162. doi: 10.1152/japplphysiol.00164.2004

Ploeger, H. E., Takken, T., de Greef, M. H., and Timmons, B. W. (2009). The effects of acute and chronic exercise on inflammatory markers in children and adults with a chronic inflammatory disease: a systematic review. *Exerc. Immunol. Rev.* 15, 6–41.

Praet, S. F., Jonkers, R. A., Schep, G., Stehouwer, C. D., Kuipers, H., Keizer, H. A., et al. (2008). Long-standing, insulin-treated type 2 diabetes patients with complications respond well to short-term resistance and interval exercise training. *Eur. J. Endocrinol.* 158, 163–172. doi: 10.1530/EJE-07-0169

Racz, I., Nadal, X., Alferink, J., Baños, J. E., Rehnelt, J., Martín, M., et al. (2008). Interferon-gamma is a critical modulator of CB(2) cannabinoid receptor signaling during neuropathic pain. *J. Neurosci.* 28, 12136–12145. doi: 10.1523/JNEUROSCI.3402-08.2008

Ren, K., and Dubner, R. (2010). Interactions between the immune and nervous systems in pain. *Nat. Med.* 16, 1267–1276. doi: 10.1038/nm.2234

Rittner, H. L., Brack, A., and Stein, C. (2008). Pain and the immune system. *Br. J. Anaesth.* 101, 40–44. doi: 10.1093/bja/aen078

Sandrow-Feinberg, H. R., Izzi, J., Shumsky, J. S., Zhukareva, V., and Houle, J. D. (2009). Forced exercise as a rehabilitation strategy after unilateral cervical spinal cord contusion injury. *J. Neurotrauma* 26, 721–731. doi: 10.1089/neu.2008.0750

Scholz, J., and Woolf, C. J. (2007). The neuropathic pain triad: neurons, immune cells and glia. *Nat. Neurosci.* 10, 1361–1368. doi: 10.1038/nn1992

Schreiber, A. K., Nones, C. F., Reis, R. C., Chichorro, J. G., and Cunha, J. M. (2015). Diabetic neuropathic pain: physiopathology and treatment. *World J. Diabetes* 6, 432–444. doi: 10.4239/wjd.v6.i3.432

Seo, T. B., Oh, M. J., You, B. G., Kwon, K. B., Chang, I. A., Yoon, J. H., et al. (2009). ERK1/2-mediated Schwann cell proliferation in the regenerating sciatic nerve by treadmill training. *J. Neurotrauma* 26, 1733–1744. doi: 10.1089/neu.2008.0711

Serpell, M. (2006). Anatomy, physiology and pharmacology of pain. *Surgery (Oxford)* 24, 350–353. doi: 10.1053/j.mpsur.2006.08.002

Shankarappa, S. A., Piedras-Rentería, E. S., and Stubbs, E. B. Jr. (2011). Forced-exercise delays neuropathic pain in experimental diabetes: effects on voltage-activated calcium channels. *J. Neurochem.* 118, 224–236. doi: 10.1111/j.1471-4159.2011.07302.x

Sharma, N. K., Ryals, J. M., Gajewski, B. J., and Wright, D. E. (2010). Aerobic exercise alters analgesia and neurotrophin-3 synthesis in an animal model of chronic widespread pain. *Phys. Ther.* 90, 714–725. doi: 10.2522/ptj.20090168

Shen, J., Fox, L. E., and Cheng, J. (2013). Swim therapy reduces mechanical allodynia and thermal hyperalgesia induced by chronic constriction nerve injury in rats. *Pain Med.* 14, 516–525. doi: 10.1111/pme.12057

Sigal, R. J., Kenny, G. P., Boule, N. G., Wells, G. A., Prud'homme, D., Fortier, M., et al. (2007). Effects of aerobic training, resistance training, or both on glycemic control in type 2 diabetes: a randomized trial. *Ann. Intern. Med.* 147, 357–369. doi: 10.7326/0003-4819-147-6-200709180-00005

Singleton, J. R., Marcus, R. L., Jackson, J. E., Lessard, M. K., Graham, T. E., and Smith, A. G. (2014). Exercise increases cutaneous nerve density in diabetic patients without neuropathy. *Ann. Clin. Transl. Neurol.* 1, 844–849. doi: 10.1002/acn3.125

Singleton, J. R., Marcus, R. L., Lessard, M. K., Jackson, J. E., and Smith, A. G. (2015). Supervised exercise improves cutaneous reinnervation capacity in metabolic syndrome patients. *Ann. Neurol.* 77, 146–153. doi: 10.1002/ana.24310

Sluka, K. A., O'Donnell, J. M., Danielson, J., and Rasmussen, L. A. (2013). Regular physical activity prevents development of chronic pain and activation of central neurons. *J. Appl. Physiol.* 114, 725–733. doi: 10.1152/japplphysiol.01317.2012

Smith, A. G., Russell, J., Feldman, E. L., Goldstein, J., Peltier, A., Smith, S., et al. (2006). Lifestyle intervention for pre-diabetic neuropathy. *Diabetes Care* 29, 1294–1299. doi: 10.2337/dc06-0224

Smith, S. A., Leal, A. K., Williams, M. A., Murphy, M. N., Mitchell, J. H., and Garry, M. G. (2010). The TRPv1 receptor is a mediator of the exercise pressor reflex in rats. *J. Physiol.* 588(Pt 7), 1179–1189. doi: 10.1113/jphysiol.2009.184952

Song, C. H., Petrofsky, J. S., Lee, S. W., Lee, K. J., and Yim, J. E. (2011). Effects of an exercise program on balance and trunk proprioception in older adults with diabetic neuropathies. *Diabetes Technol. Ther.* 13, 803–811. doi: 10.1089/dia.2011.0036

Stagg, N. J., Mata, H. P., Ibrahim, M. M., Henriksen, E. J., Porreca, F., Vanderah, T. W., et al. (2011). Regular exercise reverses sensory hypersensitivity in a rat neuropathic pain model: role of endogenous opioids. *Anesthesiology* 114, 940–948. doi: 10.1097/ALN.0b013e318210f880

Starkie, R., Ostrowski, S. R., Jauffred, S., Febbraio, M., and Pedersen, B. K. (2003). Exercise and IL-6 infusion inhibit endotoxin-induced TNF-alpha production in humans. *FASEB J.* 17, 884–886. doi: 10.1096/fj.02-0670fje

Streckmann, F., Kneis, S., Leifert, J. A., Baumann, F. T., Kleber, M., Ihorst, G., et al. (2014). Exercise program improves therapy-related side-effects and quality of life in lymphoma patients undergoing therapy. *Ann. Oncol.* 25, 493–499. doi: 10.1093/annonc/mdt568

Tesfaye, S., Boulton, A. J., and Dickenson, A. H. (2013). Mechanisms and management of diabetic painful distal symmetrical polyneuropathy. *Diabetes Care* 36, 2456–2465. doi: 10.2337/dc12-1964

Tsuchimochi, H., Yamauchi, K., McCord, J. L., and Kaufman, M. P. (2011). Blockade of acid sensing ion channels attenuates the augmented exercise pressor reflex in rats with chronic femoral artery occlusion. *J. Physiol.* 589(Pt 24), 6173–6189. doi: 10.1113/jphysiol.2011.217851

Tsukui, S., Kanda, T., Nara, M., Nishino, M., Kondo, T., and Kobayashi, I. (2000). Moderate-intensity regular exercise decreases serum tumor necrosis factor-alpha and HbA1c levels in healthy women. *Int. J. Obes. Relat. Metab. Disord.* 24, 1207–1211. doi: 10.1038/sj.ijo.0801373

Vaegter, H. B., Handberg, G., and Graven-Nielsen, T. (2014). Similarities between exercise-induced hypoalgesia and conditioned pain modulation in humans. *Pain* 155, 158–167. doi: 10.1016/j.pain.2013.09.023

Vaegter, H. B., Handberg, G., and Graven-Nielsen, T. (2015a). Isometric exercises reduce temporal summation of pressure pain in humans. *Eur. J. Pain* 19, 973–983. doi: 10.1002/ejp.623

Vaegter, H. B., Handberg, G., Jørgensen, M. N., Kinly, A., and Graven-Nielsen, T. (2015b). Aerobic exercise and cold pressor test induce hypoalgesia in active and inactive men and women. *Pain Med.* 16, 923–933. doi: 10.1111/pme.12641

Wang, H. J., Li, Y. L., Zhang, L. B., Zucker, I. H., Gao, L., Zimmerman, M. C., et al. (2011). Endogenous reactive oxygen species modulates voltage-gated sodium channels in dorsal root ganglia of rats. *J. Appl. Physiol.* 110, 1439–1447. doi: 10.1152/japplphysiol.01409.2010

Wang, J., Song, H., Tang, X., Yang, Y., Vieira, V. J., Niu, Y., et al. (2012). Effect of exercise training intensity on murine T-regulatory cells and vaccination response. *Scand. J. Med. Sci. Sports* 22, 643–652. doi: 10.1111/j.1600-0838.2010.01288.x

Wei, X. H., Na, X. D., Liao, G. J., Chen, Q. Y., Cui, Y., Chen, F. Y., et al. (2013). The up-regulation of IL-6 in DRG and spinal dorsal horn contributes to neuropathic pain following L5 ventral root transection. *Exp. Neurol.* 241, 159–168. doi: 10.1016/j.expneurol.2012.12.007

Woods, J. A., Wilund, K. R., Martin, S. A., and Kistler, B. M. (2012). Exercise, inflammation and aging. *Aging Dis.* 3, 130–140.

Xu, Z. Z., Zhang, L., Liu, T., Park, J. Y., Berta, T., Yang, R., et al. (2010). Resolvins RvE1 and RvD1 attenuate inflammatory pain via central and peripheral actions. *Nat. Med.* 16, 592–597, 591p following 597. doi: 10.1038/nm.2123

Yoon, H., Thakur, V., Isham, D., Fayad, M., and Chattopadhyay, M. (2015). Moderate exercise training attenuates inflammatory mediators in DRG of Type 1 diabetic rats. *Exp. Neurol.* 267, 107–114. doi: 10.1016/j.expneurol.2015.03.006

Zdziarski, L. A., Wasser, J. G., and Vincent, H. K. (2015). Chronic pain management in the obese patient: a focused review of key challenges and potential exercise solutions. *J. Pain Res.* 8, 63–77. doi: 10.2147/JPR.S55360

Zhang, X., Zhang, Y., Gao, X., Wu, J., Jiao, X., Zhao, J., et al. (2014). Investigating the role of backward walking therapy in alleviating plantar pressure of patients with diabetic peripheral neuropathy. *Arch. Phys. Med. Rehabil.* 95, 832–839. doi: 10.1016/j.apmr.2014.01.003

Zhuo, M., Wu, G., and Wu, L. J. (2011). Neuronal and microglial mechanisms of neuropathic pain. *Mol. Brain* 4, 31. doi: 10.1186/1756-6606-4-31

Zilliox, L., and Russell, J. W. (2011). Treatment of diabetic sensory polyneuropathy. *Curr. Treat. Options Neurol.* 13, 143–159. doi: 10.1007/s11940-011-0113-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

OPEN ACCESS Freely available online

PLOS ONE

Longitudinal Associations between Exercise and Pain in the General Population - The HUNT Pain Study

Tormod Landmark[1,2]*, Pål R. Romundstad[3], Petter C. Borchgrevink[2,4], Stein Kaasa[1,5], Ola Dale[4,6]

1 Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway, **2** National Competence Centre for Complex Symptom Disorders, St Olavs University Hospital, Trondheim, Norway, **3** Department of Public Health, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway, **4** Department of Circulation and Medical Imaging, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway, **5** Department of Oncology, St Olavs University Hospital, Trondheim, Norway, **6** Department of Anaesthesia and Emergency Medicine, St Olavs University Hospital, Trondheim, Norway

Abstract

Background: Population-based studies have reported conflicting findings on the relationship between physical activity and pain, and most studies reporting a relationship are cross sectional. Temporal relationships are therefore difficult to infer and associations may be subject to confounding from a variety of other factors. The aim of the current study was to investigate the association between exercise and pain longitudinally and to use within subjects analyses to remove between subjects confounding.

Methods: In the population-based HUNT 3 study, participants reported both pain and level of exercise. A random sub-sample of 6419 participants was in addition invited to report their last week pain and exercise every three months over a 12 month period (five measurements in total). We used multilevel mixed effects linear regression analyses to prospectively estimate the association between regular levels of exercise (measured in HUNT 3) and subsequent longitudinal reporting of pain. We also estimated within-subjects associations (i.e. the variation in pain as a function of variation in exercise, over time, within individuals) to avoid confounding from between subject factors.

Results: Among those invited to participate (N = 6419), 4219 subjects returned at least two questionnaires. Compared with subjects who reported no or light exercise, those who reported moderate levels of exercise or more at baseline, reported less pain in repeated measures over a 12 month period in analyses adjusted for age, sex,education and smoking. Adjusting for baseline level of pain distinctly attenuated the findings. Within subjects, an increase in exercise was accompanied by a concurrent reduction in intensity of pain. However, we found no indication that exercise level at one occasion was related to pain reporting three months later.

Conclusion: This longitudinal population-based study indicates that exercise is associated with lower level of pain and that this association is close in time.

Citation: Landmark T, Romundstad PR, Borchgrevink PC, Kaasa S, Dale O (2013) Longitudinal Associations between Exercise and Pain in the General Population - The HUNT Pain Study. PLoS ONE 8(6): e65279. doi:10.1371/journal.pone.0065279

Editor: Sam Eldabe, The James Cook University Hospital, United Kingdom

Received January 21, 2013; **Accepted** April 23, 2013; **Published** June 12, 2013

Funding: This work has been funded by the Research Council of Norway. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: tormod.landmark@ntnu.no

Introduction

Pain complaints are common and costly. The prevalence of current pain ranges from 27% to 49% [1,2], and the prevalence of chronic pain ranges from 11% to 64% [3–5]. Common pain conditions are major reasons for work related disability and for lost productivity in the work force [6,7]. The health care expenditures among subjects with common pain complaints have been estimated to be more than twice as high as for those without pain complaints, and they seem to continue to escalate [8–10]. Moreover, pain is associated with a substantial reduction in self reported health and functioning [11,12]. The best way of managing this public health problem is uncertain [13]. However, promoting a healthy lifestyle in the whole population may have beneficial effects on the prevention of pain complaints and its consequences [14].

Clinical studies have shown that exercise may relieve pain among patients with fibromyalgia and chronic low back pain [15,16] and prevent the recurrence of low back pain after treatment [17]. However, there is conflicting evidence whether exercise relates to the occurrence of pain in the general population [18–21]. Results are difficult to compare due to high variability in the definitions and measurements of both activity and pain and differences in study design and population. It has been suggested that significant associations may be hidden when measures are dichotomized into active vs. inactive [22], and that physical activity may be related to the severity of pain once established [23]. In a recent study, we showed that both frequency, duration and the intensity of exercise were indepen-

dently associated with a lower prevalence of chronic pain of at least moderate intensity in the general Norwegian population [24]. The cross-sectional nature of these findings limits their use in interpreting the relationship since low levels of exercise may be both a risk and a consequence of pain. Moreover, chronic pain is determined by multiple causal chains involving biological, psychological and social risk factors which may interact with or be associated with physical activity. Socioeconomic status, occupation, lifestyle and genetic makeup are factors that remain stable over time and may confound the relationship between exercise and pain. A confounder may, however, also vary across time. For example, variation in sleep, mood, or injuries may explain variations in both level of exercise and pain across time, within individuals.

In the current longitudinal population-based study, we used two separate analytical strategies to investigate the relationship between exercise and pain. First, we prospectively studied whether the baseline level of regular exercise was associated with the level of pain during 12 months of follow up. Second, we estimated the association between exercise and pain within individuals over time. When investigating the association within subjects, each individual serves as its own control and the estimates are not subject to confounding related to factors that remain stable within individuals (such as sex, socioeconomic status, occupation, genetic makeup, presence of chronic disease etc.).

Materials and Methods

Study Population

The basis for the present study is the Nord-Trøndelag Health Study (the HUNT study) conducted in the county of Nord-Trøndelag in Norway. The HUNT study consists of three cross-sectional surveys (HUNT 1, 1985–1987, HUNT 2, 1995–1997 and HUNT 3, 2006–2008). All inhabitants in Nord-Trøndelag aged 20 or more (N = 94194) was invited to participate in the HUNT 3 study. A total of 50839 (54%) participated. The response rate was higher among women (58%) than men (50%) and lowest among the youngest age groups (31% and 42% for the age groups 20–29 and 30–39 years, respectively). The study population is stable with sex and age distributions similar to the average of Norway, but with somewhat lower levels of education and income compared to national averages. The county is mostly rural and sparsely populated [25].

Participants and Procedure

A random sample of 6419 HUNT 3 participants in two municipalities (Levanger and Verdal) was mailed a questionnaire and invited to participation in the current project, which main focus is on physical activity and pain. Questionnaires were mailed every three months for the following 12 months (totally five questionnaires) to those agreeing to participate (n = 4782). Reminders were mailed to non-responders together with a copy of the questionnaire after one month. If the reminder was not returned, but the subjects had not actively withdrawn from the study, no new questionnaires were mailed until the fifth mailing at 12 months follow up.

The study was approved by the Regional Committee for Medical and Health Research Ethics Central-Norway and the Norwegian Data Inspectorate.

Questionnaire

The HUNT 3 questionnaire included three questions regarding exercise during the past year; the average number of times exercising per week *(never, less than once, once a week, 2–3 times per week or almost every day)*, the average minutes each time *(less than 15 minutes, 16–30 minutes, 30–60 minutes or more than 60 minutes)* and average intensity each time (*easy, without breaking a sweat or losing breath, lose breath and brake into sweat or near exhaustion*). The questions have shown acceptable test-retest reliability with kappa values ranging from 0.52 to 0.77 and significant correlations with VO_{2max} (ranging from 0.31 for duration) to (0.43 for frequency) in adult males [26]. In a previous HUNT 3 study [24], we showed that association between frequency of exercise and prevalence of chronic pain was u-shaped among participants in working age, whereas the association between intensity of exercise and chronic pain was linear. The associations were stronger among those above working age (65 years or more) and linear in shape. To account for the unique contribution of all three dimensions (frequency, duration and intensity) of exercise, and the divergence from linearity in the association with chronic pain, we constructed a variable as follows: Those who reported no activity, light intensity activity and activity for less than 30 minutes were defined as reference group. Those reporting moderate to vigorous physical activity of 30 minutes or more were divided into two groups; those who reported 1–3 times per week, and those who reported nearly every day.

The HUNT 3 questionnaire included one question regarding pain intensity: "How much bodily pain have you had during the past four weeks?" This is a six point verbal rating scale including the response options: None, very mild, mild, moderate, severe or very severe. It has been extensively used, among others in the various versions of the SF-36 health survey [27] and is validated as a single item measure as part of the SF-8 health surveys [28].

In the one year follow up study, each of the five mailings included the one week version of the SF-8 bodily pain scale [28]. The scale was transformed according to the scoring procedures by assigning a new value to each response category based on the US SF-36 norm data [28]. This ensured a mean score close to 50 and a standard deviation close 10 in the US normative data.

Recreational exercise was defined in the follow up questionnaires by giving the following examples: going for a walk, skiing, swimming, exercise or sports. The Borg ratings of perceived exertion (RPE) scale [29] was used as an index of exercise intensity with the following instruction: "On a scale from 6 to 20, how hard is the activity that you usually do when you exercise? (Take an average from the last week). The Borg RPE scale has been shown to be a valid measure of exercise intensity in various populations [30]. In a recent investigation using the same instruction in another subsample form the HUNT 3 study, the scale corresponded well with Peak oxygen uptake (VO_{2peak}) measured during an exercise test [31]. Responders were also asked how often they had engaged in recreational exercise during the last week, and the average duration each time. For the purpose of the current study, participants reporting no exercise or exercise of less than 15 minutes were assigned the value 5 and included in the Borg scale. This gave a variable ranging from 5 (no exercise) to 20 (very, very hard).

Information on the highest attained level of education was obtained from the National Education database (NUDB). Educational attainment was classified into three levels; primary, secondary and tertiary.

Statistical Analyses

To investigate longitudinal associations between exercise and pain, multilevel mixed effects linear regression analyses were performed using the xtmixed function in Stata version 11.0 for Windows (Stata Corporation, College Station, Texas). In longitudinal studies, mixed models accounts for the dependency of

Table 1. Characteristics of the study sample at each follow up (T1–T5) and compared to the entire HUNT 3 population.

	Study sample					Hunt 3
	T1 n = 4782	T2 n = 4219	T3 n = 3926	T4 n = 3791	T5 n = 4140	n = 50827
	%	%	%	%	%	%
Sex						
female	56.0	56.1	56.3	56.4	56.1	54.6
male	44.0	43.9	43.7	43.6	43.9	45.4
Age						
20–44 yrs	28.4	26.2	25.1	24.5	26.0	29.6
45–64 yrs	47.4	48.4	48.9	49.1	49.1	43.6
65 yrs or more	24.3	25.4	26.0	26.5	24.9	26.8
Education						
Primary	17.2	16.9	16.7	16.8	16.7	21.2
Secondary	49.7	49.5	49.7	49.6	49.7	52.7
Tertiary	33.2	33.6	33.7	33.7	33.6	26.1

doi:10.1371/journal.pone.0065279.t001

observations within subjects by inducing subject specific (random) effects into the model. Missing data are handled by using all available data for each person.

First, we prospectively studied the association of exercise measured at baseline in the HUNT 3 study with the reporting of pain in the following five subsequent measurements. That is, we estimated the difference in pain during the 12 month follow up period by different levels of exercise at baseline. The estimates from these analyses were based on the mixed effects, i.e. no attempts were made at disentangling the variation within subjects from the variation between subjects. The analyses were adjusted for sex, age, education, smoking and baseline level of pain.

Second, we used the repeated measurements in the 12 month follow up to investigate within subjects associations. To disentangle the within subjects associations from the between subjects associations, two exercise variables were computed for each person: a mean score across all five measurement occasions and a deviation from the mean at each measurement occasion. The deviation scores were used to calculate the within subjects associations. These are longitudinal in that they estimate the variation in pain as a function of variation in exercise over time, within individuals. Most cross sectional and prospective analyses address research questions on a group level, such as: "Compared with individuals who report lower level of exercise, do individuals who report higher level of exercise report less pain?" Whereas within subjects analyses address questions on and individual level, such as: "Compared with time points when they report lower level of exercise, do individuals report less pain at time points when they report higher level of exercise?" In this way, subjects function as their own controls and the analyses have the advantage of not being subject to confounding by factors that remain constant over time, such as sex, socioeconomic status, genetic makeup and presence of chronic disease. In the primary model we studied whether change in exercise was associated with a simultaneous change in pain. We then investigated whether level of exercise at one occasion was associated with pain reporting three months later.

Likelihood ratio tests were used to evaluate the interactions between exercise and age and exercise and sex. Analyses were also carried out separately for each sex.

Results

Characteristics of the Participants

Of the 6419 subjects invited to participate in the HUNT pain study, 75% (n = 4782) responded to the baseline questionnaire (table 1). Among these, 56% were women, 28% were aged 20–44 year, 47% were 45 to 64 years and 24% were 65 years or older. One third of the participants had tertiary education, 50% had secondary education, and 17% had only primary education. Compared to the HUNT 3 population, the sex distribution was similar, whereas the proportion of middle aged and individuals with higher education were higher in the HUNT pain study. Less than 15% of the participants were lost to 12 months follow up, and attrition was neither associated with sex nor education. The proportion of subjects in the youngest age group declined somewhat throughout the follow up period. The mean pain score in the SF-8 scale (49.4; sd = 9.6) and mean exercise score on the Borg scale (11.4; sd = 3.9) were similar throughout the five occasions, indicating no attrition due to the primary study variables. Intraclass correlation coefficients (ICC) for exercise was 0.55 (95% CI 0.54, 0.57) and for pain it was 0.66 (95% CI = 0.65, 0.67). Thus, 45% of the variance in exercise and 34% of the variance in pain was accounted for by within-subject variation, respectively. This implies that the measures were quite stable, and reduces the power to detect significant within subject associations.

Prospective Associations between Regular Exercise and Subsequent Pain

In the HUNT 3 study, subjects reported their level of exercise on an average week during the past year. Compared to those not reporting regular exercise in HUNT 3, those reporting at least moderate exercise 1–3 times a week reported 1.12 points less pain on the SF-8 scale (95% CI: 0.60, 1.63) during the 12 months of follow up in analyses adjusted for sex, age, education and smoking (table 2). The difference remained significant although attenuated when additionally adjusted for baseline level of pain. A similar but weaker association was seen between reports of moderate or hard exercises almost every day and subsequent level of pain. Significant interactions were seen between exercise and sex (p-value interaction$<$0.001). Stratified analyses revealed a stronger association between exercise of 1–3 times a week of at least

Table 2. Prospective associations between exercise* reported in the HUNT 3 study and subsequent reporting of pain† measured every third month during a 12 month follow up period of the HUNT pain study.

	Total sample		Women		Men	
	Coefficient	**95% CI**	**Coefficient**	**95% CI**	**Coefficient**	**95% CI**
Unadjusted						
None exercise	0	Ref	0	Ref	0	Ref
1–3 times/week	2.15	1.63, 2.67	1.95	1.25, 2.66	2.26	1.51, 3.00
≥4 times/week	1.53	0.69, 2.37	1.83	0.70, 2.96	1.13	−0.10, 2.36
Adjustment‡						
None exercise	0	Ref	0	Ref	0	Ref
1–3 times/week	1.12	0.60, 1.63	0.78	0.66, 1.50	1.47	0.72, 2.21
≥4 times/week	0.78	0.03, 1.60	0.83	−0.28, 1.94	0.63	−0.58, 1.85
Adjustment§						
None exercise	0	Ref	0	Ref	0	Ref
1–3 times/week	0.42	0.23, 0.82	0.10	−0.42, 0.64	0.81	0.21, 1.40
≥4 times/week	0.32	−0.32, 0.96	0.27	−0.58, 1.11	0.34	−0.64, 1.32

*Average number of times per week during the last year of at least 30 minutes and either lose breath and brake into sweat or near exhaustion.
†SF-8 Bodily pain scale.
‡Adjusted for age, education, smoking and sex as appropriate.
§Further adjustment for baseline pain.
doi:10.1371/journal.pone.0065279.t002

moderate exercise and subsequent pain among men than women in the adjusted models (table 2). After adjustment for baseline pain the association was no longer significant for women. Significant interactions were also seen between exercise and age, revealing a stronger relationship when age increased (p-value interaction <0.001).

Within Subjects Associations between Exercise and Pain

Within subjects associations were considered in two different temporal models (table 3). In the first model we investigated whether exercise intensity were associated with concurrently reported pain intensity (during the past week). A robust association was seen, indicating that a one point change in exercise was associated with a concurrent 0.25 point improvement in pain (95% CI: 0.21, 0.28). Thus, a 7 point increase on the scale, which would indicate a change from no to moderate exercise, would account for a simultaneous 1.75 points improvement in pain on the SF-8 scale.

A significant interaction was seen between exercise and sex ($p<0.001$). However, separate analyses revealed quite similar findings for men and women (table 3). Interaction between exercise and age ($p<0.001$) suggested a stronger association with increasing age.

In the second model, exercise at one occasion was not related to pain reported at a subsequent occasion (table 3).

Discussion

In this longitudinal population-based study, regular exercise reported at baseline was associated with less pain in repeated measures over a subsequent 12 month period. However, the associations were substantially attenuated when adjusting for baseline level of pain and remained significant only for men. The within subjects analyses revealed a significant concurrent association between exercise and pain. However, no association was seen between exercise at one occasion and pain measured three months later.

Most previous population-based studies have failed to show an association between physical activity and pain [18–21]. It has been suggested that significant associations may be hidden when measures of physical activity are dichotomized [22] and when severity of pain is not accounted for [23]. In a recent cross sectional study we found that exercise was associated with a lower prevalence of chronic pain of at least moderate intensity, especially among older subjects [24]. In a previous HUNT study lower level

Table 3. Within subjects associations between exercise* and pain† at the same time points (concurrent) and after three months (subsequent).

	Total sample		Women		Men	
	Coefficient	**95% CI**	**Coefficient**	**95% CI**	**Coefficient**	**95% CI**
Concurrent	0.25	0.21, 0.28	0.29	0.25, 0.34	0.19	0.14, 0.24
Subsequent	0.00	−0.05, 0.0)	0.01	−0.05, 0.07	−0.02	−0.08, 0.03

*Borg scale of Perceived exertion; how hard is the activity that you usually do when you exercise? (Take an average from the last week) 5 = no exercise; 20 = very, very hard.
†SF-8 Bodily pain scale.
doi:10.1371/journal.pone.0065279.t003

of physical activity at baseline was associated with higher prevalence of widespread chronic pain 11 years later [32]. However, this study failed to account for baseline pain. It is difficult to infer any temporal relationship between activity and pain from these studies since pain might have caused reduced physical activity. One previous longitudinal study showed that physical activity was associated with less pain on the SF-36 scale measured repeatedly during three years of follow up among midlife women not reporting moderate or severe pain at baseline [33].

In the within subject analyses we found evidence for a relationship between exercise and pain that is close in time. That is, subjects reported less pain at times when they reported more exercise, whereas exercise was not related to a subsequent change in pain within individuals. This close relation in time may indicate an important reciprocity of the relationship between exercise and pain. That is, a lower level of exercise may be both a risk for and a consequence of pain. This is of particular importance when interpreting cross sectional studies of the association. However, it also shows the importance of considering baseline level of pain in prospective studies. When adjusting for baseline pain in our cohort, exercise was only related to subsequent level of pain among men. This sex difference was not as evident in the unadjusted model and may indicate that a bidirectional relationship between exercise and pain is stronger among women.

The current findings show that changes in pain might be related to exercise, in particular among men, and that the relationship is independent of time invariant factors that differ between subjects such as other lifestyle factors, sex, socioeconomic status, genetic makeup, presence of chronic disease, occupation etc.

It is difficult to draw firm conclusions about the importance of these findings. Even though we found statistically significant associations, the effect sizes were small and far from what can be regarded as clinically significant [34]. However, considering the high prevalence of chronic pain [5], even low effect sizes could have public health significance. That is, if we could increase the level of physical activity in the population, chronic pain could potentially be prevented in a noticeable number of subjects. Future studies should use long term follow up with the aim at identifying the proportion of cases with significant chronic pain that might be prevented by regular exercise. Moreover, the relationship is likely to be stronger in certain clinical populations than in the population at large [15,16]. Identifying subgroups that may benefit more from exercise interventions on a population level may therefore be an objective for future investigations.

Some considerations regarding the statistical analyses need to be mentioned. When modelling within subjects associations, the factors of interest must vary within individuals. In the current study both pain and exercise were relatively stable. This may have reduced our power to study longitudinal associations as only those individuals with time related variations contributed to the within subject estimates. Still, the number of participants was substantial and the model was able to detect significant relations. Although these analyses removed the confounding of time invariant factors, factors that may vary within individuals, such as injuries, mood, sleep and anxiety could have confounded the associations. However, these factors may be part of causal chains between physical activity and pain, and including them as time-varying covariates in the analysis would require quite complex theoretical models of the relationships [35]. We assumed a liner relationship after having plotted the SF-8 pain scores against the exercise scores in a cross sectional dataset. A linear association between intensity of exercise and prevalence of chronic pain was also reported in a previous cross sectional study [24]. In that study, the frequency of exercise was not linearly associated with the prevalence of chronic pain, however, and this was also evident in our prospective analyses which indicated a weaker association with pain as the frequency of exercise exceeded 3 times a week. In the prospective analyses we adjusted for baseline pain. In some cases, when there is considerably measurement error, adjustment for baseline scores of the outcome variable might cause inflation of the association [36]. Such adjustments should therefore be done with caution. Our adjustments, on the other hand, led to an attenuation of the associations, which was in accordance to what would be expected.

In the time lag model, we did not find evidence for an association between exercise during one week and subsequent changes in pain. One possible explanation for this might be that the three month intervals between measurements were too long. That is, exercise during one week might have been related to pain during the next week, although it was not related to pain during one week three months later. However, the lack of evidence for an association in the time lag model corresponds with the attenuation of the estimates in the prospective analyses after adjustments for baseline pain, indicating that only a limited change in pain was seen over the one year course as a function of regular exercise reported at baseline.

We had to rely on self report measures. In terms of pain there is no alternative as pain per definition is a subjective experience. Even though the verbal rating scale we used to assess pain is well validated [28], it is unlikely to possess ratio qualities, i.e. equal intervals between the categories. Nevertheless, it has been increasingly recognised that parametric statistics, such as regression analyses, are valid for ordinal pain scales, at least those containing 5 categories or more [37]. Objective measures of physical fitness are likely to give more valid results than self reports of physical activity [38,39]. However, the repetition of measurements at five occasions during one year in a large population-based sample would require extensive financial resources and even though the activity may change, measures if fitness would not change in the same degree. We therefore used the Borg Scale of perceived exertion which gives detailed information on exercise intensity. The scale is well validated and, self-reports of usual exercise intensity is independently associated with VO_{2peak} in the general population [31].

Conclusion

This longitudinal population-based study gives robust evidence for an association between exercise and pain. However, the association was close in time and weak, and its importance remains open to debate.

Acknowledgments

The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine, Norwegian University of Science and Technology NTNU), Nord-Trøndelag County Council and The Norwegian Institute of Public Health. We thank all the participants, and the staff who contributed in the data collection: Berit Bjelkåsen, Vanja Strømsnes, Cinzia Marini, Ingunn Johansen, & Aleksandra Szczepanek. A special thanks to Karin Tulluan for her participation in the data collection and in administering the database.

Author Contributions

Conceived and designed the experiments: TL PRR PCB SK OD. Performed the experiments: TL PRR PCB SK OD. Analyzed the data: TL PRR. Wrote the paper: TL PRR OD. Critically commented on the mansucript: PRR OD PCB SK. Read and approved the final manuscript: PRR OD PCB SK.

References

1. Krueger AB, Stone AA (2008) Assessment of pain: a community-based diary survey in the USA. The Lancet 371: 1519–1525.
2. Gerdle B, Bjork J, Henriksson C, Bengtsson A (2004) Prevalence of current and chronic pain and their influences upon work and healthcare-seeking: a population study. J Rheumatology 31: 1399–1406.
3. Ospina M, Harstall C (2002) Prevalence of chronic pain: an overview. Alberta Heritage Foundation for Medical Research, Health Technology Assesment, 28 th Report. Edmonton, Canada.
4. Watkins EA, Wollan PC, Melton LJ 3rd, Yawn BP (2008) A population in pain: report from the Olmsted County health study. Pain Med 9: 166–174.
5. Landmark T, Romundstad P, Dale O, Borchgrevink PC, Kaasa S (2012) Estimating the prevalence of chronic pain: Validation of recall against longitudinal reporting (the HUNT pain study). Pain 153: 1368–1373.
6. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R (2003) Lost productive time and cost due to common pain conditions in the US workforce. JAMA 290: 2443–2454.
7. Gjesdal S, Ringdal PR, Haug K, Maeland JG (2004) Predictors of disability pension in long-term sickness absence: results from a population-based and prospective study in Norway 1994–1999. Eur J Public Health 14: 398–405.
8. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, et al. (2008) Expenditures and health status among adults with back and neck problems. JAMA 299: 656–664.
9. Eriksen J, Sjogren P, Ekholm O, Rasmussen NK (2004) Health care utilisation among individuals reporting long-term pain: an epidemiological study based on Danish National Health Surveys. Eur J Pain 8: 517–523.
10. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA (2009) Comprehensive review of epidemiology, scope, and impact of spinal pain. Pain Physician 12: E35–70.
11. Mantyselka PT, Turunen JH, Ahonen RS, Kumpusalo EA (2003) Chronic pain and poor self-rated health. JAMA 290: 2435–2442.
12. Sprangers MAG, de Regt EB, Andries F, van Agt HME, Bijl RV, et al. (2000) Which chronic conditions are associated with better or poorer quality of life? J Clin Epidemiol 53: 895–907.
13. Dagenais S, Roffey DM, Wai EK, Haldeman S, Caro J (2009) Can cost utility evaluations inform decision making about interventions for low back pain? Spine J 9: 944–957.
14. Bergman S (2007) Public health perspective–how to improve the musculoskeletal health of the population. Best Pract Res Clin Rheumatol 21: 191–204.
15. Hayden JA, van Tulder MW, Malmivaara A, Koes BW (2005) Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev: CD000335.
16. Busch AJ, Barber KA, Overend TJ, Peloso PM, Schachter CL (2007) Exercise for treating fibromyalgia syndrome. Cochrane Database Syst Rev: CD003786.
17. Choi BK, Verbeek JH, Tam WW, Jiang JY (2010) Exercises for prevention of recurrences of low-back pain. Cochrane Database Syst Rev: CD006555.
18. Heneweer H, Staes F, Aufdemkampe G, van Rijn M, Vanhees L (2011) Physical activity and low back pain: a systematic review of recent literature. Eur Spine J 20: 826–845.
19. Hildebrandt VH, Bongers PM, Dul J, van Dijk FJ, Kemper HC (2000) The relationship between leisure time, physical activities and musculoskeletal symptoms and disability in worker populations. Int Arch Occup Environ Health 73: 507–518.
20. Hoogendoorn WE, van Poppel MN, Bongers PM, Koes BW, Bouter LM (1999) Physical load during work and leisure time as risk factors for back pain. Scand J Work Environ Health 25: 387–403.
21. Bell JA, Burnett A (2009) Exercise for the primary, secondary and tertiary prevention of low back pain in the workplace: a systematic review. J Occup Rehabil 19: 8–24.
22. Heneweer H, Vanhees L, Picavet HSJ (2009) Physical activity and low back pain: A U-shaped relation? Pain 143: 21–25.
23. Jacob T, Baras M, Zeev A, Epstein L (2004) Physical activities and low back pain: a community-based study. Med Sci Sports Exerc 36: 9–15.
24. Landmark T, Romundstad P, Borchgrevink PC, Kaasa S, Dale O (2011) Associations between recreational exercise and chronic pain in the general population: Evidence from the HUNT 3 study. Pain 152: 2241–2247.
25. Holmen J, Midthjell K, Krüger Ø, Langhammer A, Holmen TL, et al. (2003) The Nord-Trøndelag Health Study 1995–1997 Norsk Epidemiologi 13: 19–32.
26. Kurtze N, Rangul V, Hustvedt BE, Flanders WD (2008) Reliability and validity of self-reported physical activity in the Nord-Trondelag Health Study: HUNT 1. Scand J Public Health 36: 52–61.
27. Ware JE Jr., Sherbourne CD (1992) The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. Medical Care 30: 473–483.
28. Ware JE, Kosinski M, Dewey JE, Gandek B (2001) How to score and interpret single item health status measures: A manual for users of the SF-8 health Survey. Lincoln: Quality Metric Incorporated.
29. Borg G (1998) Borg's perceived exertion and pain scales. Champaign, Ill. Human Kinetics. viii, 104 s., [104] bl. p.
30. Chen MJ, Fan X, Moe ST (2002) Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. J Sports Sci 20: 873–899.
31. Nes BM, Janszky I, Aspenes ST, Bertheussen GF, Vatten LJ, et al. (2012) Exercise Patterns and Peak Oxygen Uptake in a Healthy Population: The HUNT-Study. Med Sci Sports Exerc 44: 1881–1889.
32. Holth HS, Werpen HK, Zwart JA, Hagen K (2008) Physical inactivity is associated with chronic musculoskeletal complaints 11 years later: results from the Nord-Trondelag Health Study. BMC Musculoskelet Disord 9: 159.
33. Dugan SA, Everson-Rose SA, Karavolos K, Sternfeld B, Wesley D, et al. (2009) The impact of physical activity level on SF-36 role-physical and bodily pain indices in midlife women. J Phys Act Health 6: 33–42.
34. Sloan JA, Cella D, Hays RD (2005) Clinical significance of patient-reported questionnaire data: another step toward consensus. Journal of Clinical Epidemiology 58: 1217–1219.
35. Kraemer H, FauStice E, D Offord, Kupfer D (2001) How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. Am J Psychiatry 158: 848–856.
36. Glymour MM, Weuve J, Berkman LF, Kawachi I, Robins JM (2005) When Is Baseline Adjustment Useful in Analyses of Change? An Example with Education and Cognitive Change. Am J Epidemiol 162: 267–278.
37. Von Korff M, Jensen MP, Karoly P (2000) Assessing global pain severity by self-report in clinical and health services research. Spine 25: 3140–3151.
38. Shephard RJ (2003) Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med 37: 197–206.
39. Heneweer H, Picavet HS, Staes F, Kiers H, Vanhees L (2012) Physical fitness, rather than self-reported physical activities, is more strongly associated with low back pain: evidence from a working population. Eur Spine J 21: 1265–1272.

Central modulation of exercise-induced muscle pain in humans

Chester A. Ray[1] and Jason R. Carter[2]

[1]*Heart and Vascular Institute, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA*
[2]*Department of Exercise Science, Health and Physical Education, Michigan Technological University, Houghton, MI 49931, USA*

The purpose of the current study was to determine if exercise-induced muscle pain is modulated by central neural mechanisms (i.e. higher brain systems). Ratings of muscle pain perception (MPP) and perceived exertion (RPE), muscle sympathetic nerve activity (MSNA), arterial pressure, and heart rate were measured during fatiguing isometric handgrip (IHG) at 30% maximum voluntary contraction and postexercise muscle ischaemia (PEMI). The exercise trial was performed twice, before and after administration of naloxone (16 mg intravenous; $n = 9$) and codeine (60 mg oral; $n = 7$). All measured variables increased with exercise duration. During the control trial in all subjects ($n = 16$), MPP significantly increased during PEMI above ratings reported during IHG (6.6 ± 0.8 to 9.5 ± 1.0; $P < 0.01$). However, MSNA did not significantly change compared with IHG (7 ± 1 to 7 ± 1 bursts $(15\ s)^{-1}$), whereas mean arterial blood pressure was slightly reduced (104 ± 4 to 100 ± 3 mmHg; $P < 0.05$) and heart rate returned to baseline values during PEMI (83 ± 3 to 67 ± 2 beats min^{-1}; $P < 0.01$). These responses were not significantly altered by the administration of naloxone or codeine. There was no significant relation between arterial blood pressure and MSNA with MPP during either IHG or PEMI. A second study ($n = 8$) compared MPP during ischaemic IHG to MPP during PEMI. MPP was greater during PEMI as compared with ischaemic IHG. These findings suggest that central command modulates the perception of muscle pain during exercise. Furthermore, endogenous opioids, arterial blood pressure and MSNA do not appear to modulate acute exercise-induced muscle pain.

(Received 11 July 2007; accepted after revision 24 September 2007; first published online 11 October 2007)
Corresponding author C. A. Ray: Penn State College of Medicine, The Milton S. Hershey Medical Center, Heart & Vascular Institute H047, 500 University Drive, Hershey, PA 17033-2390, USA. Email: caray@psu.edu

Pain is an emotional and subjective experience that involves both peripheral and central mechanisms. Modulation of pain is a complex system in which processing can occur in both ascending and descending pathways. Nociceptors of the periphery sense pain and relay this perception of pain to the central nervous system via group III and IV afferent fibres (Besson, 1999; Millan, 2002). Nociceptive afferents synapse primarily in the dorsal horn of the spinal cord where the nociceptive signals are processed and transmitted to supraspinal brain areas (Millan, 2002). Several supraspinal sites have been implicated in nociceptive processing, but the most recognized are the hypothalamus, periaqueductal grey (PAG), rostral ventrolateral medulla (RVM) and dorsolateral pontomesencephalic tegmentum (DLPT).

Although central processing of pain has been extensively studied, one area that has received little attention is central modulation of exercise-induced pain in humans. Several studies indicate an analgesic effect during exercise, but the mechanisms underlying this phenomenon are poorly understood (Cook *et al.* 1997). In a previous study, Cook *et al.* (2000) examined the role of the endogenous opioid system on forearm muscle pain by recording muscle pain perception during dynamic handgrip after administration of either codeine (an opioid agonist), naltrexone (an opioid antagonist) or placebo. Ratings of muscle pain perception were not different among trials, indicating the endogenous opioid system does not alter muscle pain perception during exercise (Cook *et al.* 2000). However, the experimental design of this study by Cook *et al.* (2000) could not definitively assess if pain perception during exercise was centrally modulated by higher brain systems (i.e. central command) because endogenous opioid receptors are found on peripheral (group III and IV afferents) and central (PAG, RVM and DLPT) sites involved in pain processing (Millan, 2002). Furthermore, it has been demonstrated that central motor command can inhibit group III muscle afferent input to

 DOI: 10.1113/jphysiol.2007.140509

the dorsal horn (Degtyarenko & Kaufman, 2003). Thus, it is possible that central command may have interacted with afferent feedback from the muscle and the opioid system to modulate pain perception.

Therefore, the primary purpose of this study was to examine the effect of central command on muscle pain perception during exercise. Muscle pain perception was compared during isometric handgrip (IHG) and postexercise muscle ischaemia (PEMI) because IHG engages central command whereas PEMI does not. Central command affects both cardiovascular and ventilatory control during exercise (Williamson *et al.* 2006); thus we hypothesized that central command may also influence the perception of exercise-induced muscle pain. Specifically, it was hypothesized that perception of exercise-induced muscle pain would be augmented during PEMI when central command is minimal. A secondary purpose was to test the hypothesis that endogenous opioids alter central modulation of muscle pain. Our results suggest that central command attenuates muscle pain perception during exercise and that endogenous opioids, arterial blood pressure and MSNA do not appear to influence this central modulation of pain.

Methods

Subjects

Twenty-four healthy men and women (18 men and 6 women; age 25 ± 1 years, height 176 ± 2 cm, weight 77 ± 4 kg) volunteered to participate in the study. Subjects abstained from nicotine, alcohol and caffeine for a minimum of 8 h prior to the experiment. The Institutional Review Board at The Pennsylvania State University College of Medicine approved the study and the written informed consent form. All participants signed the informed consent form after verbal explanation of the testing procedures.

Experimental design

Study 1. Subjects performed two bouts of exercise. The first exercise bout was designated as the control trial because no drug intervention was performed. The second exercise bout was performed after administration of either naloxone ($n = 9$) or codeine ($n = 7$). Naloxone was infused intravenously (16 mg) over 2 min into the non-exercising arm 20 min after the control exercise bout. Another 2 min elapsed between final infusion and the start of the second exercise bout to allow systemic distribution of the drug. During the codeine trial, subjects received a 60 mg capsule of codeine immediately following the control exercise bout. After 1 h, the exercise protocol was repeated. The timing of codeine administration was based on previous reports that peak plasma concentrations occur ∼1 h after a single oral does of 60 mg (Quiding *et al.* 1986). Subjects were randomly assigned to naloxone and codeine groups, and both the investigator and the subjects were blinded with regard to the drug intervention until analysis of data was completed.

During each exercise bout, subjects performed IHG (30% maximum voluntary contraction) to fatigue, followed by 2 min of PEMI before (control) and after administration of either naloxone or codeine. Maximal voluntary contraction was established using the peak force generated from three maximal handgrip efforts. PEMI was induced 5 s prior to the cessation of exercise by inflating a blood pressure cuff on the arm to suprasystolic levels (240 mmHg). Each exercise trial began and ended with a 3 min baseline and recovery period. Forearm muscle pain and exertion ratings were obtained every 15 s of IHG and PEMI.

Study 2. To determine if cuff inflation during PEMI influenced pain perception, a second study ($n = 8$) was performed in which subjects performed IHG during muscle ischaemia induced by the same cuff compression used during PEMI. This ischaemic IHG was followed by 2 min of PEMI. IHG was performed until subjects reached fatigue or until a pain perception score of seven or greater was reported. This number was selected to match the level of pain reached by subjects in the first study. Forearm muscle pain and exertion ratings were obtained every 15 s of ischaemic IHG and PEMI.

Pain and exertion assessment

Forearm muscle pain perception was assessed using a category scale with ratio properties. The pain intensity scale ranged from 0 (no pain at all) to 10 (extremely intense pain, almost unbearable). If the subjective intensity increased above 10, the subject chose any number larger in proportion to 10 that described the proportional growth of the sensation. Prior work has provided evidence for the validity and reliability of this scale for quantifying naturally occurring muscle pain during exercise (Cook *et al.* 1997). Ratings of perceived exertion were assessed during and after exercise by using Borg's 6–20 category scale (Borg, 1978).

Measurements

Multifibre recordings of MSNA were made by inserting a tungsten microelectrode into the peroneal nerve at the head of the fibula of a resting leg. A reference electrode was inserted subcutaneously 2–3 cm from the recording electrode. Both electrodes were connected to a differential preamplifier, and then to an amplifier (total gain between 40 000 and 80 000) where the nerve signal was band-pass filtered (700–2000 Hz), and integrated (time constant, 0.1s) to obtain a mean voltage display of the nerve

Table 1. Preexercise baseline values during the control, naloxone, and codeine trials

Variable	Control ($n = 16$)	Naloxone ($n = 9$)	Codeine ($n = 7$)
MSNA (bursts (15 s)$^{-1}$)	4 ± 1	3 ± 1	4 ± 1
MSNA, total	191 ± 26	189 ± 25	193 ± 52
HR (beats min^{-1})	65 ± 2	65 ± 3	61 ± 3
SAP (mmHg)	120 ± 2	122 ± 4	120 ± 2
DAP (mmHg)	63 ± 2	63 ± 3	65 ± 2
MAP (mmHg)	79 ± 2	80 ± 3	80 ± 2
MPP (a.u.)	0 ± 0	0 ± 0	0 ± 0
RPE (a.u.)	6 ± 0	6 ± 0	6 ± 0

Values are mean ± S.E.M.; MSNA, muscle sympathetic nerve activity; HR, heart rate; SAP, systolic arterial pressure; DAP, diastolic arterial pressure; MAP, mean arterial pressure; MPP, muscle pain perception; RPE, rating of perceived exertion; a.u., arbitrary units. Baseline values for all three trials were not different from each other; $P > 0.05$. Subjects were randomly assigned to the naloxone and codeine groups.

activity. Satisfactory recordings of MSNA were defined by spontaneous, pulse synchronous bursts that increased during end-expiratory apnoea, and did not change during stroking of the skin or auditory stimulation (yell).

Arterial blood pressure and heart rate (HR) were recorded using a Finapres (Ohmeda, Louisville, CO, USA) positioned on the middle digit of the subject's non-exercising hand. A pneumograph bellows was wrapped around the subject's chest to monitor respiratory rate and to ensure subjects avoided a Valsalva manoeuver during IHG. Force output from handgrip and all other recorded variables were routed and recorded to an on-line computer (MacLab 8E, ADInstruments, Milford, MA, USA).

Data analysis

As control trials were not different in the naloxone and codeine groups, these results were combined for data presentation. The naloxone and codeine trials were analysed using a two-within-factor (intervention (placebo *versus* drug) × exercise bout) repeated analysis of variance. Pearson correlations were used to examine the relations between systolic arterial pressure, MSNA and muscle pain perception. Muscle pain perception was reported every 15 s, and the highest values reported during IHG and PEMI were used for data analysis. Significance was accepted at the $P < 0.05$ level. All data are presented as mean ± S.E.M.

Results

Study 1: Isometric handgrip

Preexercise baseline values before (control) and after administration of either naloxone or codeine are presented in Table 1. Naloxone and codeine did not change baseline values of MSNA, heart rate and arterial blood pressure. Subjects reported no muscle pain or perceived exertion during baseline.

Heart rate and mean arterial blood pressure (MAP) significantly increased during IHG for all trials (Fig. 1). During PEMI, MAP remained elevated from baseline, but MAP decreased slightly compared with IHG. Heart rate returned to baseline levels during PEMI. MSNA increased during IHG and PEMI for all trials, but increases in MSNA during IHG and PEMI were not different (Fig. 1).

Muscle pain perception significantly increased during IHG and PEMI with and without any drug intervention (Fig. 2). During PEMI, muscle pain perception was significantly greater than IHG values during the control and codeine trials ($P < 0.01$) and tended to increase in the naloxone trial ($P < 0.09$). Increases in muscle pain perception were not correlated to changes in arterial blood pressure during IHG and PEMI for any of the trials (Fig. 3). Similarly, muscle pain perception was not correlated to changes in MSNA during IHG (total activity, $R^2 = 0.01$; burst frequency, $R^2 = 0.16$) or PEMI (total activity, $R^2 = 0.0005$; burst frequency, $R^2 = 0.006$). Ratings of perceived exertion increased as a function of exercise duration and were not different between trials (peak value 19 ± 0 units).

Study 2: Ischaemic isometric handgrip

Heart rate (60 ± 2 to 78 ± 6 beats min^{-1}) and MAP (93 ± 2 to 113 ± 4 mmHg) increased during ischaemic IHG. During PEMI, MAP (111 ± 9 mmHg) remained elevated from baseline, but heart rate (64 ± 3 beats min^{-1}) returned to baseline levels. Muscle pain perception significantly increased during ischaemic IHG (5.9 ± 0.9 units), but increased even further during PEMI (8.4 ± 1.1 units, $P < 0.05$; Fig. 4).

Discussion

This study identifies three novel findings: (1) muscle pain perception increases during PEMI compared with IHG; (2) endogenous opioids do not modulate muscle

pain perception during either forearm exercise or muscle ischaemia; and (3) muscle pain perception during forearm exercise or PEMI is not correlated to changes in arterial blood pressure or MSNA. Since PEMI reduces central command but not muscle afferent feedback, our results suggest that central command attenuates muscle pain perception during exercise and thus serves as a modulator of acute exercise-induced muscle pain.

During exercise, several reflexes are simultaneously engaged, including the muscle metaboreflex, muscle mechanoreflex, arterial baroreflex and central command (Rowell & O'Leary, 1990). A method commonly used to specifically examine the effect of muscle metaboreflex during exercise is PEMI. During PEMI, the exercising forearm is occluded to prevent removal of the metabolic by-products of exercise. In addition, PEMI eliminates the muscle mechanoreflex and greatly reduces the input from central command. Therefore, any responses suppressed by central command during exercise should be observed during PEMI. To test if central command influences the perception of muscle pain, we induced PEMI after fatiguing IHG. Our subjects reported an increase in muscle pain perception during IHG and a further increase during PEMI. This greater increase in pain perception during

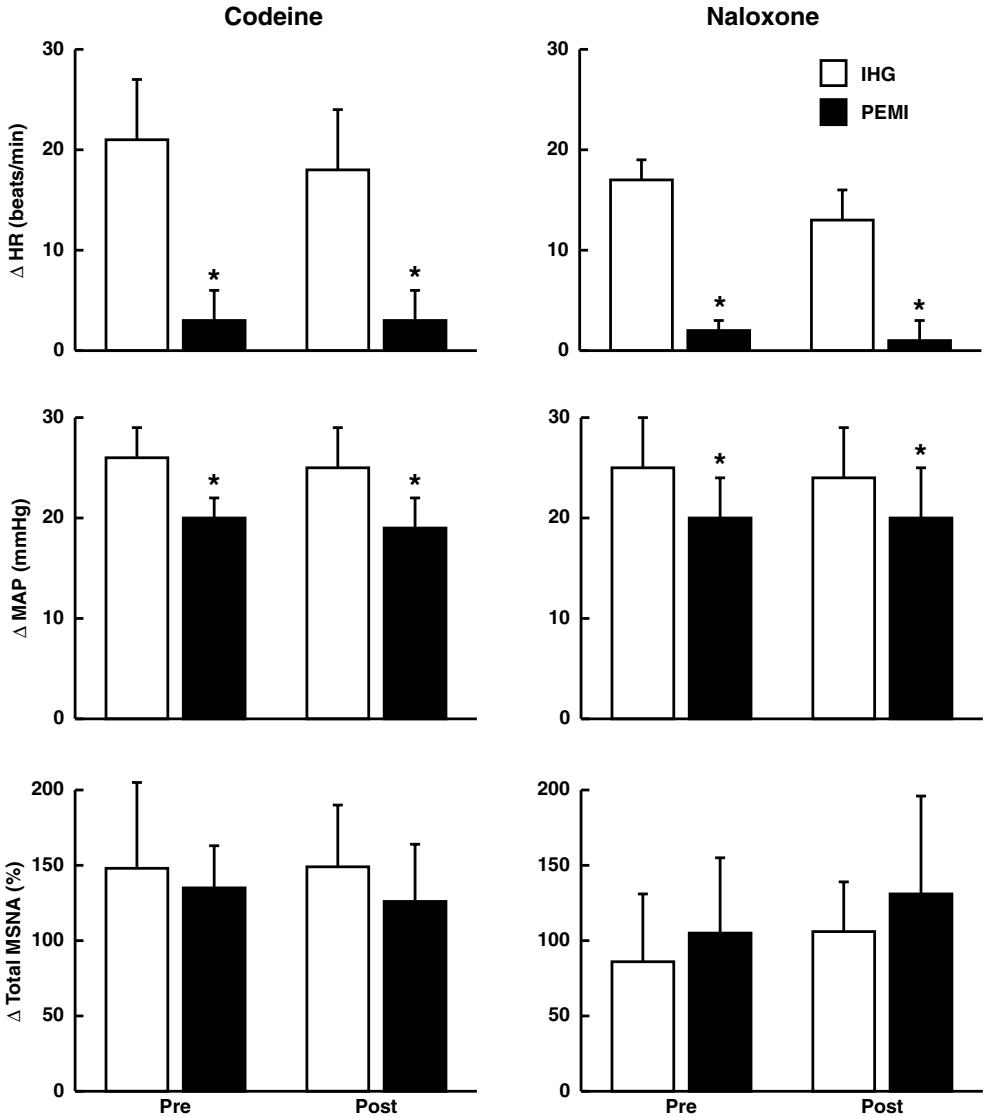


Figure 1. Change in heart rate (HR), mean arterial pressure (MAP) and muscle sympathetic nerve activity (MSNA; total activity) during isometric handgrip (IHG) and postexercise muscle ischaemia (PEMI) before (Pre) and after (Post) codeine and naloxone

MSNA and MAP significantly increased during all trials, while HR only increased during IHG trials. Administration of codeine ($n = 7$) and naloxone ($n = 9$) did not alter HR, MAP or MSNA responses to IHG or PEMI. $^*P < 0.05$ *versus* IHG.

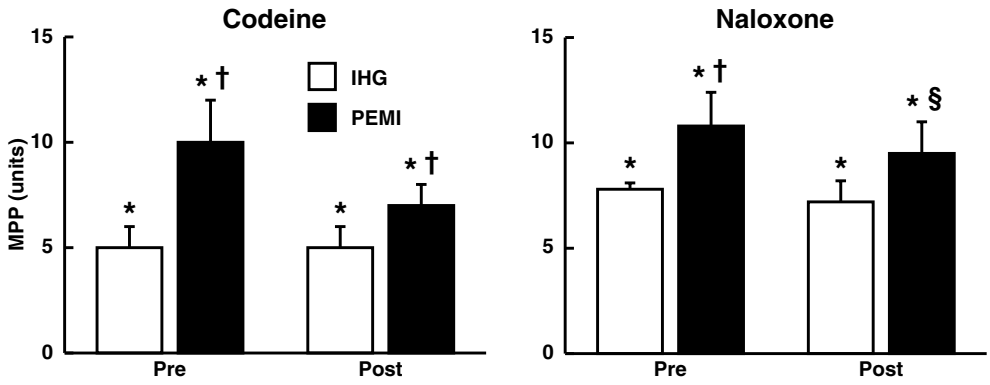


Figure 2. Muscle pain perception during isometric handgrip (IHG) and postexercise muscle ischaemia (PEMI) before (Pre) and after (Post) codeine and naloxone
Muscle pain perception increased during IHG for all trials but was greater during PEMI. *$P < 0.01$ *versus* baseline; †$P < 0.05$ *versus* IHG; §$P < 0.09$ *versus* IHG.

PEMI strongly suggests that the pain was masked centrally during the IHG trial. These results indicate that central command attenuates the perception of muscle pain.

In Study 1, PEMI reduced one stimulus (central command), but added a new stimulus (cuff compression). Thus, the increase in pain rating could have been mediated by (1) the reduction of central influences or (2) the

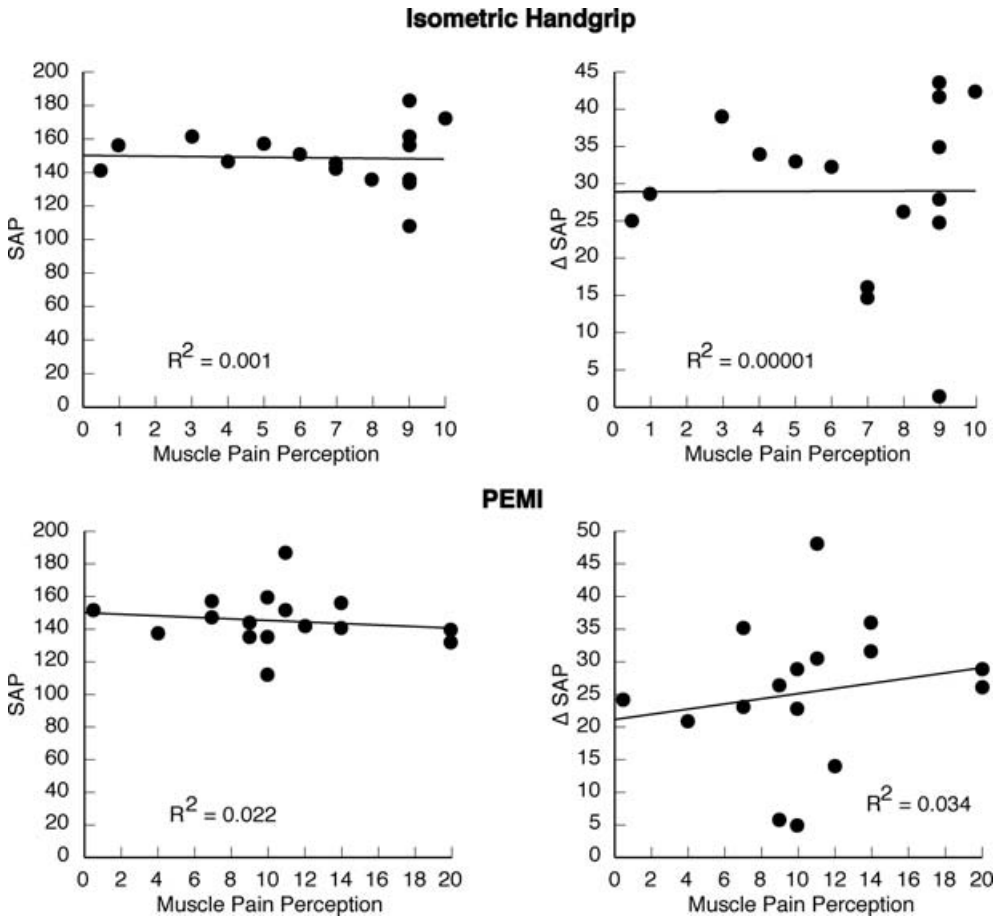


Figure 3. Correlations of muscle pain perception and systolic arterial pressure (SAP) during isometric handgrip and postexercise muscle ischaemia (PEMI) in the control trial for Study 1
There is no correlation between muscle pain perception and SAP during either isometric handgrip or PEMI ($n = 16$).

addition of cuff compression. Therefore, we designed a second control study that permitted us to determine if the elevation in muscle pain perception during PEMI was due to the added cuff compression. In Study 2, fatiguing IHG was performed during muscle ischaemia and then followed by PEMI. Using this design, the cuff compression was present throughout the experiment and any change in pain perception during PEMI would be due to withdrawal of central command. The data from Study 2 also demonstrate an increase in muscle pain perception during PEMI, thus supporting the results from Study 1. Collectively, both studies indicate that muscle pain perception is modulated by central command during exercise.

What central mechanisms could mediate the attenuation of pain perception during muscle contraction? Pain can be modulated at peripheral and central sites. A number of possible areas of the brain modulate pain perception, including the thalamus, hypothalamus, nucleus tractus solitarius, RVM, dorsal reticular nucleus, parabrachial nucleus, periaqueductal grey and amygdala (Millan, 2002). It is also possible that GABA and glycine release in the spinal cord may play an important role in the suppression of muscle afferent activity by central command (Degtyarenko & Kaufman, 2003). The current study does not permit us to determine which of these areas is most prominent in attenuating the pain perception during exercise, but our results clearly demonstrate that central modulation is occurring during IHG. This modulation could help explain the analgesic effects observed during exercise.

Although pain processing by the central nervous system is a complex process, the endogenous opioid system has been recognized as a powerful modulator of pain perception (Kanjhan, 1995; Stein, 1995; Urban & Gebhart, 1999). Endogenous opioid receptors are located on nociceptive afferent fibres and several centres of the brain stem that are involved with pain processing (Millan, 2002). Activation of opioid receptors have well-established analgesic actions, including decreasing the sensitivity of pain perception in humans.

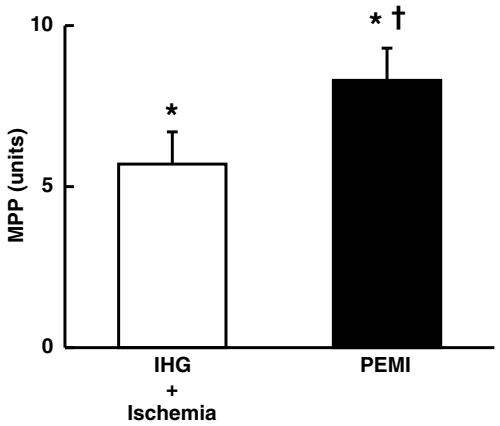


Figure 4. Muscle pain perception during ischaemic isometric handgrip (IHG) and postexercise muscle ischaemia (PEMI)
Muscle pain perception increased during ischaemic IHG and increased further during PEMI ($n = 8$). *$P < 0.01$ *versus* baseline; †$P < 0.05$ *versus* IHG.

Cook *et al.* (2000) previously reported that the opioid agonist codeine and the opioid antagonist naltrexone do not alter the perception of muscle pain during exercise. However, this study could not definitively assess whether muscle pain perception during exercise was altered by central mechanisms (i.e. higher brain systems). It is possible that central command may have interacted with afferent feedback from the exercising muscle to modulate pain perception. In the current study, muscle pain perception during PEMI was significantly increased from the corresponding IHG value during the control and codeine trials ($P < 0.01$) and tended to increase in the naloxone trial ($P < 0.09$). These results suggest that administration of codeine and naloxone had little influence on perception of muscle pain. The current study answers an important question that could not be answered in our first study; opioids do not appear to centrally modulate muscle pain perception. Collectively, these findings indicate that the endogenous opioid system does not alter the perception of acute exercise-induced muscle pain.

Previous studies suggest a relation between pain perception and arterial blood pressure (Randich & Maixner, 1984; Ghione *et al.* 1988; Lovick, 1993; Schobel *et al.* 1998). Specifically, several studies report that hypertensive subjects have a higher pain threshold compared with nomotensive subjects, suggesting that increased levels of arterial blood pressure are associated with diminished perception of pain (Ghione *et al.* 1988; Schobel *et al.* 1998). It has been suggested that the decreased pain perception reported in hypertensive subjects may be modulated by the arterial baroreflexes and the release of endogenous opioids (Randich & Maixner, 1984). Thus, it is reasonable to speculate that the increased arterial blood pressure during exercise may decrease the perception of pain to exercise.

Our results reveal that muscle pain perception during exercise is not correlated with changes in arterial blood pressure. This finding suggests that increased arterial blood pressure during exercise is not modulating the perception of pain. In Study 1, PEMI increased muscle pain perception and slightly decreased arterial blood pressure when compared with IHG, but muscle pain perception was not correlated to changes in arterial blood pressure. Study 2 demonstrated that PEMI increased muscle pain perception but did not change arterial pressure when compared with ischaemic IHG. Collectively, these results indicate that increases in arterial blood pressure during exercise are not associated with alterations in muscle pain perception. This also indicates that muscle pain perception during exercise is not modulated by arterial baroreflexes.

These findings support the concept that the suppression of muscle pain perception during exercise is modulated by central command.

It has also been suggested that there may be a potential relation between MSNA and pain perception. Specifically, Knardahl *et al.* (1998) demonstrated an increased pain threshold that paralleled increases in MSNA after acupuncture, suggesting that pain may be attenuated by increased MSNA. However, Cook *et al.* (2000) reported no correlation between muscle pain perception and MSNA during IHG. Our results support the findings of Cook *et al.* (2000) and extend them by demonstrating no correlation between muscle pain perception and MSNA during PEMI. Furthermore, the naloxone and codeine trials also revealed no correlation between muscle pain perception and MSNA during IHG or PEMI. Ray & Pawelczyk (1994) had previously demonstrated that naloxone did not modulate MSNA during IHG or PEMI, but muscle pain perception was not recorded. To our knowledge, this is the first study that has examined the relation between muscle pain perception and MSNA during both IHG and PEMI. The results of the current study, coupled with prior work (Victor *et al.* 1987; Ray & Pawelczyk, 1994; Cook *et al.* 2000), support the concept that pain is not correlated to MSNA during exercise.

The current study has three potential limitations. First, we cannot guarantee that the ischaemic contractions of Study 2 did not alter group III and IV muscle afferents. Kaufman *et al.* (1984) demonstrated that some group III and IV muscle afferents are stimulated more during ischaemic static contraction than during non-ischaemic contraction in cats. However, our data from Study 2 (ischaemic IHG) parallel the data from Study 1 (non-ischaemic IHG); thus we do not believe this limitation affects our conclusions. Second, our results indicate that central command modulates exercise-induced muscle pain, but we do not offer a mechanism of action. However, we suggest the dismissal of arterial blood pressure and MSNA as potential mechanisms because muscle pain perception during IHG and PEMI was not correlated to changes in arterial blood pressure or MSNA. Third, our results suggest that the central modulation is not influenced by endogenous opioids, but do not exclude other potential modulators of pain at the spinal level (Jordan *et al.* 1978, 1979). Furthermore, it must be noted that there are several factors that can interfere with afferent traffic, including presynaptic inhibition or primary afferent depolarization induced by higher brain centres (Lundberg *et al.* 1962; Lundberg & Voorhoeve, 1962).

In summary, this study demonstrates that muscle pain perception increases during exercise and further increases with PEMI. The augmentation of muscle pain perception during PEMI was not related to changes in arterial blood pressure or MSNA. Furthermore, endogenous opioids do not appear to modulate muscle pain perception during isometric forearm exercise. These findings suggest that central command, not an increase in arterial blood pressure or MSNA, modulates the perception of muscle pain during exercise, and reinforces the concept that endogenous opioids do not modulate acute exercise-induced muscle pain.

References

Besson JM (1999). The neurobiology of pain. *Lancet* **353**, 1610–1615.

Borg G (1978). Subjective aspects of physical and mental load. *Ergonomics* **21**, 215–220.

Cook DB, O'Connor PJ, Eubanks SA, Smith JC & Lee M (1997). Naturally occurring muscle pain during exercise: assessment and experimental evidence. *Med Sci Sports Exerc* **29**, 999–1012.

Cook DB, O'Connor PJ & Ray CA (2000). Muscle pain perception and sympathetic nerve activity to exercise during opioid modulation. *Am J Physiol Regul Integr Comp Physiol* **279**, R1565–R1573.

Degtyarenko AM & Kaufman MP (2003). Bicuculline and strychnine suppress the mesencephalic locomotor region-induced inhibition of group III muscle afferent input to the dorsal horn. *Neuroscience* **118**, 779–788.

Ghione S, Rosa C, Mezzasalma L & Panattoni E (1988). Arterial hypertension is associated with hypalgesia in humans. *Hypertension* **12**, 491–497.

Jordan LM, Kenshalo DR Jr, Martin RF, Haber LH & Willis WD (1978). Depression of primate spinothalamic tract neurons by iontophoretic application of 5-hydroxytryptamine. *Pain* **5**, 135–142.

Jordan LM, Kenshalo DR Jr, Martin RF, Haber LH & Willis WD (1979). Two populations of spinothalamic tract neurons with opposite responses to 5-hydroxytryptamine. *Brain Res* **164**, 342–346.

Kanjhan R (1995). Opioids and pain. *Clin Exp Pharmacol Physiol* **22**, 397–403.

Kaufman MP, Rybicki KJ, Waldrop TG & Ordway GA (1984). Effect of ischemia on responses of group III and IV afferents to contraction. *J Appl Physiol* **57**, 644–650.

Knardahl S, Elam M, Olausson B & Wallin BG (1998). Sympathetic nerve activity after acupuncture in humans. *Pain* **75**, 19–25.

Lovick TA (1993). Integrated activity of cardiovascular and pain regulatory systems: role in adaptive behavioural responses. *Prog Neurobiol* **40**, 631–644.

Lundberg A, Norrsell U & Voorhoeve P (1962). Pyramidal effects on lumbo-sacral interneurones activated by somatic afferents. *Acta Physiol Scand* **56**, 220–229.

Lundberg A & Voorhoeve P (1962). Effects from the pyramidal tract on spinal reflex arcs. *Acta Physiol Scand* **56**, 201–219.

Millan MJ (2002). Descending control of pain. *Prog Neurobiol* **66**, 355–474.

Quiding H, Anderson P, Bondesson U, Boreus LO & Hynning PA (1986). Plasma concentrations of codeine and its metabolite, morphine, after single and repeated oral administration. *Eur J Clin Pharmacol* **30**, 673–677.

Randich A & Maixner W (1984). Interactions between cardiovascular and pain regulatory systems. *Neurosci Biobehav Rev* **8**, 343–367.

Ray CA & Pawelczyk JA (1994). Naloxone does not affect the cardiovascular and sympathetic adjustments to static exercise in humans. *J Appl Physiol* **77**, 231–235.

Rowell LB & O'Leary DS (1990). Reflex control of the circulation during exercise: chemoreflexes and mechanoreflexes. *J Appl Physiol* **69**, 407–418.

Schobel HP, Handwerker HO, Schmieder RE, Heusser K, Dominiak P & Luft FC (1998). Effects of naloxone on hemodynamic and sympathetic nerve responses to pain in normotensive vs. borderline hypertensive men. *J Auton Nerv Syst* **69**, 49–55.

Stein C (1995). The control of pain in peripheral tissue by opioids. *N Engl J Med* **332**, 1685–1690.

Urban MO & Gebhart GF (1999). Central mechanisms in pain. *Med Clin North Am* **83**, 585–596.

Victor RG, Seals DR & Mark AL (1987). Differential control of heart rate and sympathetic nerve activity during dynamic exercise. Insight from intraneural recordings in humans. *J Clin Invest* **79**, 508–516.

Williamson JW, Fadel PJ & Mitchell JH (2006). New insights into central cardiovascular control during exercise in humans: a central command update. *Exp Physiol* **91**, 51–58.

Acknowledgements

We thank Noelle P. Dahl and Charity L. Sauder for their technical assistance with this project. This study was supported by the National Institutes of Health (DC-006459, HL-58503 and P01HL077670), NASA (NAG 9-1034), National Space Biomedical Research Institute (CA00207), and a National Institutes of Health sponsored General Clinical Research Center (M01-RR-10732 and C06-RR-016499).

Journal of Pain Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Influence of exercise on visceral pain: an explorative study in healthy volunteers

This article was published in the following Dove Press journal:
Journal of Pain Research
28 December 2016
Number of times this article has been viewed

Laura JGM van Weerdenburg[1]
Christina Brock[2,3]
Asbjørn Mohr Drewes[2]
Harry van Goor[1]
Marjan de Vries[1]
Oliver HG Wilder-Smith[4]

[1]Department of Surgery, Radboud University Medical Centre, Nijmegen, the Netherlands; [2]Mech-Sense, Department of Gastroenterology, Aalborg University Hospital, Aalborg, [3]Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark; [4]Department of Anesthesiology, Pain and Palliative Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands

Background and objectives: Contradictory results have been found about the effect of different exercise modalities on pain. The aim of this study was to investigate the early effects of aerobic and isometric exercise on different types of experimental pain, including visceral pain, compared to an active control condition.

Methods: Fifteen healthy subjects (6 women, mean [standard deviation] age 25 [6.5] years) completed 3 interventions consisting of 20 minutes of aerobic cycling, 12 minutes of isometric knee extension and a deep breathing procedure as active control. At baseline and after each intervention, psychophysical tests were performed, including electrical stimulation of the esophagus, pressure pain thresholds and the cold pressor test as a measure for conditioned pain modulation. Participants completed the Medical Outcome Study Short-Form 36 and State-Trait Anxiety Inventory prior to the experiments. Data were analyzed using two-way repeated measures analysis of variance.

Results: No significant differences were found for the psychophysical tests after the interventions, compared to baseline pain tests and the control condition.

Conclusion: No hypoalgesic effect of aerobic and isometric exercise was found. The evidence for exercise-induced hypoalgesia appears to be not as consistent as initially thought, and caution is recommended when interpreting the effects of exercise on pain.

Keywords: motor activity, breathing exercises, pain measurement, pain perception

Abbreviations

ANOVA: analysis of variance, BP: blood pressure, CPM: conditioned pain modulation, EIH: exercise-induced hypoalgesia, HR: heart rate, MCS: mental component score, MOS SF-36: medical outcomes study short-form-36 health survey, PCS: physical component score, PPT: pressure pain threshold, STAI: State-Trait Anxiety Inventory, VAS: visual analogue scale, VO_{2max}: maximal oxygen uptake.

Introduction

The modulatory effect of physical exercise on pain perception has been widely studied. Many studies found a favorable effect in healthy volunteers on somatic pain indicated with the term EIH, which is manifested as increased pain thresholds and pain tolerance levels and decreased evoked pain ratings during and immediately after exercise, persisting for 10–30 minutes post exercise.[1] This effect is seen with several types of exercise, including aerobic exercise,[2–4] isometric exercise[5–10] and dynamic resistance exercise.[11,12] EIH has been shown in healthy individuals, as well as in patients with chronic low back pain,[13,14] shoulder myalgia,[15] fibromyalgia[16–18] and chronic musculoskeletal pain,[19]

Correspondence: Laura JGM van Weerdenburg
Department of Surgery (route 618), Radboud University Medical Centre, PO Box 9101, 6500 HB Nijmegen, the Netherlands
Tel +31 6 5790 7591
Email laura.van.weerdenburg@student.ru.nl

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/JPR.S121315

although large variations were demonstrated between chronic pain syndromes.[20]

In a meta-analysis, Naugle et al[1] combined different experimental pain threshold effect sizes from several studies, which were averaged for each exercise type and pain testing method and adjusted for sample size. They calculated effect sizes using Cohen's *d* as a standardized mean difference between the control condition and the exercise condition and reported moderate effect size of 0.43 for aerobic exercise (4 studies), a large effect size of 1.05 for isometric exercise (9 studies) and 0.83 for dynamic resistance exercise (2 studies). Furthermore, the effect sizes for pain intensity ratings reported by the participants varied from 0.64 (7 studies) to 0.72 (7 studies) to 0.75 (2 studies) for the 3 exercise types, respectively.[1]

However, not all studies found positive effects. Some studies found the hypoalgesic effect only in women, not in men,[4,5] some found only trivial effects at lower exercise intensities and durations[8,21,22] and others used pain testing methods with more variability in the effect size such as thermal stimulation.[23]

This contradicting evidence from the literature is not surprising due to many methodological variations. Studies have used different types of exercise and different exercise intensities, durations and measures to control the intensity. Moreover, varying methods of pain testing were used, including pressure, electrical and thermal stimulation, which were applied to different body sites and yielded pain thresholds, suprathreshold intensity ratings or general pain intensity ratings.[1] Another method of pain testing used is CPM, the ability to influence the incoming pain signals from the periphery via descending pain inhibition from brainstem centers. It has been shown that CPM can induce a transient hypoalgesic effect, which involves a neural network comprising the nucleus tractus solitarius and brainstem nuclei.[24–26] The most frequently studied stimuli are cold water immersion as conditioning stimulus and PPTs as test stimulus, which have shown good inter- and intrasession reliability.[27] The dissimilarities between study methods make comparisons between studies and interpretations of the results difficult.

Another limitation in most study designs is the lack of a control condition. Only a few studies used quiet rest for this purpose,[2,3,10] which is not adequate, since it does not control for attention differences and cardiovascular changes during exercise. In this explorative study, a deep breathing procedure was used as active control to take into account the increased breathing rate and attention. Moreover, deep breathing controls for the increased HR in exercise conditions by causing an HR reduction through parasympathetic activation.

Furthermore, so far known, all studies evaluated the effect of exercise on somatic pain, thereby disregarding visceral pain as a common cause of chronic pain. Visceral pain is difficult to characterize in contrast to somatic pain, mainly due to diffuse termination of afferents and poor corticotropic organization.[28] This makes treatment often challenging for physicians and alternative treatments very relevant. To obtain detailed information about the visceral pain response, experimental pain models can be used to induce visceral pain in a controlled manner, while psychophysical and neurophysiological measures are carried out. In this explorative study, an experimental model of acute visceral pain was used, by delivering single pulse electrical stimuli in the esophagus. To measure the effect of exercise on other pain modalities than those of visceral origin, PPTs and CPM were also assessed. The hypothesis was that both aerobic and isometric exercise would induce hypoalgesia on experimentally induced pain, based on the psychophysical measurements. Hence, the aim of this study was to investigate the immediate effect of aerobic and isometric exercise compared with deep breathing as active control condition on visceral pain sensitivity, PPTs and induction of descending inhibition.

Methods

Participants

Fifteen participants (9 men and 6 women, mean [SD] age 25 [6.5] years) were recruited in Region North Jutland in Denmark. These healthy volunteers had no history of cardiovascular, gastrointestinal or neurological disorders that could interfere with the exercise interventions and pain measurements. The study protocol was approved by the Regional Ethics Committee of Northern Jutland, Denmark (N-200900), and all participants signed informed consent. Participants were instructed to refrain from any pain-modifying medication, alcohol and physical exercise 24 hours prior to the experimental procedure. Additionally, to minimize the unpleasantness of the esophageal tube, food, drinks, nicotine and caffeine were restrained 2 hours prior to insertion.

Study design

The crossover study with a randomized order of interventions was carried out at Mech-Sense, Department of Gastroenterology at Aalborg University Hospital. An overview of study procedures can be seen in Figure 1. Baseline pain measurements were conducted, including esophageal electrical stimulation, pressure algometry and cold pressor test. Within 5 minutes thereafter, 3 interventions: aerobic

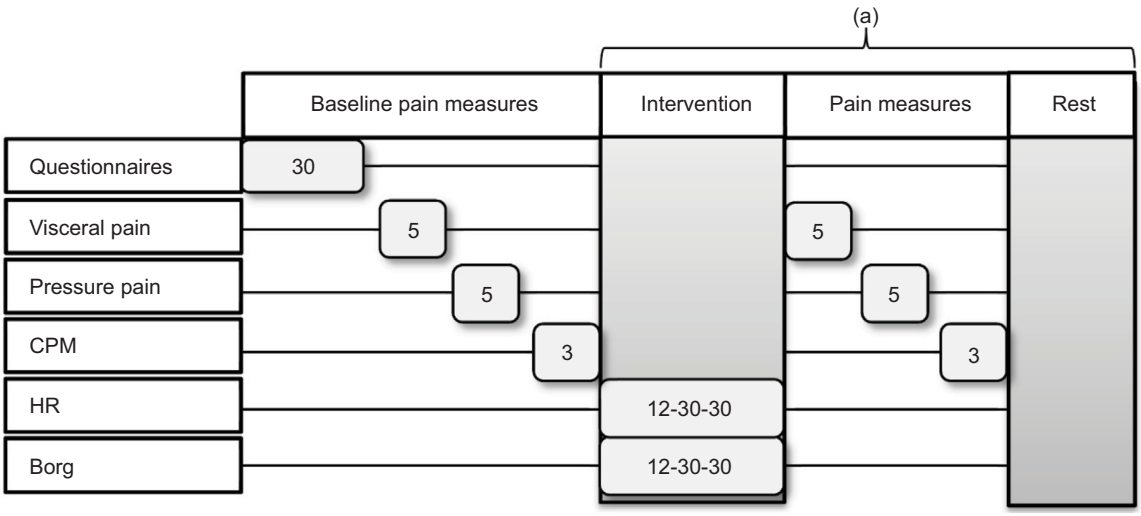


Figure 1 Timeline study procedures.
Notes: Interventions consisted of aerobic cycling, isometric knee extension and deep breathing as active control condition, followed by the pain measures and 30 minutes of rest. The numbers indicate the time in minutes. [a]This block was repeated twice.
Abbreviations: CPM, conditioned pain modulation; HR, heart rate; Borg, Borg's Rate of Perceived Exertion scale.

bicycling exercise, isometric knee extensions and a control condition were randomly performed to avoid bias of period effects and order effects. The randomization list was generated from http://www.randomisation.com. Directly after every intervention, the pain measurements were repeated, followed by a resting period of 30 minutes.

Questionnaires

The participants filled out the Danish MOS SF-36, a general health survey of 36 questions. It produces a profile of 8 scales, addressing several health aspects, and 2 composite summary scores of physical health (PCS) and mental health (MCS).[29] Furthermore, they filled out the Y1 and Y2 form of the Danish translation of STAI, which evaluates general emotional, cognitive and behavioral aspects of anxiety. The Y1 form was about anxiety "at this moment" and the Y2 form about anxiety "in general."[30]

Visual analogue scale

A modified VAS comprising ratings of non-painful (1–5) and painful sensations (5–10) was used to rate the sensation of electrical stimulation in the esophagus. This scale was used as strong pain stimuli to the esophagus carry the risk of excessive vomiting, which makes it difficult to use a pure pain VAS. It has previously been used for more than 50 studies of the gastrointestinal tract, where it has shown to be robust and reliable.[31,32]

The following anchor words were used to further assist in rating on the scale. 1, vague perception of mild sensation; 2, definite perception of mild sensation; 3, vague perception of moderate sensation; 4, definite perception of moderate sensation; 5, pain detection threshold; 6, slight pain; 7, moderate pain; 8, medium pain intensity; 9, intense pain and 10, unbearable pain.

For the cold pressor test, a pure pain VAS was used, where "0" indicated no pain, "5" moderate pain and "10" the worst pain imaginable. For the pressure algometry, VAS scores were used to clarify at which VAS level participants indicated their PPT.

Psychophysical tests

Visceral pain sensitivity

For electrical stimulation of the esophagus, a 2.6 mm diameter probe was used with 2 bipolar platinum ring electrodes attached to it, at 8.0 and 9.0 cm from the distal end (Gaeltec transducer; Gaeltec Ltd., Isle of Skye, Scotland, UK). The probe was inserted through the mouth until the interelectrode space was positioned at 34 cm from the frontal teeth and taped to the skin. Before stimulation, the impedance was checked and kept <3 kΩ by giving some water or by changing the participants' position. During stimulations, a 3-lead electrocardiogram was recorded to monitor the heart. Single pulse electrical stimulation of 2 ms was provided by a computer-controlled current stimulator, which started at an intensity of 0 mA and was increased with steps of 0.5 mA, with a predefined maximum of 60 mA. The participants scored the sensation with the modified VAS, indicating when they reached 1, 3, 5 and 7 on the VAS. At a VAS score

corresponding to 7, corresponding to moderate pain, electrical stimulation was stopped.

Pressure algometry

PPTs were measured using a hand held algometer with a standard probe tip of 1 cm^2 (SBMEDIC Electronics, Solna, Sweden). The algometer was pressed on 5 locations at the dominant site, namely on the medial part of the trapezius muscle, the dorsal T10 dermatome, the thenar muscle, the rectus femoris muscle and the abductor halluces muscle as shown in Figure 2. Starting at 0 kPa, the pressure was gradually increased with 30 kPa/s. The participants were asked to indicate the moment the sensation changed from pressure to pain, whereupon pressure testing was stopped immediately and the maximal reached pressure was noted as PPT. The mean PPT of the 5 locations was calculated for every participant in every intervention. All measurements were performed by the same investigator.

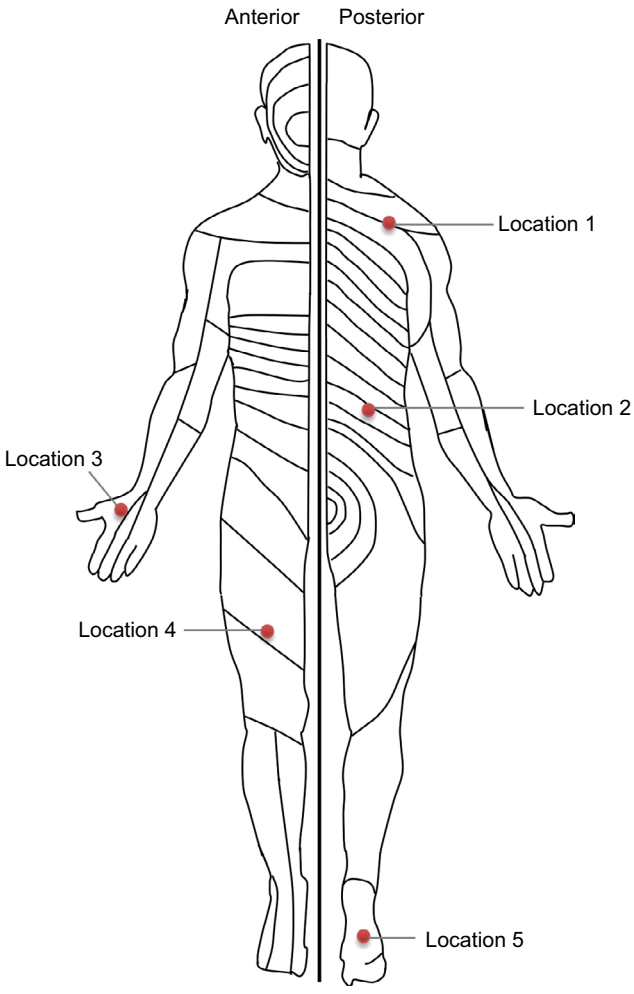


Figure 2 Locations pressure algometry.
Note: Pressure algometry locations, including the trapezius muscle (location 1), T10 dermatome (location 2), thenar muscle (location 3), rectus femoris muscle (location 4) and adductor halluces muscle (location 5).

Cold pressor test

CPM was examined with the cold pressor test, studying the ability of descending inhibitory modulation. The participant immersed their nondominant hand up to the wrist with the fingers spread in a water bath containing cold circulating water with a temperature of 2°C (±0.1°C). They kept the hand in the water for 2 minutes, or less if the pain was unbearable and reached the maximum VAS score of 10 on the pure pain VAS. Before and immediately after the test, the PPT on the quadriceps muscle at the nondominant site was examined. Furthermore, the participant rated the pain every 30 seconds during the test and immediately after, with the VAS.[33] The relative change between the PPT before and the PPT after the cold water test was calculated in percentages, as well as the mean VAS scores during the immersion.

Interventions

Aerobic bicycling exercise

After warming up for 10 minutes at a self-selected cycling intensity, the participants bicycled 20 minutes at 75%–88% of their HR_{max}, which corresponds to 60%–80% of their VO_{2max}. The individual HR that matches this intensity was calculated with the Karvonen formula, which is related to the age-predicted HR_{max} but allows for differences in resting HR: Target HR=[(220–age–resting HR)×%Intensity]+resting HR.[34] The participants had visual feedback of the HR on the oximeter (Nellcor™ OxiMax N-65; Tyco Healthcare Group LP, Pleasanton, CA, USA) and were encouraged to keep their HR in the 75%–88% range by cycling faster or adjusting the resistance of the bicycle.

Isometric knee extension

The participants performed isometric knee extension of the quadriceps muscle. They sat straight with 90° flexion in the hip joint and in 0° extension in the knee joint. A weight strap of 0.75 kg was attached around the ankle at the dominant side, to obtain the same strenuous intensity in all participants. They were instructed to extend the knee, without lifting the upper leg from the bed, for a maximum of 12 minutes or to exhaustion.

Control condition

Deep breathing was used as active control condition. The participants executed a deep breathing procedure for 30 minutes, consisting of 10 rounds. For 1 minute in every round, the participants inhaled quickly applying diaphragmatic

(abdominal) breathing and hold their breath for a total of 4 seconds. Then, they exhaled to their forced expiratory vital capacity for 6 seconds, through pursed lips. This produces breathing at a frequency of 0.1 Hz, corresponding to 6 breaths per minute. After this minute, there was a period with normal breathing until the HR and breathing had normalized and the participant was ready for the next round.[35]

Exercise measurements

HR

To measure the cardiovascular reaction on exercise, HR was measured before and every 5 minutes during the isometric knee extension. In the aerobic cycling exercise, HR was measured every 5 minutes. In the deep breathing intervention, HR measurements were used to monitor the parasympathetic nervous system activation, indicated by a decrease in HR. To demonstrate a vagal activation, the starting and lowest HR in every round of deep breathing were noted.

Borg

The Borg's Rate of Perceived Exertion scale ranges from 6 to 20 to follow the general HR of a healthy adult by multiplying with 10. In this scale, "6" means no exertion at all and "20" means maximal exertion. The participants were told to focus on the overall feeling of exertion and not just to 1 factor, such as muscle pain. The score was asked every 5 minutes during aerobic cycling exercise and every 2.5 minutes during isometric knee extension. The Borg scale was not used in the control condition, as exertion was not applicable to this intervention.

Statistics

The absolute outcomes and baseline-corrected outcomes (baseline values subtracted from the pain measurements) were compared between the interventions and control condition using two-way RM-ANOVA. For the visceral stimulation, the factors intervention (3 levels) and VAS score (4 levels) were analyzed, for the pressure algometry, the factors intervention (3 levels) and location (5 levels). If an overall difference was found, post hoc analyses (Student's *t*-test compared with Bonferroni corrected *p*-values) were used to describe the differences within the pain measurements. One-way ANOVA was used to compare the differences between interventions (3 levels) regarding the change in pain thresholds before and after the cold pressor test. *p*-values <0.05 were considered significant. In this explorative study, effect sizes were calculated using Cohen's *d*, which is a standardized mean difference. The effect sizes were calculated with the use of the baseline-corrected data, as the mean for the deep breathing condition minus the mean for the 2 exercise interventions, divided by the pooled standard deviation. It was calculated for the visceral pain sensitivity at moderate pain (VAS 7), the mean PPT from pressure algometry and the mean relative increase in PPT after CPM.

Results

Baseline characteristics and questionnaires

The baseline characteristics of the study participants are presented in Table 1.

Psychophysical tests

The data are presented as mean (SD) in the text and in Table 2.

Visceral pain sensitivity

There was no significant difference between the baseline-corrected mean (SD) of the control condition and the exercise interventions for the esophageal stimulation ($F(2, 78)=2.0$; $p=0.15$), as shown in Figure 3. The effect size at moderate visceral pain for aerobic cycling was $d=-0.39$ and for the isometric exercise $d=-0.18$. These results indicate that exercise induced no visceral hypoalgesia.

Pressure algometry

When comparing the baseline-corrected means of the PPTs on the 5 locations as shown in Figure 4, no significant difference was found between the control condition and the exercise interventions ($F(2, 112)=0.37$; $p=0.7$). The effect size for aerobic cycling was $d=-0.09$ and for isometric exercise $d=-0.06$. These data suggest that no hypoalgesia was induced by exercise.

Table 1 Baseline characteristics of the healthy volunteers (n=15)

Characteristics	Mean (SD)
Age, years	25 (6.5)
Gender, M:F	9:6
Height, m	1.79 (0.08)
Weight, kg	73 (9.7)
Body mass index, kg/m^2	22.6 (2.0)
STAI Y1 score[a]	27.67 (5.5)
STAI Y2 score[a]	31.20 (9.7)
MOS SF-36 PCS	54.02 (3.4)
MOS SF-36 MCS	51.93 (5.0)

Note: [a]Y1: state, score at this moment; Y2: trait, score in general.
Abbreviations: M, male; F, female; STAI, State-Trait Anxiety Inventory; MOS SF-36, Medical Outcome Short-Form 36; PCS, physical component summary; MCS, mental component summary.

Table 2 Outcomes of the pain measurement at baseline and after aerobic cycling exercise, isometric knee extension and the control condition

Pain test	Baseline, mean (SD)	Aerobic cycling, mean (SD)	Isometric extension, mean (SD)	Control condition, mean (SD)
Visceral stimulation (mA)				
VAS 1	9.1 (3.6)	9.2 (4.6)	7.8 (2.9)	7.9 (4.6)
VAS 3	11.7 (3.4)	13.4 (6.5)	11.3 (3.5)	11.4 (4.6)
VAS 5	17.9 (7.1)	18.6 (8.2)	17.1 (6.0)	16.0 (5.2)
VAS 7	22.0 (8.4)	22.2 (8.7)	20.8 (7.2)	19.8 (4.8)
Pressure algometry (kPa)				
Trapezius muscle	491 (118)	461 (143)	426 (111)	433 (184)
T10 dermatome	549 (142)	509 (124)	496 (114)	528 (145)
Thenar muscle	493 (116)	462 (77)	445 (85)	445 (100)
Rectus femoris muscle	641 (179)	669 (154)	671 (185)	598 (134)
Adductor halluces muscle	552 (141)	515 (98)	559 (115)	548 (157)
CPM				
PPT before (kPa)	677 (148)	688 (161)	679 (137)	619 (139)
PPT after (kPa)	724 (154)	756 (198)	777 (145)	775 (227)
VAS 30 seconds	4.6 (1.7)	4.9 (2.4)	5.1 (2.1)	5.2 (1.9)
VAS 60 seconds	6.5 (1.6)	6.3 (2.1)	6.4 (1.6)	6.8 (1.5)
VAS 90 seconds	7.5 (1.4)	7.1 (1.5)	7.4 (1.4)	7.5 (1.3)
VAS 120 seconds	7.9 (1.2)	7.7 (1.4)	8.0 (1.3)	8.1 (1.2)
VAS overall	6.9 (1.2)	6.9 (1.6)	7.1 (1.5)	7.3 (1.4)

Notes: PPT before: pressure pain threshold before cold pressor test, PPT after: pressure pain threshold after cold pressor test and VAS overall: mean VAS scores during the immersion.
Abbreviations: SD, standard deviation; VAS, visual analogue scale; CPM, conditioned pain modulation.

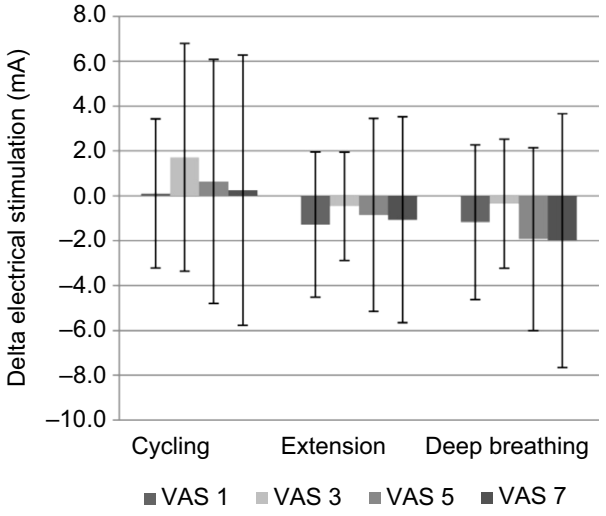


Figure 3 Visceral stimulation.
Note: Baseline-corrected mean intensities (mA) where participants rated 1, 3, 5 and 7 on the visual analogue scale (VAS) after aerobic cycling, isometric knee extension and the control condition.

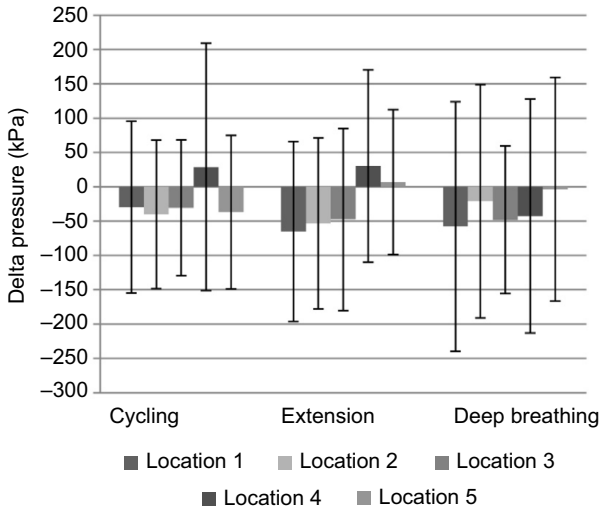


Figure 4 Pressure algometry.
Note: Baseline-corrected mean pressure pain thresholds at the trapezius muscle (location 1), T10 dermatome (location 2), thenar muscle (location 3), rectus femoris muscle (location 4) and adductor halluces muscle (location 5) after aerobic cycling, isometric knee extension and the control condition.

CPM

Except for one, all participants were able to complete the 2-minute cold pressor tests. An overall increase in PPTs was found after CPM induction (F(1, 42)=14.8; p=0.002; Figure 5). The mean (SD) relative increase for baseline was 9.3% (20.1), for aerobic cycling 10.2% (15.4), for isometric knee extension 16.1% (15.9) and for deep breathing 25.5% (22.5). However, no significant difference in CPM effect was found between the conditions (F(2, 28)=2.9; p=0.07). The effect size for the aerobic cycling was d=0.81 and for isometric exercise d=0.49. No significant differences were found between mean VAS scores during the cold pressor test, which were 6.7 (1.8) at baseline, 6.6 (2.1) after aerobic cycling, 6.8 (1.9) after isometric extension and 7.0 (1.8) after deep breathing (F=1.0; p=0.4).

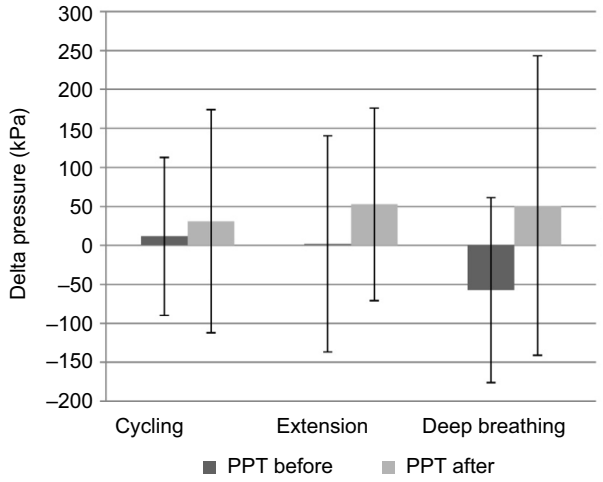


Figure 5 Conditioned pain modulation.
Note: Baseline-corrected pressure pain thresholds (PPT) before and after the cold pressor test after aerobic cycling, isometric extension and the control condition.

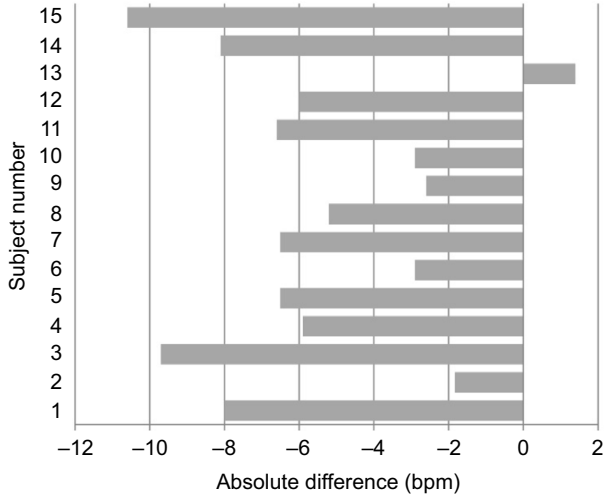


Figure 6 Heart rate during deep breathing.
Notes: Mean absolute difference between start heart rate (HR) and minimum HR in every round of deep breathing in the control condition. Data are presented for each participant individually.
Abbreviation: bpm, beats per minute.

Table 3 Heart rate (HR) during aerobic cycling, isometric extension and the control condition

HR	Aerobic cycling, mean (SD)	Isometric extension, mean (SD)	Control condition, mean (SD)
Rest[a]	63 (7.2)	76 (13.4)	69 (7.9)
End[b]	162* (9.6)	86 (11.3)	64 (7.6)
Mean[c]	164* (5.7)	87 (11.4)	66 (7.6)

Notes: [a]HR rest was measured before the interventions. [b]HR end was measured at the end of aerobic cycling and isometric knee extension. In the control condition, "HR end" was the minimum HR in every round of deep breathing. [c]HR mean was the mean heart rate over the complete intervention. *Significant increase; $p<0.001$.

Cardiovascular responses and exertion during interventions

Data are presented in Table 3.

Aerobic cycling exercise

The mean HR of the participants during the 20 minutes of cycling ranged between the target intensities 75% to 88%. There was a significant increase from HR during rest to HR during exercise ($F(2, 42)=801$; $p<0.001$). Borg scores increased significantly (from 6 (0.5) to 16 (2.2); $F(2, 42)=175$; $p<0.001$).

Isometric knee extension

There was no significant increase in HR during isometric knee extension. However, Borg scores increased significantly (from 6 (0.7) to 15 (1.6); $F(2, 42)=137$; $p<0.001$).

Control condition

Figure 6 shows the mean absolute difference between the starting HR and the lowest HR in every round of deep breathing for every individual. Except for one participant, HR decreased, with a mean (SD) of 5.4 beats per minute (4.9). However, no significant decrease was found when comparing the mean values of all participants.

Discussion

The aim of this study was to evaluate the effect of different exercise modalities on visceral and somatic pain sensitivity and CPM, compared to deep breathing as active control condition. Unexpectedly, no significant effects of aerobic and isometric exercise were found on any of the pain tests. Furthermore, no differences between the exercise conditions and control condition were found. These results suggest that exercise may not change pain evoked in healthy subjects.

Our findings contradict other studies, which found higher PPTs during and after exercise, using similar types of aerobic and isometric exercise.[2,3,6,9] However, this effect was not consistently found, likely due to many methodological variations and the absence of a control condition in many previous studies. Another difference when comparing the literature is that our study was performed on 1 day to maintain similar physiological and emotional states. To minimize period and carry over effects, the interventions were randomized and a 30-minute washout period was held between the end of pain measurements and the beginning of the next intervention. The number of participants included in the study is comparable to previous studies with similar exercise interventions; however, insufficient reliable input assumptions were available to perform a prospective power analysis. The relative small

sample size and low statistical power could have influenced the nonsignificant results.

In this study, deep breathing was used as active control condition, compared to no control condition or quiet rest in previous studies.[2,3,10] Positive characteristics of deep breathing are the control for increased breathing intensity, which occurs also during exercise, without the physical exercise and increase in HR. Furthermore, distraction during the interventions is taken into account, as participants are supposed to focus on their breathing. However, 2 main concerns have to be considered for the use of deep breathing as active control condition. First, the response to deep breathing differs among individuals and it is difficult to objectively measure the largely unknown variations in this response. Second, the deep breathing could have induced a hypoalgesic effect of itself, which makes the interpretation of the study effects complicated. It has been shown that slow, deep breathing results in lower heat pain intensity ratings[36] and increased thermal pain thresholds,[37] induces hypoalgesia for suprathreshold electrical stimulations[38] and prevents the development of acid-induced esophageal hypersensitivity.[35] It is thought that HR variability and thus parasympathetic activity during deep breathing might contribute to the hypoalgesic effect by shared cardiorespiratory and nociceptive neurophysiological pathways,[35,37] although this is not consistently found.[38]

Various parameters for determining exercise intensities at which hypoalgesia would occur have been investigated in healthy individuals. Naugle et al[9] showed a dose–response effect between cycling exercise intensity and hypoalgesic effect. According to American College of Sports Medicine, intensities corresponding to 60% to 80% of the VO_{2max} are favorable for developing cardiovascular fitness and thus often used for training. Corresponding to this intensity, Swain et al[39] recommended the use of 75%–88% of HR maximum, which is a more practical method of measuring the intensity. Therefore, in our study, HR was used to monitor the exercise intensity, using the Karvonen formula to calculate the individual target HR, which takes the resting HR and age-related maximum HR into account.[9] However, this monitored intensity could only be used for the aerobic exercise and not for isometric exercise, which makes it impossible to compare the physiological stress between the exercise conditions.

During isometric exercise, the strongest effect of hypoalgesia has been shown at low-to-moderate intensity held for longer durations, as high-threshold motor units become increasingly activated to maintain the required force. Consequently, a plausible explanation is that in order to evoke hypoalgesia, high-threshold motor units need to be recruited.[1,8] Synergistically, central inhibitory pathways might be activated, as studies showed an extrasegmental hypoalgesic effect, thus not restricted to the contracting muscle. In the same line, the hypoalgesic effect on heterotopic body parts was shown to be comparable to that on the contracting muscle.[1] Our nonsignificant results could not reproduce these previous findings. The isometric knee extension was performed for 12 minutes with a 0.75 kg weight attached around the ankle.[6] This produced the same strenuous intensity for every participant and therefore this was preferred over other methods, in which a dynamometer is used to assess the maximal voluntary contraction.

In this research, 2 different exercise types were used to evaluate different cardiovascular responses. An inverse relationship between resting BP and pain perception has been found,[40] and a few studies investigated the interaction between exercise, BP and hypoalgesia. There is some evidence for the hypothesis that an interaction exists between pain modulatory and cardiovascular systems, involving the same neuropeptides (e.g., opioids), neurotransmitters (e.g., monoamines) and brain stem nuclei (e.g., nucleus tractus solitarius and locus coeruleus).[1,20,40–42] The HR increased significantly during aerobic cycling and not during isometric extensions, thus the cardiovascular responses was dissimilar. However, no differences between the hypoalgesic effects were found.

An acute experimental pain model was used to induce visceral pain in healthy volunteers. In patients, pain is a subjective experience, influenced by many factors, for example, emotional and psychological aspects, genetics and cultural background. This makes it difficult to characterize pain mechanisms and hypoalgesic effects. The use of an experimental pain model prevents some of this bias and facilitates a controlled frequency, duration, intensity and localization of the pain stimuli. To mimic the clinical setting as much as possible, different pain modalities can be used, such as mechanical, thermal, electrical and chemical stimuli, and the pain perception can be assessed both subjectively (using the VAS) and objectively (e.g., with nociceptive reflexes or cerebral evoked potentials).[43] With these characteristics, experimental pain models help reduce the gap between preclinical studies and clinical trials.

There are some limitations inherent in this study. First, only electrical stimulations were used to induce visceral pain in the experimental pain model, due to ethical and practical motives. Electricity stimulates afferent nerves directly, therefore bypassing receptors. Furthermore, the 4-hour long position of the esophageal probe during exercise was not

visually controlled, as endoscopy was avoided to minimize the unpleasantness. Even though impedance was controlled before stimulation, it was not checked after the interventions, and hence this may affect pain measurements. As innervation and nerve density of the esophagus are unevenly distributed, minor changes in probe position could in themselves lead to differences.[44] It is challenging to measure visceral pain sensitivity objectively as it is difficult to characterize for both patients and investigators. However, it remains important to study this pain type, as it is a common cause of chronic pain with limited treatment possibilities.[45]

Conclusion

This explorative study was the first to investigate the effect of aerobic and isometric exercise on visceral and somatic pain in an experimental pain model, compared to deep breathing as an active control condition. No significant differences were found for the psychophysical tests after the 2 exercise interventions compared to the control condition, although methodological problems cannot be excluded. The hypoalgesic effect of exercise appears to be less stable than initially thought. Further studies are recommended to increase our knowledge about the effect of exercise and deep breathing on pain perception, including comparisons of the effect of exercise on different types of pain between exercise interventions and an equivalent control condition.

Acknowledgments

The authors thank Anne Grave Pedersen, Isabelle M Larsen and Annie Baunwall for their help with recruitment of participants and data collection in the laboratory.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Naugle KM, Fillingim RB, Riley JL 3rd. A meta-analytic review of the hypoalgesic effects of exercise. *J Pain*. 2012;13(12):1139–1150.
2. Gurevich M, Kohn PM, Davis C. Exercise-induced analgesia and the role of reactivity in pain sensitivity. *J Sports Sci*. 1994;12(6):549–559.
3. Koltyn KF, Garvin AW, Gardiner RL, Nelson TF. Perception of pain following aerobic exercise. *Med Sci Sports Exerc*. 1996;28(11):1418–1421.
4. Sternberg WF, Bokat C, Kass L, Alboyadjian A, Gracely RH. Sex-dependent components of the analgesia produced by athletic competition. *J Pain*. 2001;2(1):65–74.
5. Koltyn KF, Trine MR, Stegner AJ, Tobar DA. Effect of isometric exercise on pain perception and blood pressure in men and women. *Med Sci Sports Exerc*. 2001;33(2):282–290.
6. Kosek E, Lundberg L. Segmental and plurisegmental modulation of pressure pain thresholds during static muscle contractions in healthy individuals. *Eur J Pain*. 2003;7(3):251–258.
7. Koltyn KF, Umeda M. Contralateral attenuation of pain after short-duration submaximal isometric exercise. *J Pain*. 2007;8(11):887–892.
8. Hoeger Bement MK, Dicapo J, Rasiarmos R, Hunter SK. Dose response of isometric contractions on pain perception in healthy adults. *Med Sci Sports Exerc*. 2008;40(11):1880–1889.
9. Naugle KM, Naugle KE, Fillingim RB, Riley JL 3rd. Isometric exercise as a test of pain modulation: effects of experimental pain test, psychological variables, and sex. *Pain Med*. 2014;15(4):692–701.
10. Vaegter HB, Handberg G, Graven-Nielsen T. Isometric exercises reduce temporal summation of pressure pain in humans. *Eur J Pain*. 2015;19(7):973–983.
11. Koltyn KF, Arbogast RW. Perception of pain after resistance exercise. *Br J Sports Med*. 1998;32(1):20–24.
12. Focht BC, Koltyn KF. Alterations in pain perception after resistance exercise performed in the morning and evening. *J Strength Cond Res*. 2009;23(3):891–897.
13. Hoffman MD, Shepanski MA, Mackenzie SP, Clifford PS. Experimentally induced pain perception is acutely reduced by aerobic exercise in people with chronic low back pain. *J Rehabil Res Dev*. 2005;42(2):183–190.
14. Meeus M, Roussel NA, Truijen S, Nijs J. Reduced pressure pain thresholds in response to exercise in chronic fatigue syndrome but not in chronic low back pain: an experimental study. *J Rehabil Med*. 2010;42(9):884–890.
15. Lannersten L, Kosek E. Dysfunction of endogenous pain inhibition during exercise with painful muscles in patients with shoulder myalgia and fibromyalgia. *Pain*. 2010;151(1):77–86.
16. Newcomb LW, Koltyn KF, Morgan WP, Cook DB. Influence of preferred versus prescribed exercise on pain in fibromyalgia. *Med Sci Sports Exerc*. 2011;43(6):1106–1113.
17. Kadetoff D, Kosek E. The effects of static muscular contraction on blood pressure, heart rate, pain ratings and pressure pain thresholds in healthy individuals and patients with fibromyalgia. *Eur J Pain*. 2007;11(1):39–47.
18. Hoeger Bement MK, Weyer A, Hartley S, Drewek B, Harkins AL, Hunter SK. Pain perception after isometric exercise in women with fibromyalgia. *Arch Phys Med Rehabil*. 2011;92(1):89–95.
19. Cook DB, Stegner AJ, Ellingson LD. Exercise alters pain sensitivity in Gulf War veterans with chronic musculoskeletal pain. *J Pain*. 2010;11(8):764–772.
20. Nijs J, Kosek E, Van Oosterwijck J, Meeus M. Dysfunctional endogenous analgesia during exercise in patients with chronic pain: to exercise or not to exercise? *Pain Physician*. 2012;15(3 suppl):ES205–ES213.
21. Hoffman MD, Shepanski MA, Ruble SB, Valic Z, Buckwalter JB, Clifford PS. Intensity and duration threshold for aerobic exercise-induced analgesia to pressure pain. *Arch Phys Med Rehabil*. 2004;85(7):1183–1187.
22. Umeda M, Newcomb LW, Koltyn KF. Influence of blood pressure elevations by isometric exercise on pain perception in women. *Int J Psychophysiol*. 2009;74(1):45–52.
23. Ruble SB, Hoffman MD, Shepanski MA, Valic Z, Buckwalter JB, Clifford PS. Thermal pain perception after aerobic exercise. *Arch Phys Med Rehabil*. 2005;86(5):1019–1023.
24. Vaegter HB, Handberg G, Graven-Nielsen T. Similarities between exercise-induced hypoalgesia and conditioned pain modulation in humans. *Pain*. 2014;155(1):158–167.
25. Ellingson LD, Koltyn KF, Kim JS, Cook DB. Does exercise induce hypoalgesia through conditioned pain modulation. *Psychophysiology*. 2014;51(3):267–276.
26. Lemley KJ, Hunter SK, Hoeger Bement MK. Conditioned pain modulation predicts exercise-induced hypoalgesia in healthy adults. *Med Sci Sports Exerc*. 2015;47(1):176–184.
27. Kennedy DL, Kemp HI, Ridout D, Yarnitsky D, Rice AS. Reliability of conditioned pain modulation: a systematic review. *Pain*. Epub 2016 Aug 22.
28. Lottrup C, Olesen SS, Drewes AM. The pain system in oesophageal disorders: mechanisms, clinical characteristics, and treatment. *Gastroenterol Res Pract*. 2011;2011:910420.

29. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol*. 1998;51(11):903–912.
30. Spielberger CD, Gorsuch RL, Lushene RE. *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press; 1997.
31. Drewes AM, Gregersen H, Arendt-Nielsen L. Experimental pain in gastroenterology: a reappraisal of human studies. *Scand J Gastroenterol*. 2003;38(11):1115–1130.
32. Drewes AM, Schipper KP, Dimcevski G, et al. Multi-modal induction and assessment of allodynia and hyperalgesia in the human oesophagus. *Eur J Pain*. 2003;7(6):539–549.
33. Yarnitsky D, Arendt-Nielsen L, Bouhassira D, et al. Recommendations on terminology and practice of psychophysical DNIC testing. *Eur J Pain*. 2010;14(4):339.
34. Naugle KM, Naugle KE, Fillingim RB, Samuels B, Riley JL 3rd. Intensity thresholds for aerobic exercise-induced hypoalgesia. *Med Sci Sports Exerc*. 2014;46(4):817–825.
35. Botha C, Farmer AD, Nilsson M, et al. Preliminary report: modulation of parasympathetic nervous system tone influences oesophageal pain hypersensitivity. *Gut*. 2015;64(4):611–617.
36. Zautra AJ, Fasman R, Davis MC, Craig AD. The effects of slow breathing on affective responses to pain stimuli: an experimental study. *Pain*. 2010;149(1):12–18.
37. Chalaye P, Goffaux P, Lafrenaye S, Marchand S. Respiratory effects on experimental heat pain and cardiac activity. *Pain Med*. 2009;10(8):1334–1340.
38. Martin SL, Kerr KL, Bartley EJ, et al. Respiration-induced hypoalgesia: exploration of potential mechanisms. *J Pain*. 2012;13(8):755–763.
39. Swain DP, Abernathy KS, Smith CS, Lee SJ, Bunn SA. Target heart rates for the development of cardiorespiratory fitness. *Med Sci Sports Exerc*. 1994;26(1):112–116.
40. Koltyn KF, Umeda M. Exercise, hypoalgesia and blood pressure. *Sports Med*. 2006;36(3):207–214.
41. Ray CA, Carter JR. Central modulation of exercise-induced muscle pain in humans. *J Physiol*. 2007;585(pt 1):287–294.
42. Bruehl S, Chung OY. Interactions between the cardiovascular and pain regulatory systems: an updated review of mechanisms and possible alterations in chronic pain. *Neurosci Biobehav Rev*. 2004;28(4):395–414.
43. Olesen AE, Andresen T, Staahl C, Drewes AM. Human experimental pain models for assessing the therapeutic efficacy of analgesic drugs. *Pharmacol Rev*. 2012;64(3):722–779.
44. Woodland P, Sifrim D, Krarup AL, et al. The neurophysiology of the esophagus. *Ann N Y Acad Sci*. 2013;1300:53–70.
45. Sikandar S, Dickenson AH. Visceral pain: the ins and outs, the ups and downs. *Curr Opin Support Palliat Care*. 2012;6(1):17–26.

Rheumatology 2008;47:670–678
Advance Access publication 27 March 2008
doi:10.1093/rheumatology/ken021

Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes

L. Scascighini[1], V. Toma[1], S. Dober-Spielmann[2] and H. Sprott[1]

Objectives. To provide an overview of the effectiveness of multidisciplinary treatments of chronic pain and investigate about their differential effects on outcome in various pain conditions and of different multidisciplinary treatments, settings or durations.
Methods. In this article, the authors performed a systematic review of all currently available randomized controlled trials (RCTs) fulfilling the inclusion criteria, by using a recently developed rating system aimed to assess the strength of evidence with regard to the methodological quality of the trials.
Results. Compared with other non-disciplinary treatments, moderate evidence of higher effectiveness for multidisciplinary interventions was shown. In contrast to no treatment or standard medical treatment, strong evidence was detected in favour of multidisciplinary treatments. The evidence that comprehensive inpatient programmes were more beneficial that outpatient programmes was moderate. Fibromyalgia and chronic back pain patients tended to profit more substantially than patients with diverse origins or chronic pain diagnoses. No evidence was found that treatment variables, such as duration or programme components, were influential for the success of the intervention.
Conclusion. A standard of multidisciplinary programmes should be internationally established to guarantee generally good outcomes in the treatment of chronic pain. Our results highlight the lack of quality of design, execution or reporting of many of the RCTs included in this article. Future studies should more specifically focus on differential effects of treatment components and patient variables, allowing the identification of subgroups, which most probably would profit from multidisciplinary pain programmes.

Key words: Back pain, Chronic pain, Fibromyalgia, Multidisciplinary treatment, Systematic review.

Introduction

Chronic pain symptoms cause major medical and socioeconomical problems in industrialized countries due to high direct and indirect costs and are the most common cause of long-term disability in middle-aged people [1]. A great variety of treatment strategies suggest difficulties to treat these patients effectively. Knowing that chronic pain and disability are not only influenced by somatic pathology, but also by psychological and social factors, multidisciplinary interventions for chronic pain have become more accepted in various comprehensive approaches and have rapidly increased in number over the last few decades [2–4]. These are currently based on a cognitive-behavioural principle aimed at reducing disability through the modification of both cognitive processes and environmental contingencies. While cognitive treatment is aimed at modifying maladaptive cognitions on pain and its control, operant-behavioural treatment is designed to support healthy behaviours by reinforcement of those behaviours and through withdrawal of attention from pain behaviour. Time-contingent instead of pain-contingent drug use may be a part of this strategy as well, as is the involvement of the spouse. A third approach focuses on the physiological response system and aims at reducing muscular tension by providing the patient with a model of the relationship between tension and pain and teaching him/her relaxation techniques. It is mostly combined with cognitive techniques. A further common method is 'the graduated activity exposure or pacing, which is an operant-strategy used in the management of chronic pain conditions, to enable patients to control exacerbations in pain by learning to regulate the activity and once a regime of paced activity is established, to gradually increase their activity level' [5].

A comprehensive treatment approach for chronic pain patients includes one or more of these four methods combined with therapies such as physiotherapy, pain management by medication, patient education and ergonomic training. Multidisciplinary treatment has been acknowledged in the past few decades and now finds further expansion [6]. It has been evaluated in many studies and some reviews do exist, but they have their specific limitations.

The first meta-analysis [2] retrieved in our literature search included non-controlled clinical trials. More recent reviews or meta-analyses are either restricted to chronic low back pain [7, 8], fibromyalgia (FM) [9, 10] or investigated behavioural treatment alone and not multidisciplinary approaches [11–15]. Others have not been updated in the last 5 yrs [3], or included different intervention modalities for FM (i.e. pharmacological approach) [16].

For those reasons, the aims of this systematic comprehensive review on multidisciplinary treatment of chronic pain first is to give an overview on multidisciplinary treatment for chronic non-malignant pain in general, second, to compare the results for different pain diagnoses and third, to find out whether a conclusion may be drawn about the efficacy of different kinds of multidisciplinary treatments, settings or durations.

Methods

The updated guidelines for systematic reviews of the Cochrane Collaboration Back Review Group were consulted to determine the inclusion criteria, as well the methods, used in this systematic review [17], some aspects (i.e. quality assessment) were tailored according to the recent literature [18].

Publications were retrieved by comprehensive, computer-aided search on the Cochrane Central Register of Controlled Trials, MEDLINE, CINAHL, EMBASE, PEDro, PSYCINFO and PSYNDEX up to September 2006. A specific search strategy was developed for each database by using the Cochrane methodological filter for randomized controlled trials (RCTs) and combing MeSH keywords and other relevant terms including: 'multidisciplinary, interdisciplinary, patient care team, back pain, fibromyalgia, chronic pain syndrome,' exploded when necessary.

[1]Department of Rheumatology and Institute of Physical Medicine, University Hospital Zurich and [2]Federal Invalidity Insurance, Bern, Switzerland.

Submitted 10 October 2007; revised version accepted 9 January 2008.

Correspondence to: H. Sprott, Department of Rheumatology and Institute of Physical Medicine, University Hospital Zurich, Gloriastr. 25, CH-8091 Zurich, Switzerland. E-mail: haiko.sprott@usz.ch

TABLE 1. GRADE quality assessment criteria [22]

Quality of evidence	Study design	Lower if[a]	Higher if[a]
High	Randomized trial	Study quality −1—serious limitation −2—very serious limitations −1—important inconsistency Directness −1—some uncertainty −2—major uncertainty −1 Sparse data −1 High probability of Reporting bias	Strong association +1—strong, no plausible confounders, consistent and direct evidence +2—very strong, no major threats to validity and direct evidence +1—evidence of a dose response gradient +1—all plausible confounders would have reduced the effect
Moderate	Quasi-randomized trial		
Low	Observational study		
Very low	Any other evidence		

[a]1 or 2: move up or down one/two grade/s.

TABLE 2. Levels of evidence [17]

Strong evidence	Moderate evidence	Limited evidence	No evidence
• Multiple high-quality RCTs with consistent findings	• One high-quality RCT and one or more low-quality RCTs with consistent findings	• One high-quality RCT or • Multiple low-quality RCTs with consistent findings or • Contradictory outcomes of studies with high and low quality	• Only one low-quality RCT or • Contradictory outcomes of studies of the same quality

The secondary search strategy was performed by contacting experts in this field, screening of references of the RCTs included and relevant reviews.

Abstract selection and eligibility criteria

In order to optimize agreement between the two reviewers (L.S. and V.T.), all assessment tools were independently pretested using a few studies and comparing the results. After this pilot stage, L.S. and V.T. inspected the titles and abstracts of all the references retrieved by our search strategy. L.S., V.T. independently assessed the abstracts of relevant papers using a structured form to determine whether the inclusion criteria were fulfilled. In doubtful cases, the article was retrieved in full length and evaluated before making any decision. In case of uncertainties, a third reviewer (H.S.) was consulted.

RCTs were exclusively included. The original study had to deal with adult patients (>18-yrs old) with chronic non-specific musculoskeletal pain (e.g. chronic low back or back pain, FM). At least one study group had to be treated in a multidisciplinary approach in a group setting. To rank as a multidisciplinary treatment, at least three out of the following categories of psychotherapy (PS), physiotherapy, relaxation techniques, medical treatment or patient education, vocational therapy, needed to be part of the programme. At least 2 of the 12 following domains had to be covered: pain, emotional strain, quality of life, disability, coping, physical capacity, return to work, sick leave, use of medicaments, use of the health care system, pain behaviour or subjective overall success. A follow-up (FUP) of at least 3 months had to have been conducted. The studies had to be published in full length in any language and no publication date restrictions were made. To note, we focused, as mentioned in the introduction, on cognitive-behavioural, operant, psychological response system and graded exposure pain management programmes, excluding work-hardening programmes, which have partially the same contents but are otherwise weighted and have generally other primary outcomes.

Methodological quality assessment and levels of evidence

Even though there is still limited empirical evidence [19] of a relationship between specific methodological criteria and bias, it cannot be excluded that methodological flaws, which affect the internal validity of a study may introduce some bias in its results. All trials selected were judged according to a 10-item checklist by two independent reviewers (L.S., V.T.) to describe the methodological quality. 'Assessing the quality of trials in the field of this systematic reviews is faced up to differences regarding pharmacological trial (e.g. influence of experience of the care givers, blinding of the patients not always possible) and therefore specific instruments should be used' [20]. Hence, a recently developed checklist to evaluate reports of non-pharmacological trials (CLEAR NPT [18, 21]) was utilized to assess the methodological quality of the studies included in this systematic review. This checklist was specifically developed to assess the reporting of RCTs assessing non-pharmacological treatment [18, 21]. Many validity questionnaires include the items about comparability of the different groups at baseline and eligibility criteria. As those items are not part of the CLEAR NPT, indeed, we decided to introduce two supplementary items [(11) Comparability; (12) Eligibility criteria]. To draw a conclusion on the quality of evidence, we followed the criteria of the modified GRADE quality assessment, as described elsewhere (Table 1) [22].

We based our conclusions on the effectiveness of the various therapeutic interventions and on the strength of scientific evidence using a rating system with four different levels based on the quality of the studies (Table 2) [17].

Data extraction

Two reviewers (L.S., V.T.) independently extracted data according to a pre-defined protocol and a final version of the data extraction was developed by consensus. The majority of the studies measured various outcomes and our decision about primary and secondary outcomes was somewhat arbitrary. In accordance with the literature, we considered the following domains as primary outcomes: psychological strain, disability in everyday life, health-related quality of life and pain, as well as more appropriate coping strategies, which seem to account for these changes [23]. Physical capacity, return to work rate, sick leave, the use of the health care system, medication, pain behaviour, quality of sleep and other domains (e.g. subjective improvement) were considered as secondary outcomes. Furthermore, we extracted data regarding duration of the multidisciplinary pain programme (weeks and hours), type of interventions of the pain programmes and treatment components, setting and follow-up length.

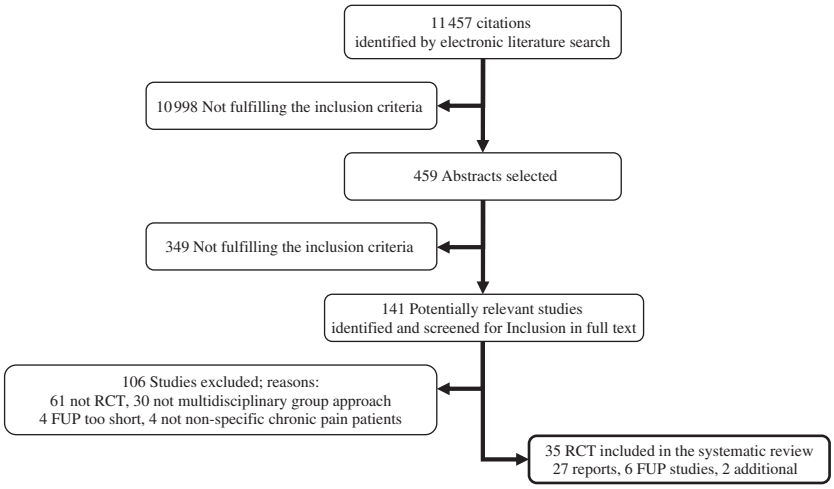


FIG. 1. Flow diagram of included and excluded studies.

Determination of success

Most chronic pain patients have a long clinical history of more or less successful treatments and the goals of therapy have to be realistically adapted to each individual situation. A multidisciplinary treatment was considered as successful if it was more effective than a control treatment [treatment as usual (TAU), waiting list control (WLC), placebo (attention control) or a treatment that did not fulfil our criteria for a multidisciplinary treatment (e.g. either physiotherapy, PS or relaxation techniques solely)]. The higher effectiveness had to be demonstrated in at least two out of the five primary outcomes, or at least in one of the primary and two of the secondary outcomes.

Results

Study selection

We retrieved 11 457 articles with our search strategy. Thereafter, 459 abstracts were selected on the basis of the title, abstract and keywords. Of those 459 abstracts, 141 articles were obtained in full-text version. Finally, we selected 35 articles by personal searching and use of references. The flow chart through the study is reported in Fig. 1.

Upon evaluation, 27 studies did qualify for entry into this review [24–50], 6 FUP studies [51–56] and 2 studies with additional analysis (Table 3) [57, 58]. Of these studies, 21 included patients with chronic low back or back pain [24–28, 32–35, 37, 38, 42, 43, 45, 48, 51–54, 56, 58], 9 included patients with FM [29–31, 36, 39, 41, 46, 49, 57] and 5 included mixed chronic pain patients [40, 44, 47, 50, 55]. Three studies had treatment programmes devoted to women only (two for chronic back pain [33, 35] and one for FM [29] with additional analysis study [57]).

Description of included studies

The number of patients of the studies included, varied between 15 and 214 (median = 86), totalling 2407 patients. The size of the individual treatment groups varied between 3 and 10 patients, but was mostly between 5 and 7.

Eighteen of 27 programmes were performed in an outpatient setting [25, 26, 29–32, 35–37, 39–43, 45–47, 49], five of 27 took place as an inpatient setting [24, 27, 33, 38, 48] (one of these with an outpatient post-treatment after inpatient treatment [38]) and four compared an inpatient with an outpatient setting [28, 34, 44, 50].

The duration of the programmes varied between 4 and 15 weeks for outpatient programmes over 15–135 h (median = 31 h) and between three and eight weeks for inpatient programmes over up to 200 h (median = 150 h). Based on the available data, the median duration of all treatments was 45 h. In order to obtain a better comparability, we tried to classify the multidimensional treatments into treatments with cognitive-behavioural approaches (CBT) and operant-behavioural approaches (OBT), although the authors called it integrated or multidisciplinary group therapy.

Central elements of multidisciplinary therapy

As study settings, populations, interventions and control groups were heterogeneous, we decided not to pool effect sizes in a meta-analysis. Details of the intervention administered were made in all reports (100.0%, Table 4). The 27 studies comprised of 74 groups including 39 with multidisciplinary treatment regimens, 20 with non-multidisciplinary treatment strategies and 15 with WLC or TAU (Table 3).

CBTs are the most common interventions and are used in all studies and in 38 of the 74 treatment groups (48%). OBT is part of the programme in 14 studies [24, 30–32, 35, 37–41, 44, 48, 49, 56]. PS is mainly administered in groups. Individual PS is part of the programme in four studies [26, 38, 47, 56]. This part usually covers 1 or 1.5 h/week, but increases to up to 6 h of group therapy per week. Aerobic exercises were used to foster endurance in 10 studies [24, 26–29, 34, 36, 39, 44, 45, 56] and muscle stretching techniques were part of the physical program in 9 studies [26–29, 31, 34, 36, 37, 50, 56]. Exercise therapy to improve activity tolerance and strengthening were part of 17 studies [24, 26–28, 30, 32–40, 47, 48, 50, 56] and back-education was taught in 4 studies [35, 42, 43, 56]. Hydrotherapy or swimming was used in nine studies [29–32, 34, 42–45].

Biofeedback training was performed in six studies [24, 38, 44, 46, 48, 49]. Progressive muscle relaxation [59] was part of the programme in eight studies [25, 30, 32, 35, 37, 41–43] and 'autogenic training' [60] was part of the programme in one study [36]. Twelve studies used other less common techniques (e.g. applied relaxation) [24, 27–29, 33, 39, 40, 44, 45, 47, 49, 50].

A medical doctor was part of the team in eight studies [31, 32, 34, 35, 39, 40, 44, 47]. His/her task was mostly the adaptation and/or reduction of the medication, as well as information about the patho-physiological processes of chronic pain.

Patient education was often an integral part of the therapy. In 16 studies, some sort of patient education was conducted [24–27, 30–33, 37, 39, 41, 44–46, 49, 50, 56]. Other elements that were part of the therapy were ergonomic training [25, 32, 35, 37], vocational therapy or occupational therapy [24, 27, 30, 34, 35, 38, 40, 44, 47, 56], nutritional counselling [31, 39, 42, 43] or therapeutic massage [39].

TABLE 3. Characteristics of studies included ($n = 35$)

Reference; FUP study	Diagnosis	Total *n* in the study	*n* in the treatments groups	FUP (months)	Multidimensional treatment: setting, duration in weeks (w) and total hours (h)	Non-multidimensional control treatment: setting, duration in weeks (w) and total hours (h)	Domains of measurements (Bold shows significant results at post measurement, cursive at FUP)											Quality of evidence according to GRADE [22]
							Pain[a]	Emotional strain[b]	Quality of life[c]	Disability[d]	Coping[e]	Physical capacity[f]	Sick leave[g]	Drug intake[h]	Use of HCP[i]	Pain behaviour[j]	Other[k]	
Linton and Gotestan [40]	Mixed CP	15	5	9	OBT + AR: out, 4 w, ?80 h	AR out, 5 w, 7.5 h WLC 4 w	*NRS*	BDI VAS		*ADL*	–	–	–	*X*	–	–	27	Low
Peters and Large [44] FUP in: Peters *et al.* [55]	Mixed CP	22	6–10	12	CBT + OBT: in, 4 w, 200 h CBT: out, 9 w, 18 h	TAU	*VAS* MPQ PD	BDI	GHQ	SIP	–	–	–	X	–	*PBC* Video	*21.25* 29	Low
Nicholas *et al.* [42]	CBP	58	5	12	2 CBT groups with/without relaxation + PT out, 5 w, 17.5 h 2 OBT groups with/without relaxation + PT ut, 5 w, 17.5 h	PT + discussion out, 5 w, 17.5 h PT out, 5 w, 17.5 h	PRC	*BDI STAI*	–	*SIP*	*CSQ PBQ*	–	–	*D*	X	–	–	Moderate
Altmaier *et al.* [24] FUP in: Patrick *et al.* [54]	CBP	45	?	6	OBT + CBT + TAU in, 3 w,? h	TAU in, 3 w ? h	*MPQ*	*WHYMPI*	–	*LBPRS*	SE	–	–	–	–	–	–	Moderate
Nicholas *et al.* [43]	CBP	20	5	6	CBT + PT out, 5 w, 17.5 h	Attention control + PT out, 5 w, 17.5 h	PRC	BDI	–	*SIP*	*CSQ* PBQ *PSEQ*	–	–	X	X	–	27	Moderate
Burckhardt *et al.* [29] Lomi *et al.* [57]	FM women	99	5–6	6	CBT + PT out, 6 w, 15 h	PE; out, 6 w, 9 h WLC, 12 w	FIQ	BDI	QOLS	*FIQ*	FAI SELF ASES	Div.	–	–	–	–	*23*	Low
Vlaeyen *et al.* [48]	CBP	71	4	12	OBT + CBT; in, 8 w, ?h	OBT; in, 8 w, ? OBT + AR, in, 8w, ?	VAS	*BDI*	–	–	*PCL*	–	–	–	–	*CHIP BAT*	–	Low
Bendix *et al.* [26] FUP in: Bendix *et al.* [56] Bendix *et al.* [52] Bendix *et al.* [51]	CBP	132	6-8	60	CBT + physical training out, 6 w, 135 h	Physical training out, 6 w, 24 h PS + physical training, out, 6 w, 24 h	*NRS*	–	–	*NRS*	–	–	–	X	*X*	–	*21.22*	Low
Bendix *et al.* [27] FUP in: Bendix *et al.* [52] Bendix *et al.* [51]	CBP	106	7	24	CBT + physical training in, 3 w,117 h	TAU	*NRS*	–	–	*NRS*	–	–	–	X	*X*	–	–	Low
Vlaeyen *et al.* [49]	FM	131	6	12	CBT + OBT; out, 6 w, 42 h PE + discussion out, 6 w, 42 h	WLC; 8 w	MPQ	BDI FSS-III-R MOCI	–	–	CSQ PCL MPCL	–	–	–	–	UAB CHIP BAT	24	Moderate
Williams *et al.* [50]	Mixed CP	121	10	12	CBT; in, 4 w, 140 h CBT; out, 8 w, 28 h	WLC	VAS MPI	*BDI* STAI	–	SIP	PSEQ *CSQ* PCQ	*Div.*	–	X	X	–	27	Moderate
Basler *et al.* [25]	CBP	94	5–8	6	CBT + PT + medical treatment out, 12 w, 30 h	TAU	*D*	–	–	*DDS*	HCS	–	–	–	–	–	22	Moderate
Keller [37]	CBP	65	9	6	OBT + CBT out, 6 w, 45 h	WLC	*NRS*	CES-D	*WBQ*	PDI	SE	*Div.*	–	–	–	*Video*	–	Low

Rose [45]	CBP	102	5–10	6	6 CBT groups; comparison of individual and group therapy and of 1 w (15 h, 30 h) or 1.5 w (60 h) out	No non-multidimensional control treatment	VAS	ZDI	–	*RMDQ*	PLOC PSEQ	–	–	–	–	–	*26*	Low
Jensen *et al.* [33]	CBP women	63	?	18	2 CBT groups; both: in, 5 w, 200 h	No non-multidimensional control treatment	VAS	*BDI*	GSI	*DRI*	*CSQ RAI*	–	–	–	–	–	22	High
Nicassio *et al.* [41]	FM	86	3–7	6	OBT + CBT out, 10 w, 15 h	PE + discussion out, 10w, 15 h	FIQ MPQ	*CES-D*	QWB	–	RAI PMI	–	–	–	–	PBCL OPB	23	Low
Keel *et al.* [36]	FM	32	8	3	CBT out, 15 w, 30 h	Autogenic training out, 15 w, 30 h	*D*	–	–	–	LOC	–	–	D	D	–	21	Moderate
Kole-Snijders [38] Spinhoven *et al.* [58][m] Non-standardized OBT WLC	CBP	148	5	12	OBT + CBT In, 5 w + out, 3 w), 160 h	OBT + group discussion, in, 5 w + out 3 w	VAS	*BDI FSS-III-R*	–	–	*CSQ MPLC PCL*	BAT	–	–	–	*PBS CHIP*	25	High
Gowans *et al.* [31]	FM	41	?	6	CBT + OBT out, 6 w, 18 h	WLC	–	–	–	*FIQ*	*ASES*	*Div.* RPE	–	–	–	–	*24*	Moderate
Bendix *et al.* [28]	CBP	127	?	12	CBT + physical training In?, 3 w, 117 h	Physical training Out, 8 w, 36 h	NRS	–	–	*ADL*	–	X	–	X	–	–	*21*	Moderate
Jensen *et al.* [32] FUP in: Jensen *et al.* [53]	CBP	214	4–8	36	CBT + PT; out, 4 w, 134 h	TAU PT; out, 4 w, 80 h CBT; out, 4 w, 54 h	–	–	*SF-36*	–	–	–		–	X	–	–	High
Soares and Gross [46]	FM	53	3–5	6	PE Out, 10 w, 102 h	CBT + AR out, 10 w, 120 h WLC	D MPQ PQ	*SCL-90R*	–	FIQ	*CSQ ASES*	–	–	X	–	–	30	Low
Turner-stokes *et al.* [47]	Mixed CP	113	8–10	12	CBT; out, 8 w, 32 h	Individual PS; out, 8 w, 8 h		*BDI WHYMPI* STAI	–	–	–	–	–	X	–	–	–	Moderate
Jousset *et al.* [34]	CBP	86	?	6	PT + OT + medical treatment In, 5 w, 150 h	Individual PT out, 5 w, 15 h	*VAS*	*HAD*	*DPQ*	*QBPD*	–	*Div.*		*X*	–	–	–	Moderate
Cedraschi *et al.* [30]	FM	164	8–10	6	CBT + OBT out, 6 w, 18 h	WLC	RPS	–	*PGWB* SF-36	*FIQ*	–	–	–	–	–	–	23, 27	High
Lemstra and Olszynski [39]	FM	79	?	15	CBT + OBT + PT; out, 6 w, 31 h?	TAU	*VAS*	*BDI*	*PDI*	–	–	–		*X*	–	–		High
Kääpä *et al.* [35]	CBP women	120	6–8	24	CBT + OBT out, 8 w, 70 h	Individual PT; out, 6–8 w, 10 h	*NRS*	DEPS	WBQ	*OSW*	–	–			*X*	–	*28*	High

[a]Pain: VAS: Visual Analogue Scale; D: diary; NRS: Numeric Rating Scale; FIQ: Fibromyalgia Questionnaire; MPQ: McGill Pain Questionnaire; PRC: pain rating chart; PD: pain drawings; RPS: regional pain score. [b]Emotional Strain: BDI: Beck Depression Inventory; MSPQ: Modified Somatic Perception Questionnaire; ZDI: Zung Depression Inventory; ADS: Allgemeine Depressivitätsskala: CES-D: Center for Epidemiological Studies – Depression Scale; STAI: State-Trait Anxiety Inventory; FSS-III-R: Fear Survey Schedule; MOCI: Maudsley Obsessive Compulsive Inventory; VAS-D: VAS for Depression; WHYMPI: West Haven Yale Multidimensional Pain Inventory; POMS: Profile of Mood States; HAD: Hospital Anxiety Depression. [c]Quality of Life: GSI: Global Self Rating Index; WBQ: Well-Being Questionnaire; QWB: Quality of Well-Being Scale; QOLS: Quality of Life Scale; GHQ: General Health Questionnaire; DPQ: Dallas Pain Questionnaire. [d]Disability: NRS: Numeric Rating Scale; RMDQ: Roland and Morris Disability Questionnaire; DRI: Disability Rating Scale; PDI: Pain Disability Index; DDS: Düsseldorf Disability Scale; SIP: Sickness Impact Profile; FIQ: Fibromyalgia Impact Questionnaire; LBPPS: Low Back Pain Rating Scale; QBPD: Quebec Back Pain Disability. [e]Coping: CSQ: Coping Strategies Questionnaire; MPLC: Multidimensional Pain Locus of Control Scale; PLOC: Pain Locus of Control Scale; PCL: Pain Cognition List; LOC: Locus of Control Scale; PSEQ: Pain Self Efficacy Questionnaire; RAI: Rheumatology Attitudes Index; SE/SELF: Self Efficacy Scale; ASES: Arthritis Self-efficacy Scale; HCS: Heidelberg Coping Scale; PMI: Pain Management Inventory; FAI: Fibromyalgia Attitudes Index; PBQ: Pain Beliefs Questionnaire [f]Physical capacity: Div.: Diverse Tests; RPE: Rate of perceived exertion. [g]Return to work/sick leave: X: not specified. [h]Drug consummation: D: Diary; X: not specified. [i]Consultation of HCP: D: Diary; X: not specified. [j]Pain Behaviour: PBS: Pain Behaviour Scale; CHIP: Checklist for Interpersonal Pain Behaviour; PBCL: Pain Behaviour Check List; OPB: Observed Pain Behaviour; UAB: University of Alabama at Birmingham Pain Behaviour Scale; BAT: Behavioural Approach Test. [k]Other: 21: subjective improvements; 22: days of absence at work; 23: Tender Points; 24: Knowledge (FM); 25: Activity; 26: MSPQ: Modified Somatic Perception Questionnaire; 27: Satisfaction/Expectancy; 28: Subjective working ability; 29: ISCRG: Illness Self-construct repertory grid; 30: KSQ: Karolinska Sleep Questionnaire. [l]Burckhardt *et al.* [29]; Lomi *et al.* [57], same study sample, in Lomi *et al.* [57] additional analysis of the ASES. [m]Kole *et al.* [38], Spinhoven *et al.* [58], same sample, in Spinhoven *et al.* [58] additional analysis. AR: applied relaxation; PT: physiotherapy; OT: occupational therapy; PE: patient education; In: inpatient setting; Out: outpatient setting. mixed CP: groups with patients with pain of mixed localization or origin; HCP: health care professionals; RTW: return to work; ?: No detailed information in the original article.

TABLE 4. Numbers (%) of the rated articles ($n=27$) without FUPs and additional analysis[a] with corresponding CLEAR NPT [18, 21] (modified) items

Items of the CLEAR NPT [18, 21] (modified)	Yes	
	n	%
(1) Was the generation of allocation sequences adequate?	21/27	77.8
(2) Was the treatment allocation concealed?	7/27	25.9
(3) Were details of the intervention administered to each group made available ?	27/27	100.0
(4) Were care providers' experience or skill in each arm adequate?	13/27	48.1
(5) Was participants (e.g., patients) adherence assessed quantitatively?	9/27	33.3
(6) Were participants adequately blinded?	3/27	11.1
(6.1) If participants were not adequately blinded were all other treatments and care (i.e., co-interventions) the same in each randomized group?[b]	4/24	16.7
(6.2) Were withdrawals and lost to FUP the same in each randomized group?[b]	6/24	25.0
(7) Were care providers or persons caring for the participants adequately blinded?	3/27	11.1
(7.1) If care providers were not adequately blinded were all other treatments and care (i.e., co-interventions) the same in each randomized group?[b]	4/24	16.7
(7.2) Were withdrawals and lost to FUP the same in each randomized group?[b]	6/24	25.0
(8) Were outcome assessors adequately blinded to assess the primary outcomes?	7/27	25.9
(8.1) If outcome assessors were not adequately blinded, were specific methods used to avoid ascertainment bias (systematic differences in outcome assessment)?[b]	0/20	0.0
(9) Was the FUP schedule the same in each group?	25/27	92.6
(10) Were the main outcomes analysed according to the intention-to-treat principle?	10/27	37.0
(11) Comparability at baseline	25/27	92.7
(12) Eligibility criteria	25/27	92.6

[a]FUP studies ($n=6$) and additional analysis ($n=2$) studies not included. [b]Item (6.1), (6.2), (7.1), (7.2), (8.1): If main item 'Yes', those questions are not to be answered.

Methodological quality of the studies

Table 4 shows the items of the CLEAR NPT of the 27 studies included without FUP studies and the additional analysis studies. The generation of allocation sequences was considered adequate in 77.8% of the trials and only in 25.9% of the reports, the treatment allocation was concealed. Based on the information available in the text, we judged the care providers' experience or skill in each arm being adequate in 48.1% of the included studies, though the information were rather scarcely reported. The participants' adherence was assessed quantitatively in just 33.3% of the included reports. Blinding was adequately reported for the participants in only 11.1% of the studies, for care providers in 11.1% and for the outcome assessors in 25.9%. When the blinding criterion was not fulfilled, co-interventions were the same in each randomized group in 16.7% of the studies. Withdrawals and losses to follow-up were the same in each randomized group in 25.0% of the studies. In most of the papers included, there was insufficient information to make a decision for the items 6.1–6.2 and 7.1–7.3 ('Unclear' 75.0–25.0% and, respectively, 75.0–29.1%).

No specific methods were used to avoid ascertainment bias (0.0%). The FUP schedule was the same in each group in almost all studies (92.6%). The median of the length of the FUP is 12 months. An intention-to-treat analysis was calculated in 37.0% of the articles. The baseline comparability was fulfilled in 25 articles (92.7%). The same results were shown for the declaration of the eligibility criteria (92.7%).

Table 3 shows the overall design quality of the studies included. Only six studies were ranked as high-quality studies [30, 33, 35, 38, 39, 61] according to the GRADE definition (Table 1) [22].

Comparison of multidisciplinary treatment vs WLC or TAU

Fifteen studies comparing multidisciplinary treatment *vs*. WLC or TAU [24, 25, 27, 29–32, 37–40, 44, 46, 49, 50] showed strong evidence that a multidisciplinary treatment is superior to a standard medical treatment or WLC (Table 5). Thirteen studies reported positive results [25, 27, 29–32, 37–40, 44, 49, 50], and two did not demonstrate positive results [24, 46]. Results of long-term FUPs were not available for this comparison in all studies, as patients from waiting lists often entered the treatment programme after the post-assessment, due to ethical reasons. The differences after treatment were maintained at FUP in those studies where results were described.

Comparison of multidisciplinary treatment vs other control group treatments

Fifteen studies comparing multidisciplinary treatment *vs* non-multidisciplinary control group treatment (e.g. physiotherapy with discussion group, patient education) were identified [26, 28, 29, 34–36, 38, 40–43, 46–48, 61]. Together they showed moderate evidence that a multidisciplinary treatment is more effective. In five studies, the results indicated no significant difference between the groups [28, 35, 41, 46, 47]. Where success was recorded, it was maintained at FUP (Table 5).

Comparison inpatient vs outpatient programmes

Four studies directly compared inpatient and outpatient programmes [28, 34, 44, 50]. Three of them demonstrated moderate evidence for superior long-term effects of intensive inpatient programmes. One study showed no differences [28]. Notably, the inpatient programmes were much more intensive than the outpatient programmes (Table 3).

Comparison of effects for groups with different pain diagnoses

There is moderate evidence that a multidisciplinary programme is more effective than no treatment or non-multidisciplinary treatment for chronic back pain patients. Six of seven studies comparing it with a WLC or TAU had positive results [25, 27, 32, 37, 38, 48], as well as the 8 of 11 studies comparing it with another treatment showed moderate evidence that a multidisciplinary treatment is more effective [26, 32–34, 38, 42, 43, 48]. In five studies, no differences were shown between the groups [24, 28, 35, 45, 61].

In FM, there is moderate evidence that a multidisciplinary programme is more effective than no treatment. Three studies [30, 31, 49] showed positive results for a multidisciplinary treatment *vs* a WLC, on the other hand two studies did not show any difference [29, 46].

In two studies, the comparisons with other treatments did not show any difference [29, 41]. Only two studies showed a superiority of the multidisciplinary group [36, 39].

There was limited evidence that a multidisciplinary programme for mixed chronic pain patients was more effective compared with TAU or WLC [40, 44, 50]. No difference were shown for other treatment strategies [47].

TABLE 5. Results according to the determination of success

Study	Success at post measurement	Success at FUP	Success at long-term FUP
Linton and Gotestam [40]	AR, AR + OBT > WLC	AR > AR + OBT	–
Peters and Large [44] FUP in: Peters *et al.* [55]	CBT in > TAU; CBT out > TAU CBT in = CBT out	CBT in > CBT out > TAU	CBT in > CBT out > TAU
Nicholas *et al.* [42]	CBT + AR + PT, CBT + PT, OBT + AR + PT, OBT + PT > discussion + PT, PT OBT + AR + PT, OBT + PT > CBT + AR + PT, CBT + PT	BT + AR + PT, CBT + PT, OBT + AR + PT, OBT + PT > discussion + PT, PT	–
Altmaier *et al.* [24] FUP in: Patrick *et al.* [54]	OBT + CBT = TAU	OBT + CBT = TAU	OBT + CBT = TAU improvements maintained
Nicholas *et al.* [43]	CBT > attention control	CBT > attention control	–
Burckhardt *et al.* [29] Lomi *et al.* [57]	CBT + PT > WLC Patient education > WLC CBT + PT = patient education	CBT + PT = patient education	–
Vlaeyen *et al.* [48]	OBT + CBT, OBT, OBT + AR > WLC OBT + CBT, OBT + AR > OBT	OBT + CBT, OBT, OBT + AR > WLC OBT + CBT, OBT + AR > OBT OBT + CBT > OBT, OBT + AR	–
Bendix *et al.* [26] FUP in: Bendix *et al.* [56] Bendix *et al.* [52] Bendix *et al.* [51]	CBT > physical training CBT > PS and physical training	CBT > physical training CBT > PS and physical training	CBT + OBT + PE > PS + physical training, physical training
Bendix *et al.* [27] FUP in: Bendix *et al.* [52] Bendix *et al.* [51]	No results in the article	CBT > TAU	CBT > TAU
Vlaeyen *et al.* [49]	CBT + OBT = PE + discussion > WLC	CBT 1 = CBT 2 > WLC	–
Williams *et al.* [50]	CBT in > CBT out > WLC	CBT in > CBT out	–
Basler *et al.* [25]	CBT > TAU	CBT > TAU	–
Keller *et al.* [37]	CBT + OBT > WLC	Improvements maintained	–
Rose *et al.* [45]	Individual = group; 15 h = 30 h = 60 h all CBT groups successful	Individual = group; 15 h = 30 h = 60 h Improvements maintained	–
Jensen *et al.* [33]	CBT women > CBT general	CBT women > CBT general	–
Nicassio *et al.* [41]	OBT + CBT = patient education + discussion	OBT + CBT = patient education + discussion	–
Keel *et al.* [36]	CBT = autogenic training	CBT > autogenic training	–
Kole-Snijders *et al.* [38] Spinhoven *et al.* [58]	OBT + CBT = OBT + discussion OBT + CBT, OBT + discussion > PS + PT OBT + CBT, OBT + discussion > WLC	OBT + CBT = OBT + discussion OBT + CBT, OBT + discussion > PS + PT	–
Gowans *et al.* [31]	CBT + OBT > WLC	CBT + OBT > WLC	–
Bendix *et al.* [28]	No results in the article	CBT + physical training = physical training	–
Jensen *et al.* [32] FUP in: Jensen *et al.* [53]	No results in the article	CBT + OBT, CBT, PT = TAU CBT, PT > TAU (women) CBT + OBT, CBT > TAU (women)	CBT + OBT > CBT, PT > TAU (women)
Soares *et al.* [46]	CBT > PE = WLC	CBT = PE = WLC	–
Turner-stokes *et al.* [47]	CBT = individual PS	CBT = individual PS	–
Jousset *et al.* [34]	No results in the article	PT + OT + medical treatment > PT PT + OT + medical treatment = PT	–
Cedraschi *et al.* [30]	No results in the article	CBT + OBT > WLC	–
Lemstra Olzynski [39]	CBT + OBT + PT > WLC	CBT + OBT + PT > WLC	–
Kääpä *et al.* [35]	CBT + OBT = PT	CBT + OBT = PT	–

>: first group has significantly better results than the second group; =: no significant difference between the two groups. Bold shows significant results in at least two of the primary outcomes or in at least one primary and two secondary outcomes

Comparison of different multidisciplinary programmes

Four studies compared different kinds or duration of multidisciplinary treatments [33, 42, 45, 49]. There is no evidence that a special kind, duration or setting of multidisciplinary treatment as described in the evaluated studies is superior to any of the other study regimens (Table 5).

Success in connection with measurements

The range of instruments to assess the various domains of interest is very broad. In fact, in each domain, 6–12 different instruments were administered. There is no tendency that special domains or certain instruments show successful results more often and are more sensitive than others (Table 3). Most of the RCTs used instruments to assess coping strategies (16/27; 59.3%), emotional strain (19/27; 70.4%), health-related quality of life (10/27; 37.0%) and/or disability outcomes (19/27; 70.4%). Remarkably, pain measurement was rarely reported as a primary outcome (88.9%).

Discussion

This article provides the most current and comprehensive review of the existing evidence of the efficacy of multidisciplinary pain programmes and represents an unique evaluation with a detailed overview of the outcome instruments and intervention in multidisciplinary pain programmes. With reference to our first aim, it seems that a minimum standard of multidisciplinary therapy can be currently established from these data, namely ideally: specific individual exercising, regular training in relaxation techniques, group therapy led by a clinical psychologist (1.5 h) per week, patient education sessions once a week, two physiotherapy treatments per week (CBT) for pacing strategies, medical training therapy and neuro-physiology information given by trained physician.

The efficacy of such programmes is not only better than standard medical treatment, but also better than other non-multidisciplinary treatments. Therefore, the set-up of multidisciplinary programmes for chronic pain patients appears to be reasonable and patients should be referred to adequately specialized institutions, instead of being sent to various individual medical specialists sequentially.

In relation to our second aim, the results seen in patients with mixed chronic pain are definitely less beneficial as compared with the promising studies with FM and chronic back pain patients, and should be a question of further investigation. FM as well chronic back pain are different but share some similarities. In fact, both musculoskeletal disorders are strongly associated to a behavioural component, i.e. fear avoidance, over-under activity, passive coping strategies, etc. Additionally, we observed that both diagnostic groups have maladaptive beliefs about the explanation of the pain (catastrophizing behaviour, structural damage, kinesiophobic disturbs, high level of depression, distress).

Our third aim was to assess different kinds of multidisciplinary programmes. Intensive inpatient programmes seem to be more effective, which is consistent with the findings of Guzman *et al*. [7]. Such programmes may be justified for patients with more severe disabilities. Regarding treatment components or duration, there is no evidence for a superior effect of a special treatment regimen. However, a final conclusion cannot be drawn due to the low number of studies comparing this aspect.

The overall methodological quality of the studies was found to be rather low. Some requirements, such as the blinding of care provider and patients, may not be met by multidisciplinary therapy. Other requirements, such as coverage of the method of randomization or concealment of treatment allocation, were insufficiently reported. An important point to consider is the small study population in some investigations. As a consequence, some studies were underpowered and some effects may not have been detected. For physicians it is fundamental to apply the evidence from systematic reviews only if the results are judged as clinically relevant and applicable. Thus, in accordance with the criteria recommended from Malmivaara *et al*. [62], we can state that generally the papers included are to be considered as clinically relevant and applicable.

Our systematic review is (as any review or meta-analysis) bound to publication bias and we cannot exclude that we may have missed some relevant trials, despite the fact that we used a highly sensitive search strategy, we did not have any language restrictions and consulted an experienced librarian, as recommended in Crumley *et al*. [63]. We did not apply a quantitative pooling of effect sizes but decided to summarize the findings by strength of evidence. Regarding the large heterogeneity of the studies, this seemed to us the more appropriate way to report the results. The decision to include or exclude some articles fulfilling the inclusion criteria, but not with the main focus on pain programme, is questionable. Our decision was based on the content of the programme and depending on the primary outcome measurements.

Multidisciplinary treatments are effective, but it is still not known which treatment components are really important and whether all patients (with different diagnoses, age, duration of pain, social background, etc.) would profit from all components. Future studies should compare different methods, settings and durations of multidisciplinary treatments and examine their connection with patient characteristics in more detail in order to detect differential effects. In order to achieve these demanding goals, multicentre studies may be useful. Further studies are needed to establish determinants or prognostic indicators of success, and to also define the therapeutic potential for a successful rehabilitation. As an upshot of this systematic review, we would recommend a stronger observance of methodological guidelines and the use of internationally accepted outcome measures in order to make studies more comparable, due to the extensive heterogeneity among the outcome measurements. An important task for the future will be the realization of more cost–benefit analyses in order to see which of the treatments are really worth being carried out. Health care insurances should finance and promote high quality of pain programmes that fulfil the minimal recommendations mentioned, representing the state of the art for multidisciplinary pain programmes.

In summary, this work may be helpful, especially for practising physicians in their daily work, in setting priorities more on disabilities and health-related quality of life in the treatment of chronic pain patients and also for researchers to optimally plan the outcome measurements and intervention modalities of future clinical trials.

Rheumatology key messages

- Multidisciplinary pain programme represent the state of the art of the management of complex, chronic, non-malignant pain patients.
- A standard requirement for a multidisciplinary pain programme is discussed.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995;62:233–40.
2 Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. Pain 1992;49:221–30.
3 Karjalainen K, Malmivaara A, van Tulder M *et al*. Multidisciplinary rehabilitation for fibromyalgia and musculoskeletal pain in working age adults. Cochrane Database of Systematic Reviews 1999, Issue 3. DOI: 10.1002/14651858.CD001984.
4 Gatchel RJ, Okifuji A. Evidence-based scientific data documenting the treatment and cost-effectiveness of comprehensive pain programs for chronic non-malignant pain. J Pain 2006;7:779–93.
5 Harding V, Watson PJ. Increasing activity and improving function in chronic pain management. Physiotherapy 2000;86:619–30.
6 Vlaeyen JW, Morley S. Cognitive-behavioural treatments for chronic pain: what works for whom? Clin J Pain 2005;21:1–8.
7 Guzmán J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002;CD000963.
8 van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic non-specific low back pain. A systematic review of randomized controlled trials of the most common interventions. Spine 1997;22:2128–56.
9 Sim J, Adams N. Systematic review of randomized controlled trials of nonpharmacological interventions for fibromyalgia. Clin J Pain 2002;18:324–36.
10 Oliver K, Cronan T, Walen H. A review of multidisciplinary interventions for fibromyalgia patients: where do we go from here? J Musculoskeletal Pain 2001;9:63–80.
11 Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1–13.
12 van Tulder MW, Ostelo R, Vlaeyen JW, Linton SJ, Morley SJ, Assendelft WJ. Behavioural treatment for chronic low back pain: a systematic review within the framework of the Cochrane Back Review Group. Spine 2000;25:2688–99.
13 Keefe FJ. Behavioural assessment and treatment of chronic pain: current status and future directions. J Consult Clin Psychol 1982;50:896–911.
14 Ostelo RWJG, van Tulder MW, Vlaeyen JWS, Linton SJ, Morley SJ, Assendelft WJJ. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005;CD002014.
15 Pearce S. A review of cognitive-behavioural methods for the treatment of chronic pain. J Psychosom Res 1983;27:431–40.
16 Rossy LA, Buckelew SP, Dorr N *et al*. A meta-analysis of fibromyalgia treatment interventions. Ann Behav Med 1999;21:180–91.
17 van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. Spine 2003;28:1290–9.
18 Boutron I, Moher D, Tugwell P *et al*. A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus. J Clin Epidemiol 2005;58:1233–40.
19 van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for spinal disorders. Spine 1997;22:2323–30.
20 Ravaud P, Boutron I. Primer: assessing the efficacy and safety of nonpharmacologic treatments for chronic rheumatic diseases. Nat Clin Pract Rheumatol 2006;2:313–9.
21 Jacquier I, Boutron I, Moher D, Roy C, Ravaud P. The reporting of randomized clinical trials using a surgical intervention is in need of immediate improvement: a systematic review. Ann Surg 2006;244:677–83.

22 Atkins D, Best D, Briss PA *et al.* Grading quality of evidence and strength of recommendations. Brit Med J 2004;328:1490.
23 Jensen MP, Turner JA, Romano JM, Karoly P. Coping with chronic pain: a critical review of the literature. Pain 1991;47:249–83.
24 Altmaier EM, Lehmann TR, Russell DW, Weinstein JN, Kao CF. The effectiveness of psychological interventions for the rehabilitation of low back pain: a randomized controlled trial evaluation. Pain 1992;49:329–35.
25 Basler H, Jakle C, Kroner-Herwig B. Incorporation of cognitive-behavioural treatment into the medical care of chronic low back patients: a controlled randomized study in German pain treatment centers. Patient Educ Couns 1997;31:113–24.
26 Bendix AF, Bendix T, Ostenfeld S, Bush E, Andersen. Active treatment programs for patients with chronic low back pain: a prospective, randomized, observer-blinded study. Eur Spine J 1995;4:148–52.
27 Bendix AF, Bendix T, Vaegter K, Lund C, Frølund L, Holm L. Multidisciplinary intensive treatment for chronic low back pain: a randomized, prospective study. Cleve Clin J Med 1996;63:62–9.
28 Bendix T, Bendix AF, Busch E, Jordan A. Functional restoration in chronic low back pain. Scand J Med Sci Sports 1996;6:88–97.
29 Burckhardt CS, Mannerkorpi K, Hedenberg L, Bjelle A. A randomized, controlled clinical trial of education and physical training for women with fibromyalgia. J Rheumatol 1994;21:714–20.
30 Cedraschi C, Desmeules J, Rapiti E *et al.* Fibromyalgia: a randomised, controlled trial of a treatment programme based on self management. Ann Rheum Dis 2004;63:290–6.
31 Gowans SE, deHueck A, Voss S, Richardson M. A randomized, controlled trial of exercise and education for individuals with fibromyalgia. Arthritis Care Res 1999;12:120–8.
32 Jensen IB, Bergström G, Ljungquist T, Bodin L, Nygren AL. A randomized controlled component analysis of a behavioral medicine rehabilitation program for chronic spinal pain: are the effects dependent on gender? Pain 2001;91:65–78.
33 Jensen IB, Dahlquist C, Nygren Å, Royen E, Stenberg M. Treatment for 'helpless' women suffering from chronic spinal pain: a randomized controlled 18-month follow-up study. J Occup Rehabil 1997;7:225–38.
34 Jousset N, Fanello S, Bontoux L *et al.* Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. Spine 2004;29:487–93.
35 Kääpä EH, Frantsi K, Sarna S, Malmivaara A. Multidisciplinary group rehabilitation versus individual physiotherapy for chronic non-specific low back pain: a randomized trial. Spine 2006;31:371–6.
36 Keel PJ, Bodoky C, Gerhard U, Müller W. Comparison of integrated group therapy and group relaxation training for fibromyalgia. Clin J Pain 1998;14:232–8.
37 Keller S, Ehrhardt-Schmelzer S, Herda C, Schmid S, Basler HD. Multidisciplinary rehabilitation for chronic back pain in an outpatient setting: a controlled randomized trial. Eur J Pain 1997;1:279–92.
38 Kole-Snijders AM, Vlaeyen JW, Goossens ME *et al.* Chronic low-back pain: what does cognitive coping skills training add to operant behavioural treatment? Results of a randomized clinical trial. J Consult Clin Psychol 1999;67:931–44.
39 Lemstra M, Olszynski WP. The effectiveness of multidisciplinary rehabilitation in the treatment of fibromyalgia: a randomized controlled trial. Clin J Pain 2005;21:166–74.
40 Linton SJ, Gotestam KG. A controlled study of the effects of applied relaxation and applied relaxation plus operant procedures in the regulation of chronic pain. Br J Clin Psychol 1984;23(Pt 4):291–9.
41 Nicassio PM, Radojevic V, Weisman MH *et al.* A comparison of behavioural and educational interventions for fibromyalgia. J Rheumatol 1997;24:2000–7.
42 Nicholas MK, Wilson PH, Goyen J. Operant-behavioural and cognitive-behavioural treatment for chronic low back pain. Behav Res Ther 1991;29:225–38.
43 Nicholas MK, Wilson PH, Goyen J. Comparison of cognitive-behavioural group treatment and an alternative non-psychological treatment for chronic low back pain. Pain 1992;48:339–47.
44 Peters JL, Large RG. A randomised control trial evaluating in- and outpatient pain management programmes. Pain 1990;41:283–93.
45 Rose MJ, Reilly JP, Pennie B, Bowen-Jones K, Stanley IM, Slade PD. Chronic low back pain rehabilitation programs: a study of the optimum duration of treatment and a comparison of group and individual therapy. Spine 1997;22:2246–51.
46 Soares JJF, Grossi G. A randomized, controlled comparison of educational and behavioural interventions for women with fibromyalgia. Scand J Occup Therapy 2002;9:35–45.
47 Turner-Stokes L, Erkeller-Yuksel F, Miles A, Pincus T, Shipley M, Pearce S. Outpatient cognitive behavioural pain management programs: a randomized comparison of a group-based multidisciplinary versus an individual therapy model. Arch Phys Med Rehabil 2003;84:781–8.
48 Vlaeyen JW, Haazen IW, Schuerman JA, Kole-Snijders AM, van Eek H. Behavioural rehabilitation of chronic low back pain: comparison of an operant treatment, an operant-cognitive treatment and an operant-respondent treatment. Br J Clin Psychol 1995;34(Pt 1):95–118.
49 Vlaeyen JW, Teeken-Gruben NJ, Goossens ME *et al.* Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I. Clinical effects. J Rheumatol 1996;23:1237–45.
50 Williams AC, Richardson PH, Nicholas MK *et al.* Inpatient vs. outpatient pain management: results of a randomised controlled trial. Pain 1996;66:13–22.
51 Bendix AE, Bendix T, Haestrup C, Busch E. A prospective, randomized 5-year follow-up study of functional restoration in chronic low back pain patients. Eur Spine J 1998;7:111–9.
52 Bendix AF, Bendix T, Labriola M, Boekgaard P. Functional restoration for chronic low back pain: two-year follow-up of two randomized clinical trials. Spine 1998;23:717–25.
53 Jensen IB, Bergström G, Ljungquist T, Bodin L. A 3-year follow-up of a multidisciplinary rehabilitation programme for back and neck pain. Pain 2005;115:273–83.
54 Patrick LE, Altmaier EM, Found EM. Long-term outcomes in multidisciplinary treatment of chronic low back pain: results of a 13-year follow-up. Spine 2004;29:850–5.
55 Peters J, Large RG, Elkind G. Follow-up results from a randomised controlled trial evaluating in- and outpatient pain management programmes. Pain 1992;50: 41–50.
56 Bendix AF, Bendix T, Lund C, Kirkbak S, Ostenfeld S. Comparison of three intensive programs for chronic low back pain patients: a prospective, randomized, observer-blinded study with one-year follow-up. Scand J Rehab Med 1997;29:81–9.
57 Lomi C, Burckhardt C, Nordholm L, Bjelle A, Ekdahl C. Evaluation of a Swedish version of the arthritis self-efficacy scale in people with fibromyalgia. Scand J Rheumatol 1995;24:282–7.
58 Spinhoven P, Ter Kuile M, Kole-Snijders AM, Hutten Mansfeld M, Den Ouden DJ, Vlaeyen JW. Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. Eur J Pain 2004;8:211–9.
59 Jacobson E. Entspannung als Therapie: progressive relaxation in theorie und praxis. 4th edn. Munich (D): Pfeiffer, 1999.
60 Schultz J. Das autogenen training: Konzentrative Selbsentspannung. 19th edn. Stuttgart (D): Thieme Verlag, 1991.
61 Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. J Consult Clin Psychol 2001;69:655–62.
62 Malmivaara A, Koes BW, Bouter LM, van Tulder MW. Applicability and clinical relevance of results in randomized controlled trials: the Cochrane review on exercise therapy for low back pain as an example. Spine 2006;31:1405–9.
63 Crumley ET, Wiebe N, Cramer K, Klassen TP, Hartling L. Which resources should be used to identify RCT/CCTs for systematic reviews: a systematic review. BMC Med Res Methodol 2005;5:24.

Published in final edited form as:
J Pain. 2012 December ; 13(12): 1139–1150. doi:10.1016/j.jpain.2012.09.006.

A meta-analytic review of the hypoalgesic effects of exercise

Kelly M. Naugle, **Roger B. Fillingim**, and **Joseph L. Riley III**
Comprehensive Center for Pain Research, University of Florida

Abstract

The purpose of this article was to examine the effects of acute exercise on pain perception in healthy adults and adults with chronic pain using meta-analytic techniques. Specifically, studies using a repeated measures design to examine the effect of acute isometric, aerobic, or dynamic resistance exercise on pain threshold and pain intensity measures were included in this metaanalysis. The results suggest that all three types of exercise reduce perception of experimentally induced pain in healthy participants, with effects ranging from small to large depending on pain induction method and exercise protocol. In healthy participants, the mean effect size for aerobic exercise was moderate (d_{thr} =0.41, d_{int} =0.59), while the mean effect sizes for isometric exercise (d_{thr} =1.02, d_{int} =0.72) and dynamic resistance exercise (d_{thr} =0.83, d_{int} =0.75) were large. In chronic pain populations, the magnitude and direction of the effect sizes were highly variable for aerobic and isometric exercise and appeared to depend on the chronic pain condition being studied as well as the intensity of the exercise. While trends could be identified, the optimal dose of exercise that is needed to produce hypoalgesia could not be systematically determined with the amount of data available.

Index words

hypoalgesia; analgesia; aerobic exercise; isometric exercise; resistance exercise; pain

Introduction

Physical exercise is an important component in the treatment and rehabilitation of many patients with chronic pain, as well as vital to the overall health and wellbeing of any individual. Importantly, laboratory studies report that acute exercise reduces sensitivity to painful stimuli in healthy individuals, indicative of a hypoalgesic response. This phenomenon has been termed exercise-induced analgesia or exercise-induced hypoalgesia (EIH).[36,37] However, the methodology of studies investigating exercise-induced hypoalgesia is diverse and the results are not always consistent. A comprehensive understanding of how exercise influences pain perception is necessary to optimize the clinical utility of exercise as a method of pain management.

Corresponding Author: Kelly M. Naugle, Comprehensive Center for Pain Research, College of Dentistry, University of Florida, Phone: 352-273-5975, Fax: 352-273-5985, knaugle@dental.ufl.edu.

Conflict of Interest
There are no conflicts of interest, or any financial interests, to report with regard to this work for any of the authors.

Numerous experimental studies have examined the effect of acute exercise on responses to experimentally induced noxious stimulation. These studies have included a variety of exercise modalities, as well as a variety of pain induction techniques and measurement procedures. For example, exercise modalities have included aerobic exercise, isometric exercise, and dynamic resistance exercises. Aerobic exercises have typically included stationary cycling, running, or step exercise. Isometric and dynamic resistance exercises are both a form of strength training. Isometric exercise involves a static contraction in which the joint angle does not change, whereas dynamic resistance exercise involves muscle contractions that do produce joint movement. These exercise modes have differed across many dimensions including the type, intensity, and duration of exercise. Furthermore, techniques of pain induction have included electrical, pressure, thermal, and other forms of noxious stimulation. These stimuli also differ across many dimensions, including site of bodily application and temporal parameters of the stimulation. Pain measures have most commonly included pain thresholds (i.e., the point at which noxious stimulation is first perceived as painful) and/or suprathreshold pain intensity ratings during and following exercise. EIH has also been investigated in healthy and clinical populations, including fibromyalgia syndrome (FMS), chronic fatigue syndrome (CFS), chronic low back pain (CLB), chronic musculoskeletal pain (CMP), and shoulder myalgia. These collective differences between studies have made direct comparisons across studies difficult.

While several narrative reviews have elegantly summarized the exercise-induced hypoalgesia literature[36,37], to our knowledge no quantitative review of the acute exercise literature has been published. Meta-analytic methods offer an alternative method to study the impact of acute exercise on pain in terms of the magnitude and direction of effect. Therefore, to extend and update the work in the previous reviews, the present study used meta-analysis methodology to answer the following questions: 1) Is there a hypoalgesic effect of acute bouts of exercise using measures of pain intensity and/or threshold 2) If there is a hypoalgesic effect, what is its magnitude using the effect size metric? 3) Does the magnitude of the effect vary by exercise mode (aerobic, isometric, dynamic resistance)? 4) Is a hypoalgesic effect of exercise on experimental pain observed in healthy and chronic pain populations?

Methods

Sample of Studies

Acute exercise studies that used an outcome measure involving pain were located on computer based searches conducted on PubMED, Medline, PsychINFO, and Academic Search Premier databases from 1900 to May 2012. The key words included ‘pain’, ‘exercise’, ‘contraction’, ‘hypoalgesia’, ‘analgesia’, and ‘isometric’. These searches were extended by examining reference sections from published articles identified from the databases. We believe that these studies represent a comprehensive selection of empirical studies. Only published research was included in the analysis, which may have biased the results as non-significant results are less likely to be published than those with significant findings. When studies did not provide adequate statistical information for the calculation of effect sizes, means and standard deviations were estimated from figures and authors were contacted via electronic mail. Studies were included if they met the following criteria: 1) study was performed on healthy adults or a chronic pain population, 2) a repeated measures, with-in subject design was used, 3) pain threshold and/or intensity measures were used, 4) exercise protocol was standardized, 5) pain induction protocol was standardized. The literature search located 50 total studies. Eleven studies did not provide adequate information for the calculation of effect sizes[6,7,16,32,33,34,43,44,52,53,56], three studies did not include pain threshold or intensity measures[2,10,31], two studies did not implement standardized exercise[1,66], two studies did not standardize the method of pain induction[18,54],

two studies combined exercise with another manipulation[25,70], and one study did not use a repeated measures, within subjects design[4]. Thus, a total of 25 studies met criteria to be included in the analysis, consisting of 622 participants (437 healthy, 185 chronic pain) and 118 effects (88 healthy, 28 chronic pain).

Statistical Analysis

The effect size (ES) for each study was calculated using Cohen's *d*, defined as the mean for the control condition minus the mean for the exercise condition, divided by the pooled within group standard deviation ($\underline{d}=[X_{control} - X_{exercise}]$/pooled standard deviation). Thus, $\underline{d}$ is a standardized mean difference that can be interpreted in the same manner as any standard score. If data were reported separately for men and women, the effects were averaged into one[13]. Effect sizes were calculated so that reductions in pain sensitivity resulted in positive effect sizes. Due to the within subjects designs of the studies, the effect sizes were adjusted as recommended by Portney and Watkins.[58]

The mean effect size of $\underline{d}$ was calculated using the pooled effect sizes within each exercise mode for measures of threshold and intensity ratings (across pain stimuli). Due to the variation in sample sizes, it has been argued that not all studies in meta-analyses should be given equal weight. Hedges[19,20], noting the bias in estimates of $\underline{d}$ when weighting for sample size, developed a weighted estimator of effect size (d) which is asymptotically efficient and appropriate for group sizes greater than 10:

$$d = \Sigma wd/\Sigma w \text{ where } w=2N/8 + d^2$$

In sum, we report the mean of the raw effect size $\underline{d}$, standard deviation of $\underline{d}$, and weighted mean effect size (d). The effect sizes of healthy adults and those with chronic pain were analyzed separately, and thus are presented separately.

Results

Division of Studies

The studies were first divided by type of exercise, with 12 studies implementing isometric exercise (healthy: N=267, 59 effects; chronic pain: N= 84; 21 effects), 11 studies using aerobic exercise (healthy: N= 136 participants, 23 effects; chronic pain: N=101, 10 effects), and 2 studies using dynamic resistance training (healthy: N=34, 8 effects). They were further subdivided by threshold and intensity pain measures. Of the 12 isometric studies, ten measured threshold (healthy: 42 effects; chronic pain: 18 effects) and seven measured intensity (healthy: 17 effects; chronic pain: 2 effects). Ten studies used pressure stimuli as the method of pain induction, one study used thermal heat, and one used electric stimulation. Of the eleven aerobic studies, six used threshold (healthy: 8 effects; chronic pain: 6 effects) and nine used intensity (healthy: 15 effects; chronic pain: 4 effects) measures. Six studies used pressure stimuli as the pain induction method, three used thermal heat stimuli, and three used cold stimuli. The two dynamic resistance exercise studies measured threshold (4 effects) and intensity (4 effects) of pain induced by pressure stimuli.

Healthy Adults

Aerobic Exercise

Table 1 presents the results for the eight studies involving aerobic exercise and measuring pain threshold and/or intensity. This table shows that aerobic exercise reduced pain sensitivity across all types of pain stimuli and exercise type, with the largest effects found for studies using pressure stimuli and the smallest effects on average for those using cold and heat stimuli. The summary results (the mean of effect size $\underline{d}$, standard deviation of $\underline{d}$,

and weighted mean effect size (*d*) averaged within each exercise type and stimuli) for pain threshold and intensity are shown in Table 2 and Table 3, respectively. When averaged across pain stimuli, the effect size for pain threshold was positive and moderate at 0.48 and when adjusted for sample size and bias, 0.43. Two studies, Meeus et al.[48] and Koltyn et al.[39], used pressure pain thresholds to test for EIH and reported moderate effect sizes, *d* =0.58. Koltyn et al. found that pain threshold continued to be reduced 15 minutes post exercise, with an effect size of 0.79. One study, Ruble et al.[62], tested pain thresholds using hot and cold thermal stimuli and found trivial effects, $\underline{d}$'s= 0.04. Ruble et al. also found no effect of thermal stimuli 30 minutes post exercise, $\underline{d}$'s=0.21–0.25. However, Kemppainen et al. reported a moderate and positive effect of 0.48 using cold stimuli.[31]

Averaged across stimuli, the average effect size for pain intensity was positive and slightly greater in magnitude than for pain threshold at 0.68, and 0.64 when adjusted for sample size. Once again, the effect size pooled within pressure stimuli was greater at 0.69 (3 studies - 5 effects) than those for heat stimuli, *d*=0.59 (2 studies – 2 effects), and cold stimuli *d*=0.61 (3 studies – 3 effects). Two studies took follow-up pain intensity measures 30 minutes post exercise, with an average effect size of 0.33 (SD= 0.12).[26,67]

The pre-post exercise measurement design involving repeated tests before and after exercise is commonly used in the EIH literature. This study design without the inclusion of a resting control condition for comparison is flawed by the possibility that post-exercise pain ratings are influenced by pre-exercise pain tests. Two studies, Koltyn et al.[39] and Vierck et al.[72], compared pain measures assessed during an exercise condition to a resting control condition. Importantly, these studies actually found positive and larger effect sizes ($\underline{d}$'s = 0.83–1.18) than studies employing pre-post designs without a resting control comparison condition, with the exception of Gurevich et al.[17] Two studies, Gurevich et al.[17] and Ruble et al.[62], conducted reliability testing in which pain measures were assessed pre and post quiet rest. These studies found no significant changes in pain ratings from pre to post, with effect sizes ranging from −0.14 to 0.16.

Isometric Exercise

Table 4 presents the results for the 11 studies assessing pain threshold and/or intensity immediately following or during isometric exercise. This table shows that isometric exercise reduced pain perception across all pain stimuli and exercise protocols, with the exception of the pain intensity measure in Umeda et al. 2009.[69] The summary results for threshold and intensity measures (the mean of effect size $\underline{d}$, standard deviation of $\underline{d}$, and weighted mean effect size (*d*) averaged within each exercise type and stimuli) are shown in Table 2 and Table 3, respectively. The average effect size for pain threshold (9 studies – 43 effects, all studies used pressure stimuli) was positive and large at 1.27, with the weighted mean value of 1.05. Three studies measured pain threshold during the contraction and reported large positive effect sizes (14 effects: $\underline{d}$=1.76, *d*=1.69)[30,45,46], while six studies measured pain threshold immediately after exercise reporting moderate to large effects (14 effects: $\underline{d}$=0.70, *d*=0.69).[22,40,42,68,69] Three studies also measured pain threshold 15 minutes post contraction (14 effects)[30,45,46], with values of 0.58 and 0.43 for $\underline{d}$ and *d*, respectively.

The effect size for pain intensity measures averaged across stimuli was also positive and large at 0.83, while the unbiased effect size was 0.72 (7 studies – 17 effects). Five studies used pressure stimuli to test for EIH, with an unbiased effect size of 0.73. One study, Staud et al.[65], tested pain sensitivity using thermal heat stimuli and found a mean effect size of 1.35 (2 effects). The study using electrical stimulation, Ring et al.[61], reported a medium mean effect size of 0.40 (2 effects). Two studies measured pain intensity during the contraction with an average effect of 0.87 and an unbiased effect of 0.67 (4 effects).[61,65] The average effect size for the studies measuring pain intensity immediately following the

contraction was similar at 0.81, with an unbiased effect of 0.72 (12 effects). [22,40,42,68,69] No studies conducted follow-up (i.e., 15–30 minutes post exercise) pain intensity tests.

Five studies assessed pain measures on the contracting body area, as well as on a remote body area (often contralateral to contracting body part) following isometric exercise.[30,42,45,46,65] The average effect size for pain threshold (6 effects) assessed on the contracting body area was 1.74 (SD= 0.53), and almost identical on the remote body area at 1.73 (SD=0.82). The average effect size for pain intensity (2 effects) assessed on the contracting body area was 2.02 (SD=1.13), and slightly lower on the remote body area at 1.54 (SD=0.08). Thus, isometric exercise appears to exert a generalized pain inhibitory response.

The magnitude of the effect of isometric exercise on pain threshold and intensity generally increased for contractions of longer duration. Contractions of 1 minute or less had an average effect size of 0.51 (SD=0.27, 2 effects) for threshold and 0.87 (SD=0.72, 4 effects) for intensity. Contractions of 2–3 minutes had an average effect size of 0.96 (SD=0.36, 6 effects) for threshold and 0.83 (SD=1.00, 6 effects) for intensity, while contractions 5 minutes or greater were even larger at 1.74 (SD=0.75, 15 effects) and 1.70 (SD=0.13; 2 effects) for threshold and intensity, respectively. Examination of contraction intensity reveals the largest positive effects at moderate intensity contractions. Those at 40–50% MVC had an average effect size of 1.75 (SD=0.99, 3 effects) for intensity and 1.12 (SD= 0.14, 3 effects) for threshold, while those for the 10–25% MVC contractions were 0.67 (SD= 0.51, 11 effects) and 1.13 (SD=0.72, 16 effects) for intensity and threshold, respectively. Contractions at 80%–100% MVC had the smallest effect on pain intensity (M=0.50, SD=0.29, 3 effects) and threshold (M=0.57, SD=0.33, 3 effects) measures.

Few isometric exercise studies included a resting control condition in the experimental design. Umeda et al. applied a pressure stimulus to the forefinger for 2 minutes following isometric exercise and quiet rest.[69] Interestingly, the effect sizes were generally smaller in magnitude compared to the other isometric studies, ranging from −0.16 to 0.54. Ring et al. compared pain intensity measures during 15 and 25% MVC contractions to a 1% MVC control condition and reported moderate effect sizes (0.31–0.41).[61] Hoeger Bement et al. found trivial changes in the pain measures during reliability testing consisting of 30 minutes of quiet rest (threshold = −0.03, intensity= 0.04).[22]

Dynamic Resistance Exercise

Two studies measured pain threshold and intensity immediately following dynamic resistance exercise (See Table 5).[9,38] The mean effect size for pain threshold was 0.99 (SD=0.18) and the weighted mean effect size was 0.83. The mean effect size for pain intensity was 0.83 (SD=0.37) and the weighted mean effect size was 0.75. Both studies took follow-up measures at 15 minutes post exercise, with the unbiased average effect size of 0.21 for threshold and 0.18 for intensity. Koltyn & Arbogast included a quiet rest condition, which showed no significant changes from pre to post immediately following exercise, d= −0.118, or 15 minutes post exercise, d=0.04.[38]

Chronic pain populations

Aerobic Exercise

As a reminder, the effect size data presented for chronic pain populations represent subjects' responses to experimental pain and not subjects' assessments of their pre-existing chronic pain. Table 6 presents the results for the five studies involving chronic pain subjects and aerobic exercise. As shown in the table, the effects sizes were highly variable, ranging from - 1.13 to 1.50. When averaged across chronic pain syndromes, the effect for pain threshold

was positive and small at 0.19 (SD=0.52) and when adjusted for sample size and bias, 0.15. The effect sizes for pain intensity were highly variable with an average effect size of 0.42 (SD=1.53) and the adjusted effect size similar at 0.43. The two studies investigated FMS reported contrasting effect sizes which were likely due to differing aerobic and pain testing protocols.[51,72] Newcomb et al. found that cycle ergometry at a self-selected intensity increased PPTs, $\underline{d}$=1.11, and decreased pressure pain intensity ratings, $\underline{d}$=0.64.[51] Cycle ergometry at a prescribed intensity of 60–75% of HRmax had no effect on pressure pain threshold, $\underline{d}$=0.01, and a moderate pain reducing effect on pain intensity, $\underline{d}$=0.55. In contrast, Vierck et al. reported that temporal summation of pain was increased following maximal treadmill exercise, with an effect size of - 1.59.[72] One study investigated CFS and found reduced PPTs following submaximal aerobic exercise, $\underline{d}$=−0.45.[48] Two studies examining CLB found pain reducing effects of submaximal cycle ergometry.[27,48] Meeus et al. reported a small hypoalgesic effect on PPTs, $\underline{d}$= 0.11, while Hoffman and colleagues reported a large hypoalgesic effect on pressure pain intensity ratings 2 and 32 minutes following exercise, $\underline{d}$=1.50 and 1.14, respectively. One study investigating CMP reported small to minimal effects of submaximal cycle ergometry on pressure and heat pain thresholds, with values of 0.07 and 0.31, respectively.[3]

Isometric Exercise

Table 7 presents the results for the four studies assessing EIH in chronic pain populations using isometric exercise. This table primarily shows that isometric exercise reduces pain perception for individuals with shoulder myalgia, but increases pain perception for individuals with FMS. Across chronic pain conditions, the average effect size for pain threshold was 0.40 (SD=1.43), while the unbiased effect size was 0.17. The average effect size for pain intensity was −1.94 (SD=0.36), with the unbiased effect size −1.92. Three studies assessed PPTs in individuals with FMS following[24] or during[30,46] isometric contractions, with an unbiased effect size of −0.20 (11 effects). Two of these studies also took threshold measures 10–15 minutes post isometric exercise, with values of 0.37 and 0.18 (8 effects) for $\underline{d}$ and d, respectively. One study of FMS patients measured pain intensity using thermal stimuli during isometric exercise and found large hyperalgesic effects on the contracting and contralateral forearms, with values of −1.68, and −2.2, respectively.[65] One study assessed EIH in individuals with should myalgia using pressure pain thresholds.[46] When subjects contracted the affected shoulder, PPTs assessed on that shoulder were lower indicating a hyperalgesic effect, $\underline{d}$=−0.94. However, a hypoalgesic effect (average effect size of 1.25) was found 1) when PPTs were assessed on resting muscles during contraction of the affected shoulder and 2) during contractions of the knee when PPTs were assessed at the contracting knee, resting knee, and affected shoulder.

Discussion

The impact of acute exercise on experimentally induced noxious stimulation was evaluated with meta-analytic techniques. Effect sizes were derived from studies that measured pain perception following or during aerobic, isometric, and dynamic resistance exercise. The results suggest that all three types of acute exercise reduce perception of experimentally induced pain in healthy participants, with the largest effect sizes found following isometric exercise. In addition, pain response measures of threshold and intensity ratings were similar in healthy adults, with threshold differences somewhat larger for isometric and dynamic resistance exercise and intensity differences larger for aerobic exercise. The size and direction of the effects for chronic pain conditions depended on the type of medical condition being studied.

Aerobic Exercise in Healthy Adults

The overall effect for aerobic EIH for pain threshold was moderate at 0.43 and somewhat larger for pain intensity ratings at 0.64. The magnitude of the effect was variable, ranging from 0.11 to 1.18 for intensity and from 0.04 to 1.47 for threshold. This broad range was likely a function of several factors including pain induction techniques and intensity and duration of exercise. Additionally, alterations in pain perception after exercise appeared to last up to 15 minutes post exercise[39], with trivial to small effects at 30 minutes post exercise.[29,62]

The average effect size for the four studies assessing pain threshold and/or intensity using pressure pain was moderate, with the results suggesting a dose response relationship between the intensity and duration of exercise and its hypoalgesic effect. The largest effect sizes were found when exercise was performed at a high intensity (i.e., 75% of VO_{2max}) and relatively longer duration (> 10 minutes). Thirty minutes of exercise performed at 50% $VO_{2\ max}$ produced a comparatively smaller effect, but still in the moderate range, while 10 minutes of high intensity exercise produced a small effect.[62] Given that this dose-response hypothesis is based on only a small number of effects, more work is needed to confirm this relationship and determine whether it applies to other pain stimuli.

The four studies using thermal stimulation showed considerable variability in the magnitude of the effect of exercise on pain perception, ranging from 0.04 – 1.17. Ruble et al. found small and trivial effects (0.04–0.20) of 30 minutes of aerobic exercise performed at 75% VO_{2max} when hot and cold thermal stimuli were delivered using a thermode placed on the thenar eminence of the hand.[62] In contrast, Sternberg et al. reported a moderate effect of 10 minutes of treadmill running at 85% VO_{2max} on intensity of cold pressor ratings.[67] However, this effect separated by gender revealed a large effect for women (0.88) and no effect for men (0.01). Additionally, Kemppainen et al. found moderate to large effects of 24–32 minutes of incremental cycling exercise using a cold pressor task in male fighter pilots without neck pain.[31] In contrast to Ruble's thermal heat results, Vierck et al. revealed a large effect of treadmill running to exhaustion on temporal summation of late pain responses to heated thermal stimulation.[72] Temporal summation of second pain is related to C-fiber mediated processes, whereas suprathreshold first pain measures are mediated by A-delta fibers.[71] Research has shown that exercise activates endogenous opioid mechanisms, and A-delta mediated pain is less susceptible to opioid inhibition.[64,71] As such, the source of nociceptive input may be a potentially important factor to consider when testing the effect of exercise on thermal pain responses. Furthermore, it has been suggested that the mixed results for thermal stimulation could be attributed to changes in skin and body temperature during exercise, causing hot and cold thresholds to be obtained at higher stimulation temperatures following exercise.[35,36,53] However, evidence has also shown that heat pain thresholds are not impacted by skin or body temperature.[35] Nevertheless, future research is needed to determine the magnitude of aerobic EIH with thermal stimulation techniques and whether the effect differs depending on the type of measure (i.e., first pain responses vs. second pain responses).

It should be noted that a substantial number of studies using aerobic exercise had to be excluded from this meta-analysis because of either of a lack of information to calculate effect size, not using intensity or threshold measures, or not standardizing exercise. All eight of the studies excluded for a lack of information to calculate effect sizes found a hypoalgesic effect of exercise (N=63) in healthy adults, with either an increase in pain thresholds or a decrease in pain ratings following cycling exercise. Seven out of eight of these studies found a reduction in pain using electrical dental pulp stimulation techniques. Thus, inclusion of these studies in this metaanalysis would likely have confirmed or even strengthened the hypoalgesic effect of aerobic exercise, while also extending it to an additional pain induction

technique. Additionally, three of the excluded studies showed attenuation of pain responses to heat and cold pressor pain following exercise.[10,66,67] For example, Sternberg and colleagues found reduced pain responses on a cold pressor test after participants competed in basketball, track, and fencing competitions.[66,67] Additionally, Robinson and Fuller found lower discriminability measures on a heat pain perception task following the completion of a 6 mile outdoor road course.[9] However, these studies did not control for the intensity and duration of exercise and the competition within the exercise bouts provided a potential confounding variable when determining the interaction between exercise and thermal pain perception.

In sum, aerobic exercise has shown to be an effective means to reduce pain perception in healthy adults among a variety of pain induction techniques. EIH appeared to be the strongest when exercise was performed at a moderate to high intensity pace. Additionally, hypoalgesia following exercise was found more consistently in studies that used pressure stimuli to produce pain compared to studies that used thermal stimulation. Due to the small number of studies, conclusions regarding differences in the magnitude of aerobic EIH among pain induction techniques remain tenable.

Isometric Exercise in Healthy Adults

The magnitude of EIH for isometric exercise was generally moderate to large for both threshold and intensity measures taken immediately after or during exercise, regardless of the contraction location, intensity or duration, as well as the pain induction stimulus and location. However, within the moderate to large effect size range, subtle patterns did emerge with the hypoalgesic effect tending to be larger for contractions at a low to moderate intensity held for longer durations. This finding was supported by Hoeger Bement et al. who investigated the dose response of isometric contractions on pain perception and found the greatest changes in pain threshold and intensity following long duration (i.e., until task failure, ~ 5–9 minutes), low intensity contractions compared to low intensity contractions held for a relatively shorter duration (2 minutes) and high intensity contractions held for 3–5 seconds.[22] During a longduration static muscle contraction, active motor units eventually become fatigued and higher threshold motor units become increasingly recruited to maintain the required force.[8,11] Thus, the authors explained their findings by suggesting that high-threshold motor units need to be recruited during isometric contractions to elicit a significant hypoalgesic response. However, this is likely not a complete explanation because other studies have found moderate to large hypoalgesic effects following contractions of shorter duration (i.e., 2 minutes or less).[40,65,68]

In regards to pain induction technique, only two studies have investigated changes in pain perception following isometric contractions using pain induction techniques other than pressure stimuli. The study employing electrical stimulation of the sural nerve found a moderate effect of low-intensity handgrip contractions held for 4–5 minutes in men.[61] Staud et al. found very large effects of low intensity handgrip contractions on pain ratings of 5 s supra-threshold heat stimuli applied to the forearm in women.[65] While the results of these two studies are promising, additional evidence is needed to confirm the efficacy of isometric contractions in producing hypoalgesia with experimental pain induction techniques other than pressure.

Several studies assessed the effect of isometric exercise on the contracting body part, as well as on the contralateral and a distant body part to the contracting one.[30,45,46,65] Importantly, the hypoalgesic effect of isometric exercise was multisegmental and not isolated to the contracting muscle. Moreover, the pain reducing effects of isometric exercise on the contralateral and distant body parts were similar in magnitude to the local body part. These results suggest that a central widespread inhibitory mechanism is activated by static muscle

contractions. As discussed by Kosek and Lundenburg, these central mechanisms may include increased secretion of β-endorphins, attention mechanisms, activation of diffuse noxious inhibitory controls, or an interaction of the cardiovascular and pain regulatory systems.[45]

Duration of the hypoalgesic effect of exercise has important implications for the use of exercise as a method to manage clinical pain symptoms. The data suggest that isometric contractions produce moderate to large pain reducing effects during the contraction and immediately following the contraction, with the effects attenuating over time. For example, Lannersten and Kosek assessed pain thresholds 10 minutes post contraction and EIH had almost completely dissipated in the non-contracting body areas, but moderate effects still existed in the contracting muscle.[46] Kosek and Lundenburg revealed small effects of isometric contractions in the contracting muscle 30 minutes post contraction and no effect in the non-contracting body areas.[45] This result is similar to the aerobic and dynamic resistance exercise literature showing no EIH 30 minutes after the cessation of exercise.[9,26,38,62]

Dynamic Resistance Exercise in Healthy Adults

Only two dynamic resistance exercise studies measuring threshold and/or intensity were included in the analysis.[9,38] Both measures showed large effect sizes when assessed one to five minutes after the dynamic resistance exercise session and small effects when assessed 15 minutes post exercise. Dynamic resistance exercise sessions were identical in each study, including 10 repetitions of four different exercises performed at 75% 1RM. As such, the threshold of dynamic resistance exercise required to produce EIH still needs to be determined. For example, would completion of only one of the exercises produce the same effect? Additionally, both studies used pressure stimuli to induce pain; therefore, whether EIH elicited by dynamic resistance exercise generalizes to other types of pain stimuli remains unknown. Importantly though, these studies showed that intermittent exercise, and not just continuous exercise, is capable of producing medium to large EIH effects.

EIH in Chronic Pain Populations

The effect sizes for pain threshold and intensity measures from studies examining EIH in chronic pain populations were highly variable for both aerobic and isometric exercise. The type of chronic pain condition partially explained this variability. For example, studies examining CLB found EIH effects similar to healthy individuals.[27,48] Meeus et al. even found that incremental cycle ergometry had pain reducing effects on PPTs at multiple body sites, including the back. Furthermore, Hoffman et al. demonstrated that this effect is still large 30 minutes post exercise.[27] In individuals with shoulder myalgia, isometric contractions of the quadriceps muscle elicited large hypoalgesic effects.[46] Indeed, PPTs assessed at the chronically painful shoulder even increased, with a large effect. However, during contractions of the shoulder with myalgia, PPTs were lower at that shoulder. These studies suggest that exercise of nonpainful muscles for individuals with regional chronic pain conditions produce a hypoalgesic effect and may be an effective method to temporarily relieve pain in painful muscles. Importantly, future research needs to determine the effects of acute exercise on pre-existing clinical pain.

Several studies indicated that moderate submaximal isometric exercise and vigorous aerobic exercise have a moderate to large hyperalgesic effect on experimental pain in FMS.[46,65,72] However, aerobic exercise performed at a preferred intensity or a prescribed moderate intensity elicited EIH in individuals with FM, with large to moderate effects.[51] Furthermore, submaximal isometric contractions performed at a low intensity (~10%) increased PPTs of the deltoid muscle in FM patients, also with a large effect.[30] These results suggest that EIH in FM patients may only be elicited in response to low to moderate intensity exercise, which

is in contrast to the results for healthy adults. However, additional studies are needed to confirm this hypothesis. Moderate to vigorous intensity aerobic exercise also had a moderate hyperalgesic effect on PPTs in individuals with CFS with chronic widespread pain[48] and minimal effects on heat and pressure thresholds in Gulf War veterans with CMP.[3] The mechanisms underlying exercise-induced hyperalgesia in response to moderate or vigorous exercise in these chronic widespread conditions remain unknown, but have been suggested to be caused by abnormal descending inhibition or excessive activation of muscle nociceptive afferents.[65,72]

Conclusions

The analysis from this study provides quantitative evidence to address the question of the magnitude of exercise-induced hypoalgesia in response to experimentally induced pain. We found the average effect size to range from moderate to large in healthy adults depending on pain induction method and exercise protocol. Importantly, all three types of exercise were capable of producing large effects in healthy adults, although the effects were generally transient. Also, while trends could be identified, the optimal dose of exercise that is needed to produce hypoalgesia could not be systematically determined with the amount of data available. We also found small to large EIH effects in individuals with regional chronic pain conditions at the painful muscle when a distant muscle was being exercised and in individuals with FMS when exercising at a low to moderate intensity. However, EIH was nonexistent in individuals with chronic widespread pain when exercising at a moderate to high intensity, with exercise often exacerbating experimental pain.

Although the exact mechanisms remain unknown, several have been proposed to explain exercise-induced hypoalgesia. Perhaps the most widely considered mechanism is that the activation of the endogenous opioid system during exercise reduces pain perception following exercise. Exercise of sufficient intensity and duration results in the release of peripheral and central beta-endorphins which have been associated with changes in pain sensitivity.[15,55,57,64] However, animal research has provided the most consistent support for this hypothesis[23,64], while the human data has been mixed.[6,18,31,52] Animal data also shows that non-opioid systems exist (e.g., endocannabinoid, neurotransmitters such as serotonin and norephinephrine) and that parameters of the exercise (i.e., duration of session, continuous vs. intermittent, and varying water temperature for swim protocols) may determine which system is activated.[5,28,50]

Another potential mechanism involves an interaction between pain modulatory and cardiovascular systems (See Koltyn & Umeda for a review).[40] For example, pain regulation and blood pressure control involve the same brain stem nuclei[47,74], neurotransmitters (e.g., monoamines) and neuropeptides (e.g., opioids).[12,60] Additionally, blood pressure and heart rate increase significantly during aerobic and isometric exercise and these elevations have been reported in conjunction with alterations in sensitivity to painful stimuli.[12,13,63] However, only a few studies have systematically tested the relationship between blood pressure and exerciseinduced hypoalgesia producing equivocal results.[61,68,69] Other potential mechanisms with mixed support include activation of ascending (e.g., activation of muscle afferent A delta and C fibers)[49] and descending (e.g., exercise acting as a distraction and altering attention away from the pain stimulus)[73] pain inhibition pathways by exercise. The conflicting evidence for the causal mechanisms of EIH illustrates the complexity of this phenomenon and suggests that EIH is likely caused by a combination of factors.

Experimental rigor is an important factor which can influence the magnitude of effect sizes, with poorly designed studies having the potential to inflate or yield smaller effects. An important study design characteristic includes the inclusion of a control condition. Few studies in this meta-analysis compared pain perception during an exercise condition to a

control condition. The aerobic exercise studies that included a control condition found greater effect sizes than those without a control condition, suggesting that this factor likely did not lead to the overestimation of the overall effect of aerobic exercise on pain perception. The two isometric exercise studies with a control comparison condition reported considerably smaller effects than the overall average effect size for isometric exercise. As such, it is essential that future studies include a resting control condition so that valid estimations of EIH elicited by isometric exercise can be estimated. Importantly, several studies did perform reliability testing and found small and trivial effects of repeated pain testing on the pain measures, indicating that the hypoalgesic effect was produced by exercise and not pain pre-testing.

Acknowledgments

Disclosures

This work was funded by National Institutes of Health training grant T32NS045551-06 (K.M.N.).

References

1. Bartholomew JB, Lewis BP, Linder DE, Cook DB. Post-exercise analgesia: replication and extension. J Sports Sci. 1996; 14:329–334. [PubMed: 8887212]
2. Bulbulian R. Endogenous opioid effects on motoneuron pool excitability: potential analgesic effect of acute exercise. J Manipulative Physiol Ther. 2002; 25:209–215. [PubMed: 12021739]
3. Cook DB, Stegner AJ, Ellingson LD. Exercise alters pain sensitivity in Gulf War veterans with chronic musculoskeletal pain. J Pain. 2010; 11:764–772. [PubMed: 20338824]
4. Dannecker EA, Koltyn KF, Riley JL 3rd, Robinson ME. The influence of endurance exercise on delayed onset muscle soreness. J Sports Med Phys Fitness. 2002; 42:458–465. [PubMed: 12391441]
5. Dietrich A, McDaniel WF. Endocannabinoids and exercise. Br J Sports Med. 2004; 38:536–541. [PubMed: 15388533]
6. Droste C, Meyer-Blankenburg M, Greenlee MW, Roskamm H. Effect of physical exercise on pain thresholds and plasma beta-endorphins in patients with silent and symptomatic myocardial ischemia. Eur Heart J. 1988; 9:25–33. [PubMed: 3246253]
7. Droste C, Greenlee M, Schreck M. Experimental pain thresholds and plasma beta-endorphin levels during exercise. Med Sci Sports Exerc. 1991; 23:334–42. [PubMed: 2020272]
8. Fallentin N, Jorgensen K, Simonsen EB. Motor unit recruitment during prolonged isometric contractions. Eur J Appl Physiol Occup Physiol. 1993; 67:335–341. [PubMed: 8299601]
9. Focht B, Koltyn K. Alterations in pain perception after resistance exercise performed in the morning and evening. J Strength Cond Res. 2009; 23:891–897. [PubMed: 19387388]
10. Fuller A, Robinson M. A test of exercise analgesia using signal detection theory and a within-subjects design. Percept Mot Skills. 1993; 76:1299–1310. [PubMed: 8337083]
11. Garland SJ, Enoka RM, Serrano LP, Robinson GA. Behavior of motor units in human biceps brachii during a submaximal fatiguing contraction. J Appl Physiol. 1994; 76:2411–2419. [PubMed: 7928865]
12. Ghione S. Hypertension-associated hypalgesia: evidence in experimental animals and humans, pathological mechanisms, and potential clinical consequences. Hypertension. 1996; 28:494–504. [PubMed: 8794839]
13. Ghione S, Rosa C, Panattoni E, Nuti M, Mezzasalma L, Giuliano G. Comparison of sensory and pain threshold in tooth pulp stimulation in normotensive man and essential hypertension. J Hypertens. 1985; (Supple 3):S113–S115.
14. Glesser, LJ.; Olkin. Stochastically dependent effect sizes. In: Cooper, H.; Hedges, LV., editors. The handbook of research synthesis. New York: Sage; 1994. p. 339-354.
15. Goldfarb AH, Jamurtas AZ. Beta-endorphin response to exercise. An update. Sports Med. 1997; 24:8–16. [PubMed: 9257407]

16. Guieu R, Blin O, Pouget J, Serratrice G. Nociceptive threshold and physical activity. Can J Neurol Sci. 1992; 19:69–71. [PubMed: 1562910]
17. Gurevich M, Kohn PM, Davis C. Exercise-induced analgesia and the role of reactivity in pain sensitivity. J Sports Sci. 1994; 12:549–559. [PubMed: 7853451]
18. Haier RJ, Quaid K, Mills JSC. Naloxone alters pain perception after jogging. Psychiatry Res. 1981; 5:231–232. [PubMed: 6945616]
19. Hedges L. Distribution theory for Glass's estimator of effect size and related estimators. J Educat Stat. 1981; 6:490–499.
20. Hedges L. Estimation of effect size from a series of independent experiments. Psychol Bull. 1982; 92:490–499.
21. Hoeger Bement M, Dicapo J, Rasiarmos R, Hunter S. Dose response of isometric contractions on pain perception in healthy adults. Med Sci Sports Exerc. 2008; 40:1880–1889. [PubMed: 18845975]
22. Hoeger Bement M, Rasiarmos RL, DiCapo JM, Lewis A, Keller ML, Harkins AL, Hunter SK. The role of the menstrual cycle phase in pain perception before and after an isometric fatiguing contraction. Eur J Appl Physiol. 2009; 106:105–112. [PubMed: 19189119]
23. Hoeger Bement M, Sluka KA. Low-intensity exercise reverses chronic muscle pain in the rat in a naloxone-dependent manner. Arch Phys Med Rehabil. 2005; 86:1736–1740. [PubMed: 16181935]
24. Hoeger Bement M, Wyer A, Hartley S, Drewek B, Harkins AL, Hunter SK. Pain perception after isometric exercise in women with fibromyalgia. Arch Phys Med Rehabil. 2011; 92:89–95. [PubMed: 21187210]
25. Hoeger Bement M, Wyer A, Hartley S, Yoon T, Hunter SK. Fatiguing exercise attenuates pain-induced corticomotor excitability. Neuroscience Letters. 2009; 452:209–213. [PubMed: 19383441]
26. Hoffman MD, Shepanski M, Ruble S, Valic Z, Buckwalter J, Clifford P. Intensity and duration threshold for aerobic exercise-induced analgesia to pressure pain. Arch Phys Med Rehabil. 2004; 85:1183–7. [PubMed: 15241771]
27. Hoffman MD, Shepanski MS, Mackenzie SP, Clifford PS. Experimentally induced pain perception is acutely reduced by aerobic exercise in people with chronic low back pain. J Rehabil Res Dev. 2005; 42:183–190. [PubMed: 15944883]
28. Hoffman P, Skarphedinsson JO, Delle M, Thoren P. Electrical stimulation of the gastrocnemius muscle in the spontaneously hypertensive rat increases the pain threshold: Role of different serotonergic receptions. Acta Physiol Scand. 1990; 138:125–131. [PubMed: 2138403]
29. Janal MN, Colt EWD, Clark WC, Glusman M. Pain sensitivity, mood and plasma endocrine levels in man following long distance running: Effects of Naloxone. Pain. 1984; 19:13–25. [PubMed: 6330643]
30. Kadetoff D, Kosek E. The effects of static muscular contraction on blood pressure, heart rate, pain ratings and pressure pain thresholds in healthy individuals and patients with fibromyalgia. European Journal of Pain. 2007; 11:39–47. [PubMed: 16480906]
31. Kemppainen P, Hämäläinen O, Könönen M. Different effects of physical exercise on cold pain sensitivity in fighter pilots with and without the history of acute in-flight neck pain attacks. Med Sci Sports Exerc. 1988; 30:577–582. [PubMed: 9565940]
32. Kemppainen P, Pertovaara A, Huopaniemi T, Johansson G, Karonen S. Modification of dental pain and cutaneous thermal sensitivity by physical exercise in man. Brain Res. 1985; 360:33–40. [PubMed: 3000534]
33. Kemppainen P, Pertovaara A, Huopaniemi T, Johansson G. Elevation of dental pain threshold induced in man by physical exercise is not reversed by cyproheptadine-mediated suppression of growth hormone release. Neurosci Lett. 1986; 70:388–92. [PubMed: 3774237]
34. Kemppainen P, Paalasmaa P, Pertovaara A, Alila A, Johansson G. Dexamethasone attenuates exercise-induced dental analgesia in man. Brain Res. 1990; 519:329–32. [PubMed: 2168784]
35. Kojo I, Pertovaara A. The effects of stimulus area and adaptation temperature on human heat pain and warm thresholds. Int J Neurosci. 1987; 32:875–880. [PubMed: 3596930]
36. Koltyn K. Analgesia following exercise. Sports Med. 2000; 29:85–98. [PubMed: 10701712]

37. Koltyn K. Exercise-induced hypoalgesia and intensity of exercise. Sports Med. 2002; 32:477–87. [PubMed: 12076175]
38. Koltyn K, Arbogast R. Perception of pain after resistance exercise. Br J Sports Med. 1998; 32:20–24. [PubMed: 9562159]
39. Koltyn K, Garvin A, Gardiner R, Nelson T. Perception of pain following aerobic exercise. Med Sci Sport Exerc. 1996; 28:1418–21.
40. Koltyn K, Trine M, Stegner A, Tobar D. Effect of isometric exercise on pain perception and blood pressure in men and women. Med Sci Sport Exerc. 2001; 33:282–90.
41. Koltyn K, Umeda M. Exercise, hypoalgesia, and blood pressure. Sports Med. 2006; 36:207–214. [PubMed: 16526833]
42. Koltyn K, Umeda M. Contralateral attenuation of pain after short-duration submaximal isometric exercise. J Pain. 2007; 8:887–92. [PubMed: 17681886]
43. Kosek E, Ekholm J. Modulation of pressure pain thresholds during and following isometric contractions. Pain. 1995; 61:481–486. [PubMed: 7478692]
44. Kosek E, Ekholm J, Hansson P. Modulation of pressure pain thresholds during and following isometric contractions in patients with fibromyalgia and in healthy controls. Pain. 1996; 64:415–423. [PubMed: 8783304]
45. Kosek E, Lundberg L. Segmental and plurisegmental modulation of pressure pain thresholds during static muscle contractions in healthy individuals. Eur J Pain. 2007; 7:251–8. [PubMed: 12725848]
46. Lannersten L, Kosek E. Dysfunction of endogenous pain inhibition during exercise with painful muscles in patients with shoulder myalgia and fibromyalgia. Pain. 2010; 151:77–86. [PubMed: 20621420]
47. Lovick TA. Integrated activity of cardiovascular and pain regulatory systems: role in adaptive behavioral responses. Prog Neurobiol. 40:631–644. [PubMed: 8484005]
48. Meeus M, Roussel N, Truijen S, Nijs J. Reduced pressure pain thresholds in response to exercise in Chronic Fatigue Syndrome but not in Chronic Low Back Pain: An experimental study. J Rehab Med. 2010; 42:884–90.
49. Mense, S.; Simons, DG. Muscle Pain: Understanding its nature, diagnosis, and treatment. Philadelphia, PA: Lippincott Williams & Wilkins; 2001. p. 1-385.
50. Mogil JS, Belknap JK. Sex and genotype determine the selective activation of neurochemically distinct mechanisms of swim stress-induced analgesia. Pharmacol Biochem Behav. 1997; 56:61–66. [PubMed: 8981610]
51. Newcomb LW, Koltyn K, Morgan WP, Cook DB. Influence of preferred versus prescribed exercise on pain in fibromyalgia. Med Sci Sports Exerc. 2011; 43:1106–1113. [PubMed: 21085031]
52. Olausson B, Eriksson E, Ellmarker L. Effects of naloxone on dental pulp pain threshold following muscle exercise and low frequency transcutaneous nerve stimulation: A comprehensive study in man. Acta Physiol Scand. 1986; 126:299–305. [PubMed: 3486546]
53. Paalasmaa P, Kemppainen P, Pertovaara A. Modulation of skin sensitivity by dynamic and isometric exercise in man. Eur J Appl Physiol. 1991; 62:279–285.
54. Padawer WJ, Levine FM. Exercise-induced analgesia: fact or artifact? Pain. 1992; 48:131–135. [PubMed: 1589230]
55. Paulev PE, Thorboll JE, Nielsen U, Kruse P, Jordal R, Bach FW, Fenger M, Pokorski M. Opioid involvement in the perception of pain due to endurance exercise in trained man. Jpn J Physiol. 1989; 38:507–517. [PubMed: 3148778]
56. Pertovaara A, Huopaniemi T, Virtanen A, Johansson G. The influence of exercise on dental pain thresholds and the release of stress hormones. Physiol Behav. 1984; 33:923–6. [PubMed: 6100393]
57. Pertovaara A, Kemppainen P, Johansson G, Karonen SL. Ischemic pain nonsegmentally produces a predominant reduction of pain and thermal sensitivity in man: a selective role for endogenous opioids. Brain Res. 1982; 251:83–92. [PubMed: 6293648]
58. Portney, LG.; Watkins, MP. Foundations of clinical research: applications to practice. East Norwalk, CT: Appleton and Lang; 1993.

59. Price D, Von der Gruen A, Miller J, Rafii A, Price C. A psychophysical analysis of morphine analgesia. Pain. 1985; 22:261–269. [PubMed: 2993984]
60. Randich A, Maixner W. Interactions between cardiovascular and pain regulatory systems. Ann N Y Acad Sci. 1986; 467:371–384. [PubMed: 3524385]
61. Ring C, Edwards L, Kavussanu M. Effects of isometric exercise on pain are mediated by blood pressure. Biol Psychol. 2008; 78:123–128. [PubMed: 18316153]
62. Ruble S, Hoffman M, Shepanski M, Valic Z, Buckwalter J, Clifford P. Thermal pain perception after aerobic exercise. Arch Phys Med Rehabil. 2005; 86:1019–23. [PubMed: 15895351]
63. Schobel HP, Handwerker HO, Schmieder RE, Heusser K, Dominiak P, Luft FC. Effects of naloxone on hemodynamic and sympathetic nerve responses to pain in normotensive vs. borderline hypertensive men. J Auton Nerv Syst. 1998; 69:49–55. [PubMed: 9672123]
64. Stagg NJ, Mata HP, Ibrahim MM, Henriksen EJ, Porreca F, Vanderah TW, Malan PT. Regular exercise reverses sensory hypersensitivity in a rat neuropathic pain model: role of endogenous opioids. Anesthesiology. 2011; 114:940–948. [PubMed: 21386701]
65. Staud R, Robinson M, Price D. Isometric exercise has opposite effects on central pain mechanisms in fibromyalgia patients compared to normal controls. Pain. 2005; 118:176–184. [PubMed: 16154700]
66. Sternberg W, Bailin D, Grant M, Gracely RH. Competition alters the perception of noxious stimuli in male and female athletes. Pain. 1998; 76:231–238. [PubMed: 9696478]
67. Sternberg W, Bokat C, Kass L, Alboyadjian A, Gracely R. Sex-dependent components of the analgesia produced by athletic competition. J Pain. 2001; 2:65–74. [PubMed: 14622787]
68. Umeda M, Newcomb L, Ellingson L, Koltyn K. Examination of the dose-response relationship between pain perception and blood pressure elevations induced by isometric exercise. Biol Psychol. 2010; 85:90–96. [PubMed: 20594949]
69. Umeda M, Newcomb L, Koltyn K. Influence of blood pressure elevations by isometric exercise on pain perception in women. Int J Psychophysiol. 2009; 74:45–52. [PubMed: 19616592]
70. Vecchiet L, Marini I, Colozi A, Feroldi P. Effects of aerobic exercise on muscular pain sensitivity. Clin Therapist. 1984; 6:354–363.
71. Vierck CJ, Cannon RL, Fry G, Maixner W, Whitsel BL. Characteristics of temporal summation of second pain sensations elicited by brief contact of glabrous skin by a preheated thermode. J Neurophysiol. 1997; 78:992–1002. [PubMed: 9307129]
72. Vierck C, Staud R, Price D, Cannon R, Mauderli A, Martin A. The effect of maximal exercise on temporal summation of second pain (windup) in patients with Fibromyalgia Syndrome. J Pain. 2001; 2:334–44. [PubMed: 14622813]
73. Villemure C, Bushnell MC. Cognitive modulation of pain: how do attention and emotion influence pain processing? Pain. 2002; 95:195–199. [PubMed: 11839418]
74. Zamir N, Maixner W. Interactions between cardiovascular and pain regulatory systems. Ann NY Aca Sci. 1986; 467:371–384.

Perspective: This article presents a quantitative review of the exercise-induced hypoalgesia literature. This review raises several important questions that need to be addressed while also demonstrating that acute exercise has a hypoalgesic effect on experimentally-induced pain in healthy adults, and both a hypoalgesic and hyperalgesic effect in adults with chronic pain.

Table 1

Studies examining pain perception following acute bouts of aerobic exercise in healthy participants

Author, year	Sample Size (M/F)	Pain induction stimulus/location	Mode of exercise	Exercise intensity	Exercise Duration	Pooled ES (intensity, threshold)
Gurevich et al., 1994	15/0	Pressure/index finger	step exercise	75% VO_{2max}	12 minutes	0.92, NA
Koltyn et al., 1996	14/2	Pressure/forefinger	cycle ergometer	65–75% VO_{2max}	30 minutes	1.18, 1.47
Kemppainen et al., 1988	8/0	Cold pressor/hand	cycle ergometer	Incremental 50–200 Watt	24–32 minutes	1.17, 0.46
Sternberg et al., 2001	10/10	Cold pressor/arm	Treadmill	85% VO_{2max}	10 minutes	0.64, NA
Vierck et al., 2001	10/10	Heat thermal/glabrous skin of hand	Treadmill	To exhaustion	11–12 minutes	0.83, NA
Hoffman et al., 2004	5/7	Pressure/index finger	Treadmill	75% VO_{2max}	10 minutes	0.08, NA
Hoffman et al., 2004	5/7	Pressure/index finger	Treadmill	75% VO_{2max}	30 minutes	0.65, NA
Hoffman et al., 2004	5/7	Pressure/index finger	Treadmill	50% VO_{2max}	30 minutes	0.51, NA
Ruble et al. 2005	6/8	Thermal heat/glabrous skin of hand	Treadmill	75% VO_{2max}	30 minutes	0.20, 0.04
Ruble et al., 2005	6/8	Thermal cold/glabrous skin of hand	Treadmill	75% VO_{2max}	30 minutes	0.11, 0.04
Meeus et al., 2010	21/10	Pressure/arm, hand, back, calf	cycle ergometer	Incremental 20–130 Watt	22–29 minutes	NA, 0.23

Note. NA=Not available.

Table 2

Summary of pain threshold effect sizes by exercise and pain stimuli type in healthy participants

Pain response measure	Number of studies	Subjects	Mean effect size	Effect size SD	Unbiased effect size[a]
Aerobic					
Pressure	2	47	0.84	0.89	0.58
Thermal heat	1	10	0.04		0.04
Thermal cold	2	18	0.34	0.43	0.30
All aerobic Threshold	4	57	0.48	0.61	0.43
Isometric					
Pressure	9	222	1.27	0.78	1.05
Dynamic resistance					
Pressure	2	4	0.86	0.18	0.83

[a]Weighted by sample size.

Table 3

Summary of pain intensity effect sizes by exercise and pain stimuli type in healthy participants

Pain response measure	Number of studies	Subjects	Mean effect size	Effect size SD	Unbiased effect size[a]
Aerobic					
Pressure	3	43	0.67	0.42	0.69
Thermal heat	2	38	0.55	0.40	0.59
Thermal cold	3	42	0.81	0.80	0.61
All aerobic Intensity	7	105	0.68	0.50	0.64
Isometric					
Pressure	5	140	0.81	0.76	0.73
Thermal heat	1	12	1.35		1.35
Electrical	1	24	0.40		0.40
All isometric Intensity	7	176	0.83	0.71	0.72
Dynamic resistance					
Pressure	2	34	0.83	0.37	0.75

[a]Weighted by sample size.

J Pain. Author manuscript; available in PMC 2013 December 01.

Table 4

Studies examining pain perception following acute bouts of isometric exercise in healthy participants

Author, year	Sample Size (M/F)	Pain induction stimulus/location	Exercise location	Exercise intensity	Exercise Duration	Pooled ES (intensity, threshold)
Koltyn et al., 2001	15/16	Pressure/forefinger	hand grip	max	5 seconds	0.83, 0.93
Koltyn et al., 2001	15/16	Pressure/forefinger	hand grip	40–50% MVC	2 minutes	0.84, 1.28
Kosek & Lundberg, 2003	12/12	Pressure/contracting MQ	knee extension	1 kg on ankle	Exhaust (~12 min)	NA, 1.77
Kosek & Lundberg, 2003	12/12	Pressure/resting MQ	knee extension	1 kg on ankle	Exhaust (~12 min)	NA, 1.68
Kosek & Lundberg, 2003	12/12	Pressure/contralateral MI	knee extension	1 kg on ankle	Exhaust (~12 min)	NA, 0.99
Kosek & Lundberg, 2003	12/12	Pressure/contracting MI	elbow flexion	1 kg on wrist	Exhaust (~12 min)	NA, 1.48
Kosek & Lundberg, 2003	12/12	Pressure/resting MI	elbow flexion	1 kg on wrist	Exhaust (~12 min)	NA, 2.64
Kosek & Lundberg, 2003	12/12	Pressure/contralateral MQ	elbow flexion	1 kg on wrist	Exhaust (~12 min)	NA, 1.12
Staud et al., 2005	0/11	Heat thermode/ip forearm	hand grip	30%MVC	90 seconds	1.21, NA
Staud et al., 2005	0/11	Heat thermode/co forearm	hand grip	30%MVC	90 seconds	1.48, NA
Kadetoff & Kosek, 2007	0/17	Pressure/MQ	knee extension	39 N, ~10%MVC	Exhaust (~10 min)	NA, 2.0
Kadetoff & Kosek, 2007	0/17	Pressure/deltoideus	knee extension	39 N, ~10%MVC	Exhaust (~10 min)	NA, 2.2
Koltyn & Umeda, 2007	0/14	Pressure/ ip forefinger	hand grip	40–50% MVC	2 minutes	2.81, 0.99
Koltyn & Umeda, 2007	0/14	pressure/ co forefinger	hand grip	40–50% MVC	2 minutes	1.59, 1.09
Hoeger Bement et al., 2008	11/11	Pressure/ index finger	elbow flexor	25% MVC	task failure	0.68, 0.85
Hoeger Bement et al., 2008	11/11	Pressure/ index finger	elbow flexor	25% MVC	2 minutes	0.22, 0.32
Hoeger Bement et al., 2008	11/11	Pressure/ index finger	elbow flexor	80% MVC	task failure	0.36, 0.27
Hoeger Bement et al., 2008	11/11	Pressure/ index finger	elbow flexor	Max	2–3 seconds	0.31, 0.51
Ring et al., 2008	24/0	Electrical/Sural nerve	hand grip	15% MVC	4.5 minutes	0.31, NA
Ring et al., 2008	24/0	Electrical/Sural nerve	hand grip	25% MVC	4.5 minutes	0.49, NA
Hoeger Bement et al., 2009	0/20	Pressure/ index finger	elbow flexor	25% MVC	task failure	NA, 0.72
Umeda et al., 2009	0/23	Pressure/forefinger	handgrip	25% MVC	1 minute	−0.16, 0.33
Umeda et al., 2009	0/23	Pressure/forefinger	handgrip	25% MVC	3 minutes	0.11, 0.54
Umeda et al., 2010	25/25	Pressure/forefinger	handgrip	25% MVC	1 minute	0.94, 0.70
Umeda et al., 2010	25/25	Pressure/forefinger	handgrip	25% MVC	3 minutes	0.95, 0.76
Umeda et al., 2010	25/25	Pressure/forefinger	handgrip	25% MVC	5 minutes	1.06, 0.57
Lannersten & Kosek, 2010	0/21	Pressure/contracting MI	shoulder rotation	20–25% MVC	5 minutes	NA, 2.56
Lannersten & Kosek, 2010	0/21	Pressure/resting MI	shoulder rotation	20–25% MVC	5 minutes	NA, 1.85

Author, year	Sample Size (M/F)	Pain induction stimulus/location	Exercise location	Exercise intensity	Exercise Duration	Pooled ES (intensity, threshold)
Lannersten & Kosek, 2010	0/21	Pressure/contralateral MQ	shoulder rotation	20–25% MVC	5 minutes	NA, 1.09
Lannersten & Kosek, 2010	0/21	Pressure/contracting MQ	knee extension	20–25% MVC	5 minutes	NA, 1.61
Lannersten & Kosek, 2010	0/21	Pressure/resting MQ	knee extension	20–25% MVC	5 minutes	NA, 2.05
Lannersten & Kosek, 2010	0/21	Pressure/contralateral MI	knee extension	20–25% MVC	5 minutes	NA, 1.60

Note. M=males; F=females; ES= effect size; MQ= m. quadriceps; MI=m. infraspinatus; ip=ipsilateral; co=contracting; min=minutes; NA=Not available

Table 5

Studies examining pain perception following acute bouts of dynamic resistance exercise in healthy participants

Author, year	Sample Size (M/F)	Pain induction stimulus/location	Mode of exercise	Exercise intensity	Exercise Duration	Pooled ES (intensity, threshold)
Koltyn & Arbogast	7/6	Pressure/middle finger	bench press, leg press pull downs, arm ext	3 sets of 10, 75% 1RM	45 minutes	1.08, 0.99
Focht & Koltyn, 2009	21/0	Pressure/middle finger	bench press, leg press pull downs, arm ext	3 sets of 10, 75% 1RM	45 minutes	0.56, 0.74

Note. RM=repetition max; ext = extensions; NA=Not available

Table 6

Studies examining pain perception following aerobic exercise in chronic pain populations

Author, year	Participants (M/F, pain cond.)	Pain induction stimulus/location	Mode of exercise	Exercise intensity	Exercise Duration	Pooled ES (intensity, threshold)
Newcomb et al., 2011	0/21, FMS	Pressure/index finger	cycle ergometer	Self-selected	20 minutes	0.45, 0.79
Newcomb et al., 2011	0/21, FMS	Pressure/index finger	cycle ergometer	60–75% HR_{max}	20 minutes	0.39, 0.01
Vierck et al., 2001	0/10, FMS	Heat thermal/glabrous skin of hand	Treadmill	To exhaustion	11–12 minutes	−1.13, NA
Meeus et al., 2010	21/5, CFS	Pressure/arm, hand, back, calf	cycle ergometer	Incremental 20–130 Watt	22–29 minutes	NA, −0.32
Meeus et al., 2010	11/10, CLB	Pressure/arm, hand, back, calf	cycle ergometer	Incremental 20–130 Watt	22–29 minutes	NA, 0.08
Hoffman et al. 2005	4/4, CLB	Pressure/index finger	cycle ergometer	50–70% VO_{2max}	25 minutes	1.50, NA
Cook et al. 2010	15/0, CMP	Heat thermal/glabrous skin of hand	cycle ergometer	70% VO_{2max}	30 minutes	NA, 0.31
Cook et al. 2010	15/0, CMP	Pressure/middle finger	cycle ergometer	70% VO_{2max}	30 minutes	NA, 0.07

Note. Cond=condition; FMS=Fibromyalgia Syndrome; CFS=Chronic Fatigue Syndrome; CLB=Chronic Low Back Pain; M=males; F=females; ES= effect size; NA=Not available; CMP=Chronic musculoskeletal pain.

Table 7

Studies examining pain perception following isometric exercise in chronic pain populations

Author, year	Participants (M/F, pain cond.)	Pain induction stimulus/location	Mode of exercise	Exercise intensity	Exercise Duration	Pooled ES (intensity, threshold)
Lannersten & Kosek, 2010	0/20, sh myalgia	Pressure/contracting MI	iso. shoulder rotation	20–25% MVC	5 minutes	NA, −0.67
Lannersten & Kosek, 2010	0/20, sh myalgia	Pressure/resting MI	iso. shoulder rotation	20–25% MVC	5 minutes	NA, 0.15
Lannersten & Kosek, 2010	0/20, sh myalgia	Pressure/contralateral MQ	iso. shoulder rotation	20–25% MVC	5 minutes	NA, 0.68
Lannersten & Kosek, 2010	0/20, sh myalgia	Pressure/contracting MQ	iso. knee extension	20–25% MVC	5 minutes	NA, 0.96
Lannersten & Kosek, 2010	0/20, sh myalgia	Pressure/resting MQ	iso. knee extension	20–25% MVC	5 minutes	NA, 1.48
Lannersten & Kosek, 2010	0/20, sh myalgia	Pressure/contralateral MI	iso. knee extension	20–25% MVC	5 minutes	NA, 1.62
Lannersten & Kosek, 2010	0/20, FMS	Pressure/contracting MI	iso. shoulder rotation	20–25% MVC	5 minutes	NA, −0.55
Lannersten & Kosek, 2010	0/20, FMS	Pressure/resting MI	iso. shoulder rotation	20–25% MVC	5 minutes	NA, −0.95
Lannersten & Kosek, 2010	0/20, FMS	Pressure/contralateral MQ	iso. shoulder rotation	20–25% MVC	5 minutes	NA, −0.46
Lannersten & Kosek, 2010	0/20, FMS	Pressure/contracting MQ	iso. knee extension	20–25% MVC	5 minutes	NA, −0.97
Lannersten & Kosek, 2010	0/20, FMS	Pressure/resting MQ	iso. knee extension	20–25% MVC	5 minutes	NA, −0.05
Lannersten & Kosek, 2010	0/20, FMS	Pressure/contralateral MI	iso. knee extension	20–25% MVC	5 minutes	NA, −0.14
Staud et al. 2005	0/12, FMS	Heat thermode/ip forearm	hand grip	30% MVC	90 seconds	−1.68, NA
Staud et al. 2005	0/12, FMS	Heat thermode/co forearm	hand grip	30% MVC	90 seconds	−2.20, NA
Hoeger Bement et al. 2011	0/15, FMS	Pressure/index finger	iso elbow flexor	25% MVC	Task failure	NA, 0.15
Hoeger Bement et al. 2011	0/15, FMS	Pressure/index finger	iso elbow flexor	25% MVC	2 minutes	NA, −0.36
Hoeger Bement et al. 2011	0/15, FMS	Pressure/index finger	iso elbow flexor	Max	3–5 seconds	NA, 0.02
Kadetoff & Kosek, 2007	0/17, FMS	Pressure/MQ	knee extension	39 N, ~15%MVC	Exhaust (~8 min)	NA, 1.16
Kadetoff & Kosek, 2007	0/17, FMS	Pressure/deltoideus	knee extension	39 N, ~15%MVC	Exhaust (~8 min)	NA, 2.7

Note. Cond=condition;. iso= isometric; FMS=Fibromyalgia Syndrome;M=males; F=females; ES= effect size; MQ= m. quadriceps; MI=m. infraspinatus; ip=ipsilateral; co=contracting; min=minutes; NA=Not available; sh=shoulder.

Cochrane Library

Cochrane Database of Systematic Reviews

Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews (Review)

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH.
Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews.
Cochrane Database of Systematic Reviews 2017, Issue 4. Art. No.: CD011279.
DOI: 10.1002/14651858.CD011279.pub3.

www.cochranelibrary.com

TABLE OF CONTENTS

[Overview of Reviews]

Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews

Louise J Geneen[1], R Andrew Moore[2], Clare Clarke[3], Denis Martin[4], Lesley A Colvin[5], Blair H Smith[1]

[1]Division of Population Health Sciences, University of Dundee, Dundee, UK. [2]Pain Research and Nuffield Department of Clinical Neurosciences (Nuffield Division of Anaesthetics), University of Oxford, Oxford, UK. [3]Division of Population Health Sciences, University of Dundee, Dundee, UK. [4]Institute of Health and Social Care, Teesside University, Middlesbrough, UK. [5]Anaesthesia & Pain Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK

Contact address: Louise J Geneen, Division of Population Health Sciences, University of Dundee, Dundee, UK. louise.geneen@dph.ox.ac.uk.

Editorial group: Cochrane Pain, Palliative and Supportive Care Group.
Publication status and date: Edited (no change to conclusions), published in Issue 4, 2017.

Citation: Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. *Cochrane Database of Systematic Reviews* 2017, Issue 4. Art. No.: CD011279. DOI: 10.1002/14651858.CD011279.pub3.

ABSTRACT

Background

Chronic pain is defined as pain lasting beyond normal tissue healing time, generally taken to be 12 weeks. It contributes to disability, anxiety, depression, sleep disturbances, poor quality of life, and healthcare costs. Chronic pain has a weighted mean prevalence in adults of 20%.

For many years, the treatment choice for chronic pain included recommendations for rest and inactivity. However, exercise may have specific benefits in reducing the severity of chronic pain, as well as more general benefits associated with improved overall physical and mental health, and physical functioning.

Physical activity and exercise programmes are increasingly being promoted and offered in various healthcare systems, and for a variety of chronic pain conditions. It is therefore important at this stage to establish the efficacy and safety of these programmes, and furthermore to address the critical factors that determine their success or failure.

Objectives

To provide an overview of Cochrane Reviews of adults with chronic pain to determine (1) the effectiveness of different physical activity and exercise interventions in reducing pain severity and its impact on function, quality of life, and healthcare use; and (2) the evidence for any adverse effects or harm associated with physical activity and exercise interventions.

Methods

We searched the *Cochrane Database of Systematic Reviews* (CDSR) on the Cochrane Library (CDSR 2016, Issue 1) for systematic reviews of randomised controlled trials (RCTs), after which we tracked any included reviews for updates, and tracked protocols in case of full review publication until an arbitrary cut-off date of 21 March 2016 (CDSR 2016, Issue 3). We assessed the methodological quality of the reviews using the AMSTAR tool, and also planned to analyse data for each painful condition based on quality of the evidence.

We extracted data for (1) self-reported pain severity, (2) physical function (objectively or subjectively measured), (3) psychological function, (4) quality of life, (5) adherence to the prescribed intervention, (6) healthcare use/attendance, (7) adverse events, and (8) death.

Due to the limited data available, we were unable to directly compare and analyse interventions, and have instead reported the evidence qualitatively.

Main results

We included 21 reviews with 381 included studies and 37,143 participants. Of these, 264 studies (19,642 participants) examined exercise versus no exercise/minimal intervention in adults with chronic pain and were used in the qualitative analysis.

Pain conditions included rheumatoid arthritis, osteoarthritis, fibromyalgia, low back pain, intermittent claudication, dysmenorrhoea, mechanical neck disorder, spinal cord injury, postpolio syndrome, and patellofemoral pain. None of the reviews assessed 'chronic pain' or 'chronic widespread pain' as a general term or specific condition. Interventions included aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi.

Reviews were well performed and reported (based on AMSTAR), and included studies had acceptable risk of bias (with inadequate reporting of attrition and reporting biases). However the quality of evidence was low due to participant numbers (most included studies had fewer than 50 participants in total), length of intervention and follow-up (rarely assessed beyond three to six months). We pooled the results from relevant reviews where appropriate, though results should be interpreted with caution due to the low quality evidence.

Pain severity: several reviews noted favourable results from exercise: only three reviews that reported pain severity found no statistically significant changes in usual or mean pain from any intervention. However, results were inconsistent across interventions and follow-up, as exercise did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point.

Physical function: was the most commonly reported outcome measure. Physical function was significantly improved as a result of the intervention in 14 reviews, though even these statistically significant results had only small-to-moderate effect sizes (only one review reported large effect sizes).

Psychological function and quality of life: had variable results: results were either favourable to exercise (generally small and moderate effect size, with two reviews reporting significant, large effect sizes for quality of life), or showed no difference between groups. There were no negative effects.

Adherence to the prescribed intervention: could not be assessed in any review. However, risk of withdrawal/dropout was slightly higher in the exercising group (82.8/1000 participants versus 81/1000 participants), though the group difference was non-significant.

Healthcare use/attendance: was not reported in any review.

Adverse events, potential harm, and death: only 25% of included studies (across 18 reviews) actively reported adverse events. Based on the available evidence, most adverse events were increased soreness or muscle pain, which reportedly subsided after a few weeks of the intervention. Only one review reported death separately to other adverse events: the intervention was protective against death (based on the available evidence), though did not reach statistical significance.

Authors' conclusions

The quality of the evidence examining physical activity and exercise for chronic pain is low. This is largely due to small sample sizes and potentially underpowered studies. A number of studies had adequately long interventions, but planned follow-up was limited to less than one year in all but six reviews.

There were some favourable effects in reduction in pain severity and improved physical function, though these were mostly of small-to-moderate effect, and were not consistent across the reviews. There were variable effects for psychological function and quality of life.

The available evidence suggests physical activity and exercise is an intervention with few adverse events that may improve pain severity and physical function, and consequent quality of life. However, further research is required and should focus on increasing participant numbers, including participants with a broader spectrum of pain severity, and lengthening both the intervention itself, and the follow-up period.

PLAIN LANGUAGE SUMMARY

Physical activity and exercise for chronic pain in adults - an overview of Cochrane Reviews

Background

Chronic (long-term) pain is pain that has lasted beyond the body's usual healing time. It is often described as pain that has lasted for at least three months. Chronic pain causes many problems, beyond the pain itself, including fatigue, anxiety, depression, and a poor quality of life.

In the past, people with chronic pain were told to rest. However, general advice now is to keep active - whether to affect the pain directly or to combat the other problems associated with it. Therefore, research studies have attempted to examine the effect of physical activity in people with chronic pain.

This overview aimed to bring together and analyse any reviews published by Cochrane that looked at physical activity and exercise studies in any chronic pain condition, including arthritis, back and neck pain, and menstrual (period) pain.

Key results and quality of the evidence

In January 2016, we identified 21 Cochrane Reviews which covered 10 different diagnoses (osteoarthritis (a joint disease), rheumatoid arthritis (joint pain and swelling), fibromyalgia (widespread pain condition), low back pain, intermittent claudication (cramping pain in the legs), dysmenorrhoea (period pain), mechanical neck disorders (neck pain), spinal cord injury, postpolio syndrome (a condition occurring in people who have had polio), patellofemoral pain (pain at the front of the knee)). The physical activity or exercise programme used in the trials ranged in frequency, intensity, and type, including land- and water-based activities, those focusing on building strength, endurance, flexibility and range of motion, and muscle activation exercises.

The quality of the evidence was low. This was mostly due to the small numbers of people with chronic pain who participated in each reviewed study. Ideally, a study should have hundreds of people assigned to each group, whereas most of the studies included in the review process here had fewer than 50 people in total.

There was evidence that physical activity reduced the severity of pain, improved physical function, and had a variable effect on both psychological function and quality of life. However, these results were not found in all studies. The inconsistency could be due to the quality of the studies or because of the mix of different types of physical activity tested in the studies. Additionally, participants had predominantly mild-to-moderate pain, not moderate-to-severe pain.

Conclusions

According to the available evidence (only 25% of included studies reported on possible harm or injury from the intervention), physical activity did not cause harm. Muscle soreness that sometimes occurs with starting a new exercise subsided as the participants adapted to the new activities. This is important as it shows physical activity in general is acceptable and unlikely to cause harm in people with chronic pain, many of whom may have previously feared it would increase their pain further.

Future studies should focus on increasing participant numbers, including a wider range of severity of pain (more people with more severe pain), and lengthening both the intervention (exercise programme) itself, and the follow-up period. This pain is chronic in nature, and so a long-term intervention, with longer periods of recovery or follow-up, may be more effective.

BACKGROUND

Description of the condition

Chronic pain has been defined as pain lasting beyond normal tissue healing time, generally taken to be 12 weeks (International Association for the Study of Chronic Pain; Merskey 2011). It contributes to disability, anxiety and depression, sleep disturbances, poor quality of life, and healthcare costs (Leadley 2014; Moore 2014a; Park 2012).

Chronic pain has a weighted mean prevalence in adults of 20% (Breivik 2006; Moore 2014a), which increases as the population ages (32% of adults aged 25 to 34 years, 62% of adults over 75

years; Abdulla 2013; Elliott 1999). This is a greater proportion than people with asthma (To 2012) or diabetes (IDF 2012) in the same population (van Hecke 2013a). The World Health Organization (WHO) recognises chronic pain as a public health problem throughout the world, with one systematic review assessing the growing evidence that the prevalence of chronic pain in the general population is high internationally (34% in low-income countries and 30% in high-income countries; Elzahaf 2012). Chronic painful conditions comprise four of the 10 highest ranking conditions for years lived with disability in 2013 (Vos 2015), and are responsible for considerable loss of quality of life and employment, and increased healthcare costs (Moore 2014b). Despite this, the term 'chronic pain' was only added as a MeSH term in MEDLINE in January 2012 (National Library of Medicine), highlighting the relatively small proportion of specific research dedicated to this population.

Certain factors can contribute to an increased risk of chronic pain (female gender, older age, lower socioeconomic status, geographical and cultural background, and genetics; Smith 2007; van Hecke 2013b). Other factors associated with chronic pain conditions are modifiable, such as smoking status, alcohol intake, nutrition, obesity, comorbidities, employment status and occupational factors, and physical activity level (Smith 2007; van Hecke 2013a).

A review of current issues in the treatment of chronic pain strongly suggests that health professionals traditionally focus on biomedical views of pain, utilising pharmacology first and foremost, and sometimes not addressing potential non-pharmacological approaches such as physical activity and changing attitudes towards chronic pain (Schofield 2011). Guidance often suggests that lifestyle advice is important: for example, the National Institute for Health and Care Excellence (NICE) osteoarthritis guidelines state that "exercise should be a core treatment ... irrespective of age, comorbidity, pain severity and disability. Exercise should include: local muscle strengthening [and] general aerobic fitness" (NICE 2014). Non-pharmacological treatments have been developed, investigated, and implemented, with Cochrane Reviews and protocols evaluating the available evidence for psychological, physical, and other non-medical interventions (e.g. cognitive behavioural and behavioural therapy, Eccleston 2014; Williams 2012; TENS, Nnoaham 2008; low-impact/intensity movement/exercise therapy, Wieland 2013; dietary, Straube 2015; and patient education, Engers 2008; Gross 2009). While evidence for the effectiveness of these interventions is of variable quantity and quality, the 2013 Scottish Intercollegiate Guideline Network (SIGN) guidelines on the management of chronic pain made strong recommendations on the use of exercise, based on evidence drawn from randomised controlled trials (RCTs), stating: "exercise and exercise therapies, regardless of their form, are recommended in the management of patients with chronic pain" (SIGN 2013).

Description of the interventions

Physical activity has been defined by the WHO as "any bodily movement produced by skeletal muscles that requires energy expenditure, including activities undertaken while working, playing, carrying out household chores, travelling, and engaging in recreational pursuits" (WHO 2015). WHO also states that "exercise ... is a sub-category of physical activity that is planned, structured, repetitive, and aims to improve or maintain one or more components of physical fitness" (WHO 2015).

Physical activity for health can take many different forms: it can be structured exercise, such as in classes, gym-based, or a DVD or programme performed at home; or unstructured and involve adding just a few small activities each day (activities of daily living). Physical activity and exercise can also vary in intensity, duration, and type: aerobic (such as walking) or more focused on increasing flexibility, strength, or balance. Physical activity and exercise can also be taught (or led) by another individual such as an exercise professional, or initiated and maintained through the person's own initiative and motivation.

Both physical activity and exercise can be performed on land or in the water, and can range from whole-body to localised (body site-specific) training. Most forms of exercise can also be modified to be performed where there is restricted movement (e.g. in a chair, a bed, or another assistive device).

How the intervention might work

Physical activity and exercise can be adapted for an individual, and is something people can do to help themselves. It is likely to be associated with minimal adverse effects, such as interactions with medication and potential for abuse in adults with chronic pain, when compared to pharmaceutical and surgical interventions. It is therefore an attractive option to help manage an individual's pain if the systematic reviews show benefit. However, current evidence suggests that simply giving an individual advice to exercise is insufficient to bring about significant change (SIGN 2013), and a badly prescribed intervention that does not consider the individual's conditions and present state of health and fitness, such as one that does not incorporate pacing or gradual progression, may bring about adverse events such as pain 'flare-ups', or lead to cardiac or respiratory events (American College of Sports Medicine 2007). This suggests that supervised or structured interventions may be more fruitful, though this is currently unconfirmed.

Since the 1980s, primary care physician advice for treating pain has changed, moving away from "rest", to minimising or eliminating bedrest and instead remaining active (back pain, Waddell 1987). Exercise may have specific benefits in reducing the severity of chronic pain, as well as more general benefits associated with improved overall physical and mental health, and physical functioning of people with chronic pain, as depression (Finan 2013), deconditioning (Bousema 2007), and obesity are commonly observed in these people (headache/migraine, Bigal 2012; fibromyalgia, Ursini 2011). For example, studies have revealed that a sin-

gle bout of exercise increases the production of endogenous opioids, leading to transient anti-nociception in both animals and humans, and repeated exercise produces long-lasting anti-nociception in otherwise untreated animals (Stagg 2011). Aerobic exercise is also strongly linked to weight loss (Messier 2013), which in turn has implications for the management of chronic pain as the pressure on joints is reduced. Alternatively, resistance exercise, or other forms of strength training, can improve the person's capacity to support bone and cartilage through improved musculature supporting movement around a joint, with potential to relieve stiffness (Mayer 2008) and bringing about some pain relief. Resistance training through repetitive full range-of-motion exercise around the lumbar spine (in chronic low back pain) may affect disc metabolism itself, with the possibility that the exercise programme could improve metabolic exchange in the lumbar discs and aid in repair (Mooney 2006). Training to improve balance and flexibility also has benefits as it reduces the risk of falls, and the potential for further pain or injury (Harvard 2013).

Why it is important to do this overview

If physical activity and exercise interventions are shown to effectively and safely reduce pain intensity or frequency (or both), they are likely to be a preferable alternative or adjunct therapy to pharmacological/surgical treatments for chronic pain. The interventions could promote personal involvement of individuals in the management of their pain, thus increasing self-efficacy and the ability to self-manage. In turn this could lead to an increase in overall quality of life and a consequent reduction in healthcare use. In addition, exercise is of great importance for cardiovascular (Vigorito 2014) and bone health (Sakuma 2012). Reduced physical function and consequent lack of mobility in people with chronic pain is associated with increased all-cause and cardiovascular mortality (Nüesch 2011), with other studies linking severe chronic pain to general increased all-cause mortality (Moore 2014a; Torrance 2010).

Physical activity and exercise programmes are increasingly being promoted and offered in various healthcare systems (American College of Sports Medicine (ACSM) 'Exercise is Medicine' global pledge at the Inaugural World Congress 2010) and for a variety of chronic pain conditions, including arthritis (Fransen 2014; Silva 2010), fibromyalgia (Busch 2013), and dysmenorrhoea (Brown 2010). At this stage it is important to establish the efficacy and safety of these programmes, and furthermore to address the critical factors that determine their success or failure.

It is therefore important to identify whether (and how) exercise interventions can be effectively and safely applied in people with chronic pain.

With a number of systematic reviews published by Cochrane evaluating the effectiveness of exercise in various painful conditions, it is timely and important to bring together all relevant published information to evaluate the current evidence, and identify the availability and quality of evidence-based exercise interventions. This overview will determine the extent to which the published systematic reviews have accurately assessed the evidence for exercise in chronic pain conditions/syndromes, which will help to direct future guidelines and identify current research gaps.

OBJECTIVES

To provide an overview of Cochrane Reviews of adults with chronic pain to determine (1) the effectiveness of different physical activity and exercise interventions in reducing pain severity and its impact on function, quality of life, and healthcare use; and (2) the evidence for any adverse effects or harm associated with physical activity and exercise interventions.

METHODS

Criteria for considering reviews for inclusion

We included only systematic reviews of RCTs of physical activity and exercise in participants with chronic pain, and published in the *Cochrane Database of Systematic Reviews*. The included reviews had to fulfil the following criteria:

Participants

Adults (aged 18 years and over) reporting chronic non-cancer pain, including persistent (e.g. chronic back pain, fibromyalgia) and intermittent (e.g. migraine, dysmenorrhoea) pain, for at least three months (12 weeks) in any body site.

Intervention

Reviews of RCTs assessing physical activity or exercise as the intervention (any reviews where that assessed physical activity or exercise as a stand-alone intervention). This included physical activity interventions that could be initially taught by an exercise professional, or involve periodical/ongoing supervision.

Exclusions

Interventions not deemed physical activity or exercise using the WHO definition, such as manipulation, mobilisation, or passive movement. Any multi-modal interventions were excluded if physical activity/exercise could not be assessed for effect (the effect of exercise must have been measured distinctly).

Comparison

Usual care, waiting list control, placebo/sham treatment, other treatment, or a combination of treatments (as long as the effect of exercise could be measured distinctly).

Primary outcome

- self-reported pain (severity).

This could be presented and analysed as change on a continuous scale, the proportion of participants who 'responded', or, ideally, in a dichotomised format as the proportion of participants in each group who achieved a predetermined threshold of improvement (e.g. outcome in individual participants of at least 50% pain intensity reduction, or no worse than mild pain, at the end of the trial, with at least 30% pain intensity reduction as a secondary outcome, or recovery; Moore 2013).

Secondary outcomes

- Physical function (objectively or subjectively measured).
- Psychological function.
- Quality of life.
- Adherence to the prescribed intervention.
- Healthcare use/attendance.
- Adverse events (not death).
- Death.

Reviews may not always report specifically on activity or exercise for chronic pain in adults. We anticipated two possible circumstances which might have arisen.

- A review included some interventions of interest or reported only some outcomes of interest. In this case we extracted the interventions and outcomes of interest, but we did not include interventions or outcomes outside the scope of this overview.
- Reviews occasionally included papers that included children and adults together, but the results for adults were not reported or analysed separately in the included papers or the review. In this case we made a judgement as to whether the review could be included based on the proportion of adults. Our intention was to include only those reviews where more than 80% of participants were adults.

Search methods for identification of reviews

We searched the *Cochrane Database of Systematic Reviews* (CDSR), 2016, Issue 1, on the Cochrane Library for relevant reviews using the search strategy: (*pain or migraine or headache*) and (*exercise or activity or physical*). We did not seek non-Cochrane reviews.

Data collection and analysis

Two overview authors (LG, CC) independently carried out searches and selected reviews for inclusion. Disagreements were resolved through discussion, and a third overview author (RAM) acted as arbitrator where necessary.

Two overview authors (independently carried out assessment of methodological quality (LG, CC), and extracted data (LG, RAM). Any disagreements were resolved through discussion, or involving a third overview author if necessary (DM).

One overview author (LG) tracked results of the search for the most up to date version of each review and protocol that fulfilled the inclusion criteria.

Selection of reviews

Included reviews assessed RCTs of the effects of exercise for pain management in adults (as defined by individual reviews), compared with any of the listed comparators, and included:

- a clearly defined clinical question;
- details of inclusion and exclusion criteria;
- details of databases searched and relevant search strategies;
- participant-reported pain severity (primary outcome measure);
- summary results for at least one other desired outcome.

Data extraction and management

Two overview authors (LG, RAM) independently extracted data from the included review using a standardised data extraction form and checked for agreement prior to entry into Microsoft Excel for Windows. We did not extract data from reports included in the reviews again, neither did we undertake any re-analysis of data from reviews. Data were not entered for analysis into Cochrane's statistical software due to the lack of relevant and comparable data (RevMan 2014).

We collected the following information (where available) from the reviews:

- number of included studies and participants;
- intervention (exercise or activity type) and dose (frequency/intensity);
- comparator;
- condition treated;
- time of assessment;
- duration of follow-up;
- relevant outcomes.

Where possible we extracted risk ratio (RR), number needed to treat for an additional beneficial outcome (NNTB), mean difference (MD), and standardised mean difference (SMD), and other relevant statistical data for the primary and secondary outcomes. This included:

- obtaining 50% pain relief (participant-reported);

- obtaining any other measure of 'improvement' (participant-reported);
- adverse events;
- death;
- withdrawals.

Assessment of methodological quality of included reviews

Quality of included reviews

Two overview authors (LG, CC) independently assessed each included review to see if it satisfied the criteria specified in the 'assessment of multiple systematic reviews' (AMSTAR) measurement tool (Shea 2007), for rigorous methodological quality. Arbitration by a third overview author (DM) was necessary for some fields. High quality reviews were required to fulfil each of the established AMSTAR criteria (further criteria to fulfil each field is listed in Table 1).

For each review we also planned to assess the likelihood of publication bias by calculating the number of participants in studies with zero effect (relative benefit of one) that would be needed to give an NNTB too high to be clinically relevant (Moore 2008). In this case we would have considered an NNTB of 10 or greater for the outcome of participant-reported pain relief of 30% or greater to be the cut-off for clinical relevance. This method is used as statistical tests for the presence of publication bias have been shown to be unhelpful (Thornton 2000). However, assessment of publication bias was not possible due to the lack of specificity of the populations included within the reviews, and so we were unable to extract comparable data.

Quality of evidence in included reviews

We planned to use two main indicators for the quality of evidence: all included reviews must have used only primary studies that were both randomised and double-blind, so minimising the risk of bias from these items; and all included reviews must have included only people with at least moderate pain intensity at baseline (visual analogue scale greater than 30/100, categorical rating scale greater than 1/3, and numerical rating scale greater than 3/10, Collins 1997), providing a sensitive assay of intervention efficacy. Subsequently, we planned to analyse data for each painful condition in three tiers, according to outcome and freedom from known sources of bias.

- The first tier used data meeting current best standards, where studies reported the outcome of at least 50% pain intensity reduction from baseline (where 50% was the cut-off for a dichotomous (yes/no) outcome: was a 50% reduction in pain observed?), or its equivalent, without using last observation carried forward (LOCF) or other imputation method for dropouts, reported an intention-to-treat (ITT) analysis, lasted eight or more weeks, had a parallel-group design, and had at least 200 participants (preferably at least 400) in the comparison (Moore 2010). These top-tier results were usually reported first.
- The second tier used any available data, but where one or more of these conditions were not met, for example reporting at least 30% pain intensity reduction, using LOCF or a completer analysis, lasting four to eight weeks, and where the numbers of participants were at least 200.
- A third tier of evidence related to small amounts of data (fewer than 200 participants), or short studies of less than four weeks, or where there was obvious major heterogeneity between studies, or where there were other shortcomings in allocation concealment, considerable attrition, and incomplete outcome data. For this third tier of evidence, no data synthesis was reasonable, and may have been misleading, but an indication of beneficial effects might be possible.

This overview examined the quality of all included reviews according to current best standards for reporting in pain. These included the attempt and ability of the reviews to identify studies/interventions with the maximum evidence of effectiveness, and minimum risk of bias, including the reporting of the following.

- Outcomes in trials of the proportion of participants obtaining at least 50% pain intensity reduction, or no worse than mild pain, at the end of the trial (with at least 30% pain intensity reduction as a secondary outcome). We did not consider the use of mean changes in pain scores as high quality because responses to pain interventions are not Gaussian, and few people have the mean response.
- Duration of included studies of eight weeks or longer.
- Imputation method of baseline observation carried forward (BOCF), LOCF, or worst observation carried forward (WOCF) if adverse event withdrawals were similar in active and control groups.
- At least 200 participants per treatment group in included studies, with at least two trials, as a minimum criterion for trustworthiness of any analysis. Pooled analysis of small studies may be considered good quality if at least 400 participants were involved, but we regarded these as being potentially subject to bias.

We extracted the 'Risk of bias' as assessed by the original review authors from included reviews. Counts of low risk of bias were extracted from relevant studies in the included reviews and tabulated under the following headings to evaluate the proportion of studies achieving a low risk of bias for each:

- random sequence generation (selection bias);
- allocation concealment (selection bias);
- blinding of participants and personnel (performance bias);
- blinding of outcome assessment (detection bias);
- incomplete outcome data (attrition bias);
- selective reporting (reporting bias);
- sample size;

- any other biases.

Data synthesis

Additional quantitative analyses were not required, since we only considered results from properly conducted (Cochrane) reviews. The aim was to concentrate on specific outcomes such as the proportion of participants with at least 50% pain relief, all-cause or adverse event discontinuations, or serious adverse events, and to explore how these can be compared across different treatments for the same condition. We planned to compare only like with like (where possible); for example in study duration, which can be an additional source of bias if insufficient in length (Moore 2010).
However due to the limited data available, we were unable to directly compare and analyse interventions, and have instead reported the evidence qualitatively only. We had also planned to employ subgroup analyses assessing age, condition, and intervention type/intensity, though this was not feasible using the available data from included reviews. For this reason we have also been unable to include a 'Summary of findings' table as planned and stated in the protocol.
Importantly, we have tried to highlight issues of low trial quality, inadequate size, and whether trials were truly valid for the particular condition in making between-therapy comparisons.
We approached each review with four main questions/focus, and extracted data accordingly.

- Did they report exercise versus non-exercise studies?
- Did the review or studies included in the review (or both) have low risk of bias?
- Did they have our main outcome?
- What were the actual intervention/s included in the review?

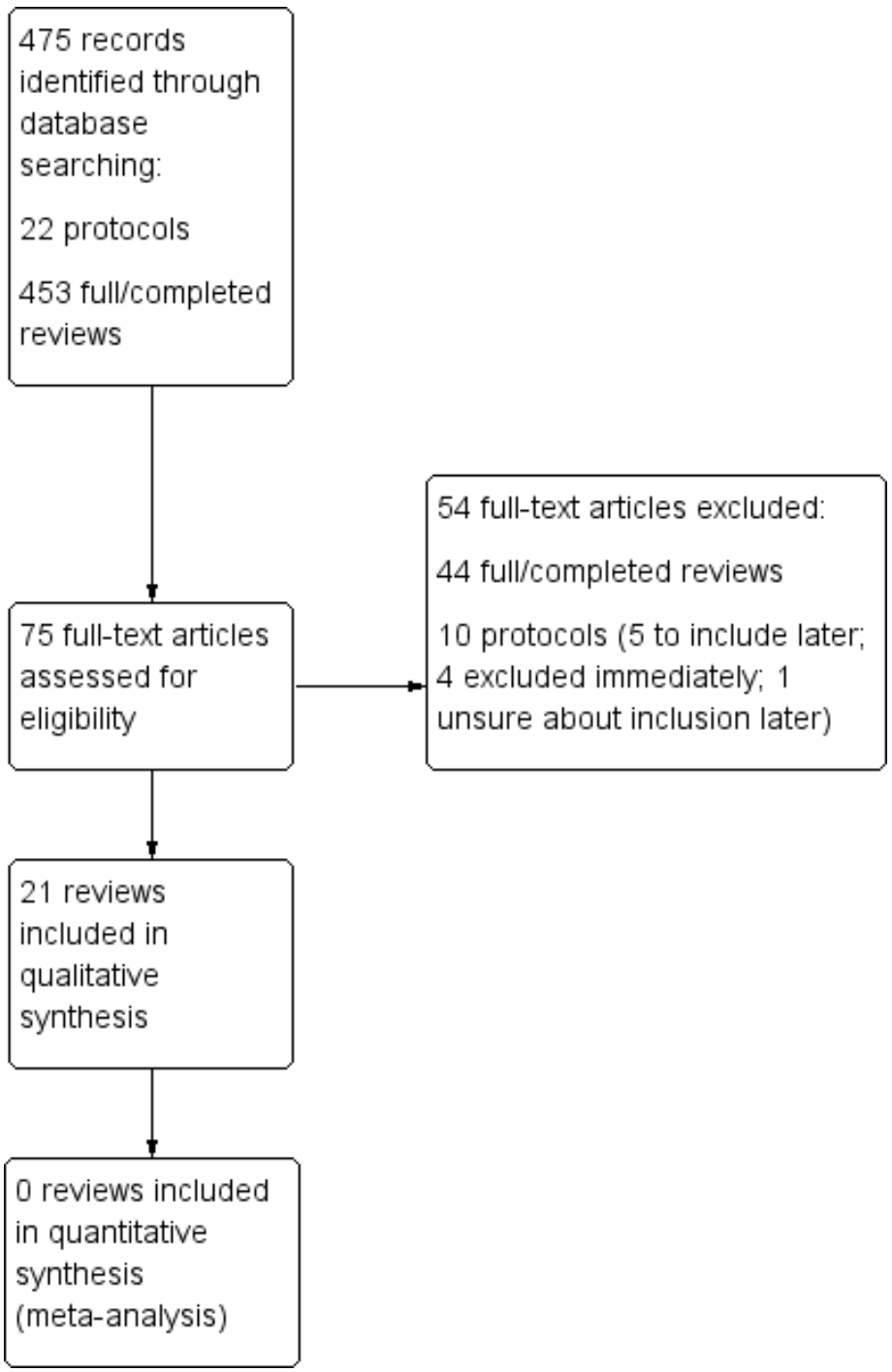
RESULTS

We included 21 reviews with 381 included studies, totalling 37,143 participants. Of these, 264 studies (19,642 participants) examined exercise versus no exercise/minimal intervention in adults with chronic pain (the focus of this overview) and so were used in the qualitative analysis.

Description of included reviews

The search strategy was performed in the Cochrane Library only, and revealed 475 potentially relevant titles, of which 75 were assessed as full papers.
The search was undertaken on 31 January 2016 (CDSR 2016, Issue 1), after which any included reviews were tracked for updates, and protocols were followed in case of full review publication until 21 March 2016 (CDSR 2016, Issue 3).
All extracted data and methodological quality assessment were taken from the most recent published version of the full review.
Ultimately, of the 75 titles requiring further assessment, 10 were reviews at protocol stage only (five of which have potential to be included once published as a full review, one which was unclear, and four that were excluded based on information within the protocol). Hence, we excluded 54 titles (10 protocols and 44 full reviews; Figure 1), reasons for which are listed in Table 2.

Figure 1. Study flow diagram.



Detailed information about the included reviews is available in Table 3. Trial and participant number, age, and gender distribution is reported in Table 4.

Specificity of chronic pain condition of included reviews

Following abstract and full paper assessment, 21 reviews fulfilled the inclusion criteria: four in rheumatoid arthritis (Cramp 2013; Han 2004; Hurkmans 2009; Silva 2010), four in osteoarthritis (Bartels 2007; Fransen 2014; Fransen 2015; Regnaux 2015), three in fibromyalgia (Bidonde 2014; Busch 2007; Busch 2013), three in low back pain (Hayden 2005; Saragiotto 2016; Yamato 2015), two in intermittent claudication (Lane 2014; Lauret 2014), one in dysmenorrhoea (Brown 2010), one in mechanical neck disorder (Gross 2015a), one in spinal cord injury (Boldt 2014), one in postpolio syndrome (Koopman 2015), and one in patellofemoral pain (van der Heijden 2015). None of the included reviews assessed 'chronic pain' or 'chronic widespread pain' as a general term or specific condition.

The 21 included reviews were published by five different Cochrane Review groups: 11 from the Cochrane Musculoskeletal Group (Bartels 2007; Bidonde 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Han 2004; Hurkmans 2009; Regnaux 2015; Silva 2010); four from the Cochrane Neck and Back Group previously the Cochrane Back Group) (Gross 2015a; Hayden 2005; Saragiotto 2016; Yamato 2015); two from the Cochrane Peripheral Vascular Diseases Group (Lane 2014; Lauret 2014); one from the Cochrane Menstrual Disorders and Subfertility Group (Brown 2010); one from the Cochrane Injuries Group (Boldt 2014); one from the Cochrane Neuromuscular Group (Koopman 2015); and one from the Cochrane Bone, Joint and Muscle Trauma Group (van der Heijden 2015).

Protocols that may be included in updates of this overview focus on osteoarthritis (Østerås 2013 from the Cochrane Musculoskeletal Group), migraine (Brønfort 2015 from the Cochrane Pain, Palliative and Supportive Care Group), chronic low back pain (Hayden 2012 from the Cochrane Back Group), ankylosing spondylitis (Regnaux 2014 from the Cochrane Musculoskeletal Group), and temporomandibular disorders (Craane 2006 from the Cochrane Oral Health Group).

Exercise and physical activity interventions implemented in the included reviews

Interventions assessed included: any specified style of land-based exercise or physical activity such as one designed to improve strength, range of movement, aerobic capacity, or a combination of these (Boldt 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hurkmans 2009; Koopman 2015; Regnaux 2015; van der Heijden 2015); a single style of land-based exercise only (tai chi only: Han 2004, walking only: Lauret 2014, walking or jogging only: Brown 2010; Lane 2014, balance training only: Silva 2010, motor control exercise only: Saragiotto 2016, Pilates method only: Yamato 2015); any pool-based or aquatic therapy (Bartels 2007; Bidonde 2014; Cramp 2013), or "any exercise therapy" (Hayden 2005).

Aquatic exercise

Any exercise performed in water. This can include swimming, though many studies will be referring to exercises performed vertically in the water (not horizontally), either using the water to support the body through the exercise, or as resistance against the body.

Range of motion and flexibility exercise

Can be performed in water or on land. The intention is to increase the range of motion around a joint through progressive stretching and mobilising of the muscles around and crossing the joint. For the purposes of this overview, we only included active movement where the movement was brought about by the participant, and not passively moved by an external force such as a therapist.

Aerobic exercise

Can be performed in water or on land. Exercise usually performed continuously to raise the heart rate and breathing rate for a prolonged period. Examples include walking, jogging, running, cycling, and swimming. Often presented as a percentage of the participant's heart rate max (HRmax) - the highest heart rate reached when performing at their absolute maximum. Similarly it may be presented as a percentage of VO_2max or VO_2peak (a proportion of the maximum amount of oxygen the muscle can take up per minute), or as an absolute value (mL/kg/minute).

Strength/resistance exercise

Can be performed in water or on land. Exercise performed against a progressive resistance with the intention of improving muscle strength, muscle endurance, muscle power, or a combination of these. Resistance can come from fixed or free weights, elastic bands, body weight (against gravity), and water resistance. It may also involve static or isometric strength (holding a position or weight without moving against it). Often presented as a percentage of the participant's one repetition maximum (1-RM) - the maximum weight they can lift/move if they only have to do it once.

Motor control exercise

Can be performed in water or on land. Exercise to bring about activation of the deep trunk muscles, targeting the restoration of control and co-ordination of these 'core muscles' (Saragiotto 2016).

Balance (proprioceptive) training

Can be performed in water or on land (water may be used initially for support). Exercise emphasises the maintenance of balance during visual and perturbation challenges with eyes open or closed, range of motion, and maintaining stability over reduced areas of support and unstable surface (Silva 2010), that is improving balance in increasingly unstable situations.

Tai chi

An ancient Chinese discipline developed from martial arts, involving a continuous series of very controlled (and usually slow) movements designed to improve physical and mental wellbeing.

Yoga

Arising out of Hindu philosophy. Exercise includes breath control, simple meditation, and the adoption of specific bodily postures. It is widely practised for health, relaxation, and control (physically and mentally). Incorporates stretching and flexibility training with isometric strength training (holding certain poses, with no movement against a resistance).

Pilates

Developed by Joseph Pilates in the 20th Century, it is a system of exercises (often using special apparatus) designed to improve physical strength, flexibility, and posture, and enhance mental awareness.

Duration and dose (frequency/intensity) of the exercise and physical activity interventions

A detailed breakdown of each review can be seen in Table 5.

Duration of intervention

Interventions assessed by the included reviews varied in length from a single session (Fransen 2015) to 30 months (Fransen 2015). Only five reviews enforced a minimum intervention period to reduce risk of bias, and were able to attribute any effects to the intervention (Brown 2010; Busch 2013; Gross 2015a; Hurkmans 2009; Silva 2010).

Frequency

There was large variation in the exercise or physical activity intervention being implemented, ranging from just once a week (Bidonde 2014; Busch 2007; Fransen 2014; Fransen 2015; Han 2004; Saragiotto 2016), to twice a day (Boldt 2014), and some performing a short series of exercises (two-minute duration) every 15 minutes during the day (Gross 2015a). However, when reported, most included studies in the reviews implemented the programme twice a week (or stated at least twice a week).

Intensity

Few studies quantified the intensity of each session. Baseline intensity was often accepted as low/moderate, with the aim to progress over the intervention period to 70% to 85% of HRmax or heart rate reserve (HRR) for aerobic interventions (Brown 2010; Cramp 2013; Hurkmans 2009), 70% to 80% of an individual's 1-RM, or 50% to 70% maximum voluntary contraction (Koopman 2015) in strength/resistance training programmes (Busch 2013; Hurkmans 2009). In other reviews, intensity was described more loosely as "variable" or "low intensity (very light) to maximum effort (vigorous)" (Bidonde 2014; Fransen 2014; Lane 2014; Regnaux 2015), "low intensity" (Fransen 2014; Gross 2015a; Han 2004; Silva 2010), or "moderate or moderate-to-high" (Cramp 2013; Fransen 2015).

Duration (per session)

Individual sessions varied in length from two minutes (Gross 2015a), to 90 minutes (Busch 2013; Cramp 2013; Han 2004) or 120 minutes (Boldt 2014), but mostly situated around 45 to 60 minutes. However, it is important to note that the shorter sessions were often performed more regularly than longer sessions. With more information it would have been possible to calculate total volume of exercise or physical activity (session duration × frequency per week × number of weeks), for a more accurate and detailed analysis.

Intervention specificity for chronic pain in the included reviews

The focus of this overview was exercise versus no-exercise interventions with the intention of answering the original question: is exercise beneficial, detrimental, or ineffective for people with chronic pain when compared to inactivity? Two of the 21 reviews did not include/locate any studies that examined simply exercise versus no exercise (Lauret 2014; Silva 2010). However, many of the included reviews compared varying exercise modality, duration, intensity, and frequency. The "no-exercise" intervention referred to the control group where there was a minimal intervention (such as sham exercise or education) or wait-list control/no treatment (see Table 3 for more information on control group activity).

Time points reported

Four of the 19 reviews that reported data, reported results at a single time point only ('post-intervention': Bidonde 2014; Busch 2007; Cramp 2013; Han 2004). Reviews also analysed outcome measures immediately post-intervention and at one or more follow-up points. Each review defined short-, intermediate-, and long-term follow-up according to their own assessment, so when the time period was not mentioned explicitly, we grouped the reviews according to the review authors' own classification only, and where a time period (weeks, month, years) was explicitly listed but not defined by the authors, we grouped them as short-term (follow-up as under six months), intermediate-term (six to 12 months), and long-term (longer than 12 months): short-term: Busch 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Lane 2014; Regnaux 2015; Saragiotto 2016; intermediate-term: Bartels 2007; Fransen 2015; Gross 2015a; Hayden 2005; Lane 2014; Regnaux 2015; Saragiotto 2016; long-term: Gross 2015a; Hayden 2005; Regnaux 2015; Saragiotto 2016. Five reviews did not report "post-intervention" but at short-term, mid/intermediate-term, and long-term postrandomisation (short, mid, and long term: Boldt 2014; short and intermediate term: Koopman 2015; Yamato 2015; short and long-term: Hurkmans 2009; van der Heijden 2015). One review assessed participants in an ongoing fashion "over three menstrual cycles" (Brown 2010).

Long-term follow-up

Of the seven reviews claiming to report "long term" follow-up, one classed long-term as longer than six weeks (intermediate term as one to six weeks' follow-up) (Boldt 2014). The remaining six reviews defined long-term follow up as over 12 months (one year) post-intervention (Gross 2015a; Hayden 2005; Hurkmans 2009; Regnaux 2015; Saragiotto 2016; van der Heijden 2015).

Methodological quality of included reviews

AMSTAR quality assessment of included reviews

No review achieved a perfect score of 11/11, though five achieved 10/11 (Boldt 2014; Busch 2013; Hayden 2005; Koopman 2015; Regnaux 2015) and eight scored 9/11 (Cramp 2013; Gross 2015a; Hurkmans 2009; Lane 2014; Lauret 2014; Saragiotto 2016; van der Heijden 2015; Yamato 2015). The lowest score was 6/11 (Silva 2010) though five categories were not applicable (n/a) due to there being no included studies. Quality assessment results for each individual review are presented in Table 6.

All reviews except one (Bidonde 2014) fulfilled the basic criteria (questions one to three of Table 1); to follow an 'a priori' design as Cochrane implements a system of protocol publication before undertaking the full reviews, where it also specifies dual study selection and data extraction from a comprehensive literature search. One review did not fulfil the 'a priori' design as this was an update and separation from a broader review series, and so the criteria had not been explicitly listed prior to publication for this specific title (Bidonde 2014).

Criteria which scored badly using the AMSTAR tool were characteristics of included studies (question six of Table 1), reporting of publication bias (question 10 of Table 1), and conflict of interest declarations (question 11 of Table 1).

- Included study characteristics were limited, often reporting the "inclusion criteria" used to recruit participants in the study instead of the characteristics of actual included participants, and excluding information such as participants' age, gender split, ethnicity, and disease status.

- Assessment of publication bias was omitted entirely in five reviews (Bartels 2007; Fransen 2014; Fransen 2015; Han 2004; Hurkmans 2009), and when it was assessed, it was reported using only a simple statement (with no test values, analyses used, or diagrams to demonstrate the result; Busch 2007; Koopman 2015). Two reviews mentioned in the methods as planned analyses, though was not mentioned again (Brown 2010; van der Heijden 2015), and a third review mentioned it in the methods, but appeared to use it interchangeably with reporting bias causing great confusion (Bidonde 2014).

- Conflicts of interest were sufficiently reported in only three out of 21 of the included reviews (Hayden 2005; Koopman 2015; Silva 2010). In the remaining reviews, a cursory statement was commonly made regarding the review authors' conflicts of interests, however, fulfilling the AMSTAR criteria also requires a statement to be made regarding any conflict of interest for any of the included studies.

Risk of bias in included reviews

The original review authors assessed risk of bias (see Table 7). The table shows the number of studies assessed as low risk of bias only, and excluded those that were assessed as unclear or high risk of bias.

Selection bias (randomisation and allocation concealment)

Selection bias had the largest proportion of included studies with low risk of bias (63% and 42% of studies adequately undertaking and reporting the methods used).

Performance and detection bias (blinding participants, personnel, outcome assessors)

With any exercise or physical activity intervention it is very difficult to blind both participants and personnel to the allocation, though some studies included in reviews attempted to by offering sham exercise.

Due to the difficulty of blinding participants to their group allocation, review authors assessed the risk of bias in different ways,

which may cause confusion: whereas the majority declared this lack of possible blinding to be high risk of bias or unclear, two reviews labelled such cases as low risk of bias in order not to exclude these studies unnecessarily from their analysis (Lane 2014; Lauret 2014). Without these two reviews, only a small percentage (7.8% or 18/229) of the included studies would have scored low risk of performance bias (blinding of participants and personnel), but by including them (all 35 studies from those two reviews assessed as low risk of bias) the overall proportion of studies assessed as having low risk of bias was closer to 20% (53/264).

Attrition (incomplete outcome data, withdrawals/dropouts)

About 55% (144/264) of the studies included in these reviews showed low risk of bias.

Reporting bias (selective reporting)

Reporting bias was classed as low risk in only 46% of included studies. However, it is important to note this was not due to the remainder having high risk of bias, but instead 'unclear', as trial protocols were not always published or accessible to the review authors to accurately assess/interpret.

Study/sample/group size

Sample size was not always included within the risk of bias assessment. It was therefore extracted directly from each review's table of included study characteristics by a single overview author (LG), and assessed as being low risk of bias when there was a minimum of 50 participants per arm, or 100 in total. Numbers were then separated for the proportion of studies with greater than 100 participants per arm (or 200 in total), and 200 participants per arm (or 400 in total), as this could then be considered higher tiered evidence.

Only 26 out of 264 included studies (10%) across the 21 reviews reported over 100 participants in total (or 50 per arm), a further 6% (15/264) included over 200 participants per arm. The remaining 223 studies (84%) had fewer than 50 participants per arm (or sample size was not reported), often not reaching 50 in total.

Other bias

The format for reporting bias has changed, and therefore some earlier reviews (that are yet to be updated) did not assess bias using the same format. Others reported additional criteria as 'other bias' including the similarity of baseline characteristics, and similarity of timing points.

Interpretation of results/conclusions by original review authors

For conclusions made by the original review authors, see Table 8. We assessed whether these conclusions/interpretations of the results accurately reflected the information provided within the review, and if any further information should have been included. This final assessment of the review is an important stage in determining any author bias within the review process, as many readers, funders, and policy makers will focus on the author conclusions without a full appraisal of the actual presented data.

Eleven of the 21 reviews reported appropriate conclusions based on the data available in the context of the quality of evidence (Bidonde 2014; Boldt 2014; Busch 2007; Busch 2013; Fransen 2015; Gross 2015a; Koopman 2015; Regnaux 2015; Saragiotto 2016; Silva 2010; Yamato 2015); five reviews had appropriate conclusions, did not mention quality of the evidence in the conclusion, but did discuss it in detail earlier in the review (Bartels 2007; Cramp 2013; Han 2004; Hayden 2005; Lauret 2014); two reviews had appropriate conclusions but had only limited discussion of quality or did not adequately consider the quality of the evidence in the interpretation of the results (Hurkmans 2009; Lane 2014); and three reviews needed further comment as the strength of the conclusions were not appropriate based on the available data (Brown 2010; Fransen 2014), or we were unable to agree with their interpretation due to difficulty in extracting the data (van der Heijden 2015).

Effect of interventions

We have interpreted results using data reported in the reviews, and did not return to the original studies. Where data have been reported as MDs or as an absolute or relative change score we have used the appropriate scales (where possible) to determine whether this was clinically significant. When data have only been presented as SMD, with or without 95% confidence intervals (CI), with or without level of significance (P value), we have cautiously used the interpretation by Cohen 1988 who defined effect size using the SMD as small (SMD 0.2 to 0.5), moderate (SMD 0.5 to 0.8), or large (SMD greater than 0.8).

For the purposes of clarity, we have used the term 'intervention' to refer to the exercise or physical activity intervention, and 'control' to refer to the included comparison group which did not involve any exercise or physical activity element.

Primary outcome

Self-reported pain (severity)

Part of the inclusion criteria for this overview was for pain severity to be listed as an outcome measure.

Two of the 21 reviews did not include/identify any studies that examined intervention versus control (Lauret 2014; Silva 2010). Of the remaining reviews that did report studies examining intervention versus control (no physical activity or exercise, or minimal intervention), two did not report pain as an absolute or relative score of severity, intensity, or change as a result of the intervention (Brown 2010; Han 2004), and one review assessed pain-free time and distance during exercise (they did not assess pain using a mean/usual pain scale; Lane 2014). We could not extract relevant data for one review as they compared two different exercise interventions and a control but did not report the data compared to the control (Regnaux 2015).

The remaining 15 reviews reported a mean or usual pain score for exercise (intervention) and no-exercise (control) groups (Bartels 2007; Bidonde 2014; Boldt 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Hurkmans 2009; Koopman 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015).

Reported baseline pain score

Of the 15 reviews that were able to assess pain (Table 9), only three reviews reported actual baseline pain scores (Bidonde 2014; Boldt 2014; Hayden 2005). Three reviews reported change data (Bartels 2007; Busch 2007; Busch 2013), but we were able to use control group baseline and earliest control group scores as assumed or approximate baseline measures for the intervention groups in nine reviews (Bartels 2007; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Koopman 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015). Overall, only three reviews that assessed pain did not provide baseline or control group scores for comparison (Busch 2013; Cramp 2013; Hurkmans 2009).

Intervention group at baseline	Control group at baseline	Control group at earliest follow-up
Median pain score 70.9/100 (based on 7 studies, n = 382; Bidonde 2014)	WOMAC 9.1/20 (2 studies, n = 380) VAS ~ 55/100 (3 studies, n = 117) HAQ 1.05/3 (1 study, n = 249) (Bartels 2007)	Mean pain score ~ 29/100 (9 studies, n = 549; Fransen 2014)
11.05 to 22.6 on a 0 to 150 WUSPI score (1 study, n = 35; Boldt 2014)	VAS 35/100 to 61/100 (4 studies, n = 204; Busch 2007)	44/100 (44 studies, n = 3537; Fransen 2015)
Mean pain score 46/100 (95% CI 41 to 50) (8 studies, n = 370; Hayden 2005)	-	40/100 to 60/100 (2 studies, n = 147; Gross 2015a)
-	-	44/100 SD 24 (1 study, n = 55; Koopman 2015)
-	-	range 25/100 to 56/100 (4 studies, n = 291; Saragiotto 2016)
-	-	2.1/10 to 6.0/10 (2 studies, n = 41; van der Heijden 2015)
-	-	range 18/100 to 52/100 (6 studies, n = 148; Yamato 2015)
Range: 46 to 70.9 on a 0 to 100 scale 16 studies, n = 787	**Range: 35 to 55 on a 0 to 100 scale 10 studies, n = 950**	**Range: 18 to 60 on a 0 to 100 scale 68 studies, n = 4768**

CI: confidence interval; HAQ: Health Assessment Questionnaire; n: number of participants; SD: standard deviation; VAS: visual analogue score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; WUSPI: Wheelchair User's Shoulder Pain Index

HAQ: mean of different category scores, 0 or 1 (mild to moderate disability), up to 2 or 3 (severe to very severe disability); WOMAC pain score: 5 items summed to 0 (no pain) to 20 (worst pain ever); WUSPI: 15 items of 0 to 10 VAS scores, summed to form total

CI: confidence i
ment Questionn
SD: standard de
score; WOMAC
ter Universities

(*Continued*)

of 0 (no pain) to 150 (worst pain ever)	Wheelchair User HAQ: mean of 1 (mild to mod (severe to very se score: 5 items s (worst pain ever) VAS scores, sum to 150 (worst pa

This suggests the majority of participants reviewed had mild-to-moderate pain (only one review reported a mean of severe pain (aquatic exercise for fibromyalgia, Bidonde 2014) at the commencement of each intervention (less than 30/100 mild pain, 30/100 to 60/100 moderate pain, more than 60/100 severe pain; Collins 1997), though labelling the majority as having only mild-to-moderate pain should be interpreted with caution due to the lack of specific data available - the baseline data of the intervention group would have been preferable to the proxies we have had to use.

Quality judgement/ tiered quality (first, second, third tier evidence)

Our assessment criteria stated that we would accept the information as graded evidence when reported as the number of participants achieving a 50% (first tier evidence) or 30% (second tier evidence) reduction in pain, but none of the included reviews reported results in this way, and so instead we used the reported absolute and relative change values.

None of the included reviews fulfilled the requirements for first tier evidence (at least 50% pain reduction from baseline, study duration longer than eight weeks, and more than 200 participants per arm).

Second tier evidence (at least 30% pain reduction from baseline, study duration between four and eight weeks, and more than 200 participants in total or 100 participants per arm) was also lacking in these reviews; three reviews found at least 30% reduction in pain from baseline (Busch 2007; Busch 2013; van der Heijden 2015), one of which also used long enough exercise programmes (eight to 21 weeks' intervention, Busch 2013) but totalled only 81 participants across two studies. The other two reviews did not fulfil the study duration criteria (interventions from 2.5 weeks, Busch 2007; and three weeks, van der Heijden 2015) or study size criteria.

Consequently results from relevant reviews have been pooled (all tier three quality) where appropriate, though results should be interpreted with caution due to the low quality evidence.

Treatment effect

Data that could be extracted for pain can be seen in Table 9 for all reviews. Only three reviews found no statistically significant changes in usual or mean pain from any intervention (Cramp 2013; Hurkmans 2009; Koopman 2015 (assumed due to lack of presented data)). The remaining reviews reported a statistically significant effect of the intervention at one or more time points, in at least one subgroup.

Three reviews found at least 30% pain reduction from baseline (post-intervention - strength training: Busch 2007; Busch 2013, at short-term follow-up: van der Heijden 2015). Additionally, seven reviews reported clinically significant results (minimally important difference: reduction in pain from baseline of at least 10 points on a 0 to 100 scale or an absolute improvement of at least 10% to 20%, Dworkin 2008) as a result of the exercise intervention (1.3/10 from aerobic training, Busch 2007; 12/100 (95% CI 10 to 15), Fransen 2015,; 14.9/100 (95% CI 7.39 to 22.40), Gross 2015a; 10.2/100 (95% CI 1.31 to 19.09), Hayden 2005; 2.5/10 (95% CI 1.52 to 3.48), Boldt 2014; 10.01/100 (95% CI 4.35 to 15.67), Saragiotto 2016; 14.05/100 (95% CI 9.19 to 18.91), Yamato 2015). Three reviews found statistically significant improvements as a result of the intervention, but they did not reach clinical significance (post-intervention, $P = 0.02$, Bartels 2007; "small to moderate" benefit post-intervention and at six-month follow-up, $P < 0.001$, Fransen 2014; "moderate effect" of 7% (95% CI 3 to 11) benefit post-intervention, Bidonde 2014).

Overall, results were inconsistent across interventions and follow-up (see Table 9), as exercise did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point.

Secondary outcomes

Physical function (objectively or subjectively measured)

Measures of physical function were the primary outcome measure in eight out of 21 reviews (Busch 2013; Han 2004; Hayden 2005;

Hurkmans 2009; Koopman 2015; Lane 2014; Lauret 2014; Silva 2010), and a reported (non-primary) outcome measure in nine more reviews (Bartels 2007; Bidonde 2014; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Regnaux 2015; Saragiotto 2016; van der Heijden 2015, plus some which assessed disability; Cramp 2013; Saragiotto 2016; Yamato 2015). Only Boldt 2014 and Brown 2010 did not list physical function (or disability, or activity limitation) as a potential outcome measure.

Treatment effect

Data that could be extracted for physical function are shown in Table 10. Two reviews which reported physical function had no data to extract (Lauret 2014; Silva 2010), and for one review we were unable to extract the relevant data (Regnaux 2015). Two reviews found no significant difference in physical function between the intervention and control groups (Han 2004; Hurkmans 2009, both rheumatoid arthritis, 8 studies, n = 240). The remaining 14 reviews showed that the intervention produced a statistically significant benefit over the control at a minimum of one reported time point (Bartels 2007; Bidonde 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Koopman 2015; Lane 2014; Saragiotto 2016; van der Heijden 2015; Yamato 2015; 129 studies, n greater than 9559 (exact number unknown due to some participant numbers not being reported)).

Many of these statistically significant results were of small or moderate effect size (as reported by the review authors, or using the definition by Cohen 1988 if unreported; small effect size: Bartels 2007; Bidonde 2014; Fransen 2014; Fransen 2015; Gross 2015a; Koopman 2015; Saragiotto 2016; Yamato 2015, moderate effect size: Busch 2007; Fransen 2015; Yamato 2015).

Only one review reported statistical significance and large effect size (both short-term and long-term follow-up: SMD 1.10 (95% CI 0.58 to 1.63) and 1.62 (95% CI 0.31 to 2.94), van der Heijden 2015). However, the original review authors highlighted the low to very low quality of the evidence as many studies had high or unclear risk of bias across multiple domains (van der Heijden 2015).

Psychological function

Only five out of 21 reviews assessed psychological function as mental health (Bartels 2007; Bidonde 2014; Busch 2013), anxiety (Cramp 2013), and depression (Boldt 2014; Busch 2013; Cramp 2013).

Treatment effect

Data that could be extracted for psychological function can be seen in Table 11. There were significant effects in favour of the intervention for mental health (Bartels 2007) and depression (Busch 2013) scores, and "variable effect" for depression (Cramp 2013). However, there was also no effect or no differences between control and intervention groups reported for mental health (Bidonde 2014; Busch 2013), anxiety (Cramp 2013), and depression (Boldt 2014).

Quality of life

A version of quality of life assessment was reported in nine reviews. Six were termed quality of life or health-related quality of life (HRQoL) (Bartels 2007; Boldt 2014; Fransen 2014; Fransen 2015; Gross 2015a; Lauret 2014).

Other reviews assessed global perceived effect (Gross 2015a), global wellbeing (Busch 2007), global assessment (Hayden 2005), global impression of recovery (Saragiotto 2016; Yamato 2015), health assessment questionnaire (Silva 2010), multi-dimensional function (Bidonde 2014; Busch 2013), and work status (Hayden 2005). These have been reported separately to quality of life (Table 12).

Treatment effect

Data that could be extracted for quality of life can be seen in Table 12. Four reviews found no significant difference between intervention and control groups in health-related quality of life post-intervention (9 studies, n = 556) (HRQoL: Boldt 2014; Fransen 2014; Gross 2015a, global assessment: Bidonde 2014; Gross 2015a)), three reviews did not or were unable to report any data (HRQoL: Lauret 2014, global assessment: Hayden 2005, other assessment: Silva 2010), and seven reviews found a significant improvement as a result of the intervention (34 studies, n = 2700) (HRQoL: Bartels 2007, Fransen 2015, global assessment: Busch 2007; Saragiotto 2016; Yamato 2015, other assessment: Bidonde 2014; Busch 2013).

Two reviews assessing strength/resistance training interventions found significantly large effect sizes (SMD greater than 0.8, as defined by Cohen 1988) in favour of the intervention (global wellbeing measure, SMD 1.43 (95% CI 0.76 to 2.10), Busch 2007; Fibromyalgia Impact Questionnaire, SMD 1.27 (95% CI 0.72 to 1.83), Busch 2013). Other statistically significant changes reported in the included reviews were of small-to-moderate effect size (SMD 0.2 to 0.8, Cohen 1988).

Adherence to the prescribed intervention

Only one review reported adherence to the intervention as an outcome measure (Regnaux 2015), but the authors were unable to perform an analysis on attendance as most studies did not clearly report attendance or compliance (Regnaux 2015). However, five reviews assessed withdrawals or dropouts (Bidonde 2014; Fransen 2014; Han 2004; Regnaux 2015; Saragiotto 2016), one reported all-cause attrition (Busch 2013), and another reported the discontinuation rate (Silva 2010).

Data that could be extracted for adherence, withdrawals, and attrition can be seen in Table 13. Pooling all available data for withdrawals/dropout/attrition gave an RR of 1.02 (95% CI 0.94 to 1.12) in favour of the control group (6 reviews, 30 studies, n = 2256, control withdrawal 81/1000, intervention withdrawal 82.8/1000).

One clinically controlled trial (CCT) in one review reported statistically significant improvement in enjoyment of exercise/rest (P = 0.0002) and self-reported benefit from exercise/rest (P = 0.006) at both post-intervention (end of therapy, 10 weeks) and follow-up (four months later) (n = 95, Han 2004).

Healthcare use/attendance

None of the reviews reported healthcare use/attendance.

Adverse events (not death)

Eighteen out of 21 reviews reported adverse effects (three reviews did not report adverse events as an outcome measure due to lack of studies or other undisclosed reasons; Brown 2010; Lauret 2014; Silva 2010). Two reviews only assessed a specific adverse event ("amputation" Lane 2014; "motor unit survival" Koopman 2015), one review observed "safety - pain and radiological damage" (Hurkmans 2009), and another referred to any "side-effects" (Han 2004).

Data that could be extracted for adverse events (not death) can be seen in Table 14. The total number of reported adverse events (not death) was 137 events across 39 studies out of 61 studies that had adverse events as an outcome measure (over one-third of all trials that reported them found no adverse events related to the intervention): six reviews reported no adverse events from the included trials (Bartels 2007; Busch 2013; Cramp 2013; Hurkmans 2009; Koopman 2015; Yamato 2015) though the authors questioned whether this was due to lack of reporting by the trial authors, or whether there were no adverse events.

Adverse events were largely reported as a total number per trial, though one review separately reported results for the intervention group versus the control group (Saragiotto 2016), and two others reported adverse events for the intervention group only (Boldt 2014; Regnaux 2015). Only one review calculated an RR for the adverse events, showing a reduced risk for amputation in the intervention group (two amputations in the usual care/control group: RR 0.20, 95% CI 0.01 to 4.15, based on one study in one review, Lane 2014).

Death

Only one out of 21 reviews reported death separately to other adverse events (Lane 2014). Based on five studies within the review, death had an RR of 0.71 (95% CI 0.28 to 1.78) in favour of exercise as being protective, though was not statistically significant (P = 0.47).

DISCUSSION

Specificity of the condition: despite the heterogeneous nature of chronic pain, in this overview we have combined several painful conditions covering a number of conditions and diagnoses. Regardless of aetiology, the impact of chronic pain is broadly similar across many conditions.

Summary of main results

Pain severity: there were favourable results in a number of reviews as a result of exercise: only three reviews found no statistically significant changes in usual or mean pain from any intervention. However, results were inconsistent across interventions and follow-up, as the intervention did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point. The exercise or physical activity interventions did not have a negative effect on the outcome (did not worsen the pain). A factor in the lack of statistical and clinically significant result may be the baseline pain severity of participants. The majority of the included population had an assumed mild-to-moderate pain severity score (assumed only due to lack of exact group data at baseline). This is often the desired outcome (post-intervention) of many drug therapies for pain, and it may therefore be difficult to show a clinically significant improvement in these people.

Physical function: physical function/disability was the most commonly reported outcome measure, and was the primary measure in eight out of the 21 reviews. Physical function was significantly (statistically) improved as a result of the intervention in 14 reviews, though even these statistically significant results had only small-to-moderate effect sizes in all but one review.

Psychological function and quality of life: there were variable results for psychological function and quality of life: results were either favourable to exercise (two reviews reporting significantly large effect sizes for quality of life), or showed no difference between groups. There were no negative effects.

Adherence to the prescribed intervention: could not be assessed in any included review. However, risk of withdrawal/dropout was slightly higher in the exercising group (82.8/1000 participants versus 81/1000 participants), though the group difference was not significant.

Healthcare use/attendance: not reported in any included review.

Adverse events, potential harm, and death: importantly, exercise caused no actual harm, with most adverse events being increased soreness or muscle pain, which reportedly subsided after several weeks of the intervention. One review reported a non-significant reduction in risk of death as a result of the intervention.

Overall completeness and applicability of evidence

Of the 21 included reviews, seven could be considered out of date as they were most recently assessed as up-to-date prior to

2010 such that any recent controlled trials assessing pain severity have not been included in this overview (Cochrane recommends updating reviews every two years) (Bartels 2007; Brown 2010; Busch 2007; Han 2004; Hayden 2005; Hurkmans 2009; Silva 2010). We included these reviews in the overview, but they may not be as relevant now due to the elapsed time since they were updated. One protocol that had potential to be included was published in 2006 with no full review available yet (Craane 2006).

Available data suggest that participants in the included reviews and studies would generally be characterised as having mild-moderate pain (moderate greater than 30/100 or 3/10) with only one review reporting moderate-severe pain (severe greater than 60/100 or 6/10). Therefore whether the evidence of change or no change seen here as a result of each intervention is applicable to people further along on the pain spectrum (with higher pain scores/worse pain) is debatable. However, it can be argued that those people are more likely to be assigned medical or surgical interventions than physical activity and exercise alone (where available), and as a group they may be less able to engage in exercise, and may therefore be more difficult to recruit into exercise-only studies. Having said this, the labelling of participants as having mild-moderate pain was a cautious one within this overview due to the lack of specific data available at baseline assessment; only three reviews included baseline pain scores in the intervention group, and two further reviews provided control group baseline scores.

There are still gaps in the available literature, and therefore also within this overview. None of the included reviews examined generalised or widespread chronic pain as a global condition, each instead examined specific conditions that included chronic pain as a symptom or result of the ongoing condition (rheumatoid arthritis, osteoarthritis, fibromyalgia, low back pain, intermittent claudication, dysmenorrhoea, mechanical neck disorder, spinal cord injury, postpolio syndrome, and patellofemoral pain). The pain in these cases can occur secondary to other symptoms such as fatigue, muscle stiffness, difficulty sleeping, and depression, all of which could separately (and more effectively) be influenced by the intervention. Additionally, only 25% of included studies actively reported adverse events. This may affect the completeness of the evidence as conclusions have been drawn based on the available data. The included reviews did not discuss the possible impact of this non-reporting by the original trials, and this may lead to underestimating possible adverse events from an intervention, or overestimating its safety.

The exercise interventions examined in the included reviews were broad; including aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi. Many of these interventions can be accessed in the community by the general public and people with chronic pain, either individually or in classes (yoga, Pilates, tai chi). Other exercise intervention programmes, such as the motor control exercise and proprioceptive (balance) training, requires at least initial supervision by a therapist to teach the correct techniques and provide feedback for progression.

Quality of the evidence

In assessing the quality of the evidence, we employed the AMSTAR tool to examine the reviews, extracted data on risk of bias to examine the available primary evidence, and evaluated the authors' conclusions to ensure that they were appropriate based on the available data.

The AMSTAR tool is useful in assessing the reporting of a systematic review, though it does not inform us of the actual undertaking or conduct of the review process. All 21 included reviews scored well across the AMSTAR assessment, though this is likely due to the stringent reporting guidelines implemented by Cochrane prior to publication. However, it may be necessary or advisable for the Cochrane guidelines to be further expanded and detailed with regards to reporting study characteristics, publication bias, and conflicts of interest, as these areas often did not meet the requirements laid out in the AMSTAR criteria (Table 1).

Data extracted from the reviews regarding their assessment of bias (risk of bias) showed moderate level scores at best across all included studies within the included reviews. Other than issues surrounding blinding (which are problematic in exercise intervention studies due to the nature of the intervention), the trials did not consistently and adequately report potential attrition and reporting biases, with less than half of studies within these reviews at low risk of bias.

However, the most prominent issue with regards to bias in these exercise and physical activity intervention studies is the sample size used. This subcategory is not used as standard in the assessment of bias in Cochrane Reviews, despite the increasing volume of research available suggesting that small studies of fewer than 100 participants per arm (Moore 2010; Nüesch 2010) are at increased risk of succumbing to the random effects in estimating both direction and magnitude of treatment effects (Moore 1998; Turner 2013) due to greater heterogeneity within and between small studies (IntHout 2015).

Studies within the included reviews here were very small (often fewer than 50 participants in total). For greater quality and a more reliable effect, at least 100 participants per arm should be analysed for a study to potentially be classed as tier two evidence (200 per arm for tier one); small studies are known to overestimate the treatment effect by up to 32% in comparison with larger studies (Deschartes 2013).

Assessing studies for risk of bias based on study size (total number or per arm) should be included in any review or meta-analysis in future, to adequately assess the influence of small trials on the estimated treatment effect (Nüesch 2010). Inclusion in the standard assessment process may in turn influence the design and undertaking of future research trials to increase the sample size, and produce more consistent clinically and statistically accurate results.

Of the 21 included reviews, 12 used a pain measure as their primary outcome (Bartels 2007; Boldt 2014; Brown 2010; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Regnaux 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015), and the remaining nine reviews included the measure as a secondary outcome only. Other outcomes were shared, including physical and psychological function, and quality of life. Likewise, each review team will have included studies that did not use their chosen outcome measures as the primary measure, and that were therefore powered according to a different primary outcome. On collating the evidence, some studies may appear underpowered for the outcome(s) of interest to us (Turner 2013), yet were adequately powered for the studies' primary measure. To increase the power of the results of this overview, and the intermediary reviews we have included, intervention studies that focus on painful conditions should include pain intensity as the primary outcome, or at least as a prominent secondary outcome; alternatively review authors should seek to include only those studies that were adequately powered for pain intensity as a primary outcome measure.

Intervention length ranged from a single session to regular sessions over a period of 30 months, though the majority were between eight and 12 weeks. Durations of this length are common among exercise and physical activity intervention studies to allow for physiological adaptation and familiarisation. In contrast, the follow-up period was often inadequate, as many reviews reported only a single follow-up point (immediately post-intervention), or repeated measures over the short-term (less than six months): only six of the 21 reviews planned to assess participants over the long term (over 12 months: Gross 2015a; Hayden 2005; Hurkmans 2009; Regnaux 2015; Saragiotto 2016; van der Heijden 2015). With chronic conditions, it would be advisable to include longer follow-up periods (beyond 12 months post-randomisation) as long-term solutions may be more relevant to their control or pain management. It is also possible that initial adaptation and potential benefits as a result of an exercise intervention may take longer to manifest in comparison to a 'healthy' person due to the possible limitations in exercise intensity and progression (a training threshold) beyond which any additional physical training may be detrimental to the underlying pathophysiological mechanisms (Daenen 2015) or simply be additional physical stress with no additional physical benefit (Benton 2011).

We grouped outcome measurement points in this overview into short term (less than six months), intermediate term (six to 12 months), and long term (longer than 12 months). The broad time window for 'short term' outcomes (less than six months) is a potential source of heterogeneity as the early period is the one where time of measurement is most likely to result in variable outcomes. These initial problems could be overcome by use of standard reporting periods in exercise intervention studies (suggested four-weekly within the 'short term' period to assess both neural adaptation and other physiological changes). This would allow review authors to use the data gathered closest to the time point they are assessing, for more accurate analyses. Additionally, by extending the follow-up period beyond one year (long-term follow-up), heterogeneity may be reduced further.

Reviews generally did not enforce a minimum exercise requirement for inclusion in their review. Additionally, not all exercise sessions were supervised or baseline fitness/physical ability was assessed subjectively, and consequently it was not reported whether the intervention was fulfilled as described, or whether the dose was enough to elicit a physiological response. Studies often rely on the self-report of participants as to the actual physical activity and exercise being undertaken, which can lead to a greater risk of bias, and reduced study quality as it is questionable as to whether the effect can be truly attributed to the intervention. This was examined in a previous review, where it was concluded that non-subjective physical assessment should be performed where possible (Perruchoud 2014), though these still have challenges regarding implementation.

In summary, the quality of the evidence was low (third tier): within this overview we found no tier one or tier two evidence. This is largely due to the small sample sizes and potentially underpowered studies. A number of studies within the reviews had adequately long interventions, but planned follow-up was limited to less than one year (12 months) in all but six reviews.

Interpretation of the available data, and conclusions drawn by the review authors, were appropriate, although the conclusions were sometimes stronger than warranted by the available data. Occasionally results were not discussed with regards to the quality of the evidence or risk of bias: it is important to discuss the findings in the context of the quality of the evidence, with complete transparency, as this may affect future research, and implications for patients, funders, and policy makers.

Potential biases in the overview process

While we have attempted to include all relevant reviews in the overview process, we do concede that by only searching the Cochrane Library, and including only current Cochrane Reviews we may have missed some key literature. However previous publications have referred to the higher quality grading (high AMSTAR score) in Cochrane Reviews due to the basic criteria necessary for publication at any stage (protocol or full review) suggesting they may be the most reliable source of evidence (O'Connell 2013).

Agreements and disagreements with other studies or reviews

This is a summary overview of current Cochrane Reviews, we are not aware of any overviews or reviews summarising non-Cochrane reviews.

AUTHORS' CONCLUSIONS

There is limited evidence of improvement in pain severity as a result of exercise. There is some evidence of improved physical function and a variable effect on both psychological function and quality of life. However, results are inconsistent and the evidence is low quality (tier three). Promisingly however, none of the physical and activity interventions assessed appeared to cause harm to the participants.

Implications for practice

For clinicians and people with chronic pain

The evidence in this overview suggests that the broad spectrum of physical activity and exercise interventions assessed here (aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi) are potentially beneficial, though the evidence for benefit is low quality and inconsistent. The most commonly reported adverse events were increased soreness or muscle pain, which subsided after several weeks of the intervention.

Physical activity and exercise may improve pain severity as well as physical function and quality of life.

For policy makers

The evidence showed variable results, though in some reviews there was a clinical and statistical benefit in pain relief and physical function (based on low quality evidence). The evidence suggests that physical activity or exercise is an acceptable intervention in people with chronic pain, with minimal negative adverse effects. However based on this low quality evidence, we cannot provide direction to the content of an exercise programme should clinicians decide to implement one.

Implications for research

There is a clear need for further research into exercise and physical activity for chronic pain in adults.

General implications

- Future research should report baseline values for outcome measures in both intervention and control groups, together with detailed relevant information about the participants. Knowing the baseline value is relevant to interpreting any change observed as a result of the intervention, and understanding the broader value of the intervention.
- Where possible, pain results should be reported as the number of people achieving 50%, 30%, and 10% pain relief, and the number who did not meet that point (dichotomous outcome). These are clinically important cut-offs in pain intervention research, and reporting in this way allows readers to observe the clinical effect more effectively.
- Reporting should include median and range as well as mean and standard deviation (SD) of results. This will allow readers to review the effects of any outliers that may have skewed the data, which often goes unnoticed in the reporting of mean and SD alone.
- The importance of clear intervention reporting is underestimated: often studies report both intervention and control programmes simply, where other researchers and clinicians alike are unable to replicate the trial or intervention. Recommendations for reporting are based on the Consolidated Standards of Reporting Trials (CONSORT) statement (www.consort-statement.org/), but this alone does not detail the extent of necessary intervention and control programmes reporting. The template for intervention description and replication (TIDieR) approach (Hoffman 2014) is intended as an extension to CONSORT item 5 ("The interventions for each group with sufficient details to allow replication, including how and when they were actually administered") and is a checklist for detailing the programmes using: why (rationale), what (materials and procedures), who, how, where, when, and how much.

Design

- One previous review highlighted the increased bias often present in questionnaires and other self-report measures of physical activity in people with chronic pain, and as a result made the recommendation to use objective measures instead, such as accelerometers, or the use of direct and indirect calorimetry, where possible (Perruchoud 2014), though these still have challenges regarding implementation. This would allow direct and exact comparison and analyses of actual energy expenditure and treatment effect.

Population/participants/sample

- There needs to be a focus on participants with generalised and/or widespread chronic pain, instead of (or as well as) condition-specific populations.
- Studies should include people with higher pain severity (greater than 50/100 on a 100-point visual analogue scale) at baseline. People with mild-moderate pain should still be included, but it would be advisable to separate the results for analysis, ensuring the study is adequately powered to allow this subgroup analysis in advance. This way we could determine if exercise has benefit overall, or affects one group more than another, and tailor exercise programmes according to the individual needs.
- It has been previously suggested that for 20% to 25% of participants undertaking an exercise programme there is little to

no favourable response (Timmons 2014), while a small percentage (5% to 10%) have adverse events (Bouchard 2012). It is therefore vitally important that much larger sample sizes are used: ideally *more than 200 participants per arm*, though even this number in total would increase the quality of the evidence in the first instance. In this way we may be able to learn to identify individuals who will benefit, and those who will require further intervention.

Interventions

- Different forms of exercise should be researched in detail. For the purposes of this overview, we combined all physical activity and exercise interventions under one banner to determine if there was any effect. However a number of reviews separately analysed resistance (strength) training, aerobic (endurance), and combination programmes. It is important to continue to examine different modalities, but currently there is not enough high quality evidence to exclude or prioritise one specific mode (resistance, endurance, stability) or medium (land/water based), or the proportion of a combination programme to be assigned to each, as all may have individual benefits for people with chronic pain.
- Intensity of exercise, duration of individual sessions, and frequency should be investigated. It is this dose alongside duration (of the entire intervention) and adherence that may determine the actual efficacy.
- More reviews and trials should attempt to minimise intervention heterogeneity by implementing minimum and maximum requirements. Only this way will the research community be able to determine more accurately the direction and magnitude of effect of a specific programme or intervention. Many of these important restrictions can be implemented as subgroup analyses, though if this is the case it is important to have adequate study numbers (ideally 200 participants per arm or subgroup).
- Due to the chronicity and long-term nature of the condition, physiological and psychological changes may take longer to manifest. It is widely accepted that there is a delay in muscular hypertrophy as a result of exercise, and initial gains within the first few weeks of any training programme will be as a result of neural factors (Enoka 1997); this is also in line with the grading of evidence (tier two evidence or higher requires a minimum of a four-week intervention). This suggests that longer interventions may be necessary (eight weeks for tier one evidence), though assessing participants at regular intervals, including at four weeks, would be beneficial to examine the effect of the neural adaptation alone.

Measurement (end-points)

- Randomised controlled trials with long-term follow-up are needed. Chronic pain is defined by its chronic nature, and therefore long-term follow-up of results is equally important as the initial short-term effect (if not more so): outcomes should be assessed beyond one year after randomisation. In turn this will inform the direct effect of the intervention, as well as the proportion of the population who maintains the programme of exercise employed in the intervention, or something else under the guise of physical activity as a result of participation.
- The broad time window for 'short term' outcomes (less than six months) is a potential source of heterogeneity as the early period is the one where time of measurement is most likely to result in variable outcomes. These initial problems could be overcome by use of standard reporting periods in exercise intervention studies (suggested four-weekly assessment within the 'short term' period to assess both neural adaptation and other physiological changes). This would allow review authors to use the data recorded closest to the time point they are assessing, for more accurate and comparable analyses.
- Outcome measures used by researchers should be standardised across trials and studies. Recommendations for selecting the most appropriate and important outcome measures to those who live with chronic pain have previously been published (Initiatives on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) Consensus Recommendations, Dworkin 2005; Turk 2003).

Other

- It would be of interest in future research to determine the reasons for non-participation in regular physical activity or non-compliance to a prescribed exercise intervention in people with chronic pain, and how to overcome these barriers.
- Future Cochrane Reviews could include: exercise for chronic pain or chronic widespread pain (and not specific conditions such as osteoarthritis, fibromyalgia, etc.), and exercise for neuropathic pain. These areas have not been covered by Cochrane with an exercise or physical activity intervention.

ACKNOWLEDGEMENTS

Thanks to Dr Amanda Williams, Professor Chris Eccleston, and Anna Erskine at the Cochrane Pain, Palliative and Supportive Care Review Group (PaPaS), for their support and rigour in the development of the concept of this overview, title proposal, and protocol. Thanks also to peer reviewers Dr Neil O'Connell and Dr Clare Stevinson, and consumer referees Jennifer Patrick and Dr Shailen Mehta.

Cochrane Review Group funding acknowledgement: the National Institute for Health Research (NIHR) is the largest single funder of the Cochrane PaPaS Group.

Disclaimer: the views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, National Health Service (NHS), or the Department of Health.

External funding support from the National Institute of Academic Anaesthesia (BJA RCoA NIAA) and the MRC Lifelong Health and Wellbeing Grant to complete the overview as part of two larger projects.

REFERENCES

References to included reviews

Bartels EM, Lund H, Hagen KB, Dagfinrud H, Christensen R, Danneskiold-Samsøe B. Aquatic exercise for the treatment of knee and hip osteoarthritis. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD005523.pub2]

Bidonde J, Busch AJ, Webber SC, Schachter CL, Danyliw A, Overend TJ, et al. Aquatic exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011336]

Boldt I, Eriks-Hoogland I, Brinkhof MWG, de Bie R, Joggi D, von Elm E. Non-pharmacological interventions for chronic pain in people with spinal cord injury. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD009177.pub2]

Brown J, Brown S. Exercise for dysmenorrhoea. *Cochrane Database of Systematic Reviews* 2010, Issue 2. [DOI: 10.1002/14651858.CD004142.pub2]

Busch AJ, Barber KA, Overend TJ, Peloso PMJ, Schachter CL. Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD003786.pub2]

Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010884]

Cramp F, Hewlett S, Almeida C, Kirwan JR, Choy EHS, Chalder T, et al. Non-pharmacological interventions for fatigue in rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD008322.pub2]

Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2008, Issue 4. [DOI: 10.1002/14651858.CD004376.pub2]

Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: 10.1002/14651858.CD007912.pub2]

Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD004376.pub3]

Gross A, Kay TM, Paquin JP, Blanchette S, Lalonde P, Christie T, et al. Exercises for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD004250.pub5]

Han A, Judd M, Welch V, Wu T, Tugwell P, Wells GA. Tai chi for treating rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2004, Issue 3. [DOI: 10.1002/14651858.CD004849]

Hayden J, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database of Systematic Reviews* 2005, Issue 3. [DOI: 10.1002/14651858.CD000335.pub2]

Hurkmans E, van der Giesen FJ, Vliet Vlieland TPM, Schoones J, Van den Ende ECHM. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2009, Issue 4. [DOI: 10.1002/14651858.CD006853.pub2]

Kay TM, Gross A, Goldsmith CH, Rutherford S, Voth S, Hoving JL, et al. Exercises for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 2012, Issue 8. [DOI: 10.1002/14651858.CD004250.pub4]

Koopman FS, Beelen A, Gilhus NE, de Visser M, Nollet F. Treatment for postpolio syndrome. *Cochrane Database of Systematic Reviews* 2015, Issue 5. [DOI: 10.1002/14651858.CD007818.pub3]

Lane R, Ellis B, Watson L, Leng GC. Exercise for intermittent claudication. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD000990.pub3]

Lauret GJ, Fakhry F, Fokkenrood HJP, Hunink MGM, Teijink JAW, Spronk S. Modes of exercise training for intermittent claudication. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD009638.pub2]

Regnaux JP, Lefevre-Colau MM, Trinquart L, Nguyen C, Boutron I, Brosseau L, et al. High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis. *Cochrane Database of Systematic Reviews* 2015, Issue 10. [DOI: 10.1002/14651858.CD010203.pub2]

Saragiotto BT, Maher CG, Yamato TP, Costa LOP, Menezes Costa LC, Ostelo RWJG, et al. Motor control exercise for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2016, Issue 1. [DOI: 10.1002/

14651858.CD012004]
Silva KNG, Mizusaki Imoto A, Almeida GJM, Atallah ÁN, Peccin MS, Fernandes Moça Trevisani V. Balance training (proprioceptive training) for patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2010, Issue 5. [DOI: 10.1002/14651858.CD007648.pub2]
van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SMA, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD010387.pub2]
Yamato TP, Maher CG, Saragiotto BT, Hancock MJ, Ostelo RWJG, Cabral CMN, et al. Pilates for low back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 7. [DOI: 10.1002/14651858.CD010265.pub2]

References to excluded reviews

Aggarwal VR, Lovell K, Peters S, Javidi H, Joughin A, Goldthorpe J. Psychosocial interventions for the management of chronic orofacial pain. *Cochrane Database of Systematic Reviews* 2011, Issue 11. [DOI: 10.1002/14651858.CD008456.pub2]
Bierma-Zeinstra SMA, Brinks A, Verhagen AP, Van Rijn RM, Koes BW, Verhaar JAN. Interventions for lateral hip pain (tendinopathy or bursitis). *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.CD008924]
Brønfort G, Haas M, Evans RL, Goldsmith CH, Assendelft WJJ, Bouter LM. Non-invasive physical treatments for chronic/recurrent headache. *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD001878.pub3]
Brønfort G, Evans RL, Goldsmith CH, Haas M, Leininger B, Levin M, et al. Spinal rehabilitative exercise and manual treatment for the prevention of migraine attacks in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 8. [DOI: 10.1002/14651858.CD011848]
Choi BKL, Verbeek JH, Tam WWS, Jiang JY. Exercises for prevention of recurrences of low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD006555.pub2]
Craane B, De Laat A, Dijkstra PU, Stappaerts K, Stegenga B. Physical therapy for the management of patients with temporomandibular disorders and related pain. *Cochrane Database of Systematic Reviews* 2006, Issue 1. [DOI: 10.1002/14651858.CD005621]
Dagfinrud H, Hagen KB, Kvien TK. Physiotherapy interventions for ankylosing spondylitis. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD002822.pub3]
Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. *Cochrane Database of Systematic Reviews* 2010, Issue 6. [DOI: 10.1002/14651858.CD007612.pub2]
Dal Bello-Haas V, Florence JM. Therapeutic exercise for people with amyotrophic lateral sclerosis or motor neuron disease. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD005229.pub3]
de Souza RF, Lovato da Silva CH, Nasser M, Fedorowicz Z, Al-Muharraqi MA. Interventions for the management of temporomandibular joint osteoarthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD007261.pub2]
Fokkenrood HJP, Bendermacher BLW, Lauret GJ, Willigendael EM, Prins MH, Teijink JAW. Supervised exercise therapy versus non-supervised exercise therapy for intermittent claudication. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD005263.pub3]
Franke H, Fryer G, Ostelo RWJG, Kamper SJ. Muscle energy technique for non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: 10.1002/14651858.CD009852.pub2]
Green S, Buchbinder R, Hetrick SE. Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD004258]
Gross A, Aker P, Goldsmith CH, Peloso PMJ. Physical medicine modalities for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 1998, Issue 2. [DOI: 10.1002/14651858.CD000961]
Gross A, Forget M, St George K, Fraser MMH, Graham N, Perry L, et al. Patient education for neck pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. [DOI: 10.1002/14651858.CD005106.pub4]
Gross A, Langevin P, Burnie SJ, Bédard-Brochu MS, Empey B, Dugas E, et al. Manipulation and mobilisation for neck pain contrasted against an inactive control or another active treatment. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD004249.pub4]
Hayden JA, Cartwright J, van Tulder MW, Malmivaara A. Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD009790]
Heintjes EM, Berger M, Bierma-Zeinstra SMA, Bernsen RMD, Verhaar JAN, Koes BW. Exercise therapy for patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2003, Issue 4. [DOI: 10.1002/14651858.CD003472]
Henschke N, Ostelo RWJG, van Tulder MW, Vlaeyen JWS, Morley S, Assendelft WJJ, et al. Behavioural treatment for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 7. [DOI: 10.1002/14651858.CD002014.pub3]
Heymans MW, van Tulder MW, Esmail R, Bombardier C, Koes BW. Back schools for non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD000261.pub2]
Hilde G, Hagen KB, Jamtvedt G, Winnem M. Advice to stay active as a single treatment for low-back pain and sciatica. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD003632.pub2]
Hoving JL, Lacaille D, Urquhart DM, Hannu TJ, Sluiter JK, Frings-Dresen MHW. Non-pharmacological interventions

for preventing job loss in workers with inflammatory arthritis. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD010208.pub2]
Hurley M, Dickson K, Walsh N, Hauari H, Grant R, Cumming J, et al. Exercise interventions and patient beliefs for people with chronic hip and knee pain: a mixed methods review. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010842]
IJzelenberg W, Rubinstein SM, Hayden J, Koes BW, van Tulder MW. Exercise therapy for acute non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2011, Issue 10. [DOI: 10.1002/14651858.CD009365]
Jones G, Crotty M, Brooks P. Interventions for treating psoriatic arthritis. *Cochrane Database of Systematic Reviews* 2000, Issue 3. [DOI: 10.1002/14651858.CD000212]
Jordan JL, Holden MA, Mason EEJ, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD005956.pub2]
Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RWJG, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. *Cochrane Database of Systematic Reviews* 2014, Issue 9. [DOI: 10.1002/14651858.CD000963.pub3]
Karjalainen KA, Malmivaara A, van Tulder MW, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary rehabilitation for fibromyalgia and musculoskeletal pain in working age adults. *Cochrane Database of Systematic Reviews* 1999, Issue 3. [DOI: 10.1002/14651858.CD001984]
Karjalainen KA, Malmivaara A, van Tulder MW, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD002194]
Larun L, Brurberg KG, Odgaard-Jensen J, Price JR. Exercise therapy for chronic fatigue syndrome. *Cochrane Database of Systematic Reviews* 2016, Issue 2. [DOI: 10.1002/14651858.CD003200.pub4]
Liddle SD, Pennick V. Interventions for preventing and treating low-back and pelvic pain during pregnancy. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD001139.pub4]
Liu H, Yao K, Zhang J, Li L, Wu T, Brox JI, et al. Sling exercise therapy for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 9. [DOI: 10.1002/14651858.CD010689]
Miller J, Gross A, Kay TM, Graham N, Burnie SJ, Goldsmith CH, et al. Manual therapy with exercise for neck pain. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD011225]
Moi JHY, Sriranganathan MK, Edwards CJ, Buchbinder R. Lifestyle interventions for chronic gout. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD010039.pub2]
O'Brien K, Nixon S, Glazier R, Tynan AM. Progressive resistive exercise interventions for adults living with HIV/AIDS. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD004248.pub2]
O'Connell NE, Wand BM, McAuley J, Marston L, Moseley GL. Interventions for treating pain and disability in adults with complex regional pain syndrome- an overview of systematic reviews. *Cochrane Database of Systematic Reviews* 2013, Issue 4. [DOI: 10.1002/14651858.CD009416.pub2]
Page MJ, O'Connor D, Pitt V, Massy-Westropp N. Exercise and mobilisation interventions for carpal tunnel syndrome. *Cochrane Database of Systematic Reviews* 2012, Issue 6. [DOI: 10.1002/14651858.CD009899]
Page MJ, Green S, Kramer S, Johnston RV, McBain B, Chau M, et al. Manual therapy and exercise for adhesive capsulitis (frozen shoulder). *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD011275]
Peters S, Page MJ, Coppieters MW, Ross M, Johnston V. Rehabilitation following carpal tunnel release. *Cochrane Database of Systematic Reviews* 2013, Issue 6. [DOI: 10.1002/14651858.CD004158.pub2]
Preston NJ, Seers K, Mortimer PS. Physical therapies for reducing and controlling lymphoedema of the limbs. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD003141.pub2]
Proctor M, Murphy PA, Pattison HM, Suckling JA, Farquhar C. Behavioural interventions for dysmenorrhoea. *Cochrane Database of Systematic Reviews* 2007, Issue 3. [DOI: 10.1002/14651858.CD002248.pub3]
Radner H, Ramiro S, Buchbinder R, Landewé RBM, van der Heijde D, Aletaha D. Pain management for inflammatory arthritis (rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis and other spondyloarthritis) and gastrointestinal or liver comorbidity. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.CD008951.pub2]
Regnaux JP, Trinquart L, Boutron I, Nguyen C, Brosseau L, Ravaud P. High-intensity versus low-intensity physical activity or exercise in patients with hip or knee osteoarthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD010203]
Regnaux JP, Lefevre-Colau MM, Palazzo C, Poiraudeau S, Ravaud P, Boutron I. Exercise therapy for ankylosing spondylitis. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011321]
Richards BL, Whittle SL, Buchbinder R. Neuromodulators for pain management in rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.CD008921.pub2]
Riemsma RP, Kirwan JR, Taal E, Rasker HJJ. Patient education for adults with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD003688]
Schaafsma FG, Whelan K, van der Beek AJ, van der Es-Lambeek LC, Ojajärvi A, Verbeek JH. Physical conditioning as part of a return to work strategy to reduce sickness absence for workers with back pain. *Cochrane Database*

of Systematic Reviews 2013, Issue 8. [DOI: 10.1002/14651858.CD001822.pub3]
Steultjens EEMJ, Dekker JJ, Bouter LM, Schaardenburg DD, Kuyk MAMAH, Van den Ende ECHM. Occupational therapy for rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2004, Issue 1. [DOI: 10.1002/14651858.CD003114.pub2]
Stones W, Cheong YC, Howard FM, Singh S. Interventions for treating chronic pelvic pain in women. *Cochrane Database of Systematic Reviews* 2005, Issue 2. [DOI: 10.1002/14651858.CD000387]
Takken T, Van Brussel M, Engelbert RH, van der Net JJ, Kuis W, Helders PPJM. Exercise therapy in juvenile idiopathic arthritis. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD005954.pub2]
van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SMA, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010387]
van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD011142.pub2]
White CM, Pritchard J, Turner-Stokes L. Exercise for people with peripheral neuropathy. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD003904.pub2]
Zammit GV, Menz HB, Munteanu SE, Landorf KB, Gilheany MF. Interventions for treating osteoarthritis of the big toe joint. *Cochrane Database of Systematic Reviews* 2010, Issue 9. [DOI: 10.1002/14651858.CD007809.pub2]
Østerås N, Kjeken I, Smedslund G, Moe RH, Slatkowsky-Christensen B, Uhlig T, et al. Exercises for hand osteoarthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010388]

Additional references

Abdulla 2013
Abdulla A, Adams N, Bone M, Elliott AM, Gaffin J, Jones D, et al. British Geriatric Society. Guidance on the management of pain in older people. *Age and Ageing* 2013;**42**(Suppl 1):i1–57. [DOI: 10.1093/ageing/afs200; PUBMED: 23420266]

American College of Sports Medicine 2007
American College of Sports Medicine, American Heart Association. Exercise and acute cardiovascular events: placing the risks into perspective. *Medicine and Science in Sport and Exercise* 2007;**39**(5):886–97. [DOI: 10.1249/mss.0b013e3180574e0e; PUBMED: 17468590]

Benton 2011
Benton MJ, Kasper MJ, Raab SA, Waggener GT, Swan PD. Short-term effects of resistance training frequency on body composition and strength in middle-aged women. *Journal of Strength and Conditioning Research/National Strength & Conditioning Association* 2011;**25**(11):3142–9. [DOI: 10.1519/JSC.0b013e31820f505f]

Bigal 2012
Bigal ME, Rapoport AM. Obesity and chronic daily headache. *Current Pain and Headache Reports* 2012;**16**(1):101–9. [DOI: 10.1007/s11916-011-0232-0; PUBMED: 22076673]

Bouchard 2012
Bouchard C, Blair SN, Church TS, Earnest CP, Hagberg JM, Häkkinen K, et al. Adverse metabolic response to regular exercise: is it a rare or common occurrence?. *PLoS One* 2012;**7**(5):e37887. [DOI: 10.1371/journal.pone.0037887; PUBMED: 22666405]

Bousema 2007
Bousema EJ, Verbunt JA, Seelen HA, Vlaeyen JW, Knottnerus JA. Disuse and physical deconditioning in the first year after the onset of back pain. *Pain* 2007;**130**(3):279–86. [PUBMED: 17467902]

Breivik 2006
Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *European Journal of Pain* 2006;**10**(4):287–333. [PUBMED: 16095934]

Cohen 1988
Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edition. Hillsdale, NJ: L. Erlbaum Associates, 1988. [ISBN 0805802835]

Collins 1997
Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? . *Pain* 1997;**72**(1-2):95–7. [DOI: 10.1016/S0304-3959(97)00005-5; PUBMED: 9272792]

Daenen 2015
Daenen L, Varkey E, Kellmann M, Nijs J. Exercise, not to exercise, or how to exercise in patients with chronic pain? Applying science to practice. *Clinical Journal of Pain* 2015;**31**(2):108–14. [DOI: 10.1097/AJP.0000000000000099; PUBMED: 24662498]

Deschartes 2013
Deschartes A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ* 2013;**346**:f2304. [DOI: 10.1136/bmj.f2304]

Dworkin 2005
Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;**113**:9–19. [PUBMED: 15621359]

Dworkin 2008
Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Consensus Statement: Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *Journal of Pain* 2008;**9**(2):105–21. [DOI: 10.1016/j.jpain.2007.09.005; PUBMED: 18055266]

Eccleston 2014
Eccleston C, Fisher E, Craig L, Duggan GB, Rosser BA, Keogh E. Psychological therapies (Internet-delivered) for the management of chronic pain in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD010152.pub2]

Elliott 1999
Elliott AM, Smith BH, Penny KI, Smith WC, Chambers WA. The epidemiology of chronic pain in the community. *Lancet* 1999;**354**:1248–52. [PUBMED: 10520633]

Elzahaf 2012
Elzahaf RA, Tashani OA, Unsworth BA, Johnson MI. The prevalence of chronic pain with an analysis of countries with a Human Development Index less than 0.9: a systematic review without meta-analysis. *Current Medical Research & Opinion* 2012;**28**(7):1221–9. [DOI: 10.1185/03007995.2012.703132]

Engers 2008
Engers A, Jellema P, Wensing M, van der Windt DA, Grol R, van Tulder MW. Individual patient education for low back pain. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD004057.pub3; PUBMED: 18254037]

Enoka 1997
Enoka RM. Neural adaptations with chronic physical activity. *Journal of Biomechanics* 1997;**30**(5):447–55. [PUBMED: 9109556]

Exercise is Medicine
American College of Sports Medicine (ACSM). Exercise is Medicine. http://www.exerciseismedicine.org/ (accessed 27 June 2014).

Finan 2013
Finan PH, Smith MT. The comorbidity of insomnia, chronic pain, and depression: dopamine as a putative mechanism. *Sleep Medicine Reviews* 2013;**17**(3):173–83. [DOI: 10.1016/j.smrv.2012.03.003; PUBMED: 22748562]

Gross 2009
Gross A, Forget M, St George K, Fraser MM, Graham N, Perry L, et al. Patient education for neck pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. [DOI: 10.1002/14651858.CD005106.pub4; PUBMED: 22419306]

Harvard 2013
Better balance: mental and physical fitness are both essential: daily physical activity, targeted balance exercise, and steps to maintain your brain all help prevent falls. *Harvard Men's Health Watch* 2013;**17**(11):5. [PUBMED: 24000423]

Hoffman 2014
Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**. [DOI: 10.1136/bmj.g1687; PUBMED: 24609605]

IDF 2012
Guariguata L. By the numbers: new estimates from the IDF Diabetes Atlas Update for 2012. *Diabetes Research and Clinical Practice* 2012;**98**(3):524–5. [DOI: 10.1016/j.diabres.2012.11.006; PUBMED: 23217268]

IntHout 2015
IntHout J, Ioannidis JPA, Borm GF, Goeman JJ. Small studies are more heterogeneous than large ones: a meta-meta-analysis. *Journal of Clinical Epidemiology* 2015;**68**(8):860–9. [DOI: 10.1016/j.jclinepi.2015.03.017; PUBMED: 25959635]

Leadley 2014
Leadley RM, Armstrong N, Reid KJ, Allen A, Misso KV, Kleijnen J. Healthy aging in relation to chronic pain and quality of life in Europe. *Pain Practice* 2014;**14**(6):547–58. [DOI: 10.1111/papr.12125; PUBMED: 24138082]

Mayer 2008
Mayer J, Mooney V, Dagenais S. Evidence-informed management of low back pain with lumbar extensor strengthening exercises. *The Spine Journal* 2008;**8**(1):96–113. [DOI: 10.1016/j.spinee.2007.09.008; PUBMED: 18164458]

Merskey 2011
Merskey H, Bogduk N. Part III: Pain terms: a current list with definitions and notes on usage. *Classification of Chronic Pain*. 2nd Edition. Seattle: IASP Task Force on Taxonomy, IASP Press, 2011 [update from 1994].

Messier 2013
Messier SP, Mihalko SL, Legault C, Miller GD, Nicklas BJ, DeVita P, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. *JAMA* 2013;**310**(12):1263–73. [DOI: 10.1001/jama.2013.277669; PUBMED: 24065013]

Mooney 2006
Mooney V, Verna J, Morris C. Clinical management of chronic, disabling low back syndromes. In: Morris C editor(s). *Rehabilitation of the Spine: a Practitioner's Manual*. New York: McGraw-Hill, 2006.

Moore 1998
Moore RA, Gavaghanb D, Tramer MR, Collins SL, McQuay HJ. Size is everything - large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. *Pain* 1998;**78**:209–16. [PUBMED: 9870574]

Moore 2008
Moore RA, Barden J, Derry S, McQuay HJ. Managing potential publication bias. In: McQuay HJ, Kalso E, Moore RA editor(s). *Systematic Reviews in Pain Research: Methodology Refined*. Seattle: IASP Press, 2008:15–23. [ISBN: 978–0–931092–69–5]

Moore 2010
Moore RA, Moore OA, Derry S, Peloso PM, Gammaitono AR, Wang H. Responder analysis for pain relief and numbers needed to treat in a meta-analysis of etoricoxib

osteoarthritis trials: bridging a gap between clinical trials and clinical practice. *Annals of Rheumatic Diseases* 2010;**69** (2):374–9. [DOI: 10.1136/ard.2009.107805; PUBMED: 19364730]

Moore 2013
Moore RA, Straube S, Aldington D. Pain measures and cut-offs - 'no worse than mild pain' as a simple, universal outcome. *Anaesthesia* 2013;**68**(4):400–12. [DOI: 10.1111/anae.12148; PUBMED: 23347230]

Moore 2014a
Moore RA, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. *Pain Practice* 2014;**14**(1):79–94. [PUBMED: 23464879]

Moore 2014b
Moore RA, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. Pain Practice 2014; Vol. 14, issue 1:79–94. [DOI: 10.1111/papr.12050]

National Library of Medicine
National Library of Medicine. Medical Subject Headings. search "chronic pain" at www.nlm.nih.gov/mesh/MBrowser.html (accessed 22 August 2014). [MeSH unique ID: D059350]

NICE 2014
National Institute for Health and Care Excellence (NICE). Osteoarthritis. Care and management in adults (CG177), February 2014. www.nice.org.uk/guidance/cg177 (accessed 4 December 2015). [ISBN: 978–1–4731–0426–6]

Nnoaham 2008
Nnoaham KE, Kumbang J. Transcutaneous electrical nerve stimulation (TENS) for chronic pain. *Cochrane Database of Systematic Reviews* 2008, Issue 3. [DOI: 10.1002/14651858.CD003222.pub2; PUBMED: 18646088]

Nüesch 2010
Nüesch E, Trelle S, Reichenbach S, Rutjes AW, Tschannen B, Altman DG, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;**341**:c3515. [PUBMED: 20639294]

Nüesch 2011
Nüesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Jüni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011;**342**:d1165. [10.1136/bmj.d1165]

Park 2012
Park J, Hughes AK. Non-pharmacological approaches to the management of chronic pain in community-dwelling older adults: a review of empirical evidence. *Journal of the American Geriatric Society* 2012;**60**(3):555–68. [DOI: 10.1111/j.1532-5415.2011.03846.x; PUBMED: 22288789]

Perruchoud 2014
Perruchoud C, Buchser E, Johanek LM, Aminian K, Paraschiv-Ionescu A, Taylor RS. Assessment of physical activity of patients with chronic pain. *Neuromodulation* 2014;**17**:42–7. [DOI: 10.1111/ner.12036]

RevMan 2014 [Computer program]
The Nordic Cochrane Centre, Cochrane. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, Cochrane, 2014.

Sakuma 2012
Sakuma M, Endo N. Space flight/bedrest immobilization and bone. Exercise training for osteoporosis. *Clinical Calcium* 2012;**22**(12):1903–7. [DOI: CliCa121219031907; PUBMED: 23187084]

Schofield 2011
Schofield P, Clarke A, Jones D, Martin D, McNamee P, Smith BH. Chronic pain in later life: a review of current issues and challenges. *Aging Health* 2011;**7**(4):551–6. [DOI: 10.2217/ahe.11.41]

Shea 2007
Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007;**7**:10. [DOI: 10.1186/1471-2288-7-10; PUBMED: 17302989]

SIGN 2013
Scottish Intercollegiate Guidelines Network (SIGN). *Management of chronic pain. SIGN publication no. 136.* 1st Edition. Edinburgh, UK: SIGN, 2013. [ISBN: 978–1–909103–17–7]

Smith 2007
Smith BH, Macfarlane GJ, Torrance N. Epidemiology of chronic pain, from the laboratory to the bus stop: time to add understanding of biological mechanisms to the study of risk factors in population-based research?. *Pain* 2007;**127**(1-2):5–10. [DOI: 10.1016/j.pain.2006.11.001; PUBMED: 17140732]

Stagg 2011
Stagg NJ, Mata HP, Ibrahim MM, Henriksen EJ, Porreca F, Vanderah TW, et al. Regular exercise reverses sensory hypersensitivity in a rat neuropathic pain model: role of endogenous opioids. *Anaesthesiology* 2011;**114**(4):940–8. [DOI: 10.1097/ALN.0b013e318210f880; PUBMED: 21386701]

Straube 2015
Straube S, Derry S, Moore RA, McQuay HJ. Vitamin D for the treatment of chronic painful conditions in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 5. [DOI: 10.1002/14651858.CD007771.pub3; PUBMED: 25946084]

Thornton 2000
Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *Journal of Clinical Epidemiology* 2000;**53**(2):207–16. [DOI: 10.1016/S0895-4356(99)00161-4]

Timmons 2014
Timmon J, Phillips SM. Do we have a valid public health strategy for exercise prescription?. *Sport and Exercise Scientist* 2014;**39**:26–7.

To 2012
To T, Stanojevic S, Moores G, Gershon AS, Bateman ED, Cruz AA, et al. Global asthma prevalence in adults: findings from the cross-sectional world health study. *BMC Public Health* 2012;**12**:204. [DOI: 10.1186/1471-2458-12-204; PUBMED: 22429515]

Torrance 2010
Torrance N, Elliott AM, Lee AJ, Smith BH. Severe chronic pain is associated with increased 10 year mortality. A cohort record linkage study. *European Journal of Pain* 2010;**14**(4): 380–6. [DOI: 10.1016/j.ejpain.2009.07.006]

Turk 2003
Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003; **106**:337–45. [PUBMED: 14659516]

Turner 2013
Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane Reviews. *PLoS One* 2013;**8**(3):e59202. [DOI: 10.1371/journal.pone.0059202; PUBMED: 23544056]

Ursini 2011
Ursini F, Naty S, Grembiale RD. Fibromyalgia and obesity: the hidden link. *Rheumatology International* 2011;**31**(11): 1403–8. [DOI: 10.1007/s00296-011-1885-z; PUBMED: 21476098]

van Hecke 2013a
van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology - where do lifestyle factors fit in?. British Journal of Pain 2013; Vol. 7, issue 4:209–17. [DOI: 10.1177/2049463713493264]

van Hecke 2013b
van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology and its clinical relevance. *British Journal of Anaesthesia* 2013;**111**(1):13–8. [DOI: 10.1093/bja/aet123; PUBMED: 23794640]

Vigorito 2014
Vigorito C, Giallauria F. Effects of exercise on cardiovascular performance in the elderly. *Frontiers in Physiology* 2014;**5**: 51. [DOI: 10.3389/fphys.2014.00051]

Vos 2015
Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;**386**(9995):743–800. [DOI: 10.1016/S0140-6736(15)60692-4; PUBMED: 26063472]

Waddell 1987
Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine* 1987;**12**(7):632–44. [PUBMED: 2961080]

WHO 2015
World Health Organization. Physical activity. Fact sheet 385. www.who.int/mediacentre/factsheets/fs385/en/ (accessed 4 December 2015).

Wieland 2013
Wieland LS, Skoetz N, Manheimer E, Pilkington K, Vempati R, Berman BM. Yoga treatment for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 7. [DOI: 10.1002/14651858.CD010671]

Williams 2012
Williams ACDC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD007407.pub3]

References to other published versions of this review

Geneen 2014
Geneen L, Smith B, Clarke C, Martin D, Colvin LA, Moore RA. Physical activity and exercise for chronic pain in adults: an overview of Cochrane reviews. *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD011279]

* *Indicates the major publication for the study*

ADDITIONAL TABLES

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews

Criteria	Specific requirements (possible answers: yes, no, cannot answer, not applicable)
1. Was an 'a priori' design used?	The research question and inclusion criteria should be established before the conduct of the review *Note: need to refer to a protocol, ethics approval, or predetermined/a priori published research objectives to score a "yes."*

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews (*Continued*)

2. Was there duplicate study selection and data extraction?	There should be at least 2 independent data extractors and a consensus procedure for disagreements should be in place *Note: 2 people do study selection, 2 people do data extraction, consensus process or 1 person checks the other person's work.*
3. Was a comprehensive literature search performed?	At least 2 electronic sources should be searched. The report must include years and databases used (e.g. CENTRAL, MEDLINE, and Embase). Keywords or MeSH terms (or both) must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialised registers, or experts in the particular field of study, and by reviewing the references in the studies found *Note: if at least 2 sources + 1 supplementary strategy used, select "yes" (Cochrane register/ CENTRAL counts as 2 sources; a grey literature search counts as supplementary).*
4. Was the status of the publication (i.e. grey literature) used as inclusion criteria?	The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language, etc *Note: if review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished literature.*
5. Was a list of studies (included and excluded) provided?	A list of included and excluded studies should be provided. *Note: acceptable if the excluded studies were referenced. If there was an electronic link to the list but the link is no longer active, select "no. "*
6. Were the characteristics of the included studies provided?	In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions, and outcomes. The ranges of characteristics in all the studies analysed, e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported *Note: acceptable if not in table format as long as they are described as above.*
7. Was the scientific quality of the included studies assessed and documented?	'A priori' methods of assessment should be provided (e.g. for effectiveness studies if the author(s) chose to include only randomised, double-blind, placebo-controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant *Note: can include use of a quality scoring tool or checklist, e.g. Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some type of result for EACH study ("low" or "high" is acceptable, as long as it is clear which studies scored "low" and which*

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews (*Continued*)

	scored "high;" a summary score/range for all studies is not acceptable).
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations *Note: might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.*
9. Were the methods used to combine findings of studies appropriate?	For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi^2 test for homogeneity, I^2 statistic). If heterogeneity exists, a random-effects model should be used or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?), or both *Note: indicate "yes" if they mention or describe heterogeneity, i.e. if they explain that they cannot pool because of heterogeneity/variability between interventions.*
10. Was the likelihood of publication bias assessed?	An assessment of publication bias should include a combination of graphical aids (e.g. funnel plot, other available tests) or statistical tests (e.g. Egger regression test), or both *Note: if no test values or funnel plot included, score "no." Score "yes" if they mention that publication bias could not be assessed because there were fewer than 10 included studies.*
11. Was the conflict of interest stated?	Potential sources of support should be clearly acknowledged in both the systematic review and the included studies *Note: to get a "yes," must indicate source of funding or support for the systematic review AND for each of the included studies.*

Table 2. Reasons for exclusion

Review	Reason for exclusion from overview
Aggarwal 2011	Not exercise/physical activity
Brønfort 2015	Protocol stage only - possibly include when published as full review
Bierma-Zeinstra 2011	Protocol stage only - exclude when published as full review
Brønfort 2014	Withdrawn from the Cochrane Library
Choi 2010	Not chronic using definition of > 3 months
Craane 2006	Protocol stage only - possibly include when published as full review

Table 2. Reasons for exclusion (*Continued*)

Dagfinrud 2008	Physiotherapy - required therapist to perform intervention
Dahm 2010	Acute pain, not chronic. Intervention was advice
Dal Bello-Haas 2013	Malignant condition
de Souza 2012	Drug- and surgery-based interventions
Fokkenrood 2013	Did not include RCTs (excluded studies with control groups)
Franke 2015	Not exercise/physical activity
Green 2003	Physiotherapy - required therapist to perform intervention
Gross 1998	Withdrawn from the Cochrane Library
Gross 2012	Not exercise/physical activity
Gross 2015b	Not exercise/physical activity
Hayden 2012	Protocol stage only - possibly include when published as full review
Heintjes 2003	Withdrawn from the Cochrane Library January 2015
Henschke 2010	Not exercise/physical activity
Heymans 2004	Exercise could not be assessed as stand-alone intervention
Hilde 2006	Withdrawn from the Cochrane Library
Hoving 2014	No exercise intervention, and no pain outcome measure
Hurley 2013	Protocol stage only - exclude when published as full review
IJzelenberg 2011	Protocol stage only - exclude when published as full review
Jones 2000	Drug-based interventions
Jordan 2010	Intervention to improve adherence to exercise, not exercise itself
Kamper 2014	Exercise could not be assessed as stand-alone intervention
Karjalainen 1999	Exercise could not be assessed as stand-alone intervention
Karjalainen 2003	Exercise could not be assessed as stand-alone intervention

Table 2. Reasons for exclusion (*Continued*)

Larun 2016	Chronic fatigue, not chronic pain
Liddle 2015	Pain in pregnancy only, not chronic pain
Liu 2013	Protocol stage only - unsure about inclusion when published as full review
Miller 2014	Protocol stage only - exclude when published as full review
Moi 2013	Exercise could not be assessed as stand-alone intervention
O'Brien 2004	No pain outcome measure
O'Connell 2013	Overview of reviews, not systematic review
Østerås 2013	Protocol stage only - possibly include when published as full review
Page 2012	No pain outcome measure
Page 2014	Manual therapy - required therapist to perform intervention
Peters 2013	Exercise could not be assessed as stand-alone intervention
Preston 2004	No pain outcome measure
Proctor 2007	Exercise could not be assessed as stand-alone intervention
Radner 2012	Drug-based interventions
Regnaux 2014	Protocol stage only - possibly include when published as full review
Richards 2012	Not exercise/physical activity
Riemsma 2003	Not exercise/physical activity
Schaafsma 2013	No pain outcome measure
Steultjens 2004	Occupational therapy - exercise could not be assessed as stand-alone intervention
Stones 2005	Exercise cannot be assessed as stand-alone intervention
Takken 2008	Aged < 18 years - not adults
van Dessel 2014	Not chronic pain and no specific pain outcome measure
White 2004	No pain outcome measure
Williams 2012	Not exercise/physical activity

Table 2. Reasons for exclusion (*Continued*)

Zammit 2010	Surgery or required therapist to perform intervention

RCT: randomised controlled trial.

Table 3. Characteristics of included reviews

Review and Cochrane Review Group	Assessed as up to date	Chronic pain condition	Duration of pain/ diagnosis	Intervention description	Control description	Outcomes with data reported	Time points reported
Bartels 2007 Cochrane Musculoskeletal Group	Aug 2007	Hip or knee OA	Not reported	All types of exercises developed in the therapeutic/heated indoor pool (ROM, dynamics, aerobics, etc.) were permitted	No treatment or other treatment.	Function, quality of life, mental health, pain, adverse events	Post-intervention (immediate), 6-month follow-up
Bidonde 2014 Cochrane Musculoskeletal Group	Oct 2013	Fibromyalgia	12 yr (range 6 to 24)	Aquatic exercise training intervention defined as "exercise conducted in a vertical standing position."	Treatment as usual, physical activity as usual, wait list control, placebo or sham, education-only, water immersion-only, and attention only	Multi-dimensional function (wellness), self-reported physical function (wellness), pain (symptoms), stiffness (symptoms), muscle strength (physical fitness), submaximal cardiorespiratory function (physical fitness), withdrawals (safety and acceptability), adverse effects (safety and acceptability)	Post-intervention (4 to 32 wk)

Table 3. Characteristics of included reviews (*Continued*)

Boldt 2014 Cochrane Injuries Group	Mar 2011	Spinal cord injury	Mean 66 months, and 1 to 24 yr when reported	"Exercise": stretching and strengthening exercises aimed at mobilising painful shoulder joint	Wait list control or no intervention.	Pain, depression, quality of life, adverse effects	Short term (within 24 hours of last intervention, i.e. post-intervention) and intermediate term (1 to 6 wk post-intervention) and long term (> 6 wk post-intervention)
Brown 2010 Cochrane Menstrual Disorders and Subfertility Group	Aug 2009	Primary dysmenorrhoea in the majority (≥ 50%) of cycles	Ongoing/not appropriate	12-wk walk or jog training programme at an intensity of 70% to 85% of the HR range. Training for 3 days/wk and duration of aerobic phase was 30 minutes with 15-minute warm-up and cool-down periods	Asked not to exercise during the experimental period.	Pain: menstrual disorders questionnaire (MDQ) score	Ongoing - over 3 menstrual cycles
Busch 2007 Cochrane Musculoskeletal Group	Aug 2007	Fibromyalgia	Not reported	Exercise-only interventions included aerobic-only training, strength-only training, flexibility-only training, or mixed exercise-only interventions	"Untreated."	Pain, global wellbeing, objectively measured physical function	Post-intervention (strength exercise 21 wk, aerobic exercise 6 to 23 wk)
Busch 2013 Cochrane Musculoskeletal Group	Mar 2013	Fibromyalgia	mean range from 4 yrs (SD 3.1) to 12 yrs (SD 4)	Defined resistance training as exercise performed against a progressive	Untreated control conditions (treatment as usual, activity as usual, wait	Multidimensional function, self-reported physical function,	Post-intervention, follow-up (12 wk) in 1 study only

Table 3. Characteristics of included reviews (*Continued*)

				resistance on a minimum of 2 days/wk (on non-consecutive days) with the intention of improving muscle strength, muscle endurance, muscle power, or a combination of these	list control, and placebo), other types of exercise or physical activity interventions (e.g. aerobic, flexibility), and other resistance training interventions (head-to-head comparisons)	pain, tenderness, muscle strength, adverse effects, all-cause attrition	
Cramp 2013 Cochrane Musculoskeletal Group	Oct 2012	Rheumatoid arthritis	Not reported	Included pool-based therapy (twice/wk, moderate intensity, music-paced), yoga (6 wk, twice/wk, 1.5-hour sessions), dynamic strength training (home-based after inpatient programme, all main muscle groups using dumbbells and elastic bands), stationary cycling (70% HRmax, 5 minute excluding: 1-minute of rest, increased duration), low-impact aerobics (class	"Could have been placebo, an alternative intervention (pharmacological or non-pharmacological) or usual care."	Fatigue, pain, anxiety, depression, disability, tender and swollen joints, adverse events	Post-intervention (only a single time point analysed)

Table 3. Characteristics of included reviews (*Continued*)

				at fitness centre and video at home, individual HR targets), tai chi (1-hour group sessions)			
Fransen 2014 Cochrane Musculoskeletal Group	May 2013	Hip OA	Not reported	Any land-based therapeutic exercise regimens aiming to relieve the symptoms of hip OA, regardless of content, duration, frequency, or intensity. This included any exercise designed to improve muscle strength, range of joint movement or aerobic capacity (or combinations of the three). Programmes could be designed and supervised by physiotherapists or other professionals, or provided as a home programme with minimal monitoring	Wait-list control, usual care, GP education.	Self-reported pain, physical function, quality of life, withdrawal or dropouts, adverse events	post-intervention (immediate in 9/10 studies) follow-up 3 to 6 months
Fransen 2015 Cochrane Musculoskeletal Group	May 2013	Knee OA	Often not reported: some less than 1yr, oth-	"land-based therapeutic exercise." Along	No exercise: active (any no-exercise intervention) or no	Knee pain, self-reported physical function, quality of	Immediately at the end of treatment (post-

Table 3. Characteristics of included reviews (*Continued*)

			ers over 10yr	with delivery mode and content, treatment 'dosage' (duration, frequency, intensity) varied widely between studies	treatment (including waiting list)	life	treatment), 2 to 6 months after cessation of monitored study treatment and longer than six months after cessation of monitored study treatment
Gross 2015a Cochrane Back Group	May 2014	Mechanical neck disorders	"Chronic" (not subacute or acute)	Cervical stretch/ROM exercises + cervical/ scapulothoracic strengthening + static/ dynamic cervical/shoulder stabilisation	Wait list control.	Pain intensity, function, quality of life, global perceived effect, adverse effects	Immediately post-treatment (≤ 1 day), short-term follow-up (1 day to 3 months), intermediate-term follow-up (3 months up to, but not including, 1 yr), and long-term follow-up (≥ 1 yr)
Han 2004 Cochrane Musculoskeletal Group	Apr 2004	Rheumatoid arthritis	Not reported	Only trials of exercise programmes with tai chi instruction or incorporating principles of tai chi philosophy	Not reported.	Function, tender and swollen joints, ROM, strength, enjoyment, withdrawals, adverse effects	Post-intervention (8 to 10 wk)
Hayden 2005 Cochrane Back Group	Sep 2004	Non-specific low back pain	Chronic, i.e. longer than 12 wk: 5.6 yr (95% CI 3.4 to 7.8)	Exercise therapy defined as "a series of specific movements with the aim of training or developing the	No exercise: no treatment or placebo treatment, other conservative therapy, or another ex-	Pain, functional ability, work status, global assessment, adverse events	Earliest, 6 wk, 6 months, 12 months

Table 3. Characteristics of included reviews (*Continued*)

				body by a routine practice or as physical training to promote good physical health;" only 54% adequately described the exercise intervention	ercise group		
Hurkmans 2009 Cochrane Musculoskeletal Group	Jun 2009	Rheumatoid arthritis	5 to 14 yr	Dynamic exercise programmes - aerobic capacity and muscle strength training; short-term muscle strength training (high quality); short-term dynamic exercise to improve aerobic capacity (not high methodological quality); exercise frequency of at least 20 minutes twice a week. Duration of exercise programme at least 6 wk (duration < 3 months was considered short-term; duration > 3 months was considered long-term)	Not reported	Functional ability, aerobic capacity, muscle strength, safety (pain and radiological damage)	Follow-up (12 wk and 24 months)

Table 3. Characteristics of included reviews (*Continued*)

				. Exercise programme performed under supervision Aerobic exercise intensity at least 55% of the maximum HR; or intensity starting at 40% to 50% of the maximum oxygen uptake reserve or HR maximum reserve. Furthermore, the intensity was increased up to 85% during the intervention. Progressively strengthening exercise loads starting at 30% to 50% and increasing to 80% of maximum (defined as the percentage of either 1 repetition maximum, 1 MVC, maximum speed, or as maximal subjective exertion)			
Koopman 2015 Cochrane Neuromuscular Group	Jul 2014	Postpolio syndrome (PPS)	Not reported	Exercise therapy (e.g. aerobic exercise, muscle strengthening exercise, respiratory muscle	Placebo, usual care or no treatment.	Self-perceived activity limitations, muscle strength, muscle endurance, fatigue, pain, adverse events	3 and 6 months

Table 3. Characteristics of included reviews (*Continued*)

				training, warm climate training, hydro training)		(minor and serious)	
Lane 2014 Cochrane Peripheral Vascular Diseases Group	Sep-2013	intermittent claudication	not reported	Any exercise programme used in the treatment of intermittent claudication was included, such as walking, skipping and running. Inclusion of trials was not affected by the duration, frequency or intensity of the exercise programme but these issues were taken into account in the meta-analysis	Exercise was compared to six different modes of treatment, the most common being usual care or placebo. Two early trials compared exercise with placebo tablets but in more recent studies usual care was used as the control comparator. Exercise was compared with the following drug therapies: antiplatelet agents pentoxifylline, iloprost, and vitamin E. One study compared exercise with pneumatic foot and calf compression	maximal walking time, pain-free walking time, pain-free walking distance, maximum walking distance, ankle brachial index (ABI), peak exercise calf blood flow, mortality, amputation	Post-intervention, 3-month follow up, six-month follow up
Lauret 2014 Cochrane Peripheral Vascular Diseases Group	Jul 2013	Intermittent claudication	Not reported	Supervised walking programme needed to be supervised at least twice a week for a	Alternative exercise.	Maximum walking distance (METs), pain-free walking distance (METs), health-related	n/a

Table 3. Characteristics of included reviews (*Continued*)

				consecutive 6 wk of training		quality of life and functional impairment	
Regnaux 2015 Cochrane Musculoskeletal Group	Jun 2014	Hip or knee OA	> 6 months	High-intensity physical activity or exercise programme.	Low-intensity physical activity or exercise programme and control (no-exercise) group in 1 study.	Pain, physical function, quality of life, adverse effects (related to intervention), severe adverse events or withdrawal (due to intervention)	Post-intervention, intermediate term (6 to 12 months), long-term (over 12 months) follow-up
Saragiotto 2016 Cochrane Back and Neck Group	Apr 2015	Low back pain	> 12 wk	MCE: activation of the deep trunk muscles, targeting the restoration of control and co-ordination of these muscles	Placebo, no treatment, another active treatment, or when MCE was added as a supplement to other interventions. When MCE was used in addition to other treatments, it had to represent at least 50% of the total treatment programme to be included	Pain intensity and disability, function, quality of life, global impression of recovery, return to work, adverse events and recurrence	Post-intervention, short term (4 to 10 wk), intermediate term (3 to 6 months), long term (12 to 36 months)
Silva 2010 Cochrane Musculoskeletal Group	Jun 2009	Rheumatoid arthritis	No studies found	Balance training (proprioceptive training).	No intervention or other intervention.	ACR-50, pain, disease activity score (DAS), Health Assessment Questionnaire (HAQ for function), gait, adverse effects, discontinuation rate	n/a

Table 3. Characteristics of included reviews (*Continued*)

van der Heijden 2015 Cochrane Bone, Joint and Muscle Trauma Group	May 2014	Adolescents and adults with patellofemoral pain	3 wk to 8 months (as minimum requirement) ; reported pain 4 wk to 9 yr	Exercise therapy for patellofemoral pain syndrome; exercises could be performed at home or under supervision of a therapist - various descriptions in the included trials, including knee exercises, hip and knee exercises, home exercises, supervised exercises, closed kinetic chain, open kinetic chain	No treatment, placebo, or waiting list controls. This also included 'exercise therapy + another intervention (e.g. taping) versus the other intervention alone (e.g. taping).'	Pain during activity, usual pain, functional ability, recovery	4- to 12-wk follow-up (short term) and 16 wk to 12 months (long term)
Yamato 2015 Cochrane Back Group	Mar 2014	Low back pain	Acute, subacute, chronic (i.e. no minimum)	Explicitly stated as based on Pilates principles, or the therapists who provided the interventions had previous training in Pilates exercises or the therapists were described as certified Pilates instructors	No intervention, placebo, or other interventions.	Pain intensity, disability, global impression of recovery, quality of life, return to work, adverse effects	Short term (4 to 8 wk), intermediate term (3 to 6 months)

ACR: American College of Rheumatology; GP: general practitioner; HR: heart rate; MCE: motor control exercise; MET: metabolic equivalents; n/a: not applicable; OA: osteoarthritis; ROM: range of motion; wk: week; yr: year.

Table 4. Further characteristics of included reviews

Review	Number of trials included	Total number of participants	Gender distribution	Participants ages
Bartels 2007	6 (4 exercise vs no exercise)	800 (674 exercise vs no exercise)	50% to 86% Female	Means ranged from 66 to 71 yr
Bidonde 2014	16 (9 exercise vs no exercise)	881 (519 exercise vs no exercise)	513 female, 6 male	Means ranged from 46.3 to 48.3 yr
Boldt 2014	16 (3 exercise vs no exercise)	616 (149 exercise vs no exercise)	115 male, 41 female across 3 studies	Range 19 to 65 yr and mean 35 to 45 yr
Brown 2010	1	36	100% female	Not reported
Busch 2007	34 (in meta-analysis - strength training vs control: 2; aerobic training vs control: 4)	2276 total (in meta-analysis - strength: 47, aerobic: 269)	96.4% female when reported (in 2197 participants)	Range reported as 27.5 to 60.2 yr
Busch 2013	5 studies as 7 publications (exercise vs control: 3 publications, 2 studies)	219 with fibromyalgia (exercise vs control: 81)	100% female	Not reported
Cramp 2013	24 (only 6 using physical activity interventions)	2882 (physical activity interventions: 371)	"A higher percentage of females"… when reported	"Mainly within the fifth decade"
Fransen 2014	10	> 549	75% to 80% female when reported	58 to 70 yr (means) when reported
Fransen 2015	54	5362	When reported 55% to 100% female	When reported mean age 60 to 70 yr
Gross 2015a	27 (16 chronic pain)	2485	Not reported	Not reported
Han 2004	4 (3 RCTs). Pain not reported in any included study	206 total; pain not reported in any included study	Not reported	Range 38 to 72 yr
Hayden 2005	61 (43 chronic low back pain)	6390 (3907 chronic low back pain)	Chronic: 46% male (95% CI 39 to 52)	Chronic: 42 yr (95% CI 40 to 44)
Hurkmans 2009	8 RCTs (5 exercise vs no-exercise)	575	"Mainly female"	52 yr

Table 4. Further characteristics of included reviews (*Continued*)

Koopman 2015	13 (2 exercise vs no exercise)	675 (68 exercise vs no exercise) - 1 study used 3 arms (no treatment in cold, exercise in cold, exercise in warm; we have excluded the warm exercise arm as cannot compare directly to the control)	~ 25% male	Mean 58 and 65 yr
Lane 2014	30	1822 total	Not reported	Mean > 65 yr
Lauret 2014	5 (0 for exercise vs no exercise)	184 (0 for exercise vs no exercise)	n/a	n/a
Regnaux 2015	6 (1 for exercise vs no exercise) only 1 study that had a no exercise control	656 (102 for exercise vs no exercise)	79 female	62.6 yr
Saragiotto 2016	29 (7 for exercise vs no exercise/minimal intervention)	2431 (671 for exercise vs no exercise)	"Mixed"	Median 40.9 yr (IQR 11.2) (range 20.8 to 54.8)
Silva 2010	None	None	n/a	n/a
van der Heijden 2015	31 (10 for exercise vs control)	1690	0% to 100% female; equally distributed across range	Mean 25 to 50 yr
Yamato 2015	10 (6 exercise vs minimal intervention (control))	478 (265 exercise vs control)	2 trials were all female, the others included both genders	Mean 38 yr (range 22 to 50)

CI: confidence interval; GP: general practitioner; IQR: interquartile range; OA: osteoarthritis; RCT: randomised controlled trial; ROM: range of motion; wk: week; yr: year.

Table 5. Dose and duration of exercise interventions in included reviews

Review	Duration	Frequency (sessions per day/wk/month)	Intensity	Duration (per session)	Other description
Bartels 2007	Not reported	Not reported	"Muscle maintenance" and "range of motion"	Not reported	No minimum requirement for inclusion. Actual intervention only reported by 2

Table 5. Dose and duration of exercise interventions in included reviews *(Continued)*

					of 6 included studies
Bidonde 2014	17 wk (range 4 to 32)	1 to 4/wk	Very light (< 57% HRmax) to vigorous (95% HRmax), self-selected, and not specified	45 minutes (range 30 to 70)	No minimum requirement for inclusion. None of the studies met the ACSM exercise guidelines specified for aerobic or strength training. Only 1 study met the ACSM guidelines for flexibility training
Boldt 2014	12 wk to 9 months	2/day to 2/wk	Not reported	Reported for 1 study only (90 to 120 minutes)	No minimum requirement for inclusion. Stretching and strengthening exercises aimed at mobilising painful shoulder joint
Brown 2010	≥ 12 wk	3/wk	70% to 85% HRR	1 hour	No minimum requirement for inclusion.
Busch 2007	3 wk to 6 months	1 to 5/wk	Not reported	Not reported	No minimum requirement for inclusion. Assessed as whether they "met ACSM recommendations."
Busch 2013	8 to 21 wk (median 16 wk)	≥ 2/wk	> 4/10 RPE rating progressing to 70% to 80% 1RM	40 to 90 minutes	Assessed as whether they "met ACSM recommendations."
Cramp 2013	6 wk (when reported)	2/wk	"Low impact", "moderate", and 70% HRmax	1 to 1.5 hours, when reported	No minimum requirement for inclusion.
Fransen 2014	6 to 12 wk (median 8)	1 to 3/wk	"Low intensity" to "max effort"	30 to 60 minutes	No minimum requirement for inclusion. Intensity only re-

Table 5. Dose and duration of exercise interventions in included reviews (*Continued*)

					ported in 2 of 10 studies.
Fransen 2015	single session to 30 months	1 to 5/wk	"Moderate to moderately high intensity"	15 to 60 minutes	No minimum requirement for inclusion. Varied in dose and duration.
Gross 2015a	2 wk to 3 months	5/wk to every 15 minutes/day	Low intensity	2 to 20 minutes	-
Han 2004	8 to 10 wk (when reported)	1 to 7/wk (median 1/wk)	Tai chi = low intensity	1 to 1.5 hours	No minimum requirement for inclusion.
Hayden 2005	Not reported	Not reported	Not reported	Not reported	No minimum requirement for inclusion. Could not extract actual data.
Hurkmans 2009	≥ 6 wk	2/wk	Aerobic: ≥ 55% HRmax increasing to 85% HRmax strength: start 30% 1RM increasing to 80% 1RM	20 minutes	-
Koopman 2015	4 to 12 wk	Daily to 3/wk	Reported in 1 study: 50% to 70% MVC	45 minutes	No minimum requirement for inclusion. 1 study: supervised progressive resistance training consisting of 3 sets of 8 isometric contractions of the thumb muscles 1 study: combination of individual and group therapy with daily treatment in a swimming pool (45 minutes), physiotherapy, individually adapted training programme

Table 5. Dose and duration of exercise interventions in included reviews *(Continued)*

Lane 2014	3 to 12 months	≥ 2/wk	"Variable"	~ 60 minutes	No minimum requirement for inclusion.
Lauret 2014	≥ 6 wk	≥ 2/wk	Not reported	Not reported	No minimum requirement for inclusion. Must be supervised.
Regnaux 2015	8 wk	3/wk	Compared high vs low intensity vs control	30 to 50 minutes	Every 2 wk 1RM was retested and increased by 5% as tolerated in each group Supervision: an experienced therapist. 3 arms (n=34 per arm): high intensity, low intensity, control (no exercise)
Saragiotto 2016	20 days to 12 wk (median 8 wk (IQR 2.0))	1 to 5/wk (median 12 sessions (IQR 6.0))	Not reported	20 to 90 minutes (median 45 (IQR 30) minutes)	MCE is usually delivered in 1:1 supervised treatment sessions, and sometimes involves ultrasound imaging, the use of pressure biofeedback units or palpation to provide feedback on the activation of trunk muscles
Silva 2010	≥ 6 wk	2/wk	Balance training only	≥ 30 minutes	No studies found.
van der Heijden 2015	3 to 16 wk	2/wk to daily	Not reported	Not reported	No minimum requirement for inclusion. Assessed by duration (< or > 3 months), frequency (sev- eral times, or once a week), medium (land or water), etc

Table 5. Dose and duration of exercise interventions in included reviews *(Continued)*

Yamato 2015	10 to 90 days (mostly 8 wk)	2/wk (mean session number 15.3, range 6 to 30)	Not reported	1 hour	No minimum requirement for inclusion. Must be supervised (for the Pilates technique).

1RM: one repetition maximum; ACSM: American College of Sport Medicine; HRmax: maximum heart rate; HRR: heart rate reserve, IQR: interquartile range; MCE: motor control exercise; MVC: maximum voluntary contraction; RPE: rating of perceived exertion; wk: week.

Table 6. Methodological quality of included reviews using the AMSTAR tool

Review	Criteria											Total "Y"	Total "N"	Total "n/a"
	1	2	3	4	5	6	7	8	9	10	11			
Bartels 2007	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Bidonde 2014	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N	8	3	-
Boldt 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Brown 2010	Y	Y	Y	N	Y	Y	Y	Y	n/a	N	N	7	3	1
Busch 2007	Y	Y	Y	N	Y	Y	Y	Y	Y	N	N	8	3	-
Busch 2013	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Cramp 2013	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Fransen 2014	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Fransen 2015	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-

Table 6. Methodological quality of included reviews using the AMSTAR tool (*Continued*)

Gross 2015a	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Han 2004	Y	Y	Y	Y	Y	N	Y	Y	N	N	N	7	4	-
Hayden 2005	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	10	2	-
Hurkmans 2009	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9	2	-
Koopman 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	10	1	-
Lane 2014	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	9	2	-
Lauret 2014	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Regnaux 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Saragiotto 2016	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Silva 2010	Y	Y	Y	Y	Y	n/a	n/a	n/a	n/a	n/a	Y	6	0	5
van der Heijden 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9	2	-
Yamato 2015	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Total "Y"	**20**	**21**	**21**	**19**	**21**	**10**	**20**	**20**	**17**	**10**	**3**	-	-	-
Total "N"	**1**	**-**	**-**	**2**	**-**	**10**	**-**	**-**	**2**	**10**	**18**	-	-	-

Table 6. Methodological quality of included reviews using the AMSTAR tool (*Continued*)

Total "n/a"	-	-	-	-	-	1	1	1	2	1	-	-	-	-

N: no; n/a: not applicable; Y: yes; out of maximum summative score of 11.
Following arbitration, the authors removed the response "cannot answer" due to no responses as such.

Table 7. Risk of bias - studies assessed as low risk of bias

Review	Number of studies in assessment	Selection bias		Performance bias	Detection bias	Attrition bias	Reporting bias	Other bias	
		Random sequence generation (studies)	Allocation concealment (studies)	Blinding of participants and personnel (studies)	Blinding of outcome assessment (studies)	Incomplete outcome data (studies)	Selective reporting (studies)	Sample size	Other biases (studies)
Bartels 2007	6	Not reported	3	Not reported	2	3	Not reported	2, n > 100 per arm	-
Bidonde 2014	9	5	3	2	8	8	5	1, n > 50 per arm	7
Boldt 2014	3	1	1	0	1	2	3	0	1
Brown 2010	1	0	0	0	0	1	1	1, n > 50 per arm	-
Busch 2007	34	17	10	8	20	Unclear	32	5, n > 50 per arm	-
Busch 2013	5	4	2	1	2	5	3	0, n > 50 per arm	-
Cramp 2013	7	5	2	0	Not reported	6	4		1
Fransen 2014	10	8	7	0	0	7	4	1, n > 50 per arm	7
Fransen 2015	54	40	22	3	4	29	10	5, total n > 200	
Gross 2015a	16	8	8	1	0	11	0	0	11

Table 7. Risk of bias - studies assessed as low risk of bias *(Continued)*

Han 2004	4	2	0	0	0	0	Not reported	0	
Hayden 2005	43	27	22	Not reported	12	29	Not reported	10, total n > 100 + 5, total n > 200	-
Hurkmans 2009	8	8	1	-	4	5	-	1, total n > 200	1
Koopman 2015	2	1	0	0	0	0	0	0	1
Lane 2014	30	16	14	30	7	19	29	3, total n > 100	
Lauret 2014	5	4	2	5	3	4	5	1, total n > 100	4
Regnaux 2015	1	1	0	0	1	0	0	1, total n > 100	1
Saragiotto 2016	7	5	4	1	1	2	7	1, total n > 100 + 1, total n > 200	7
Silva 2010	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
van der Heijden 2015	10	8	6	0	0	6	9	2, total n > 100	10
Yamato 2015	9	5	5	2	7	7	9	0	9
Studies with low risk of bias (number)	**264**	**165**	**112**	**53**	**72**	**144**	**121**	**total n > 100: 26 total n > 200: 15 total n > 400: 0**	**71**
Studies with low risk of bias (per-	**-**	**63%**	**42%**	**20%**	**27%**	**55%**	**46%**	**total n > 100: 10% total n >**	**27%**

Table 7. Risk of bias - studies assessed as low risk of bias *(Continued)*

centage)								**200: 6% total n > 400: 0%**	

n: number of participants, n/a: not applicable.

Table 8. Interpretation of results by original review authors

Review	Review authors' conclusions	Overview authors' assessment of conclusions
Bartels 2007	"Aquatic exercise has some short-term beneficial effects on the condition of OA patients with hip or knee OA or both. The controlled and randomised studies in this area are still too few to give further recommendations on how to use this therapy... No long-term effects have been found."	Appropriate conclusions based on available data. No mention of quality/risk of bias in conclusions, though found to be high quality in results section
Bidonde 2014	"Low to moderate quality evidence relative to control suggests that aquatic training is beneficial for improving wellness, symptoms, and fitness in adults with fibromyalgia. Very low to low quality evidence suggests that there are benefits of aquatic and land-based exercise, except in muscle strength (very low quality evidence favoring land). No serious adverse effects were reported."	Appropriate conclusions based on available data.
Boldt 2014	"Evidence is insufficient to suggest that non-pharmacological treatments are effective in reducing chronic pain in people living with SCI. The benefits and harms of commonly used non-pharmacological pain treatments should be investigated in randomised controlled trials with adequate sample size and study methodology"	Appropriate conclusions based on available data.
Brown 2010	"There is a lack of available evidence to support the use of exercise in the alleviation of symptoms associated with dysmenorrhoea. The limited evidence implies that there are no adverse effects associated with exercise."	Review authors should not have commented on lack of adverse events as this was not reported in the included study. The comment on lack of adverse events contravened present Cochrane guidance
Busch 2007	"There is moderate quality evidence that short-term aerobic training (at the intensity recommended for increases in cardiorespiratory fitness) produces important benefits in people with FM in global outcome measures, physical function, and possibly pain and tender points. There is limited evidence that strength training improves a number of outcomes including pain, global wellbeing, physical function,	Appropriate conclusions based on available data.

Table 8. Interpretation of results by original review authors (*Continued*)

	tender points and depression. There is insufficient evidence regarding the effects of flexibility exercise. Adherence to many of the aerobic exercise interventions described in the included studies was poor."	
Busch 2013	"We have found evidence in outcomes representing wellness, symptoms, and physical fitness favoring resistance training over usual treatment and over flexibility exercise, and favoring aerobic training over resistance training. Despite large effect sizes for many outcomes, the evidence has been decreased to low quality based on small sample sizes, small number of randomized clinical trials (RCTs), and the problems with description of study methods in some of the included studies."	Appropriate conclusions based on available data.
Cramp 2013	"There is some evidence that physical activity interventions ... may help to reduce fatigue in RA. However, the optimal parameters and components of these interventions are not yet established."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion despite low/unclear quality score in results and discussion sections No conclusions about effect on pain (insufficient data).
Fransen 2014	"There is currently high-level evidence that land-based exercise will reduce hip pain, and improve physical function, among people with symptomatic hip osteoarthritis."	Evidence was good quality though sample sizes were often small (i.e. it is debatable if this was high level evidence as claimed by authors). Agree that results demonstrate small but significant benefit from intervention
Fransen 2015	"High-quality evidence suggests that land-based therapeutic exercise provides benefit in terms of reduced knee pain and quality of life and moderate-quality evidence of improved physical function among people with knee OA… Despite the lack of blinding we did not downgrade the quality of evidence for risk of performance or detection bias."	Appropriate conclusions based on available data. May have been generous with quality assessment but this was stated in conclusions for transparency
Gross 2015a	"…there is still no high quality evidence and uncertainty about the effectiveness of exercise for neck pain… Moderate quality evidence supports the use specific strengthening exercises as a part of routine practice … Moderate quality evidence supports the use of strengthening exercises, combined with endurance or stretching exercises may also yield similar beneficial results. However, low quality evidence notes when only stretching or only endurance type exercises … there may be minimal beneficial effects for both neck pain and function."	Appropriate conclusions based on available data.

Table 8. Interpretation of results by original review authors (*Continued*)

Han 2004	"Tai chi appears to have no detrimental effects on the disease activity of RA in terms of swollen/tender joints and activities of daily living…tai chi appears to be safe, since only 1 participant out of 121 withdrew due to adverse effects and withdrawals were greater in the control groups than the tai chi groups."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias in conclusion despite very low quality score in results section
Hayden 2005	"Evidence from randomized controlled trials demonstrates that exercise therapy is effective at reducing pain and functional limitations in the treatment of chronic low-back pain, though cautious interpretation is required due to limitations in this literature."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion despite low quality score in results and discussion sections
Hurkmans 2009	"Short-term, land-based dynamic exercise programs have a positive effect on aerobic capacity (aerobic capacity training whether or not combined with muscle strength training) and muscle strength (aerobic capacity training combined with muscle strength training) immediately after the intervention, but not after a follow-up period. Short-term, water-based dynamic exercise programs have a positive effect on functional ability and aerobic capacity directly after the intervention but it is unknown whether these effects are maintained after follow-up. Long-term, land-based dynamic exercise programs (aerobic capacity and muscle strength training) have a positive effect on functional ability, aerobic capacity, and muscle strength immediately after the intervention but it is unknown whether these effects are maintained after follow-up... Based on the evidence, aerobic capacity training combined with muscle strength training is recommended for routine practice in patients with RA."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion No conclusions regarding pain severity.
Koopman 2015	"Data from two single trials suggested that muscle strengthening of thumb muscles (very low-quality evidence) ... are safe and beneficial for improving muscle strength ... with unknown effects on activity limitations." "We found evidence varying from very low quality to high quality that ... rehabilitation in a warm or cold climate are not beneficial in PPS." "Due to a lack of good-quality data and randomised studies, it was impossible to draw definitive conclusions about the effectiveness of interventions in people with PPS."	Appropriate conclusions based on available data.
Lane 2014	"… Exercise therapy should play an important part in the care of selected patients with intermittent claudication, to improve walking times and distances. Ef-	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of stud-

Table 8. Interpretation of results by original review authors (*Continued*)

	fects were demonstrated following three months of supervised exercise although some programmes lasted over one year."	ies in conclusion No conclusions regarding pain severity.
Lauret 2014	"There was no clear evidence of differences between supervised walking exercise and alternative exercise modes in improving the maximum and pain-free walking distance of patients with intermittent claudication…. The results indicate that alternative exercise modes may be useful when supervised walking exercise is not an option for the patient."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion (in discussion)
Regnaux 2015	"We found very low- to low-quality evidence for no important clinical benefit of high-intensity compared to low-intensity exercise programs in improving pain and physical function in the short term.... The included studies did not provide any justification for the levels of intensity of exercise programs. No authors reported evidence for the minimal and maximal intensity that could be delivered."	Appropriate conclusions based on available data. This overview has only used one study of the six included as it alone included a control group, for which we could not extract data as the control comparison was not used in the analysis by the review authors
Saragiotto 2016	"There is very low to moderate quality evidence that MCE has a clinically important effect compared with a minimal intervention for chronic low back pain... As MCE appears to be a safe form of exercise and none of the other types of exercise stands out, the choice of exercise for chronic low back pain should depend on patient or therapist preferences, therapist training, costs and safety."	Appropriate conclusions based on available data.
Silva 2010	"We were not able to provide any evidence to support the application of balance exercises (proprioceptive training) alone in patients with RA."	Appropriate conclusions based on available data (no included studies)
van der Heijden 2015	"This review has found very low quality but consistent evidence that exercise therapy for patellofemoral pain syndrome (PFPS) may result in clinically important reduction in pain and improvement in functional ability."	No subgroup analysis to differentiate between acute, subacute, and chronic pain made it difficult to extract appropriate data for this review
Yamato 2015	"No definite conclusions or recommendations can be made as we did not find any high quality evidence for any of the treatment comparisons, outcomes or follow-up periods investigated. However, there is low to moderate quality evidence that Pilates is more effective than minimal intervention in the short and intermediate term as the benefits were consistent for pain intensity and disability, with most of the effect sizes being considered medium."	Appropriate conclusions based on available data. There was no subgroup analysis to differentiate between acute, subacute, and chronic pain made it difficult to extract appropriate data for this review (one included study had subacute back pain (> 6 weeks), all others were chronic back pain (> 12 weeks)) but results are presented altogether as chronic pain

FM: fibromyalgia; MCE: motor control exercise; OA: osteoarthritis; PPS: postpolio syndrome; RA: rheumatoid arthritis; SCI: spinal cord injury.

Table 9. Pain severity

Review	Number of trials (and participants) assessing 'pain severity'	Baseline pain score	Post-intervention reported result or change data (or if only one data point reported in review)	Follow-up	Overall comment/ statement
Bartels 2007 (osteoarthritis)	Hip + knee OA: Post-intervention: 4 (638) Follow-up: 1 (310) Hip only: follow-up: 1 (17) Knee only: post-intervention: 1 (46)	Control baseline: Hip + knee OA WOMAC 0 to 20 (2 studies): 9.10 (SD 3.14) VAS 0 to 100 (1 study): 55.3 (SD 24.6) HAQ 0 to 3 (1 study): 1.05 (SD 0.61) Hip only VAS 0 to 100 (1 study): 56 (SD 21.89) Knee only VAS 0 to 10 (1 study): 5.6 (SD 1.4)	Hip + knee OA A minor effect of a 3% absolute reduction (0.6 fewer points on WOMAC 0 to 20 scale) and 6.6% relative reduction SMD 0.19 (95% CI 0.04 to 0.35) (P = 0.02) Knee only SMD 0.86 (95% CI 0.25 to 1.47) (P = 0.005) Absolute difference 12% (1.2 fewer points on a 0 to 10 scale) Relative change 22% improvement	Hip + knee OA Follow-up at 6 months: SMD 0.11 (95% CI -0.12 to 0.33) (ns) No difference Hip only SMD 1.00 (95% CI -0.04 to 2.04) (P = 0.06, ns)	Statistically significant post-intervention in hip + knee OA group, but not clinically significant Knee-only OA had moderate to large effect size (statistically significant) immediately post-intervention
Bidonde 2014 (fibromyalgia)	Post-intervention: 7 (382)	Weighted mean score at baseline (all participants): 69.59 median value for pain was 70.9 in studies comparing aquatic training to control	On 100-point scale: MD -6.59 (95% CI -10.71 to -2.48) SMD -0.53 (95% CI -0.76 to -0.31) Absolute difference -7% (95% CI -11 to -3) NNTB 5 (95% CI 3 to 8)	3 studies at 12, 48, or 52 weeks' post-intervention could not be combined. 2 studies showed SMD favouring intervention at follow-up.	"We found a moderate effect favouring the aquatic exercise training for pain" …"similar improvements in pain in the low pain groups (SMD -0.60, 95% CI -0.98 to -0.23) and in the high pain groups (SMD -0.57, 95% CI -1.11 to -0.03)." Among the

Table 9. Pain severity (*Continued*)

					major wellness outcomes, none of the outcomes met the threshold for clinically relevant differences (15%)
Boldt 2014 (spinal cord injury)	Post-intervention: 3 (149)	WUSPI score 22.6 (exercise group) to 11.05 (control group) in 1 group at baseline Not reported for 2 studies	WUSPI change score: Exercise group: -7.7 (SD 19.01) Control group: 12.8 (SD 12.74) SF-36 (pain experience): -1.9 (95% CI -3.4 to -0.4) favoured exercise (P = 0.01) VAS (0 to 10): MD -2.8 (95% CI -3.77 to -1.83) favoured exercise (P < 0.00001)	1 study at 4 weeks: VAS (0 to 10): -2.50 (95% CI -3.48 to -1.52) (P < 0.00001) WUSPI: -26.40 (95% CI -37.62 to -15.18favoured exercise (P < 0.00001)	"All three studies were fraught with high overall risk of bias. In particular, the comparison with 'no treatment' or waiting lists as control interventions likely leads to an overestimation of the effectiveness of the exercise programmes provided in these studies. Consequently, no conclusion on their effectiveness can be drawn."
Busch 2007 (fibromyalgia)	Strength training: 1 (21) Aerobic training: 3 (183)	Control baseline: Aerobic: 6.1/10 (VAS) (SD 1.97) Strength: 35/100 (VAS) (SD 19)	Aerobic training: SMD 0.65 (95% CI -0.09 to 1.39) (ns) Weighted absolute change 13% (1.3 cm lower on 10-cm scale) Relative change 21% Strength training: SMD 3.00 (95% CI 1.68 to 4.32) (ns) Weighted absolute change 49% (49 points lower on 100-point scale) Relative change 140%, NNTB 2	n/a	">30% improvement was seen in the strength training group as compared to an untreated control group in pain." Aerobic training led to an improvement of 1.3/10.

Table 9. Pain severity (*Continued*)

Busch 2013 (fibromyalgia)	Post-intervention: 2 (81) Follow-up at 8 weeks, 16 weeks, 28 weeks: 1 (60)	Not reported - change data only	Change score on VAS (in cm): MD -3.30 (95% CI -6.35 to -0.26) (P = 0.03) SMD -1.89 (95% CI -3.86 to 0.07) Relative % change 44.6% (95% CI 3.5 to 85.9) favoured exercise NNTB 2 (95% CI 1 to 34)	8 weeks: MD -0.68 (95% CI -1.62 to 0.26) (ns) 16 weeks: MD -1.79 (95% CI -2.70 to -0.88) (P < 0.001) 28 weeks: MD -0.85 (95% CI -1.77 to 0.07) (P = 0.07, ns) Overall (n = 180): MD -1.12 (95% CI -1.65 to -0.58) (P < 0.0001)	> 30% improvement post-intervention.
Cramp 2013 (rheumatoid arthritis)	4 (not reported)	Not reported	In narrative only - Harkcom 1985: statistics not reported separately for pain data, but reported as improvement over time; Hakkinen 2003: "stat significant improvement in 24 months"; Evans 2012 and Wang 2008: no statistically significant effects	Not reported	"Improvement over time" with "significant improvement in 24 months." No actual data available.
Fransen 2014 (OA)	End of treatment: 9 (549) 3 to 6 months: 5 (391)	Not reported; land based exercise vs no exercise: mean pain in control group ~ 29/100 (based on 9 studies' control values)	End of treatment: SMD -0.38 (95% CI -0.55 to -0.20) "small to moderate" favoured exercise (P < 0.0001)	3 to 6 months: SMD -0.38 (95% CI -0.58 to -0.18) "small to moderate" favoured exercise (P = 0.0002)	"Small to moderate" statistically significant improvement, but only mild pain at baseline
Fransen 2015 (OA)	End of treatment: 44 (3537) Follow-up (2 to 6 months): 12 (1468) Follow-up (> 6 months): 8 (1272)	Not reported; land-based exercise vs no exercise: mean pain in control group 44/100 (based on 1 study control values)	Land-based exercise vs no exercise: Mean pain in intervention groups was 0.49 SDs lower (95% CI 0.39 to 0.59 lower). This translates to an absolute mean reduction of 12 points	2 to 6 months: SMD -0.24 (95% CI -0.35 to -0.14) favoured exercise (P < 0.00001) > 6 months: SMD -0.52 (95% CI -1.01 to -0.03) favoured exercise (P = 0.04)	Absolute improvement of 12/100 post-intervention (statistically significant)

Table 9. Pain severity (*Continued*)

			(95% CI 10 to 15) compared with control group on a 0 to 100 scale SMD -0.49 (95% CI -0.39 to -0.59) (P < 0.00001) Absolute reduction 12% (95% CI 10% to 15%) Relative change 27% (95% CI 21% to 32%) NNTB 4 (95% CI 3 to 5)		
Gross 2015a (mechanical neck disorders)	12-week treatment: 2 (147) 24 week (or 12-week treatment + 12-week follow-up) : 2 (140)	Not reported, but control scores at end of treatment 40 to 60/100 (moderate pain)	12 weeks: pooled MD -14.90 (95% CI -22.40 to -7.39) favoured exercise (P = 0.0001)	24 weeks: pooled MD -10.94 (95% CI -18.81 to -3.08) favoured exercise (P = 0.0064)	2 trials showed a moderate (statistically significant) reduction in pain post-intervention (14.9/100)
Hayden 2005 (low back pain)	Earliest follow-up: 8 (370) Follow-up (time since randomisation) Short term (6 weeks): 6 (268) Intermediate term (6 months): 5 (249) Long term (12 months): 2 (126)	"Chronic group" at baseline: mean 46/100 (95% CI 41 to 50) (moderate pain)	Earliest: MD -10.20 (95% CI -19.09 to -1.31) (P = 0.02)	Short term: MD -8.58 (95% CI -18.46 to 1.29) (P = 0.09, ns) Intermediate term: MD -12.48 (95% CI -22.69 to -2.27) (P = 0.02) Long term: MD -3.93 (95% CI -9.89 to 2.02) (P = 0.2, ns)	Reduction of ~ 10/100 at earliest measurement point.
Hurkmans 2009 (rheumatoid arthritis)	4 studies (total 188 participants) in different categories (results not combined)	Not reported	Short-term (12 weeks): Short-term land-based (aerobic and strength training) SMD -0.53 (95% CI -1.09 to 0.04) Short-term land-based (aerobic only) SMD -0.27 (95% CI -0.79 to 0.26) Short-term water-based SMD 0.06	Long-term (24 months) land-based (aerobic and strength training) SMD 0.35 (95% CI -0.46 to 1.16)	No significant difference between control and intervention.

Table 9. Pain severity (*Continued*)

			(95% CI -0.43 to 0.54)		
Koopman 2015 (postpolio syndrome)	1 (55)	Not reported, but control scores at end of treatment mean 44 (SD 24) on a 0 to 100 scale (moderate pain)	3 months post-intervention: VAS (0 to 100): MD 11.00 (95% CI -0.98 to 22.98) (P = 0.072)	n/a	No significant effect/no difference between groups.
Regnaux 2015 (OA)	Only 1 study that had a no-exercise control: 1 (68) - excluded data for control (no exercise) from analysis (n = 34)	Not reported	Post-intervention: WOMAC (0 to 20) Change data presented for high- vs low-intensity groups only, not compared to control	n/a	Actual individual study data was extracted (where possible) instead of pooled MD or SMD due to comparison this overview wishes to make (exercise vs no-exercise only) Could not extract exercise vs control data.
Saragiotto 2016 (low back pain)	Short term (< 3 months): 4 (291) Intermediate term (3 to 12 months): 4 (348) Long term (> 12 months): 3 (279)	Not reported, but control scores at follow-up range 25 to 56/100 (mild-moderate pain)	Short term: MD -10.01 (95% CI -15.67 to -4.35) favoured exercise (P < 0.001)	Intermediate term: MD -12.61 (95% CI -20.53 to -4.69) favoured exercise (P = 0.002) Long term: MD -12.97 (95% CI -18.51 to -7.42) favoured exercise (P < 0.001)	Medium effect size favouring exercise at all follow-up assessments (moderate quality evidence at short- and long-term, low quality evidence at intermediate term) Clinically important effect.
van der Heijden 2015 (patellofemoral pain syndrome)	3 studies with pain > 3 months (135 participants), 2 studies used in analysis (41 participants) Long-term follow-up: 1 (94)	Not reported, but control scores at follow-up range 2.1 to 6.0/10 (mild-moderate pain)	Short-term (4 to 8 weeks): MD for usual pain in the exercise group was 0.93 (95% CI 1.60 to 0.25) SDs lower SMD -0.93 (95% CI -1.60 to -0.25) (P = 0.008)	"Long term" (16 weeks) VAS (0 to 10): MD -4.42 (95% CI -7.75 to -0.89) favoured exercise (P = 0.01)	Reduction in pain of 4/10 at 16 weeks' follow-up.
Yamato 2015 (low back pain)	Short term: 6 (265) Intermediate term:	Not reported, but control scores at ear-	Short-term follow-up (< 3 months):	Intermediate term (3 to 12 months):	"Low quality evidence (downgraded

Table 9. Pain severity (*Continued*)

	2 (148)	liest follow-up range 18 to 52/100 (mild-moderate pain)	MD -14.05 (95% CI -18.91 to -9.19) (P < 0.001)	MD -10.54, (95% CI -18.54 to -2.62) (P = 0.009)	due to imprecision and risk of bias) that Pilates reduces pain compared with minimal intervention at short-term follow-up, with a medium effect size.. intermediate-term follow-up, two trials, provided moderate quality evidence (downgraded due to imprecision) that Pilates reduces pain compared with minimal intervention, with a medium effect size"

CI: confidence interval; HAQ: Health Assessment Questionnaire; MD: mean difference; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; ns: not significant; OA: osteoarthritis; SD: standard deviation; SF-36: 36-item Short Form; SMD: standardised mean difference; VAS: visual analogue score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; WUSPI; Wheelchair User Shoulder Pain Index.

Table 10. Physical function

Review	Outcome measure	Number of trials (and participants) used in analysis	Post-intervention result (or if only 1 result reported)	Short-term follow-up (or if only 1 follow-up point reported)	Intermediate-term follow-up	Long-term follow-up	Overall comment/ statement
Bartels 2007 (OA)	Self-reported function (WOMAC and HAQ) and walking ability, and DRI	Post-intervention Hip + knee function: 4 (648) walking ability: 2 (355) Hip only function: 1 (28) Follow-up function hip + knee: 1 (306) hip only: 1 (17)	Function (hip + knee): SMD 0.26 (95% CI 0.11 to 0.42) favoured exercise (P < 0.001) Walking (hip + knee): SMD 0.18 (95% CI -0.03 to 0.39) favoured exercise (P = 0.08, ns)	Hip only Disability, SMD 1.00 (95% CI -0.04 to 2.04) favoured exercise (P = 0.06, ns)	Hip + knee (6 months) Function, SMD 0.10 (95% CI -0.12 to 0.33) (ns)	n/a	Function was significantly improved in people with hip + knee OA immediately post-intervention only - small effect size only

Table 10. Physical function (*Continued*)

			Function (hip only): SMD 0.76 (95% CI -0.02 to 1.53) favours exercise (P = 0.06, ns)				
Bidonde 2014 (fibromyalgia)	Self-reported physical function (0 to 100 scale)	5 (285)	MD -4.35 (95% CI -7.77 to -0.94) SMD -0.44 (95% CI -0.76 to -0.11) Absolute difference -4 (95% CI -8 to -1) NNTB 6 (95% CI 3 to 22)	n/a	n/a	n/a	Small difference (improvement) in aquatic exercise group. Among the major wellness outcomes, none of the outcomes met the threshold for clinically relevant differences (15%)
Busch 2007 (fibromyalgia)	Physical function	Aerobic: 4 (253) Strength: 2 (47)	Aerobic: SMD 0.66 (95% CI 0.41 to 0.92) favoured exercise (P < 0.0001) Strength: SMD 0.52 (95% CI -0.07 to 1.10) favoured exercise (P = 0.08, ns)	n/a	n/a	n/a	Function was significantly improved from aerobic exercise training, strength training neared significance Moderate effect size.
Busch 2013 (fibromyalgia)	HAQ and SF-36 for function	3 (107)	Change score MD -6.29 (95% CI -10.45 to -2.13) favoured exercise (P < 0.01)	n/a	n/a	n/a	Significantly favourable effect of exercise.
Cramp 2013 (rheumatoid arthritis)	Disability	4 (not reported)	n/a	n/a	n/a	n/a	"Studies investigating hydrotherapy and tai chi demon-

Table 10. Physical function (*Continued*)

							strated statistically significant improvements in the intervention arm compared to the control arm between baseline and follow-up. The studies investigating strength training and Ivengar yoga did not demonstrate a statistically significant difference between study arms."
Fransen 2014 (OA)	Physical function	Post-intervention: 9 (521) Follow-up (3 to 6 months): 5 (365)	SMD -0.30 (95% CI -0.54 to -0.05) "significant benefit" favoured exercise (P = 0.02) The demonstrated effect size for exercise was equivalent to an improvement of physical function of 7 points (95% CI 1 to 12) on a 0 to 100 scale compared with a control group	SMD -0.37 (95% CI -0.57 to -0.16) favoured exercise (P < 0.001)	n/a	n/a	Statistically significant, but small effect size only.
Fransen 2015 (OA)	Physical function	Post-intervention: 44 (3913) Follow-up (2 to 6 months):	SMD -0.52 (95% CI -0.64 to -0.39) favoured exercise (P < 0.	SMD -0.15 (95% CI -0.26 to -0.04) favoured exercise (P = 0.	SMD -0.57 (95% CI -1.05 to -0.10) favoured exercise (P = 0.02)	n/a	Significant effect from exercise at every follow-up point.

Table 10. Physical function (*Continued*)

		10 (1279) Follow-up (> 6 months): 8 (1266)	0001); an improvement of 10 points (95% CI 8 to 13) on a 0- to 100-point scale	008)			Moderate effect size at short- and long-term follow-up, but only small effect at intermediate-term follow-up
Gross 2015a (mechanical neck disorders)	Physical function	12 wk: 2 (147) 24 wk: 2 (140)	12 wk treatment: pooled SMD -0.50 (95% CI -1.04 to 0.03) favoured exercise (P = 0.07, ns)	24 wk treatment (or 12 wk' treatment + 12 wk follow-up) : pooled SMD -0.40 (95% CI -0.74 to -0.06) favoured exercise (P = 0.02)	n/a	n/a	2 trials showed a moderate (statistical) improvement in function
Han 2004 (rheumatoid arthritis)	Functional assessment and 50-feet walk test	Function: 2 (52) Walk test: 2 (48)	Function: MD 0.01 (95% CI -2.94 to 2.97) (ns) Walk test: MD 0.35 seconds (95% CI -1.14 to 1.84) (ns)	n/a	n/a	n/a	No significant effect.
Hayden 2005 (low back pain)	Function	Earliest: 7 (337) Short term: 6 (268) Intermediate term: 4 (216) Long term: 2 (126)	Earliest: MD -2.98 (95% CI -6.48 to 0.53) favoured exercise (P = 0.09, ns)	Short term: MD -3.03 (95% CI -6.35 to 0.53) favoured exercise (P = 0.07, ns)	Intermediate term: MD -3.84 (95% CI -7.06 to -0.61) favoured exercise (P = 0.02)	Long term: MD -4.22 (95% CI -7.99 to -0.46) favoured exercise (P = 0.03)	Favoured exercise from the earliest measure, but only reached statistical significance at intermediate and long term after randomisation
Hurkmans 2009 (rheumatoid arthritis)	Functional ability	Land-based aerobic: 2 (66) Land-based aerobic + strength: 2 (74)	n/a	Short-term training (12 wk) Land-based aerobic only training	n/a	n/a	No significant difference between control and intervention groups

Table 10. Physical function (*Continued*)

				SMD 0.03 (95% CI -0.46 to 0.51) (ns) Land-based aerobic and strength training SMD -0.4 (95% CI -0.86 to 0.06) (ns)			
Koopman 2015 (postpolio syndrome)	Muscle strength; and activity limitation (Sunnaas ADL-index range 0 to 36; Rivermead Mobility Index (RMI) range 0 to 15)	Strength: 1 (10) Activity limitation: 1 (53)	Isometric muscle strength (postintervention): MD 39.00% (95% CI 6.12 to 71.88) Activity limitation: 3 months' postintervention: ADL-index: MD -2.70 (95% CI -4.53 to -0.87) Rivermead Mobility Index (RMI): MD -1.50 (95% CI -2.93 to -0.07)	Activity limitation: 6-months postintervention: ADL-index: MD -2.90 (95% CI -4.73 to -1.07) RMI: MD -1.80 (95% CI -3.19 to -0.41)	n/a	n/a	Activity limitation: favoured intervention at both assessment points "The baseline imbalance in favour of the usual care group probably biased these results."
Lane 2014 (intermittent claudication)	Maximal walking time and maximal walking distance	Postintervention Walking time: 12 (577) Walking distance: 9 (480) 3-month follow-up Walking time: 5 (174) Walking distance: 3 (116) 6-month follow-up Walking time:	Time: MD 4.51 minutes (95% CI 3.11 to 5.92) favoured exercise (P < 0.00001) Distance: 108.99 m (95% CI 38.20 to 179.78) favoured exercise (P = 0.003)	Time: MD 6.05 minutes (95% CI 5.47 to 6.62) favoured exercise (P < 0.00001) Distance: MD 104.46 m (95% CI -64.33 to 273.24) favoured exercise (ns)	Time: MD 3.20 minutes (2.04 to 4.36) favoured exercise (P < 0.0001) Distance: MD 138.36 m (95% CI 22.39 to 254.34) favoured exercise (P = 0.02)	n/a	Objectively measured walking time and distance showed significant improvement

Table 10. Physical function (*Continued*)

		4 (295) Walking distance: 3 (156)					
Lauret 2014 (intermittent claudication)	Maximal walking time (mins) and maximal walking distance (metres)	No relevant studies	n/a	n/a	n/a	n/a	No relevant studies.
Regnaux 2015 (OA)	WOMAC (0 to 68) disability scale, and muscle strength	1 (68) - excluded control (no-exercise data: n = 34)	n/a	n/a	n/a	n/a	Could not extract exercise vs control data - data presented for high vs low intensity groups only, not compared to control
Saragiotto 2016 (low back pain)	Disability (Oswestry Disability Index, Roland Morris Disability Questionnaire)	Short-term follow-up (< 3 months): 5 (332) Intermediate term (3 to 12 months): 4 (348) Long term (> 12 months): 3 (279)	-	MD -8.63 (95% CI -14.78 to -2.47) (P < 0.01)	MD -5.47 (95% CI -9.17 to -1.77) (P = 0.004)	MD -5.96 (95% CI -9.81 to -2.11) (P = 0.002)	Small effect sizes, favoured exercise. Short term: CI included a clinically important effect.
Silva 2010 (rheumatoid arthritis)	HAQ function	No studies found	n/a	n/a	n/a	n/a	No studies found.
van der Heijden 2015 (patellofemoral pain syndrome)	Functional ability	Short-term follow-up: 7 (483) Long-term follow-up: 3 (274)	n/a	Short-term (4 to 8 wk): SMD 1.10 (95% CI 0.58 to 1.63) favoured exercise (P < 0.0001)	n/a	SMD 1.62 (95% CI 0.31 to 2.94) favoured exercise (P = 0.02)	Significant effect of exercise. Very large effect size at short- and long-term follow-up.
Yamato 2015 (low back pain)	Disability (all measures con-	Short-term (< 3 months) fol-	n/a	MD -7.95 (95% CI -13.	MD -11.17 (95% CI -18.	n/a	"Low quality ev-

Table 10. Physical function (*Continued*)

	verted to 0 to 100 scale)	low-up: 5 (248) -Interme-diate-term (3 to 12 months) follow-up: 2 (146)		23 to -2.67) (P = 0.003)	41 to -3.92) (P = 0.0025)		idence (down-graded due to imprecision and inconsis-tency) that Pi-lates improves disability at short-term follow-up compared with minimal inter-vention, with a small effect size ... interme-diate-term fol-low-up, two trials pro-vided moder-ate quality ev-idence (down-graded due to im-precision) of a significant ef-fect in favour of Pilates, with a medium ef-fect size"

ADL: activities of daily living; CI: confidence interval; DRI: Disability Rating Index; HAQ: Health Assessment Questionnaire; MD: mean difference; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; ns: not significant; OA: osteoarthritis; SF-36: 36-item Short Form; SMD: standardised mean difference; wk: week; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index,

Table 11. Psychological function

Review	Outcome measure	Number of trials (and participants) reporting psychological function	Outcome result (postintervention or if only one measurement point)	Follow-up	Additional statement/comment
Mental health					**Mental health**
Bartels 2007	-	4 studies	SMD 0.16 (95% CI 0.01 to 0.032)	No significant difference at 6 months, 1	Very small effect size postintervention.

Table 11. Psychological function (*Continued*)

			favoured aquatic exercise	study	
Busch 2013	SF-36 - Mental health scale	1 study	-	n/a	No group differences.
Bidonde 2014	SF-36 - mental Health scale SF-12 - Mental Health scale	4 studies, n = 243	MD -3.03 (95% CI -8.06 to 2.01)	n/a	No effect.
Anxiety					**Anxiety**
Cramp 2013	Brief Symptom Inventory	1 study	"No significant effect"	n/a	-
Depression					**Depression**
Boldt 2014	CES-D	1 study, n = 34	MD -6.0 (95% CI -15.87 to 3.87) (P = 0.23)	n/a	No effect.
Busch 2013	HADS - Depression Beck Depression Index	1 study, n = 21	MD -3.70 (95% CI -6.37 to -1.03) Relative difference 57%	n/a	Significant effect, favoured resistance training.
Cramp 2013	CES-D	Not reported	"Variable effect" reported in text only	n/a	-

CES-D: Centre for Epidemiological Studies-Depression; CI: confidence interval; HADS: Hospital Anxiety and Depression Scale; MD: mean difference; n: number of participants; n/a: not applicable; SF-12: 12-item Short Form; SF-36: 36-item Short Form; SMD: standardised mean difference.

Table 12. Quality of life

Review	Outcome measure	Number of trials (and participants) reporting Quality of Life (QoL)	Outcome result	Additional statement/ comment	
(Health-related) Quality of Life					**(Health-related)**
Bartels 2007	QoL: SF-12 (Physical), PQoL, EuroQoL	Hip + knee OA (post-intervention): 3 studies, n = 599 Hip only OA (post-intervention): 1 study, n = 28	Hip + knee (post-intervention): SMD 0.32 (95% CI 0.03 to 0.61) (P = 0.028) Hip only (post-intervention): SMD 0.76 (95% CI	Significantly favoured aquatic exercise post-intervention in hip + knee OA Small effect size only (when statistically	

Table 12. Quality of life (*Continued*)

		Hip only OA (follow-up): 1 study, n = 17	-0.02 to 1.53) (ns) Hip only (follow-up): SMD 1.00 (95% CI -0.04 to 2.04) (ns)	significant).
Boldt 2014	PQoL (perceived quality of life) SQoL (subjective quality of life)	Post-intervention: 1 study, n = 34, PQoL; 1 study, n = 80, SQoL Follow-up (intermediate term): 1 study, n = 80, SQoL	Post-intervention: PQoL MD 10.8 (95% CI -4.2 to 25.8) (P = 0.16) SQoL MD 0.3 (95% CI -0.22 to 0.82) (P = 0.25) Follow-up: SQoL MD 0.5 (95% CI -0.03 to 1.03) (P = 0.07)	No difference between groups.
Fransen 2014	QoL	Post-intervention: 3 studies, n = 183	SMD 0.07 (95% CI -0.23 to 0.36) (ns)	No difference between groups.
Fransen 2015	QoL: self-report questionnaire, scale 0 to 100 (100 is maximum QoL)	Post-intervention: 13 studies, n = 1073	SMD 0.28 (95% CI 0.15 to 0.40) (P < 0.0001) Absolute difference 4% (95% CI 2% to 5%) relative difference 9% (95% CI 5% to 13%)	Statistically significant, but equates to an absolute improvement of 4 points (95% CI 2 to 5) on a 0 to 100 scale Small effect size only.
Gross 2015a	QoL: SF-36 (Physical Function subscale)	Post-intervention: 2 studies, n = 143	12-wk intervention: MD -2.22 (95% CI -5.17 to 0.72) (ns) 24-wk intervention: MD 0.06 (95% CI -4.06 to 4.17) (ns)	No significant difference between groups.
Lauret 2014	HRQoL	No relevant studies	n/a	n/a
Global assessment				**Global assessme**
Busch 2007	Global wellbeing	Strength: 2 studies, n = 47 Aerobic: 4 studies, n = 269	Strength: SMD 1.43 (95% CI 0.76 to 2.10) Aerobic: SMD 0.49 (95% CI 0.23 to 0.75)	Favoured exercise - higher score showed better QoL, Strength: very large effect size. Aerobic: small-to-moderate effect size only.
Bidonde 2014	Participant-rated global (10-cm VAS)	1 study, n = 46	MD -0.87 (95% CI -1.74 to 0.00)	No effect.
Gross 2015a	Global perceived effect	1 study, n = 70	"No significant difference"	No significant difference.
Hayden 2005	Global assessment	7 studies, n = 16	Not reported	n/a

Table 12. Quality of life (*Continued*)

Saragiotto 2016	Global impression of recovery	1 study, n = 154	Short term, MD 1.30 (95% CI 0.30 to 2.30) (P = 0.01) Intermediate term, MD 1.20 (95% CI 0.31 to 2.09) (P = 0.008) Long term, MD 1.50 (95% CI 0.61 to 2.39) (P < 0.001)	Medium effect size.
Yamato 2015	Global impression of recovery	1 study, n = 86	Short term (< 3 months): MD 1.50 (95% CI 0.70 to 2.30) Intermediate term (3 to 12 months): MD 0.70 (95% CI -0.11 to 1.51)	"Low quality evidence (downgraded due to imprecision and inconsistency), we found a significant short-term effect, with a small effect size, but not for intermediate/mid-term follow up."
Other method of assessment				**Other method o**
Bidonde 2014	Multi-dimensional function- FIQ	7 studies, n = 367	MD -5.97 (95% CI -9.06 to -2.88) SMD -0.55 (95% CI -0.83 to -0.27) Absolute difference -6 (95% CI -9 to -3) NNTB 5 (95% CI 3 to 9)	Favoured aquatic exercise - lower score showed reduced impact of pain on life "Moderate difference."
Busch 2013	Multi-dimensional function - FIQ	1 study, n = 60	SMD -1.27 (95% CI -1.83 to -0.72) Absolute difference -16.75 FIQ units (95% CI -23.31 to -10.19)	Favoured exercise - lower score showed reduced impact of pain on life Very large effect size.
Hayden 2005	Work status	9 studies, n = 21	Not reported	n/a
Silva 2010	Health Assessment Questionnaire (HAQ)	No included studies	n/a	n/a

FIQ: Fibromyalgia Impact Questionnaire; HRQoL: health-related quality of life; MD: mean difference; n: number of participants; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; OA: osteoarthritis; PQoL: perceived quality of life; QoL: quality of life; SF-36: 36-item Short Form; SMD: standardised mean difference; SQoL: subjective quality of life; VAS: visual analogue scale.

Table 13. Adherence/withdrawals

Review	Number of trials (and participants) reporting withdrawals	Number withdrawn (per 1000) - intervention group	Number withdrawn (per 1000) - control group	RR or OR
Bidonde 2014 (fibromyalgia)	8 studies, n = 472	151 (imputed from reported 38/252)	129 (imputed from reported 30/232)	RR 1.13 (95% CI 0.73 to 1.77) (P = 0.45)
Busch 2013 (fibromyalgia)	3 studies, n = 107	134 (95% CI 30 to 439)	39	RR 3.50 (95% CI 0.79 to 15.49)
Fransen 2014 (osteoarthritis)	7 studies, n = 715	59 (95% CI 30 to 114)	34	OR 1.77 (95% CI 0.86 to 3.65)
Han 2004 (rheumatoid arthritis)	4 studies, n = 189	109 (imputed from reported 11/101)	284 (imputed from reported 25/88)	RR 0.37 (95% CI 0.19 to 0.72)
Regnaux 2015 (osteoarthritis)	1 study, n = 102	44 (imputed from reported 3/68 (4%) ; all from high-intensity group)	0	Calculated RR 3.55 (95% CI 0.19 to 66.8)
Saragiotto 2016 (low back pain)	7 studies, n = 671	0	0	-
Silva 2010 (rheumatoid arthritis)	No included studies	n/a	n/a	n/a
Total	**30 studies, n = 2256**	**82.8/1000**	**81/1000**	**Calculated RR 1.02 (95% CI 0.94 to 1.12) Calculated OR 1.05 (95% CI 0.88 to 1.25)**

CI: confidence interval; n: number of participants; n/a: not applicable; OR: odds ratio; RR: risk ratio.

Table 14. Adverse events (not death)

Review	Total number of trials (and participants) in review reporting exercise vs control in chronic pain population	Number of trials (and participants) reporting adverse events	Number of adverse events	Overall statement
Bartels 2007	4 (674)	2 (148)	0	Adverse events were recorded (and reported), but none occurred

Table 14. Adverse events (not death) (*Continued*)

Bidonde 2014	9 (519)	0	0	Review stated that no included studies actively reported on adverse events (some reported withdrawal)
Boldt 2014	3 (149)	2 (115)	5 events over 2 studies	"Neck, shoulder and elbow injuries in five participants in the intervention group."
Busch 2007	34 (2276)	6 (strength training: 115, aerobic: 1264)	Strength training: 3 Aerobic training: 5	-
Busch 2013	3 (81)	2 (86 exercising participants)	0	Adverse events were recorded (and reported), but none occurred
Cramp 2013	6 (371)	3	0	Adverse events were recorded (and reported), but none occurred
Fransen 2014	10 (> 549)	5	7 events over 3 studies	-
Fransen 2015	54 (5362)	11	42 events over 8 studies	-
Gross 2015a	16 (2485)	11	41 events over 6 studies	-
Han 2004	3 (206)	2	1 event in 1 study	In narrative: "approximately one-third of the patients complained of soreness in the knee, shoulder or lower back during the first 3 weeks… pain eventually subsided for all patients… only exception was one patient, who complained of knee pain."
Hayden 2005	43 (3907)	10	23 events over 10 studies	"Negative reported: 16 events over 7 trials."
Hurkmans 2009	5 (575)	2	0	Adverse events were recorded (and reported), but none occurred

Table 14. Adverse events (not death) (*Continued*)

Koopman 2015	2 (68)	1 (10)	0	Adverse events were recorded (and reported), but none occurred "The study investigated deleterious effects of this training on motor unit survival through motor unit number estimates (MUNE). Results showed that the MUNE did not change at the end of the training."
Lane 2014	30 (1822)	1 (88 exercising participants)	2 events in control group in 1 study	RR 0.20 (95% CI 0.01 to 4.15) in favour of exercise group.
Regnaux 2015	1 (102)	1 (68 exercising participants over 2 groups: low and high resistance)	3 events in 1 study	"3 participants in high resistance group discontinued the exercise intervention due to severe knee pain."
Saragiotto 2016	7 (671)	1 (154)	5 events in 1 study	"Five patients (three from the MCE [motor control exercise] group and two from the minimal intervention group) had mild adverse effects during the study (all temporary exacerbations of pain)."
van der Heijden 2015	10 (1690)	0	0	Of the relevant studies, none actively reported on adverse events
Yamato 2015	6 (265)	1 (86)	0	Adverse events were recorded (and reported), but none occurred
Total	**246 studies (> 21,772)**	**61 studies (> 2134 participants)**	**137 events over 39 studies**	**61/246 (25%) of studies have reported on adverse events; of which 39/61 (64%) did have adverse events occur as a result of the intervention or control.**

n: number of participants; RR: risk ratio.

WHAT'S NEW

Last assessed as up-to-date: 31 January 2016.

Date	Event	Description
18 April 2017	New citation required but conclusions have not changed	Conclusions not changed; retrospective open access.
10 April 2017	Amended	See Published notes.

CONTRIBUTIONS OF AUTHORS

LG conceived the idea of the overview, and wrote the protocol and full overview.

LG, BHS, LC, and RAM developed the concept and details of the overview (participants, intervention, comparison, outcomes).

LG and CC carried out searches and selected reviews for inclusion (RAM and DM acted as arbitrators).

LG and CC carried out assessment of methodological quality using the AMSTAR tool (DM acted as arbitrator).

LG and RAM extracted data and interpreted initial findings.

LG, RAM, LC, and BHS formulated the focus of the discussion and made suggestions for future study and review authors.

All authors were involved in the interpretation of results, and in approving the final review.

LG and BHS will be responsible for future updates.

DECLARATIONS OF INTEREST

LG: none known.

RAM has received grant support from Grünenthal relating to individual patient level analyses of trial data regarding tapentadol in osteoarthritis and back pain (2015). He has received honoraria for attending boards with Menarini concerning methods of analgesic trial design (2014), with Novartis (2014) about the design of network meta-analyses, and RB on understanding pharmacokinetics of drug uptake (2015). He has received honoraria from Omega Pharma (2016) and Futura Pharma (2016) for providing advice on trial and data analysis methods.

CC: none known.

DM: none known.

LC: none known. For transparency, LC has received honoraria for speaking at educational meetings to healthcare professionals on a range of chronic pain topics (Pfizer (October 2015), Astellas (June 2014, March 2015)); editor on the *British Journal of Anaesthesia* (receives an honorarium plus a contribution toward related departmental expenses (October 2010 - to date)). LC is a medical clinician attending patients in the NHS Lothian Pain Service.

BHS: none known. BHS is a medical clinician attending patients in the NHS Tayside Pain Service.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- MRC, Lifelong Health and Wellbeing Grant, ID 91029, UK.
- National Institute of Academic Anaesthesia (BJA RCoA NIAA), UK.

NOTES

This overview review was re-published in April 2017 with retrospective open access.

ORIGINAL ARTICLE

Pain Perception After Isometric Exercise in Women With Fibromyalgia

Marie K. Hoeger Bement, PT, PhD, Andy Weyer, DPT, Sarah Hartley, DPTs, Breanna Drewek, BS, April L. Harkins, PhD, Sandra K. Hunter, PhD

ABSTRACT. Hoeger Bement MK, Weyer A, Hartley S, Drewek B, Harkins AL, Hunter SK. Pain perception after isometric exercise in women with fibromyalgia. Arch Phys Med Rehabil 2011;92:89-95.

Objective: The purpose of this study was to identify exercise protocols incorporating isometric contractions that provide pain relief in women with fibromyalgia.

Design: A before-after trial.

Setting: A physical therapy department in an academic setting.

Participants: Fifteen women (mean ± SD, 52±11y) with fibromyalgia.

Interventions: Subjects completed 4 sessions: 1 familiarization and 3 experimental. The following randomized experimental sessions involved the performance of isometric contractions with the elbow flexor muscles that varied in intensity and duration: (1) 3 maximal voluntary contractions (MVCs), (2) 25% MVC held to task failure, and (3) 25% MVC held for 2 minutes.

Main Outcome Measures: Experimental pain (pain threshold and pain rating), Fibromyalgia Impact Questionnaire, and fibromyalgia pain intensity (visual analog scale).

Results: After all 3 isometric contractions, there was considerable variability between subjects in the pain response. Based on the changes in experimental pain, subjects were divided into 3 groups (increase, decrease, no change in pain). Multiple regression analysis revealed that age, baseline experimental pain, and change in fibromyalgia pain intensity were significant predictors of the experimental pain response after the isometric contractions.

Conclusions: We identified subgroups of women with fibromyalgia based on how they perceived pain after isometric contractions. The greatest pain relief for women with fibromyalgia occurred at a younger age and in women with the greatest experimental pain before exercise. Additionally, we established a link between experimental and clinical pain relief after the performance of isometric contractions.

Key Words: Fibromyalgia; Isometric exercise; Rehabilitation

FIBROMYALGIA IS A disorder that is characterized by chronic widespread pain and tenderness with palpation. The prevalence of fibromyalgia is estimated to be as high as 5% of the general population, with the majority of persons being women.[1] The management of fibromyalgia is difficult because many patients do not respond to treatment. For example, only 30% to 50% of patients experience significant improvement with medication therapy,[2] indicating the need to identify alternative therapies for pain relief.

Exercise is one of the few therapies that may improve fibromyalgia symptoms.[3,4] The majority of the exercise studies[3,5-7] have focused on aerobic activity leading to improvements in pain, physical function, fatigue, mood, and self-efficacy. Understanding the role and potential benefit of isometric contractions in managing pain with fibromyalgia is important because even patients with limited mobility can perform an isometric contraction. Therefore, there is significant potential for self-management of fibromyalgia pain symptoms by using exercise treatments incorporating isometric contractions.

In healthy adults, both low- and high-intensity isometric contractions produce significant pain relief.[8] The dosage of isometric contractions to produce pain relief in persons with fibromyalgia is not known. Thus, the primary purpose of this study was to identify effective exercise protocols incorporating isometric contractions by measuring the pain response to various intensities and durations of isometric exercise in women with fibromyalgia.

A secondary purpose was to explore potential mechanisms for exercise-induced changes in pain perception in persons with fibromyalgia. As an index of HPA activity, we measured salivary cortisol levels before and after exercise to understand the role of the HPA axis in modulating the pain response.[9]

METHODS

Subjects

Fifteen women (mean ± SD, 52±11y; range, 19–64y) diagnosed with fibromyalgia completed this study. Subjects were screened for known cardiopulmonary and neurologic problems. Informed consent was acquired before the start of the study, and the protocol was approved by the Institutional Review Board.

List of Abbreviations

25FAIL	25% MVC held to the task failure session
25TIME	25% MVC held for a 2-minute session
FIQ	Fibromyalgia Impact Questionnaire
HPA	hypothalamus-pituitary-adrenal
MAP	mean arterial pressure
MVC	maximal voluntary contraction
RPE	rate of perceived exertion
VAS	visual analog scale

From the Departments of Physical Therapy (Hoeger Bement, Weyer, Hartley, Drewek, Hunter) and Clinical Laboratory Science (Harkins), Marquette University, Milwaukee, WI.

Supported by funding from the Arthritis Foundation.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Reprint requests to Marie Hoeger Bement, PT, PhD, Dept of Physical Therapy, Marquette University, PO Box 1881, Milwaukee, WI 53201, e-mail: *mariehoeger.bement@marquette.edu*.

0003-9993/11/9201-00434$36.00/0
doi:10.1016/j.apmr.2010.10.006

General Experimental Protocol

Subjects completed 4 sessions: 1 familiarization session and 3 experimental sessions. During the familiarization session, experimental pain was measured by using a pressure pain device before and after 30 minutes of quiet rest. Additional measurements included state and trait anxiety[10,11] and clinical pain assessments using the FIQ,[12,13] the McGill Pain (short-form) Questionnaire,[14] and a VAS to measure fibromyalgia pain intensity. Three MVCs were performed at the end of the familiarization session to determine the appropriate target forces for the experimental sessions.

During the following 3 experimental sessions, pain perception was measured before and after isometric contractions that varied in intensity and duration: 25FAIL, 25TIME, and 3 MVCs (MVC session). The isometric contractions were performed during separate randomized sessions, with approximately 1 week between sessions. Heart rate, blood pressure, and the rate of perceived exertion were monitored during the submaximal isometric contractions. Subjects completed the FIQ, the McGill Pain Questionnaire (short-form), and the state anxiety form at the beginning of each session. State anxiety was also assessed after each experimental pain assessment. Fibromyalgia pain intensity measured with a VAS was measured throughout each session at the following time points: start of the experiment, after the first pain test, immediately after the second pain test (and exercise), and 20 minutes after the second pain test.

Measurement of Force During Isometric Contractions

Subjects performed isometric contractions with the elbow flexor muscles of the left arm while seated.[8,15,16] The elbow joint was flexed to 90°, and the hand and forearm were placed midway between pronation and supination in a modified wrist-hand-thumb orthesis.[a] The forces extended by the wrist in the vertical direction were measured with a force transducer[b] and recorded online by using a Power 1401 A-D converter and Spike2 software.[c]

Sustained Submaximal Isometric Contraction

During the submaximal isometric contractions (25FAIL and 25TIME), each subject was required to match the vertical target force as displayed on the monitor. Task failure was determined as the time the force declined by 10% of the target value for 3 of 5 consecutive seconds.

Experimental Pain

A pressure pain device was used to measure pain perception before and after exercise.[8,15,16] A 100-g mass was applied to a second-class lever at a distance from the axis that was equivalent to a 0.75-kg mass on the finger. The force at the finger was applied through a Lucite edge[d] and placed on the right index for 2 minutes. Subjects were instructed to say "pain" when they first felt pain (ie, pain threshold), and pain ratings were reported every 20 seconds using a 0 to 10 scale. The scale had the following terminology: 0, no pain; 5, moderate pain; and 10, worst pain.[17] We previously established the reliability of this device.[8]

Cortisol

All experiments were conducted in the afternoon when the cortisol levels were near their diurnal trough.[18] Salivary cortisol was collected upon entering the laboratory and 10 and 20 minutes after the first and second pain measurement. Cortisol levels were quantified using an enzymatic immunosolvent assay.[e] Our intra- and interassay coefficients of variation were within standard precision levels per the recommendations of the manufacturer.

Statistical Analysis

Data are reported as mean ± SD within the text and displayed as mean ± SE in the figures. For each exercise session, repeated-measures analysis of variance was used to compare the following variables across trial and/or time: pain threshold, pain rating, fibromyalgia pain intensity, state anxiety, cortisol levels, MAP, and heart rate. Because the pain response varied considerably among the fibromyalgia subjects, the data were analyzed with the between-subjects factor for pain response (increase, decrease, and no change). Pearson correlations were calculated to determine associations between dependent variables. Stepwise multiple regression was used to analyze the contribution of several variables to the exercise-induced change in pain threshold and pain ratings across all 3 exercise sessions that were pooled together. Only predictors that had significant partial effects were reported for the final regression model. A P value $\leq .05$ was used for statistical significance.

RESULTS

Familiarization Session

Pain thresholds and pain ratings remained unchanged after 30 minutes of quiet rest ($P=.86$ and .66, respectively). State anxiety and pain intensity assessed from the "current fibromyalgia pain" VAS did not change after the first or second pain test ($P=.99$ and .90, respectively). Furthermore, state anxiety was not correlated with any of the experimental or clinical pain measures.

Pain Perception

Experimental pain. Pain threshold did not change after exercise for any of the exercise sessions (trial effect: 25FAIL, $P=.22$ [fig 1A]; 25TIME, $P=.25$ [fig 2A]; MVCs, $P=.77$ [fig 3A]). Similarly, pain ratings did not change for any of the exercise sessions (trial effect: 25FAIL, $P=.92$ [fig 1B]; 25TIME, $P=.40$ [fig 2B]; MVCs, $P=.28$ [fig 3B]). There was, however, substantial variability between subjects in the exercise-induced pain response for each session. Based on the pain response, subjects were divided into 3 groups (decrease, increase, no change in pain) by using criteria for pain subgroups that we established previously.[19] The decreased pain group reported an increase in pain threshold $\geq$ 10 seconds and/or a decrease in pain ratings by $\geq$ 2 for 2 consecutive time points during the 2-minute pain pressure test. The increased pain group reported a decrease in pain threshold $\geq$ 10 seconds and/or an increase in pain ratings by greater than or equal to 2 for 2 consecutive time points. The no change in pain group experienced a less than 10-second change in pain threshold and a less than 2 change in pain ratings for 2 consecutive time points. After the submaximal contraction held until task failure (25FAIL), there was an interaction between trial and pain response (decreased pain group [n=5], increased pain group [n=5], no change in pain group [n=5]) for both pain ratings ($P=.007$) (fig 1D, F, H) and pain threshold ($P=.01$) (fig 1C, E, G). For 25TIME, there was a trial and pain response interaction (decreased pain group [n=2], increased pain group [n=8], no change in pain group [n=5]) for the pain ratings ($P=.02$) (fig 2D, F, H) but not for the pain threshold ($P=.07$) (fig 2C, E, G). For the MVC exercise session, there was an interaction between trial and pain response (decreased pain group [n=5], increased pain group [n=5], no change in pain group [n= 5]) for both pain ratings ($P<.0001$) (fig 3D, F, H) and pain threshold ($P=.001$) (fig 3C, E, G).

Clinical pain. Fibromyalgia pain intensity measured with the VAS did not change over time when analyzing the group data (time: 25FAIL, $P=.08$; 25TIME, $P=.31$; MVCs, $P=.45$).

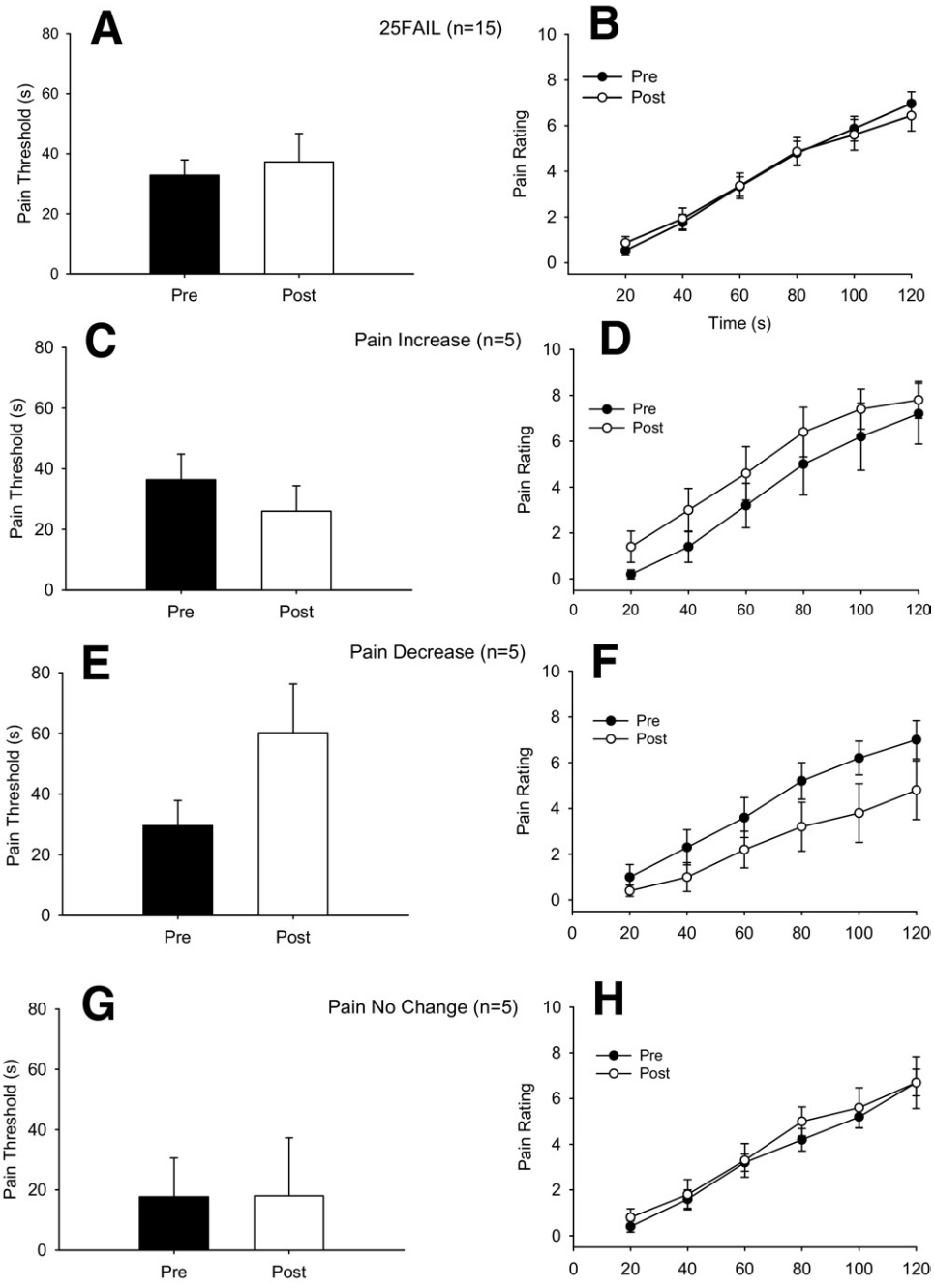


Fig 1. (A) Pain threshold and (B) pain ratings before (pre) and after (post) 25FAIL. There was no significant difference in pain threshold or pain ratings after the sustained contraction ($P>.05$). Based on the pain response, the subjects were divided into 3 groups (pain increase [C, D], pain decrease [E, F], and no change in pain [G, H]). There was a significant interaction between trial and pain response for both pain threshold and pain ratings ($P<.05$). Data are represented as the average ± standard error of the mean.

For VAS, there was no interaction between pain response and time for the submaximal contraction held for 2 minutes or the 3 MVCs (response × time, $P=.39$ and .42, respectively). However, there was an interaction for VAS between pain response and time for the submaximal contraction held until task failure ($P=.02$). Thus, after fatiguing isometric contractions, the pain subgroups were similar for experimental and clinical pain.

Time to Task Failure

Women with fibromyalgia sustained the 25% MVC held until task failure for 502 ± 262 seconds. Five subjects (33%) requested to stop the exercise session despite minimal indications of fatigue associated with force decline. The RPEs of these 5 subjects were all at a maximal of 10 when they requested to cease the contraction. There was an association between the exercise-induced change in fibromyalgia pain intensity and the time to task to failure ($r=.54$, $P=.04$) such that those subjects who had the longest time to failure experienced an increase in pain after exercise.

Anxiety

State anxiety levels did not change during any of the exercise sessions (time effect: $P>.05$). Furthermore, there was no interaction between pain response and time for state anxiety ($P>.05$).

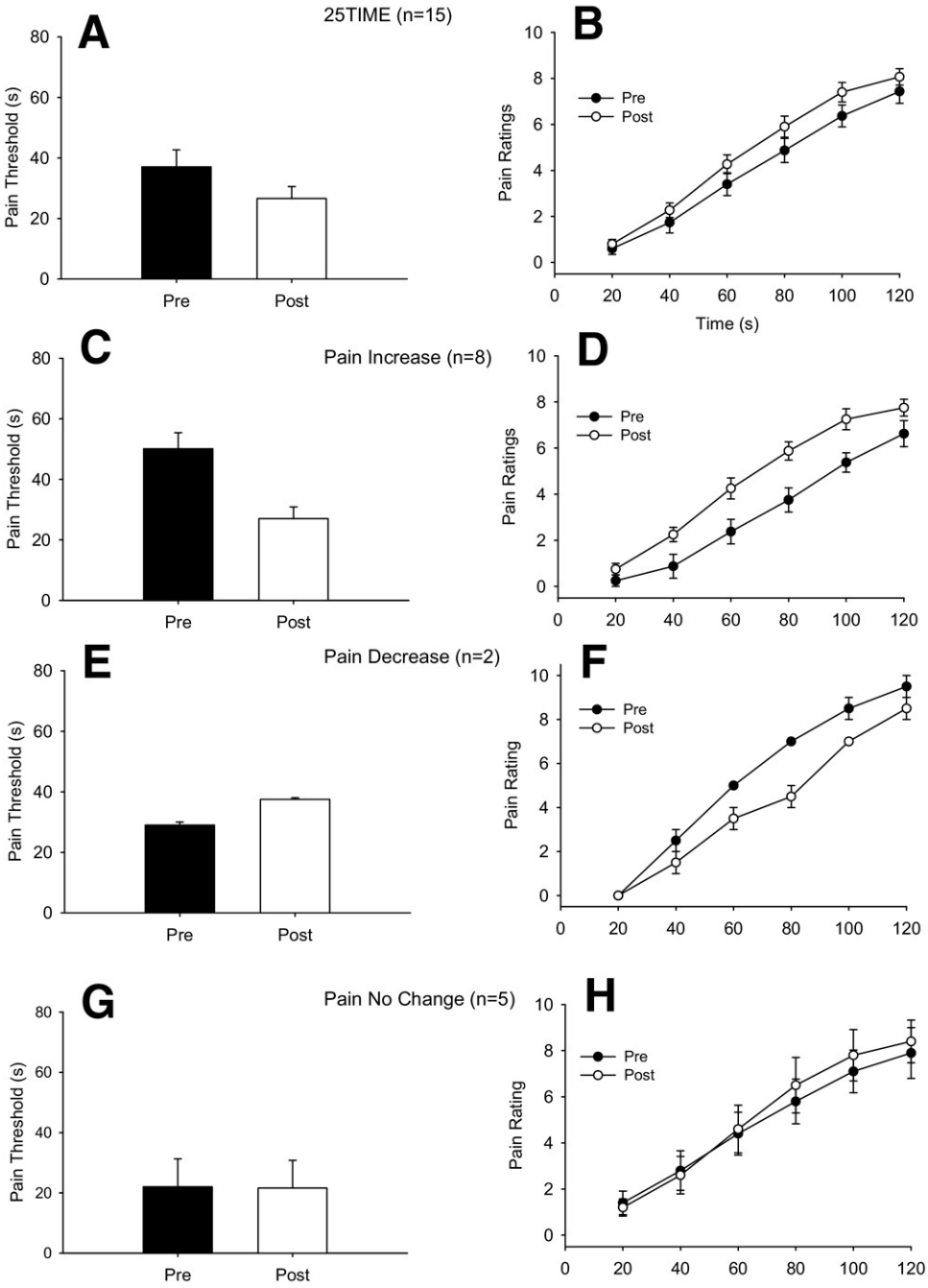


Fig 2. (A) Pain threshold and (B) pain ratings before (pre) and after (post) 25TIME. There was no significant difference in pain threshold or pain ratings after the isometric contraction ($P>.05$). Based on the pain response, the subjects were divided into 3 groups (pain increase [C, D], pain decrease [E, F], and no change in pain [G, H]). There was a significant interaction between trial and pain response for pain ratings ($P<.05$) but not for pain threshold ($P>.05$). Data are represented as the average ± standard error of the mean.

Cortisol

Salivary cortisol levels remain unchanged over time during each of the exercise sessions ($P>.05$), indicating that there was no measurable change in cortisol after the pressure pain test or the performance of isometric contractions. There was no interaction for cortisol between the pain response and time ($P>.05$).

MAP, Heart Rate, and RPE during the Submaximal Sustained Contractions

During the submaximal contraction held until task failure (25FAIL), there was a progressive increase in MAP ($P<.0001$), heart rate ($P=.001$), and RPE ($P<.0001$). There was no interaction between pain response and time for MAP, heart rate, or RPE ($P>.05$), indicating the different pain response groups responded similarly during the fatiguing contraction for these measures.

During 25TIME, there was an increase in MAP ($P<.0001$) and RPE ($P<.0001$) although heart rate showed no significant increase ($P=.27$). There was no interaction between pain response and time for MAP, heart rate, or RPE ($P>.05$).

Stepwise Multiple Regression Analysis

Stepwise regression analysis indicated that age and baseline pain threshold were the most significant predictors for the

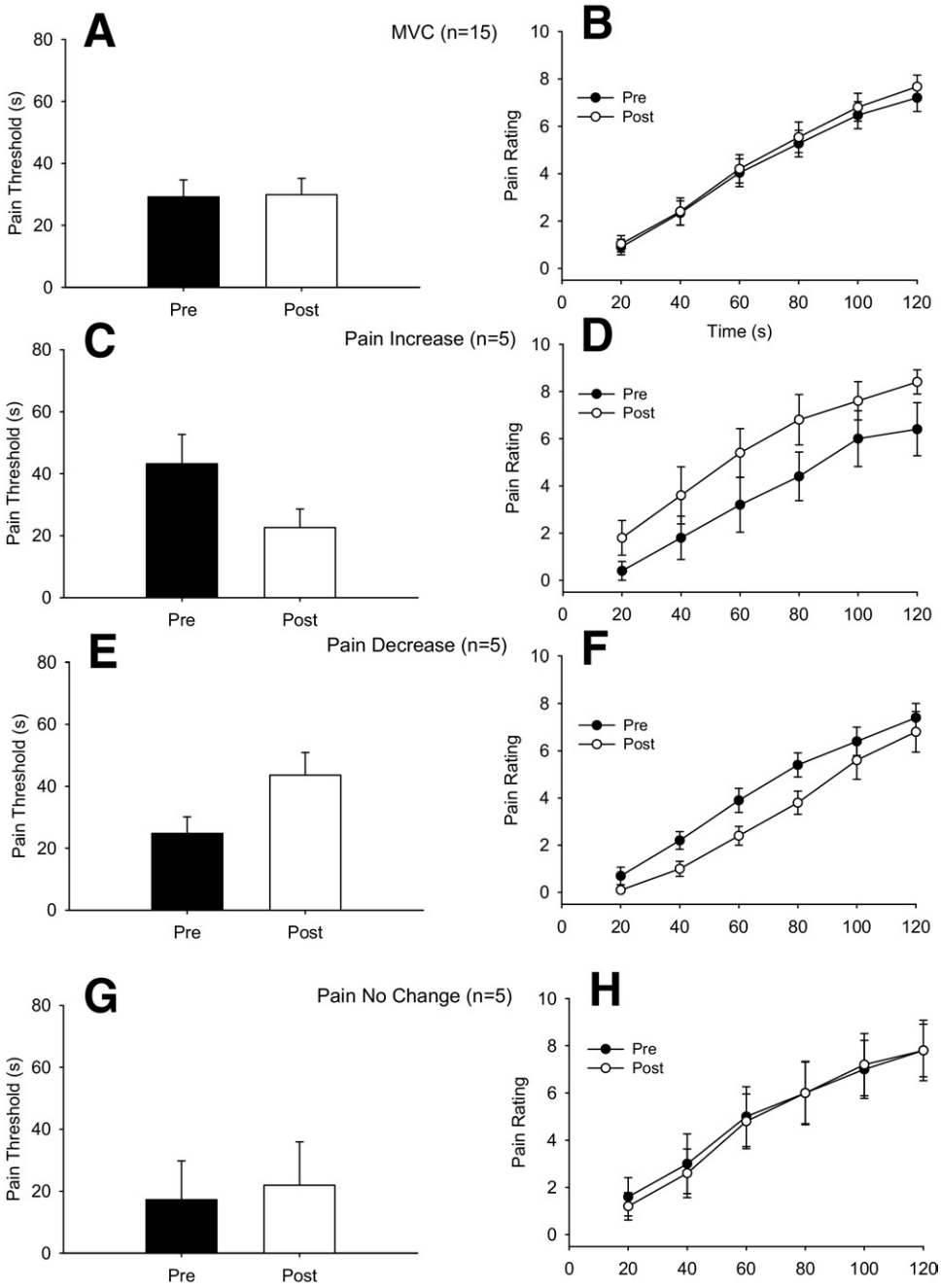


Fig 3. (A) Pain threshold and (B) pain ratings before (pre) and after (post) 3 brief MVCs. There was no significant difference (*P*>.05) in pain threshold or pain ratings after the isometric contractions. Based on the pain response, the subjects were divided into 3 groups (pain increase [C, D], pain decrease [E, F], and no change in pain [G, H]). There was a significant interaction between trial and pain response for both pain threshold and pain ratings (*P*<.05). Data are represented as the average ± standard error of the mean.

exercise-induced change in pain threshold when all 3 sessions were entered into 1 regression analysis. Together these 2 variables explained 32% (r=.56, r^2=.32, P=.001) of the variance in the pain threshold changes after exercise. These results indicate that younger women with fibromyalgia were more likely to experience pain relief after isometric exercise than older women. Also, women with lower baseline pain thresholds were more likely to experience pain relief than those with higher pain thresholds.

For changes in pain ratings, regression analysis showed that baseline pain threshold and change in fibromyalgia pain intensity were the 2 significant predictors, which accounted for 31% (r=.56, r^2=.31, P=.001) of the variance. These results indicate that greater relief in experimental pain results in greater attenuation of fibromyalgia pain intensity.

DISCUSSION

The main purpose of this study was to identify effective isometric contraction intensities and durations that elicit exercise-induced hypoalgesia in women with fibromyalgia. We first established that the application of the experimental pain test did not result in an increase in experimental pain, clinical pain, or state anxiety in women with fibromyalgia (familiarization session). These results show that any changes in pain reports in the subsequent exercise sessions were because of the exercise intervention and not the methods and execution of the pain test.

After the exercise sessions, there was no change in the pain response when analyzing the group of fibromyalgia women together regardless of the intensity or duration of the isometric contraction. However, there was considerable variability between subjects in that an equal number of subjects experienced exercise-induced increases, decreases, and no changes in pain after the isometric task held until task failure and the 3 MVC sessions. Other research studies[20-22] have also shown mixed pain responses after isometric exercise in fibromyalgia patients. Therefore, the assumption that patients with chronic pain are homogenous may explain the deficit in successful treatment options for some patients.[23]

Multiple regression analysis revealed that one of the main predictors of the change in pain threshold was the age of the woman. We found that younger women with fibromyalgia were more likely to experience pain relief than older women. Consistent with our results, for studies in which there was an exercise-induced increase in pain, the fibromyalgia subjects had an average age of 46 and 48 years old,[21,22] whereas there was no change or decrease in pain for a study in which the subjects' average age was 39 years.[20] The influence of age on the exercise response has important implications in the management of fibromyalgia because the prevalence of fibromyalgia increases with age.[1]

This study also established that baseline pain threshold predicted exercise-induced changes in both pain threshold and pain ratings. Specifically, persons with lower pain thresholds during the first experimental pain test were more likely to experience a postexercise decrease in pain compared with subjects who had higher baseline pain thresholds. Interestingly, tenderness with evoked-pain testing has been shown to be an important factor in determining subgroups.[24] However, our study is the first to show that baseline levels of experimental pain can predict pain relief and the effectiveness of exercise. One explanation for baseline levels as a predictor is that those subjects who experienced the greatest baseline pain (before exercise) may have more pain during the exercise intervention. Thus, exercise would act as a counterirritant (ie, activation of nociceptors produce an endogenous analgesic response), resulting in greater pain relief for these persons after exercise compared with those who experienced less baseline pain.[25]

Another predictor for exercise-induced changes in pain was the change in fibromyalgia pain intensity. Those subjects who had the greatest experimental pain relief also had the greatest decrease in fibromyalgia pain intensity. Thus, the experimental and clinical pain responses were similar in direction and magnitude after isometric exercise. Fibromyalgia pain intensity (VAS) was also associated with time to failure of the submaximal contraction. Those subjects who maintained the contraction for a longer duration were more likely to experience an increase in pain compared with subjects who maintained the contraction for a shorter duration. Furthermore, the time to failure was influenced by a subset of subjects (33% of the subjects) who asked to stop the submaximal contraction before the force criteria for ending the task were reached. Previous studies[22,26-29] have shown that persons with fibromyalgia report greater levels of perceived effort and higher pain levels for both static and dynamic contractions compared with healthy controls. This exertional pain may be related to exercise-induced muscle ischemia and sensitization of the peripheral nervous system.[21,30] Because we did not assess pain during the exercise sessions in this study, whether subjects requested to stop because of exercise-induced increases in pain cannot be determined.

To determine potential mechanisms responsible for exercise-induced changes in pain perception, activation of the HPA axis was assessed by measuring salivary cortisol levels. However, an acute bout of isometric contractions, regardless of intensity or duration, did not change cortisol levels. These results are similar to our previous study in healthy young women.[15] Several other studies[28,31-34] have assessed the effect of various exercise protocols on cortisol levels in fibromyalgia subjects with a wide range of results. Consequently, changes in cortisol levels after an acute bout of exercise do not appear to mediate exercise-induced hypoalgesia.

Study Limitations

The main limitation of this study was that pain perception was not measured during the exercise session. Thus, we do not know if persons who requested to stop the fatiguing exercise session did so due to pain. Another related issue is that we do not know if exercise acted as a counterirritant, therefore explaining the hypoalgesia for some subjects. Future studies warrant examining pain during exercise in patients with fibromyalgia.

CONCLUSIONS

There is considerable variability in pain reports of women with fibromyalgia after the performance of isometric contractions. The pain response was predicted by age, baseline experimental pain measures, and changes in fibromyalgia pain intensity. Additional research is needed to further characterize the subgroups of the pain response to exercise and to determine if these subgroups persist with long-term training programs.

Our findings potentially have important clinical implications for patients with fibromyalgia. First, we established that clinical and experimental pain relief can be harnessed at least in the short-term from static contractions in some patients with fibromyalgia, thereby establishing the important link between clinical and experimental pain. Furthermore, our results show that women with fibromyalgia who experience substantial pain perception before exercise may gain the greatest pain relief from isometric contractions.

References

1. Neumann L, Buskila D. Epidemiology of fibromyalgia. Curr Pain Headache Rep 2003;7:362-8.
2. Goldenberg DL. Fibromyalgia syndrome a decade later: what have we learned? Arch Intern Med 1999;159:777-85.
3. Busch AJ, Barber KA, Overend TJ, Peloso PM, Schachter CL. Exercise for treating fibromyalgia syndrome. Cochrane Database Syst Rev 2007;4:CD003786.
4. Goldenberg DL, Burckhardt C, Crofford L. Management of fibromyalgia syndrome. JAMA 2004;292:2388-95.
5. Brosseau L, Wells GA, Tugwell P, et al. Ottawa Panel evidence-based clinical practice guidelines for strengthening exercises in the management of fibromyalgia: part 2. Phys Ther 2008;88:873-86.
6. Busch AJ, Schachter CL, Overend TJ, Peloso PM, Barber KA. Exercise for fibromyalgia: a systematic review. J Rheumatol 2008; 35:1130-44.
7. Gowans SE, deHueck A. Effectiveness of exercise in management of fibromyalgia. Curr Opin Rheumatol 2004;16:138-42.
8. Hoeger Bement MK, Dicapo J, Rasiarmos R, Hunter SK. Dose response of isometric contractions on pain perception in healthy adults. Med Sci Sports Exerc 2008;40:1880-9.
9. Lundy-Ekman L. Neuroscience fundamentals for rehabilitation. 3rd ed. St. Louis: Saunders; 2007.
10. Barnes LLB, Harp D, Jung WS. Reliability generalization of scores on the Spielberger state-trait anxiety inventory. Educ Psychol Meas 2002;62:603-18.
11. Spielberg CD. Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Pr; 1983.

12. Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. Clin Exp Rheumatol 2005;23(5 Suppl 39):S154-62.
13. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol 1991;18: 728-33.
14. Melzack R. The short-form McGill Pain Questionnaire. Pain 1987;30:191-7.
15. Hoeger Bement MK, Rasiarmos RL, DiCapo JM, et al. The role of the menstrual cycle phase in pain perception before and after an isometric fatiguing contraction. Eur J Appl Physiol 2009;106:105-12.
16. Hoeger Bement MK, Weyer A, Hartley S, Yoon T, Hunter SK. Fatiguing exercise attenuates pain-induced corticomotor excitability. Neurosci Lett 2009;452:209-13.
17. McCaffery M, Pasero M. Pain: Clinical manual. 2nd ed. St. Louis: Mosby; 1999.
18. Schmidt-Reinwald A, Pruessner JC, Hellhammer DH, et al. The cortisol response to awakening in relation to different challenge tests and a 12-hour cortisol rhythm. Life Sci 1999;64:1653-60.
19. Hoeger Bement MK, Weyer A, Keller A, Harkins A, Hunter SK. Anxiety and stress can predict pain perception following a cognitive stress. Physiol Behav 2010;101:87-92.
20. Kadetoff D, Kosek E. The effects of static muscular contraction on blood pressure, heart rate, pain ratings and pressure pain thresholds in healthy individuals and patients with fibromyalgia. Eur J Pain 2007;11:39-47.
21. Kosek E, Ekholm J, Hansson P. Modulation of pressure pain thresholds during and following isometric contraction in patients with fibromyalgia and in healthy controls. Pain 1996;64:415-23.
22. Staud R, Robinson ME, Price DD. Isometric exercise has opposite effects on central pain mechanisms in fibromyalgia patients compared to normal controls. Pain 2005;118:176-84.
23. Turk DC. The potential of treatment matching for subgroups of patients with chronic pain: lumping versus splitting. Clin J Pain 2005;21:44-55; discussion 69-72.
24. Giesecke T, Williams DA, Harris RE, et al. Subgrouping of fibromyalgia patients on the basis of pressure-pain thresholds and psychological factors. Arthritis Rheum 2003;48:2916-22.
25. Willer JC, Roby A, Le Bars D. Psychophysical and electrophysiological approaches to the pain-relieving effects of heterotopic nociceptive stimuli. Brain 1984;107:1095-112.
26. Mengshoel AM, Vollestad NK, Forre O. Pain and fatigue induced by exercise in fibromyalgia patients and sedentary healthy subjects. Clin Exp Rheumatol 1995;13:477-82.
27. Miller TA, Allen GM, Gandevia SC. Muscle force, perceived effort, and voluntary activation of the elbow flexors assessed with sensitive twitch interpolation in fibromyalgia. J Rheumatol 1996; 23:1621-7.
28. Giske L, Vollestad NK, Mengshoel AM, Jensen J, Knardahl S, Roe C. Attenuated adrenergic responses to exercise in women with fibromyalgia—a controlled study. Eur J Pain 2008;12:351-60.
29. Maquet D, Croisier JL, Renard C, Crielaard JM. Muscle performance in patients with fibromyalgia. Joint Bone Spine 2002;69: 293-9.
30. Jones KD, Liptan GL. Exercise interventions in fibromyalgia: clinical applications from the evidence. Rheum Dis Clin North Am 2009;35:373-91.
31. Gursel Y, Ergin S, Ulus Y, Erdogan MF, Yalcin P, Evcik D. Hormonal responses to exercise stress test in patients with fibromyalgia syndrome. Clin Rheumatol 2001;20:401-5.
32. Paiva ES, Deodhar A, Jones KD, Bennett R. Impaired growth hormone secretion in fibromyalgia patients: evidence for augmented hypothalamic somatostatin tone. Arthritis Rheum 2002; 46:1344-50.
33. van Denderen JC, Boersma JW, Zeinstra P, Hollander AP, van Neerbos BR. Physiological effects of exhaustive physical exercise in primary fibromyalgia syndrome (PFS): is PFS a disorder of neuroendocrine reactivity? Scand J Rheumatol 1992; 21:35-7.
34. Valkeinen H, Hakkinen K, Pakarinen A, et al. Muscle hypertrophy, strength development, and serum hormones during strength training in elderly women with fibromyalgia. Scand J Rheumatol 2005;34:309-14.

Suppliers

a. Orthomerica, 505 31st Street, Newport Beach, CA 92663-3805.
b. JR-3 Force-Moment Sensor, JR-3 Inc., 22 Harter Avenue, Woodland, CA 95776.
c. Cambridge Electronics Design, 4 The Science Park, Milton Rd, Cambridge CB4 0FE, UK.
d. Romus Inc., 23W701 Maple Ave, Roselle, IL 60172.
e. Salimetrics LLC, 101 Innovation Blvd # 302, State College, PA 16803-6605.

Manual Therapy 16 (2011) 413–418

Contents lists available at ScienceDirect

Manual Therapy

journal homepage: www.elsevier.com/math

Masterclass

How to explain central sensitization to patients with 'unexplained' chronic musculoskeletal pain: Practice guidelines

Jo Nijs [a,b,c,*], C. Paul van Wilgen [e,f], Jessica Van Oosterwijck [a,b,c], Miriam van Ittersum [d,e], Mira Meeus [a,b]

[a] Department of Human Physiology, Faculty of Physical Education & Physiotherapy, Vrije Universiteit Brussel, Belgium
[b] Division of Musculoskeletal Physiotherapy, Department of Health Care Sciences, Artesis University College, Antwerp, Belgium
[c] Department of Rehabilitation and Physiotherapy, University Hospital, Brussels, Belgium
[d] Research and Innovation Group in Health Care and Nursing, Hanze University of Applied Sciences, Groningen, The Netherlands
[e] Graduate School for Health Research, University Medical Centre Groningen, University of Groningen, The Netherlands
[f] University Centre for Sport, Exercise and Health, Centre for Human Movement Sciences, University Medical Centre Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 30 November 2010
Received in revised form
21 April 2011
Accepted 23 April 2011

Keywords:
Sensitization
Chronic pain
Whiplash
Fibromyalgia
Osteoarthritis
Low back pain

ABSTRACT

Central sensitization provides an evidence-based explanation for many cases of 'unexplained' chronic musculoskeletal pain. Prior to commencing rehabilitation in such cases, it is crucial to change maladaptive illness perceptions, to alter maladaptive pain cognitions and to reconceptualise pain. This can be accomplished by patient education about central sensitization and its role in chronic pain, a strategy known as pain physiology education. Pain physiology education is indicated when: 1) the clinical picture is characterized and dominated by central sensitization; and 2) maladaptive illness perceptions are present. Both are prerequisites for commencing pain physiology education. Face-to-face sessions of pain physiology education, in conjunction with written educational material, are effective for changing pain cognitions and improving health status in patients with various chronic musculoskeletal pain disorders. These include patients with chronic low back pain, chronic whiplash, fibromyalgia and chronic fatigue syndrome. After biopsychosocial assessment pain physiology education comprises of a first face-to-face session explaining basic pain physiology and contrasting acute nociception versus chronic pain (Session 1). Written information about pain physiology should be provided as homework in between session 1 and 2. The second session can be used to correct misunderstandings, and to facilitate the transition from knowledge to adaptive pain coping during daily life. Pain physiology education is a continuous process initiated during the educational sessions and continued within both the active treatment and during the longer term rehabilitation program.

1. Introduction

Over the past decades, scientific understanding of 'unexplained' chronic pain disorders has increased substantially. It has become clear that the majority of cases of chronic musculoskeletal pain are characterized by alterations in central nervous system processing. More specifically, the responsiveness of central neurons to input from unimodal and polymodal receptors is augmented, resulting in a pathophysiological state corresponding to central sensitization, characterized by generalized or widespread hypersensitivity (Meyer et al., 1995). Central sensitization encompasses impaired functioning of brain-orchestrated descending anti-nociceptive (inhibitory) mechanisms (Meeus et al., 2008), and (over)activation of descending and ascending pain facilitatory pathways (Meeus and Nijs, 2007; Staud et al., 2007). The net result is augmentation rather than inhibition of nociceptive transmission. In addition to the switch in balance between inhibitory and facilitatory pathways, central sensitization entails altered sensory processing in the brain (Staud et al., 2007). Indeed, a modulated 'pain signature' arises in the brain of patients with central sensitization. The altered pain neuromatrix comprises of a) increased activity in brain areas known to be involved in acute pain sensations e.g. the insula, anterior cingulate cortex and the prefrontal cortex, but not in the primary or secondary somatosensory cortex (Seifert and Maihöfner, 2009); and b) brain activity in regions generally not involved in acute pain sensations e.g. various brain stem nuclei, dorsolateral frontal cortex and parietal associated cortex (Seifert and Maihöfner, 2009). 'Cognitive emotional sensitization' (Brosschot, 2002) refers to the capacity of forebrain centres in exerting powerful influences

* Corresponding author. Vrije Universiteit Brussel, MFYS-LK, Building L, Pleinlaan 2, BE 1050 Brussels, Belgium. Tel.: +3226292222.
E-mail address: Jo.Nijs@vub.ac.be (J. Nijs).

1356-689X
doi:10.1016/j.math.2011.04.005

on various nuclei of the brainstem, including the nuclei identified as the origin of the descending facilitatory pathways (Zusman, 2002). The activity in descending pathways is not constant but can be modulated, for example by the level of vigilance, attention and stress (Rygh et al., 2002).

From a musculoskeletal perspective, it is important to realize that distal/peripheral mechanisms take part in the pathophysiology of central sensitization as well. Many cases of chronic musculoskeletal pain evolve from traumatic or non-traumatic local nociceptive musculoskeletal problems characterized by a period of massive peripheral input in the (sub)acute to chronic stage (e.g. chronic whiplash associated disorders or patients reporting a history of several surgical procedures). In response the central nervous system modulates the sensitivity of the somatosensory system. In addition, once central sensitization is established in cases of chronic musculoskeletal pain, it remains highly plastic: any new peripheral injury may serve as a new source of bottom-up (peripheral) nociceptive input, which in turn sustains or aggravates the process of central sensitization (Affaitati et al., 2010). Without new peripheral input, central sensitization does not resolve quickly, but rather sustains the chronic nature of the condition.

From a clinical perspective, it remains challenging for clinicians to implement science into practice. Clinical guidelines for the recognition (Nijs et al., 2010) and treatment (Nijs and Van Houdenhove, 2009; Nijs et al., 2009) of central sensitization in patients with chronic musculoskeletal pain have been provided, yet many issues remain. For example, how should clinicians apply the science of central sensitization and chronic pain to a case of chronic whiplash where the patient is sceptical about the biopsychosocial model, and convinced that the initial neck trauma caused severe cervical damage that remains invisible to modern imaging methods? Or a patient with moderate hip osteoarthritis saying 'The cartilage of my hips is melting away due to erosion, which in turn is triggered by overuse of my lower limbs' and 'I will not participate in exercise therapy because it will worsen my hip pain and hence the erosion of my cartilage'. Likewise, a patient with fibromyalgia convinced that her pain and related symptoms are due to an undetectable or 'new' virus, is unlikely to adhere to conservative interventions.

It is clear that initiating a treatment like graded activity is unlikely to be successful in these patients. Prior to commencing treatment in such cases the gap between the perceptions of the patient and their health care professional about pain and its treatment should be narrowed. Therefore it is crucial to change the patient's maladaptive illness perceptions and maladaptive pain cognitions and to reconceptualise pain before initiating the treatment. This can be accomplished by patient education about central sensitization and its role in chronic pain, a strategy frequently referred to as 'pain (neuro)physiology education' or 'pain biology education'. Patients with 'unexplained' chronic musculoskeletal pain who are misinformed about pain, consider their pain as more threatening and demonstrate lower pain tolerance, more catastrophic thoughts and less adaptive coping strategies (Jackson et al., 2005). Treatment adherence for active treatments is often low in these patients. Therefore, education will increase motivation for rehabilitation in those with chronic musculoskeletal pain due to central sensitization. This requires a biopsychosocial assessment and an in-depth education of altered central nervous system processing of nociceptive and non-nociceptive input. This will be the focus of the present paper.

In what follows the reader is provided with a brief overview of the clinical evidence of pain physiology education in patients with chronic musculoskeletal pain. The largest part of the paper is dedicated to practice guidelines on how to apply pain physiology education in patients with chronic musculoskeletal pain.

2. Effectiveness of pain physiology education

Several groups have studied the clinical effects of pain physiology education in various chronic musculoskeletal pain populations such as chronic low back pain (Moseley, 2002, 2003b, 2004, 2005; Moseley et al., 2004; Ryan et al., 2010), fibromyalgia (Ittersum et al. submitted for publication; Ittersum et al., in press; Van Oosterwijck et al. submitted for publication), chronic whiplash associated disorders (Van Oosterwijck et al., 2011) and chronic fatigue syndrome with chronic widespread pain (Meeus et al., 2010a). In patients with chronic low back pain, pain physiology education alters pain perceptions and, in conjunction with physiotherapy, it improves functional and symptomatic outcomes (Moseley, 2002; Moseley, 2003b; Moseley et al., 2004; Moseley, 2005). A recent randomized controlled trial indicates that, in the short term, pain physiology education alone is more effective for pain relief and improving pain self-efficacy than a combination of pain physiology education and group exercise classes for patients with chronic low back pain (Ryan et al., 2010). Altered pain perceptions are directly associated with altered movement performance in those with chronic low back pain, even if there is no opportunity for the patients to be physically active during the treatment (Moseley, 2002, 2004). This implies that motor performance may be directly limited by pain perceptions. Indeed, a case series study of patients with chronic whiplash associated disorders showed improvements in illness perceptions, pain thresholds and movement performance (Van Oosterwijck et al., 2011).

In patients with chronic fatigue syndrome, pain physiology education alters pain perceptions such as catastrophizing, and pain behaviour (Meeus et al., 2010a). In another randomized controlled clinical trial, we showed that simply providing the detailed information booklet explaining pain physiology and central sensitization, did not change illness perceptions or health status in patients with fibromyalgia (Ittersum et al. submitted for publication). However, when the same written education about pain physiology was combined with two educational sessions (one face-to-face session and one by telephone) of individually-tailored pain physiology education, vitality, physical functioning, mental and general health of patients with fibromyalgia improved (Van Oosterwijck et al. submitted for publication). This indicates that physiotherapists or other health care professionals are required to provide tailored education to address individual needs rather than standardized, general written education. Written education about central sensitization and pain physiology alone is insufficient. Nevertheless, an educational booklet about pain physiology is highly appreciated by fibromyalgia patients (Ittersum et al., in press), indicating that it can be used in conjunction with face-to-face educational meetings.

From the available evidence it is concluded that face-to-face sessions of pain physiology education, in conjunction with written educational material, are effective for changing pain perceptions and health status in patients with various chronic musculoskeletal pain disorders, including those with chronic low back pain, chronic whiplash, fibromyalgia and chronic fatigue syndrome. Practice guidelines on how to apply pain physiology education in patients with chronic musculoskeletal pain are provided below (and are summarized in Fig. 1).

3. Practice guidelines for applying pain physiology education

3.1. Prior to initiating pain physiology education

Prior to commencing pain physiology education, it is important firstly to ascertain that pain physiology education is indicated in the chronic pain patient. Pain physiology education is indicated when:

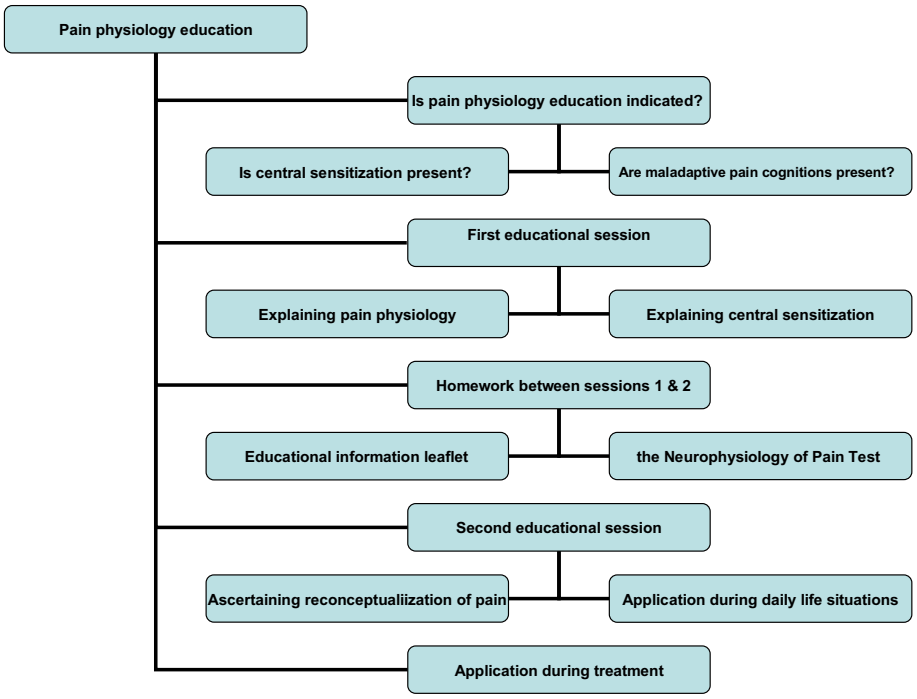


Fig. 1. Clinical guidelines for pain physiology education in patients with chronic musculoskeletal pain.

1) the clinical picture is characterized and dominated by central sensitization; and 2) maladaptive pain cognitions, illness perceptions or coping strategies are present. Both indications are prerequisites for commencing pain physiology education. Some (acute) musculoskeletal pain patients may not fulfil these requirements initially, but will do so later on during their course of treatment (e.g. a patient receiving physiotherapy for an acute muscle strain experiencing a whiplash trauma). To examine whether central sensitization is present, clinicians can use guidelines for the recognition of central sensitization in patients with chronic musculoskeletal pain (Nijs et al., 2010). In the assessment of illness perceptions patients must be asked about their perceptions about the cause of pain, the consequences, the treatment and the timeline of pain. Maladaptive pain cognitions include ruminating about pain, and hypervigilance to somatic signs, each of which can be easily assessed with short self-reported measures with excellent psychometric properties (e.g. the Pain Catastrophizing Scale[1], Pain Vigilance and Awareness Questionnaire[2], etc.) (Sullivan et al., 1995; Van Damme et al., 2002; Kraaimaat and Evers, 2003). Likewise, illness perception can be questioned or can be assessed by use of the brief Illness Perception Questionnaire[3] (Broadbent et al., 2006). This information addressing pain perceptions and coping strategies should be used by the therapist to tailor the individual education sessions (remember that pain physiology education aims to reconceptualise pain).

3.2. First educational session

It is essential for clinicians to explain the treatment rationale and discuss the practical issues of the treatment with the patient. In case of central sensitization and chronic musculoskeletal pain, explaining the treatment rationale is of prime importance. Basically, patients should understand the mechanism of central sensitization. Such education aims at altering patients' knowledge about their pain states and reconceptualising pain (Moseley, 2004). When solely cognitive and behavioural responses are encouraged, without reconceptualising pain, these responses may be counter-intuitive for chronic pain patients, because pain is still a sign of harm to them (Moseley, 2003b). Therefore education of the central sensitization model relies on deep learning, aimed at reconceptualising pain, based on the assumption that appropriate cognitive and behavioural responses will follow when pain is appraised as less dangerous (Moseley, 2003a). For example, remember the patient with chronic whiplash convinced that the initial neck trauma caused severe cervical damage that remains invisible to modern imaging methods. Simply providing education about the fear avoidance model to encourage a graded activity approach is unlikely to be beneficial. Detailed pain physiology education is required to reconceptualise pain, and to convince the patient that hypersensitivity of the central nervous system rather than local tissue damage is the cause of their presenting symptoms.

Educating patients with chronic musculoskeletal pain about central sensitization can be accomplished in one to two face-to-face educational sessions (approximately 30 min per session; depending on the change in cognitions). The aid of a booklet containing detailed written explanation and illustrations about pain physiology and central sensitization processes is recommended. The content of the education sessions can be based on the book "Explain Pain" (Butler and Moseley, 2003), covering the physiology of the nervous system in general and of the pain system in particular. Topics that should be addressed during the education sessions include the characteristics of acute versus chronic pain, the purpose of acute pain, how acute pain originates in the nervous system (nociceptors, ion gates, neurons, action potential, nociception, peripheral sensitization, synapses, synaptic gap, inhibitory/excitatory chemicals, spinal cord, descending/ascending pain pathways,

[1] http://synergytherapiesofkc.com/forms/PCS-Pain%20Catastrophizing%20Scale.pdf.

[2] The questionnaire can be obtained from the corresponding author or refer to the original publications addressing this measure.

[3] http://www.uib.no/ipq/.

role of the brain, pain memory and pain perception), how pain becomes chronic (plasticity of the nervous system, modulation, modification, central sensitization, the pain neuromatrix theory) and potential sustaining factors of central sensitization like emotions, stress, illness perceptions, pain cognitions and pain behaviour. Acute nociceptive mechanisms are typically explained first and are then contrasted with central sensitization processes i.e. in the case of chronic pain. Illustrations (e.g. Figs. 2 and 3), examples, and metaphors are frequently used (van Wilgen and Keizer, in press). The education is presented verbally (explanation by the therapist) and visually (summaries, pictures and diagrams on computer and paper). During the sessions patients are encouraged to ask questions and their input should be used to individualise the information.

3.3. Homework between sessions 1 and 2

After the face-to-face education, patients receive an educational information booklet about the neurophysiology of pain and are asked to read it carefully at home. The written information does not provide new information, it reinforces the verbal information as it tells the same story using the same drawings. Patients with central sensitization often have neurocognitive impairments, including concentration difficulties and impairments in short-term memory (Nijs et al., 2010), which implies that they can forget a number of aspects of the verbal education. Therefore additional written information that can be read afterwards is a valuable and essential part of the intervention. Sections 1, 2 and 4 from the book "Explain Pain" (Butler and Moseley, 2003) can be provided as written education to native English speakers, while a Dutch educational booklet is included in a practical guide for applying pain physiology education (van Wilgen and Nijs, 2010).

To examine whether the patient understands pain physiology, the patient version of the Neurophysiology of Pain Test[4] can be used (Moseley, 2003c; Meeus et al., 2010b). It is a valid and reliable measure for patients with chronic pain (Meeus et al., 2010b). At the end of session 1, the therapist asks the patient to fill out the Neurophysiology of Pain Test one day prior to returning to the clinic.

3.4. Second educational session

The outcome of the Neurophysiology of Pain Test can guide the clinician during the second educational session: it can identify those topics that require additional explanation. During the second session, the therapist answers and explains additional questions that arose after reading the information booklet. Based on the incorrect answers that were given on the 'Neurophysiology of Pain Test' the therapist explains these topics once again and if necessary in more detail. Hence, the clinician ascertains that the reconceptualization of pain has taken place and that illness perceptions have improved.

Next, the therapist discusses the existence of sensitization in this particular patient by giving the patient insight to somatic, psychosocial and behavioural factors associated with pain. This is followed by i.e. discussing with the patient how information provided can be applied during everyday situations. This is a crucial step in the overall treatment program, as it will set the way towards application of adaptive pain coping strategies, self-management programs and graded activity/graded exercise therapy. The therapist should start by asking the patient to explain his willingness to apply his increased understanding of his medical problem in his life for instance by setting new goals. Typical examples are stopping rumination and worrying about the aetiology and nature of their pain disorder, reducing stress, increasing physical activity levels, decreasing hypervigilance, becoming more physically active, relaxation etc. These, and other adaptive strategies, can be discussed with the patient and should lead to a clear plan of action on 'how to deal with the hypersensitive nervous system'. The transition from knowledge to adaptive pain coping can be enhanced by using the

[4] The Neurophysiology of Pain Test can be obtained from the corresponding author.

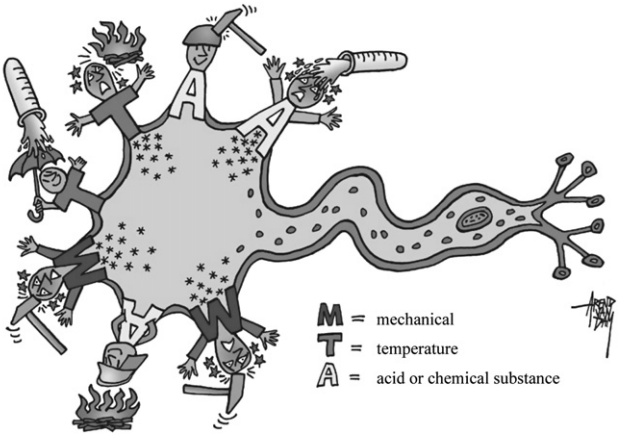


Fig. 2. Illustration used to explain basic principles of acute nociception (redrawn from reference (Butler and Moseley, 2003, p. 33) by Arend van Dam (van Wilgen and Nijs, 2010, p. 89); reproduced with permission from Bohn Stafleu van Loghum).This illustration presents a neuron, with on the left its 'sensors' which are capable of sensing temperature changes (indicated by the letter 'T'), chemical substances ('A') and mechanical pressure ('M'). Activation of such a sensor opens the corresponding ion channel in the cell membrane of the neuron. This enables an influx of sodium ions into the neuron ('positive charges enter the cell'), possibly resulting in an action potential ('the danger message'). It is important for the patient to realize that the presence of an action potential does not necessarily imply that pain is or will be experienced.

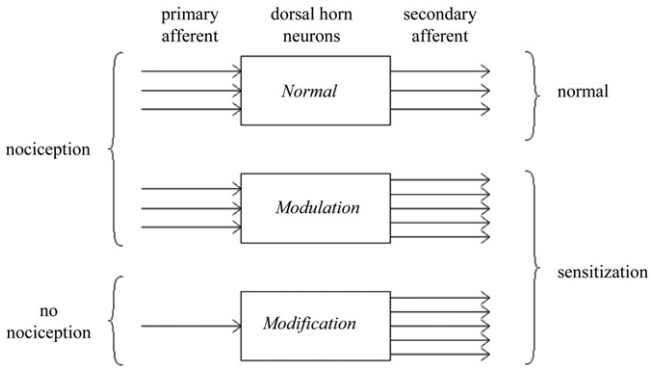


Fig. 3. Illustration used to explain the difference between acute nociception and central sensitization in chronic pain ((van Wilgen and Nijs, 2010, p. 102) reproduced with permission from Bohn Stafleu van Loghum). This illustration explains one of the essential features of central sensitization in chronic pain. The situation on top represents the normal situation, with primary afferents transporting 3 danger messages to the dorsal horn neurons, as is the case when you cut your finger. Next, the dorsal horn neurons activate the secondary afferents that transport the same 3 danger messages to the brain for processing. However, in many cases dorsal horn neurons modulate the incoming danger messages, as illustrated in the middle and below. The situation in the middle represents 'real' nociception, with 3 danger messages entering the spinal cord neurons, and 5 being sent to the brain. This implies that the incoming messages are amplified in the spinal cord prior to entering the brain. The situation below illustrates central sensitization in patients with chronic pain. Even in absence of nociception, messages from the periphery (e.g. touching the skin above the painful region or moving the affected limb) are amplified in a powerful way such that the dorsal horn neurons send several danger messages to the brain.

Pain Reaction Record (Sullivan, 2003), an easily applicable measure facilitating a cognitive approach to pain coping.

3.5. Application of pain physiology education during treatment

Pain physiology education is a continuous process initiated during the educational sessions prior to commencing active treatment (i.e. rehabilitation) and followed-up during the rehabilitation program. Indeed, pain physiology education is typically followed by various components of a biopsychosocial-oriented rehabilitation program, like stress management, graded activity and exercise therapy. It is important for clinicians to introduce these treatment components during the educational sessions, and to explain why and how the various treatment components are likely to contribute to decreasing the hypersensitivity of the central nervous system (as explained in Nijs and Van Houdenhove, 2009 and Nijs et al., 2009). Changing illness perceptions changes the patients motivation to undertake and comply with the rehabilitation program.

Likewise, long-term reconceptualization of pain, alterations in illness beliefs and adaptive pain cognitions are required at every stage of the rehabilitation program. This can be done easily by asking the patient to explain the treatment rationale of a specific treatment component. If during the treatment course any of the pain cognitions or illness beliefs have 'reset' towards maladaptive ones, then the therapist is advised to re-educate the patient. The latter can be accomplished by asking the patient to re-read the written information on pain physiology and to try to link that information with his/her current rehabilitation program. Long-term adaptive pain perceptions, and consequent adaptive pain coping strategies are required for long-term treatment compliance and continuous application of self-management strategies.

Finally, frequent side-effects and symptom fluctuations can be explained using the central sensitization model (van Wilgen and Keizer, in press). The latter should shift the patient's attention away from somatic signs towards adaptive coping strategies and reassurance. The patient's confidence in the treatment (outcome) should be a continuous treatment goal in those with chronic musculoskeletal pain.

4. Conclusion

There has been increased awareness that central sensitization provides an evidence-based explanation for many cases of 'unexplained' chronic musculoskeletal pain. Hence, rehabilitation of patients with chronic musculoskeletal pain should target, or at least take account of the process of central sensitization. Prior to commencing rehabilitation in such patients, it is crucial to change maladaptive illness beliefs, to alter maladaptive pain cognitions and to reconceptualise pain. This can be accomplished by patient education about central sensitization and its role in chronic pain, a strategy known as pain physiology education. Face-to-face sessions of pain physiology education, in conjunction with written educational material, are effective for changing pain cognitions and health status in patients with various chronic musculoskeletal pain disorders, including those with chronic low back pain, chronic whiplash, fibromyalgia and chronic fatigue syndrome. Pain physiology education comprises of a first face-to-face session explaining basic pain physiology and contrasting acute nociception versus chronic pain. Written information about pain physiology should be provided as homework in between session 1 and 2. The second session can be used to correct misunderstandings, and to facilitate the transition from knowledge to adaptive pain coping during daily life. Pain physiology education is a continuous process initiated during the two educational sessions prior to and continuing into active treatment and followed-up during the longer term rehabilitation program.

Acknowledgements

Mira Meeus is a postdoctoral research fellow of Research Foundation Flanders – FWO (Belgium). Jessica Van Oosterwijck is financially supported by grant no. OZR1596 from the research council of the Vrije Universiteit Brussel, Brussels, Belgium. The authors would like to thank Lorna Paul (PhD, PT; University of Glasgow, Scotland, UK) for editing the final version of the article.

References

Affaitati G, Costantini R, Fabrizio A, Lapenna D, Tafuri E, Giamberardino MA. Effects of treatment of peripheral pain generators in fibromyalgia patients. European Journal of Pain; 2010. doi:10.1016/j.ejpain.2010.09.002.

Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. Journal of Psychosomatic Research 2006;60:631–7.

Brosschot JF. Cognitive-emotional sensitization and somatic health complaints. Scandinavian Journal of Psychology 2002;43:113–21.

Butler D, Moseley GL. Explain pain. Adelaide: NOI Group Publishing; 2003.

Ittersum van MW, Wilgen van CP, Schans van der CP, Lambrecht L, Groothoff JW, Nijs J. Does pain physiology education change illness perceptions in patients with fibromyalgia? A multicentre randomized controlled trial with six months follow up, Submitted for publication.

Ittersum van MW, Wilgen van CP, Groothof JW, Schans van der CP. Is appreciation of written education about pain neurophysiology related to changes in illness perceptions and health status in patients with fibromyalgia? Patient Education and Counselling, in press.

Jackson T, Pope L, Nagasaka T, Fritch A, Iezzi T, Chen H. The impact of threatening information about pain on coping and pain tolerance. British Journal of Health Psychology 2005;10:441–51.

Kraaimaat FW, Evers AWM. Pain-coping strategies in chronic pain patients: psychometric characteristics of the pain-coping inventory. Journal of Behavioural Medicine 2003;10(4):343–63.

Moseley GL. Combined physiotherapy and education is efficacious for chronic low back pain. Australian Journal of Physiotherapy 2002;48:297–302.

Moseley GL. Joining forces – combining cognition-targeted motor control training with group or individual pain physiology education: a successful treatment for chronic low back pain. Journal of Manual and Manipulative Therapy 2003a;11:88–94.

Moseley GL. A pain neuromatrix approach to patients with chronic pain. Manual Therapy 2003b;8:130–40.

Moseley GL. Unraveling the barriers to reconceptualization of the problem in chronic pain: the actual and perceived ability of patients and health professionals to understand the neurophysiology. Journal of Pain 2003c;4:184–9.

Moseley GL. Evidence for a direct relationship between cognitive and physical change during an education intervention in people with chronic low back pain. European Journal of Pain 2004;8:39–45.

Moseley GL, Nicholas MK, Hodges PW. A randomized controlled trial of intensive neurophysiology education in chronic low back pain. Clinical Journal of Pain 2004;20:324–30.

Moseley GL. Widespread brain activity during an abdominal task markedly reduced after pain physiology education: fMRI evaluation of a single patient with chronic low back pain. Australian Journal of Physiotherapy 2005;51:49–52.

Meeus M, Nijs J. Central sensitization: a biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. Clinical Rheumatology 2007;26:465–73.

Meeus M, Nijs J, Van de Wauwer N, Toeback L, Truijen S. Diffuse noxious inhibitory control is delayed in chronic fatigue syndrome: an experimental study. Pain 2008;139:439–48.

Meeus M, Nijs J, Van Oosterwijck J, Van Alsenoy V, Truijen S, De Meirleir K. Pain physiology education improves pain beliefs in patients with chronic fatigue syndrome compared to pacing and self-management education: a double-blind randomized controlled trial. Archives of Physical Medicine and Rehabilitation 2010a;91:1153–9.

Meeus M, Nijs J, Elsemans KS, Truijen S, De Meirleir K. Development and properties of the Dutch neurophysiology of pain test in patients with chronic fatigue syndrome. Journal of Muscoskeletal Pain 2010b;18(1):58–65.

Meyer RA, Campbell JN, Raja SN. Peripheral neural mechanisms of nociception. In: Wall PD, Melzack R, editors. Textbook of pain. 3rd ed. Edinburgh: Churchill Livingstone; 1995. p. 13–44.

Nijs J, Van Houdenhove B, Oostendorp RAB. Recognition of central sensitization in patients with musculoskeletal pain: application of pain neurophysiology in manual therapy practice. Manual Therapy 2010;15:135–41.

Nijs J, Van Houdenhove B. From acute musculoskeletal pain to chronic widespread pain and fibromyalgia: application of pain neurophysiology in manual therapy practice. Manual Therapy 2009;14:3–12.

Nijs J, Van Oosterwijck J, De Hertogh W. Rehabilitation of chronic whiplash: treatment of cervical dysfunctions or chronic pain syndrome? Clinical Rheumatology 2009;28:243–51.

Ryan CG, Gray HG, Newton M, Granat MH. Pain biology education and exercise classes compared to pain biology education alone for individuals with chronic low back pain: a pilot randomised controlled trial. Manual Therapy 2010;15(4):382–7.
Rygh LJ, Tjolsen A, Hole K, Svendsen F. Cellular memory in spinal nociceptive circuitry. Scandinavian Journal of Psychology 2002;43:153–9.
Seifert F, Maihöfner C. Central mechanisms of experimental and chronic neuropathic pain: findings from functional imaging studies. Cellular and molecular life sciences 2009;66:375–90.
Staud R, Craggs JG, Robinson ME, Perlstein WM, Price DD. Brain activity related to temporal summation of C-fiber evoked pain. Pain 2007;129:130–42.
Sullivan MJL. 'Pain Reaction Record' in: Sullivan MJL. Client Workbook Progressive Goal Attainment Program. PDP Program Inc; 2003. p. 100.
Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychological Assessment 1995;7:524–32.
Van Damme S, Crombez G, Bijttebier P, Goubert L, Van Houdenhove B. A confirmatory factor analysis of the Pain Catastrophizing Scale: invariant factor structure across clinical and non clinical populations. Pain 2002;96:319–24.
Van Oosterwijck J, Nijs J, Meeus M, Truijen S, Craps J, Van den Keybus N, et al. Pain neurophysiology education improves cognitions, pain thresholds and movement performance in people with chronic whiplash: a pilot study. Journal of Rehabilitation Research and Development 2011;48(1):43–58.
Van Oosterwijck J, Meeus M, Paul L, Lambrecht L, Nijs J. Does pain physiology education improves diffuse noxious inhibitory control in patients with fibromyalgia? A randomized clinical trial, submitted for publication.
van Wilgen CP, Keizer D. The sensitization model to explain how chronic pain Exists without tissue damage. Pain Management Nursing, in press.
van Wilgen CP, Nijs J. Pijneducatie: een praktische handleiding voor (para)medici. [Pain education: a practical guide for (para)medics.]. The Netherlands: Bohn Stafleu van Loghum: Houten; 2010.
Zusman M. Forebrain-mediated sensitization of central pain pathways: 'non-specific' pain and a new image for MT. Manual Therapy 2002;7:80–8.

Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache

Article *in* Pain · April 1999

DOI: 10.1016/S0304-3959(98)00255-3 · Source: PubMed

3 authors, including:

Stephen Morley
University of Leeds
224 PUBLICATIONS **10,476** CITATIONS
SEE PROFILE

Amanda C de C Williams
University College London
252 PUBLICATIONS **9,739** CITATIONS
SEE PROFILE

PAIN

Pain 80 (1999) 1–13

Review Article

Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache

Stephen Morley[a,*], Christopher Eccleston[b], Amanda Williams[c]

[a]*Division of Psychiatry and Behavioural Sciences, School of Medicine, University of Leeds, Leeds, LS2 9JT, UK*
[b]*Department of Psychology, University of Bath, Bath, UK*
[c]*United Dental and Medical School, University of London, London, UK*

Received 16 March 1998; received in revised form 21 July 1998; accepted 23 October 1998

Abstract

A computer and a hand search of the literature recovered 33 papers from which 25 trials suitable for meta-analysis were identified. We compared the effectiveness of cognitive-behavioural treatments with the waiting list control and alternative treatment control conditions. There was a great diversity of measurements which we grouped into domains representing major facets of pain. Effect sizes, corrected for measurement unreliability, were estimated for each domain. When compared with the waiting list control conditions cognitive-behavioural treatments were associated with significant effect sizes on all domains of measurement (median effect size across domains = 0.5). Comparison with alternative active treatments revealed that cognitive-behavioural treatments produced significantly greater changes for the domains of pain experience, cognitive coping and appraisal (positive coping measures), and reduced behavioural expression of pain. Differences on the following domains were not significant; mood/affect (depression and other, non-depression, measures), cognitive coping and appraisal (negative, e.g. catastrophization), and social role functioning. We conclude that active psychological treatments based on the principle of cognitive behavioural therapy are effective. We discuss the results with reference to the complexity and quality of the trials.

Keywords: Systematic review; Meta-analysis; Chronic pain; Cognitive behavioural therapy

1. Introduction

Behavioural and cognitive treatments for chronic pain have become established in the 30 years since their exposition (Fordyce et al., 1968, 1973; Turk et al., 1983). There are many published open trials of treatment but fewer use control groups in which patients are randomized to treatments. Reviews, however, conclude that there is strong, if not overwhelming evidence for the efficacy of cognitive behavioural therapy (CBT) in restoring function and mood and in reducing pain and disability-related behaviour. Recently, one reviewer regretted that CBT is not provided routinely for chronic pain sufferers rather than medical and physical interventions for which there is less evidence of efficacy (Loeser, 1991). Other overviews of pain management are more critical (Ashburn, 1996). However, to date there has been no systematic review and meta-analysis of randomized controlled trials.

Of the three extant meta-analyses of CBT for chronic pain, one (Malone and Strube, 1988) combined physical and psychological treatment for chronic pain including headache and dental pain; a second (Flor et al., 1992) restricted its scope to psychological treatments and excluded headache; the most recent (Turner, 1996) selected a small sample of randomized controlled trials (RCT) of educational, behavioural and cognitive interventions for chronic low-back pain in the setting of primary care. Both meta-analyses, which included uncontrolled studies, found

* Corresponding author. Tel.: +44-113-233-2733; fax: +44-113-243-3719; e-mail: s.j.morley@leeds.ac.uk

0304-3959/99/$ - see front matter
PII: S0304-3959(98)00255-3

the largest effect sizes for treatment in outcome measures of mood, behaviour and pain ratings, and somewhat smaller ones for drug and health care use. Flor et al. (1992) concluded that: 'overall the results of this meta-analysis provide support for the conclusion that multidisciplinary pain clinics are efficacious. Even at long-term follow-up, patients who are treated in such a setting are functioning better than 75% of a sample that is either untreated or that has been treated by conventional, unimodal treatment approaches' (p. 226). Turner's (1996) findings were consistent with this excepting that the change in mood, in this case depression, was not replicated. This finding may be attributable to a floor effect as patients in her trials were mostly community volunteers and scored low on depression instruments at intake.

In this paper we report a systematic review and meta-analysis of published RCTs of CBT for chronic pain excluding headache. We sought to answer two broad questions: (1) is cognitive behavioural therapy (including behaviour therapy and biofeedback) an effective treatment for chronic pain, i.e. is it 'better' than no treatment? (2) Is cognitive behavioural therapy more effective than alternative active treatments? We chose to exclude headache due to the different emphasis in treatment, both in treatment provision and in outcomes, where pain relief is a much more realistic result of treatment than in other chronic pain. Otherwise, chronic pain was accepted as a label for a heterogeneous group of pain problems in which neither diagnosis, nor site of pain, nor medical findings are an apparent major source of variance in any of the targets of treatment (Turk, 1996; van Tulder et al., 1997). The variety of control conditions found in trials reflects the difficulties in designing suitable controls, e.g. 'inert' controls such as a waiting list can, on ethical and practical grounds, be only short-term and 'active' controls contain an unknown mixture of therapeutic ingredients (O'Leary and Borkovec, 1978; Turner et al., 1994; Schwartz et al., 1997). The comparative treatment groups were, therefore, similarly heterogeneous.

2. Methods

2.1. Search strategy

A search was conducted for published reports of randomized controlled trials of BT and CBT for adults presenting with chronic pain. A priori decisions were made to search only for studies published in full, in peer reviewed journals between 1974 and 1996. Although previous systematic reviews in pain have relied upon Medline (McQuay et al., 1996) it was recognized that the sensitivity of searches using Medline alone has been reported to be low (Adams et al., 1994; Dickersin et al., 1994). Only relevant computer based abstracting services were searched. In order to capture efficiently the maximum number of published trials a three stage plan was chosen (Jadad-Bechara, 1994; NHSCRD, 1996).

A high yield, imprecise, search-term strategy was used. The search strategy contained the word 'pain' and 22 relevant phrases (copy available from authors). Relevant Medline MeSH terms were used (e.g. behaviour therapy). This search strategy had low precision of 0.243% yielding a total of 13 598 articles. Of these 21 were relevant randomized controlled trials and 15 were relevant unrandomized trials.

Four computer abstracting services were selected and their yields compared; Medline, Psychlit, Embase and Social Science Citation Index (SSCI)

Reference lists and bibliographies were searched from all retrieved articles and relevant published reviews. The final list was cross-checked with the PARED database (Jadad et al., 1996a). Twelve additional papers were recovered from searching reference lists. This gave a total sample of 33 papers. Of the 12 that were not found by the search, ten were abstracted on Medline, five on PsychLit, ten on Embase and 12 on SSCI. Of the full set of 33 papers, searching for the specific paper by author and title, Medline has 25 abstracted, Psychlit has 24, Embase has 30 and SSCI also has 30. PARED had recorded 17. Each paper appeared in at least two databases.

SSCI and Embase covered the largest number of journals. Searching for RCTs of psychological therapy in Medline or PsychLit alone did not recover all relevant research reports due, largely, to the omission of specific journals. The 33 papers appeared in 12 journals. Of these four are not regularly abstracted for Medline and three are not regularly abstracted for PsychLit. A three step searching strategy, as employed for this study, is recommended for systematic reviews of psychological therapy.

Papers were read by each of the authors and a consensus decision was taken as to whether the paper contained data suitable for meta-analysis, i.e. contained post treatment means and variances or contrast statistics between two groups (t or F). Where this was not the case we attempted to contact the authors requesting further information about the trial and access to unpublished data. The 33 papers contained data from 30 trials, some papers reported additional or follow-up data. Five trials were excluded from the statistical analysis as the authors were unable to provide data suitable for computing effect size statistics. This left 25 controlled trials for analysis.

2.2. Coding

Development work on the first 20 papers retrieved, which included two papers not entered into the final analysis (Linton and Gotestam, 1984; Linton et al., 1985), generated coding schemes to extract information about the following features of the studies: (1) source of paper (2) the design of study (3) the participants (4) the treatments and (5) the measures and their associated effect sizes (Stock, 1994). Each paper was read to extract data for each coding scheme, i.e. a paper was read five times by each coder during the course of data extraction. Data were extracted by two or

three coders and the reliability of coding was assessed by computing Kappa or percentage agreement for categorical data, and the intraclass correlation for continuous measures. Differences between coders were resolved by consensus. As a large number of features were coded we report coding reliability data, only where necessary. Overall reliability was high.

2.3. Extracting reliability data for study measures

Our choice of meta-analytic strategy (Hunter and Schmidt, 1990) required estimates of the reliability of outcome measures for computing effect size estimates. We generated a list of all the outcome measures used in the studies and sought information about the reliability of each measure. We obtained information from a variety of sources; the study paper, references to measures contained therein, published test manuals, and unpublished data were obtained by contacting authors. In preference we used measures of test stability (test–retest) as the reliability estimate. Where this was not available we used measures of internal consistency (Cronbach's alpha) or inter-rater reliability (Kappa).

2.4. Effect size (ES) computations

We estimated the effect size using Hedges's *g* (Hedges and Olkin, 1985). The sign of the result was adjusted so that improvements on every measure were denoted as positive. Where *g* could not be computed directly from means and standard deviations given in the source paper we computed it indirectly from the available test statistics, e.g. *t*, using the formula of Rosenthal (1994). The estimates of *g* were corrected for small sample bias (Hedges and Olkin, 1985; p.79, Eq. 7) prior to further analysis. In one study (Kerns et al., 1986) outcomes were presented as *z* scores standardized on pre- and post-treatment data. Rather than eliminate the study from consideration we computed a proxy estimate of the effect size by calculating the difference between the *z* scores for the treatment and control groups. This was not corrected for bias as the distribution is not known.

2.5. Analytic rational and methods

We used the meta-analysis psychometric method of Hunter and Schmidt (1990) which assumes that the computed ES is an estimate with an associated error from which a confidence interval can be estimated. Hunter and Schmidt (1990) have provided a series of algorithms for estimating the ES and its associated variance including corrections for variations in the reliability in the dependent variable; which if uncorrected will cause variation in the ES estimate beyond the variation due to sampling error. The analytic strategy was therefore: (1) to estimate ESs correcting for measurement artefact (2) to estimate mean ES over the domain of interest and test the hypothesis that variation is due to statistical artefacts (3). Finally, if the hypothesis that $ES > 0$ cannot be rejected, to investigate the influence of study characteristics (other than those involved in measurement artefact) by disaggregation of the sample into subsamples with shared characteristics. This step is not without problems because as a sample is disaggregated the sizes of the sub-samples may become too small to yield robust estimates.

2.5.1. Comparisons: decisions concerning the multivariate nature of the data

For statistical purposes the ideal meta-analysis would be conducted on a single common measure of interest, e.g. pain intensity, or behavioural activity, extracted from every relevant study. Furthermore, each study would contribute only one effect size derived from a comparison between a single well specified treatment and a control. The current data set met neither of these criteria as most studies had both multiple measures and more than one treatment arm.

2.5.2. Multiple measures

Conducting an analysis in which all outcome measures with computed effect sizes are entered presents problems of bias and independence of measures. A composite effect size may be generated by estimating a mean ES for each study and methods for multivariate solutions of this problem are available (Raudenbusch et al., 1988; Gleser and Olkin, 1994). These methods require information about the correlations between measures and moderately large sample sizes. As these conditions were not uniformly met in the current data set we considered another, statistically simpler, solution.

On the basis of previous reviews and papers (Malone and Strube, 1988; Flor et al., 1992; Gatchel and Turk, 1996) we hypothesized that treatment outcomes would be differentially effective over different measurement domains and conducted separate analysis several domains of measurement. We identified the following domains from our knowledge of the literature and detailed cataloguing of all the measures used in the trials: pain experience; mood/affect; cognitive-coping and appraisal; pain behaviour; social role performance; biological and physical fitness measures; use of health care services; miscellaneous. Definitions of each of these are given in Table 1. The data extraction protocol enabled the assessment of the interrater agreement for assigning measures to domains. This is also given in Table 1. Although data were extracted on use of health care, biological and miscellaneous domains there were too few ESs to merit analysis. These were, therefore, excluded from the comparative analysis of treatment.

As many studies used more than one outcome measure in a given domain. We chose an analytic strategy which selected one measure from each study using the following criteria: select the most frequently occurring measure across studies, e.g. the Beck depression inventory (BDI) in preference to other measures of depression; select multi-item

Table 1

Measurement domains, inter rater agreement for allocating measures to domain and example measures

Domain	% Agree	Example measure
Pain experience: Measures of subjective pain experience captured by ratings of intensity, sensation and unpleasantness	97.5%	McGill pain questionnaire; visual analogue scales of intensity; composite diary measure of numerical ratings
Mood/affect: Primary measure of mood or affective state, but not a trait assessment (these measures were subdivided into assessment of depression and measures of other affective states)	94%	Beck depression inventory; CES-D; STAI-S
Cognitive coping and appraisal: Reports of cognitive strategies and appraisals used in attempts to manage pain (these measures were subdivided into: negative coping, measures known to be correlated with poor adjustment; and positive coping, measures associated with good adjustment)	91%	Coping strategies questionnaire and subscales, e.g. catastrophization, passive coping, active coping
Pain behaviour: Overt behavioural acts associated with pain. There were two subcategories: pain behaviour, referring to behaviour which apparently signals the presence of pain; and activity level, such as distance walked	60%	Pain behaviour, direct observational system; grimacing, guarding, bracing; activity level, distance walked
Biology/physical fitness: Assessment of biological function and physical fitness, but not including measures of behavioural activity as in previous category	85%	Vo_{max}, joint flexibility
Social role functioning: Assessments of the impact of pain on the ability of the person to function in a variety of social roles: work, leisure, marital and family	83%	Sickness impact profile, MPI-Interference, ratings of interference (VAS)
Use of health care system: Use of health care facilities, including clinic visits and drug use	100%	Outpatient medical visits, drug use
Miscellaneous: All other measures not categorized in previous categories	67%	Pain drawings, personality measures, repertory grids

measures in preference to single item (e.g. McGill pain questionnaire (MPQ) in preference to a Visual analogue scale (VAS)), since they are likely to be more reliable; select measures with known reliability coefficients wherever possible.

2.5.3. Multi-armed trials

Many (21 out of 25) studies compared more than one treatment with a control, and there were a variety of control groups used. This presented two issues to be considered: classification and combination of treatment groups and choice of comparison group for estimating ESs.

We considered pooling treatment effect sizes within studies to estimate a study effect size, but rejected this option because there is an expectation in the literature that different treatments may produce different outcomes. We, therefore, estimated overall treatment impact by including all treatment arms within a trial; acknowledging that the ES estimates in this comparison are not independent, as those drawn from a single trial will have a common control condition. We anticipated that further analyses might be possible by estimating the mean ES for treatments with common ingredients. Coding the details of treatments reported in the papers revealed wide variation between treatments described with a generic term, e.g. cognitive therapy, but there was marked variability between studies in the detail provided. We categorized the treatments into three primary types based on a consensus judgement of therapy derived from the source paper. Subcategories for several types were also coded. Details of these are given in Table 2.

We identified two classes of control group. (1) Waiting list control (WLC), where no 'new' treatment was prescribed, although the possibility that WLC patients obtained 'some' treatment, e.g. continued medication, private visits

Table 2

Treatment types

Type	Definition
Biofeedback and relaxation	Use of biofeedback and/or a form of relaxation
Behaviour therapy	Managed approach to behavioural change using the basic concepts and principles of operant psychology
Cognitive behaviour therapy	Primary focus on changing cognitive activity to achieve changes in behaviour, thought and emotion. We identified two broad groupings; coping skills training (CST) and cognitive therapy (CT) CST focuses on inculcating improved cognitively mediated coping skills CT contains CST but with additional component of Beck's cognitive therapy

to other therapists, cannot be excluded. (2) Treatment control (TC), in which a participant was allocated to a 'new' treatment for the duration of the trial. The TC conditions comprised an heterogeneous collection of treatments, including access to regular treatment provided in a pain clinic, physiotherapy, occupational therapy, and the provision of a standard educational and advice package, particularly associated with rheumatoid arthritis (Lorig, 1982). We conducted two main analyses; first a comparison of CBT and BT treatments with WLC, and a second comparison with TC. Studies also used more than one potential control group and as a result of this some studies contributed data to both sets of comparisons, i.e. treatment versus WLC and TC. Table 3 lists the studies, the treatments and the codings we allocated on the basis of which the comparisons were made.

3. Results

We report summary statistics describing the trials entered into the meta-analysis before reporting details of the effect sizes.

3.1. Trial design

Of the 25 trials suitable for meta-analysis only four (16%) provided explicit and replicable information about the method of randomization (Altman, 1996). In the remaining 21 trials the fact that randomization had occurred was simply stated in the methods section or elsewhere in the text or title. Information about the randomization procedure did not include details of whether the randomization was independent of the trialists. Nineteen trials appeared to be true randomized trials and six trials used some form of pseudo-randomization, e.g. by time period. We rated the explicitness of exclusion and inclusion criteria; seven trials gave explicit replicable exclusion criteria while 16 gave explicit inclusion criteria. Only two trials reported a priori power calculations, and four reported post hoc calculations. Eighteen trials used samples of convenience from a specified source, e.g. rehabilitation and pain clinics; two trials recruited consecutive referrals to a clinic, and information was not reported in five trials.

3.2. Participants

Only nine trials reported details of the sample size from which patients were selected, i.e. number of referrals to the trial prior to selection. When all 25 trials are considered 1672 patients were entered into the trials, 38% male, 62% female. The average age (unweighted by number in the trials) was $M = 48.35$ (SD between trials = 7.19), and the mean chronicity of the samples was 12.27 years (SD between trials = 7.47). The average number of patients entered into a trial was 67 (SD = 30.21, range = 18 to 131); the average number of subjects at the end of treatment = 57 (SD = 25.38, range = 18 to 112), giving a crude estimate of drop out rate of 14%. The primary diagnostic labels reported for the patient groups were: chronic low-back pain (36%); rheumatoid arthritis (20%); mixed, predominantly back pain (16%); osteo-arthritis (8%); upper limb pain (8%); fibromyalgia (4%); unspecified (8%).

3.3. Treatments

The modal and median number of treatment arms in the trials was 3 (14/25 trials). There were five trials each with two and four treatment arms, and one trial had six treatments. Treatment was typically delivered in groups (76%), with 20% of treatments delivered as a combination of group and individual therapy. Treatment mode was unspecified in 4% of trials. The mean treatment duration was 6.74 weeks (SD = 2.32, range = 3 to 10 weeks), and the median number of hours in treatment was 16 (range = 6 to 90; interquartile range 10 to 18 h). Sixty percent of therapists were either specifically trained for the trial or were reported as having a general training in CBT and pain. Details were not available in 24% of trials, and the remaining 16% used therapists with general training (i.e. CBT not mentioned). Twenty percent of trials used graduate students (clinical psychologist in training) as therapists; 32% used professionals qualified for more than 5 years; 20% of trials used experienced therapists drawn from several disciplines; and no details were provided in 28% of trials. Eight trials (32%) reported providing regular or some supervision given to therapists during the course of the trial (68% no details given). Only 40% reported making checks on adherence to treatment protocols. Nine trials (36%) reported that treatment was fully manualised; four (16%) referred to partial manualization; and the remaining trials (48%) reported no manualization or a general reference to a text such as Turk et al. (1983). Patients' pre-treatment expectations and the credibility of treatments were assessed in ten trials (40%) but not reported in the remaining trials. In 16 trials (64%) the therapists and treatment delivery were not confounded, i.e. each therapist delivered all treatments. Information on therapists' allegiance to therapeutic mode was not provided.

3.4. Measures

The 25 trials reported total of 221 outcome measures for which effect sizes were computable, an average of 9.21 per trial (SD = 3.59, range = 4 to 16). The majority of these outcomes were patient self-ratings (77.4%); 11% were observations made by a researcher blind to the treatment condition; 6% were made by a non-blinded researcher or therapist; and 5% were made by a spouse or family member. The outcomes were not equally distributed amongst the domains of measurement. The frequencies and percentages are shown in Table 4. The assessment of the use of health care system, biological and fitness indices were relatively

Table 3

Details of studies entered into the meta-analysis. Treatment names as given by the authors and coding used in the study

Authors	Randomization quality	Treatments and code	Patient group, location,sample size and randomization quality
Altmaier et al. (1992)	1	3.4 Psychological treatment + routine care 4.7 Control: routine care	Low-back pain: inpatient, $n = 45$
Appelbaum et al. (1988)	2	3.4 Cognitive behavioural treatment 4.7 Control: symptom monitoring	Rheumatoid arthritis: outpatients, $n = 18$
Bradley et al. (1987)	1	1.3 Cognitive behavioural treatment (BFB + relax + goal setting) 4.7 Structured group social support therapy 4.7 Control: no adjunct treatment	Rheumatoid arthritis: outpatient, $n = 68$
Flor and Birbaumer (1993)	1	1.1 EMG biofeedback 3.4 Cognitive behaviour therapy 4.7 Medical	Chronic musculoskeletal pain (low-back, temporomandibular): outpatients, $n = 78$
Keefe et al. (1990a,b)	1	3.5 Pain coping skills 4.6 Education 4.7 Standard routine care	Osteoarthritis (knee): outpatients, $n = 99$
Keefe et al. (1996)	1	3.5 Spouse assisted coping skills 3.5 Coping skills training 4.6 Education-spouse support	Osteoarthritis (knee): outpatients, $n = 88$
Kerns et al. (1986)	1	3.4 Cognitive behaviour therapy 2;- Behavioural therapy 4.8 Wait list control	Chronic pain – mixed (low back, neck, RSD, PHN, rheumatic disease, musculoskeletal): outpatient, $n = 28$
Kraaimaat et al. (1995)	2	3.4 Cognitive behaviour therapy 4.7 Occupational therapy 4.8 Wait list control	Rheumatoid arthritis: outpatients, $n = 77$
Moore and Chaney (1985)	2	3.4 Cognitive behaviour therapy (couples) 3.4 CBT (individuals) 4.8 Wait list control	Chronic pain – mixed (low-back, arm, knee, phantom limb): outpatients, $n = 43$
Newton-John et al. (1995)	3	1.1 EMG biofeedback 3.5 Cognitive behaviour therapy 4.8 Wait list control	Low-back: outpatient, $n = 44$
Nicholas et al. (1992)	1	3.5 Cognitive behavioural + physiotherapy 4.7 Physiotherapy	Low-back; outpatients, $n = 20$
Nicholas et al. (1991)	1	3.5 Cognitive treatment 3.5 Cognitive treatment + relaxation 2.- Behavioural treatment 2.- Behavioural treatment + relaxation 4.7 Attention control + physiotherapy 4.7 No attention control + physiotherapy	Low-back: outpatient, $n = 58$
O'Leary et al. (1988)	2	3.4 Cognitive behavioural therapy + bibliotherapy 4.7 Bibliotherapy	Rheumatoid arthritis: outpatient, $n = 33$
Peters and Large (1990)	1	3.5 Inpatient pain management 3.4 Outpatient pain management 4.7 Control (standard care allowed)	Chronic pain – mixed (back, arms, head, legs, chest): in- and outpatient, $n = 68$
Puder (1988)	1	3.4 Cognitive behavioural treatment (SIT) 4.8 Waiting list control	Chronic pain – mixed: outpatient, $n = 69$
Radojevic et al. (1992)	1	3.4 Behavioural treatment (CBT) with family support 3.4 Behavioural treatment (CBT) 4.7 Education family support 4.7 No treatment control	Rheumatoid arthritis: outpatient, $n = 59$
Spence (1989,1991)	1	3.5 Individual cognitive behavior therapy 3.5 Group CBT 4.8 Waiting list control	Upper limb (work related): outpatient, $n = 45$
Spence et al. (1995)	1	1.1 EMG biofeedback 1.2 Applied relaxation 1.3 Combined EMG + relaxation 4.8 Waiting list control	Musculoskeletal – cervicobrachial (work related): outpatients, $n = 48$
Turner (1982)	2	3.4 Cognitive behaviour therapy 1.2 Progressive relaxation 4.8 Waiting list control	Low-back pain: outpatients, $n = 36$

Table 3 *continued*

Details of studies entered into the meta-analysis. Treatment names as given by the authors and coding used in the study

Authors	Randomization quality	Treatments and code	Patient group, location,sample size and randomization quality
Turner and Clancy (1988)	1	2.- Operant behavioural 3.4 Cognitive behavioural 4.8 Waiting list control	Low-back pain: outpatient, $n = 81$
Turner and Jensen (1993)	1	1.2 Relaxation 3.5 Cognitive therapy 3.5 Cognitive therapy + relaxation 4.8 Wait list control	Low-back pain: outpatient, $n = 102$
Turner et al. (1990)	1	2.- Behavioural therapy + exercise 2.- Behavioural therapy 4.7 Exercise 4.8 Wait list control	Low-back pain: outpatient: $n = 96$
Vlaeyen et al. (1995)	3	2.- Operant treatment 3.4 Cognitive treatment 1.1 Respondent treatment 4.8 Wait list control	Low-back pain: outpatients, $n = 71$
Vlaeyen et al. (1996)	1	3.4 Combined cognitive/educational 4.6 Attention control (education) 4.8 Waiting list control	Fibromyalgia: outpatients, $n = 131$
Williams et al. (1996)	1	3.5 Inpatient cognitive behavioural 3.5 Outpatient cognitive behavioural 4.8 Waiting list control	Chronic pain – mixed: in- and outpatients, $n = 121$
Linton and Gotestam, (1984)[a]		–	–
Linton et al, (1985)[a]		–	–
Parker et al., (1988)[a]		–	–
Peters et al., (1992)[a]		–	–
Pilowsky et al., (1995)[a]		–	–
Strauss et al, (1986)[a]		–	–

[a]References that were retrieved in the literature search but did not contain useable data.

Randomization quality: 1, random assignment; 2, random assignment of matched pairs or counter balanced by explicit criterion; 3, random assignment compromised. *Treatment*: *first digit*; 1, biofeedback and relaxation (Bfb); 2, behaviour therapy; 3, cognitive therapy; 4, control group. *Second digit*: 1, biofeedback; 2, relaxation; 3, 1 and 2; 4, coping skills training (after Turk et al., 1983); 5, cognitive restructuring (after Beck et al., 1979); 6, education/bibliotherapy; 7, active treatment (treatment as usual – TAU); 8, waiting list control. *Patient group*: *n*, total number treated in the trial.

under-sampled. Table 4 also shows the numbers of trials which sampled each domain and the average number of measures taken per trial.

3.5. Effect sizes

Inspection of Table 4 reveals that three domains, biological, use of health care system, and miscellaneous, were sampled by very few trials. We, therefore, did not compute ESs for these domains. Inspection of the measures used in other domains led us to consider subdividing three of them on the basis of the measures within them. In domain 2 (mood/affect) there was a clear division between measures of depression (BDI, CES-D) and measures of other affective states, predominantly anxiety (STAI-S). Domain 3 (cognitive appraisal and coping) contained measures which might broadly be defined as 'negative', i.e. related to poor adjustment (catastrophizing, passive coping), and 'positive', i.e. related to good adjustment (active coping). This basic conceptual division has been substantiated in the coping literature and we, therefore, conducted separate analyses on the two components. Finally, we noted that in domain 4 (behavioural) the measures broadly tap two components of pain behaviour: (1) the behavioural expression of pain, as indicated by postural adjustments and para-vocalisations, and assessed by measures such as Keefe's pain behaviour observation system (Keefe and Block, 1982) (2) increasing activity levels, usually measured by self report, e.g. MPI-Activity (Kerns et al., 1985). Successful treatment is expected to decrease the overt expression of pain and increase behavioural activity. We therefore, divided the pain behaviour domain to reflect these differences.

3.5.1. Treatment versus waiting list control

Table 5 displays the results for the comparisons between the all treatments and the waiting list control conditions. The left side of the table shows the number of comparisons (*n*) contributing to the estimated ES and the estimates of mean effect size, weighted by the sample sizes of each contributing comparison, and corrected for unreliability in the measurements for each measurement domain. The homogeneity of each set of ESs was also computed with one or two exceptions the application of the Hunter and Schmidt (1990) '75% rule' indicated that the samples

Table 4

Distribution of outcome measures in the 25 trials entered into the meta-analysis and the percentage of trials contributing to each domain. The mean number of measures per trial (column 3) is calculated only for those trials which contributed a measure in the relevant domain

Domain	Number of trials sampling the domain	Mean number of measures per trial	SD	Range
Pain experience	25 (100%)	1.64	1.20	1–7
Mood/affect	22 (88%)	1.64	0.77	1–3
Cognitive coping and appraisal	17 (68%)	2.34	1.16	1–5
Behavioural activity	17 (68%)	1.88	1.23	1–4
Biological	9 (36)%	2.89	2.28	1–9
Social role functioning	19 (76%)	2.00	1.12	1–4
Use of health care system	3 (12%)	1.00	0.00	0
Miscellaneous	5 (20%)	1.40	0.49	1–2

were heterogeneous. We, therefore, decided to report all the data on the assumption that the ESs are heterogeneous[1]. The reported 95% confidence intervals in the both Tables 5 and 6 were calculated on this assumption as are the z values: z values ≥ 1.96 indicate that the mean ES is significantly greater than 0 at the conventional 5% (two-tailed) level, i.e. the null hypothesis that treatment is no more efficacious than the WLC condition is rejected. Without exception this hypothesis was rejected for all measurement domains. The median value of the ES for the measurement domains shown in Table 5 is 0.5, i.e. patients in receipt of treatment are, on average, improved by half a standard deviation relative to those assigned to WLC conditions.

The right side of Table 5 shows the same statistics for sub-groups of treatment types. The null hypothesis tested is that the particular treatment is no more efficacious than the WLC condition: no comparisons between treatment type are made. The number of treatment versus control comparisons, on which each of the ESs is made, is variable with most data being due to the CBT subgroup. CBT is more efficacious than the WLC control conditions for all measurement domains except the expression of pain behaviour. There were relatively fewer comparisons between behaviour therapy and WLC conditions and the estimates of mean ES are based on smaller samples. There were ES > 0 for the domains of pain experience, mood/affect (other than depression), social role functioning (reduced interference) and most markedly for the expression of pain behaviour. The number of comparisons between biofeedback and relaxation treatments and WLC conditions was also small. There were ES > 0 for pain experience, mood/affect (depression), positive and negative coping, and social role functioning.

[1] The decision to regard the ESs as heterogeneous seemed prudent given that most of the analyses indicated heterogeneity, and that where homogeneity was indicated it might have been attributable to the fact that estimates were based on samples in which individual ESs were drawn from the same study. The effect of the assumption is to increase the confidence interval which is tantamount to increasing the probability of a Type II error. We note that in no case, where homogeneity was indicated, did the assumption of heterogeneity change the significance of the result.

All three types of treatment are effective in changing pain experience, i.e. reducing pain intensity, improving social role functioning, and (accepting the single behaviour therapy comparison available) in reducing negative appraisal and coping (predominantly catastrophization).

3.5.2. Treatment versus treatment control

Summary statistics for the comparisons between treatments and active treatment controls are shown in Table 6, which has the same format as Table 5. Altogether there were fewer comparisons between treatments and ATC conditions and the majority of these comprised CBT treatments. When the overall (left side of Table 6) mean ESs are estimated, treatments are reliably more effective (ES > 0) than ATC conditions for the domains of pain experience, cognitive coping and appraisal (increasing positive coping), and pain behaviour (reducing expression of pain). There was no effect of treatment on the other domains, although it should be noted that no data were available to estimate an ES for the increasing activity component of pain behaviour.

When treatment subtypes are considered the results for the largest group, CBT correspond to the findings for the overall estimate. This is not surprising given that CBT contributes most to the overall estimate. The ES estimates for the small number of behaviour therapy comparisons are generally not >0, with two notable exceptions, a reduction in the expression of pain behaviour, and an improvement in social role functioning. The latter is notable since the estimate of the overall ES = 0.

4. Discussion

4.1. Resume

In answer to the two questions addressed by this study, we conclude that active psychological treatments based on the principles of cognitive-behavioural therapy (including behaviour therapy and biofeedback) are effective relative to waiting list control conditions. CBT produced significant changes in measures of pain experience, mood/affect, cog-

Table 5

Effect Sizes for treatments versus waiting list controls. See text for explanation of table. Figures in parenthesis in the penultimate column (95%CI) are the standard errors of a single effect size. CBT, cognitive behavioural therapy, BT, behavior therapy, BFB, biofeedback

Domain	Overall				Sub groups				
	n	Mean ES	95% CI	*z*		*n*	Mean ES	95% CI	*z*
Pain experience	28	0.40	0.22–0.58	4.28	CBT	16	0.33	0.09–0.57	2.78
					BT	5	0.32	–0.09–0.55	2.73
					BFB	7	0.74	0.28–1.20	3.17
Mood/affect depression	24	0.36	0.13–0.59	3.11	CBT	13	0.38	0.07–0.69	2.43
					BT	4	–0.03	–0.21–0.15	–0.33
					BFB	7	0.74	0.28–1.20	3.17
Mood/affect other	16	0.52	0.19–0.84	3.10	CBT	9	0.41	0.00–0.82	1.96
					BT	2	0.74	0.41–1.08	4.34
					BFB	5	0.71	–0.01–1.43	1.94
Cognitive coping and appraisal negative	16	0.50	0.27–0.73	4.20	CBT	8	0.41	0.08–0.73	2.44
					BT	1	1.41	[(±0.41)]	(3.79)
					BFB	7	0.52	0.29–0.76	4.47
Cognitive coping and appraisal positive	11	0.53	0.28–0.78	4.20	CBT	8	0.58	0.28–0.89	3.72
					BT	1	0.56	[(±0.37)]	(1.51)
					BFB	2	0.17	0.03–0.32	2.41
Behaviour expression	12	0.50	0.22–0.78	3.49	CBT	5	0.49	–0.08–1.05	1.68
					BT	5	0.45	0.31–0.59	6.28
					BFB	2	0.71	–0.03–1.45	1.89
Behaviour activity	14	0.46	0.25–0.72	4.34	CBT	7	0.48	0.20–0.77	3.31
					BT	2	0.54	0.28–0.79	4.12
					BFB	5	0.39	–0.03–0.80	1.81
Social role functioning (social role interference)	25	0.60	0.44–0.76	7.28	CBT	15	0.61	0.39–0.84	5.35
					BT	4	0.34	0.17–0.51	3.90
					BFB	6	0.85	0.58–1.31	6.05

nitive coping and appraisal (reduction of negative coping and increase in positive coping), pain behaviour and activity level, and social role function. When compared across the same range of outcomes with other treatments or control conditions, the efficacy of CBT is of a smaller size and limited to the outcomes of pain experience, positive coping and social role function. The overall effect sizes in the order of 0.5, is concordant with those from larger meta-analyses for psychological treatments for a variety of disorders (Shadish et al., 1997), and our conclusion is similar to that reached by Compas et al. (1998) in a narrative review of selected studies.

How might our conclusion be affected by our methods? We identify three areas where we made clear decisions about the treatment of the data: exclusion of unpublished trials, treatment of study measures, and not weighting studies by quality. (1) *Published trials*: the use of only published trials assumes that no unpublished trials would qualify for inclusion, but given the liberality of inclusion criteria in this review, that may well be unfounded. However, there would need to be many such trials, or ones with large samples and effects, to make a significant difference (Chalmers, 1991). Like reviewers in many other fields, we judged this to be unlikely, but the 'amnesty' for trials is wholeheartedly welcomed as they will provide a more satisfactory basis for such judgements. (2) *Measurement*: We attempted to gain control of the variability in the measurements by two means: we corrected for unreliability in the measures, and we grouped measures into reliably defined domains. We recognise that neither of these procedures is perfect. The correction for reliability was dependent on the availability of published coefficients and the degree to which these coefficients may be generalized to the samples is not always known. Our decision to analyze the outcome measures by treatment domains was pragmatic. Clearly investigators expect changes in conceptually distinct areas of measurement, which we believe are reflected in the domains used in this analysis. (3) *Weighting trials*: It is unusual to reject the trial weighting approach in the pain field since the quality of medical trials in pain is known to be associated with the likelihood of finding a positive effect (Jadad et al., 1996b). However, the practice is by no means universally endorsed (Egger and Davey Smith, 1997): the judgements of quality are necessarily subjective and the weightings arbitrary (Thompson and Pocock, 1991; Egger et al., 1997). Although there are excellent guidelines (Altman, 1996), there is no 'gold standard' (Chalmers, 1991). We decided instead to be catholic in our criteria for study inclusion and conservative in the statistical treatment of their results.

4.2. Comment on the quality of trials

Arguably most trials were statistically under powered.

Table 6

Effect sizes for treatments versus treatment controls. Details of the table are given in the text

Domain	Overall					Sub groups			
	n	Mean ES	95% CI	*z*		*n*	Mean ES	95% CI	*z*
Pain experience	22	0.29	0.11–0.46	3.21	CBT	17	0.26	0.05–0.47	2.45
					BT	4	0.33	–0.04–0.71	1.70
					BFB	1	0.52	[(±0.28)]	(1.83)
Mood/affect depression	15	–0.14	–0.32–0.04	–1.52	CBT	11	–0.14	–0.36–0.08	–1.22
					BT	4	–0.14	–0.38–0.11	–1.07
					BFB	0	–	–	–
Mood/affect other	16	0.05	–0.27–0.37	–0.30	CBT	13	0.01	–0.34–0.36	0.05
					BT	2	0.62	–0.55–1.79	1.03
					BFB	1	0.15	[(±0.28)]	(0.54)
Cognitive coping and appraisal (negative)	14	0.17	–0.08–0.42	1.35	CBT	11	0.09	–0.18–0.35	0.63
					BT	2	0.53	–0.16–1.22	1.53
					BFB	1	0.54	[(±0.28)]	(1.92)
Cognitive coping and appraisal (positive)	15	0.40	0.21–0.62	3.60	CBT	12	0.55	0.38–0.72	6.57
					BT	2	–0.13	–0.55–0.29	–0.61
					BFB	1	–0.40	[(±0.28)]	(–1.42)
Behaviour expression	11	0.27	0.08–0.47	2.76	CBT	8	0.31	0.05–0.56	2.36
					BT	2	0.06	0.02–0.10	3.11
					BFB	1	0.34	(±0.28)	(1.20)
Behaviour activity (no data)	–	–	–	–	–	–	–	–	–
Social role functioning (social role interference)	14	0.17	–0.08–0.34	1.62	CBT	10	0.10	–0.15–0.35	0.76
					BT	4	0.37	0.17–0.58	3.61
					BFB	0	–	–	–

This is unsurprising given the demands of delivering a complex multicomponent treatment with sufficient consistency for large numbers of patients over a prolonged period of time. On the other hand, some trials might be regarded as over-complex with multiple treatment and control groups. In future trialists might consider the merits of simple two armed trials with sufficient numbers comparing a treatment with suitable control. The content of and differentiation of control groups from treatment requires careful consideration (Schwartz et al., 1997). Patients assigned to a waiting list in one trial may continue to receive existing treatments (such as physical therapy or pharmacotherapy) which may be equivalent to the treatment control in another trial. Unless expectations of efficacy are monitored, it would be invidious to make assumptions about equivalence in terms of patients' experience. There was also variability within the class of treatment controls from continuing previously ineffective treatments to starting a treatment with demonstrated benefit, such as arthritis education (e.g. Lorig, 1982; Keefe et al., 1996). The distinction between the content of active treatment and control condition can also be a fine one. We surmize that being allocated to a control condition would have different psychological consequences to being allocated to an active treatment, even though that treatment is based on predominantly non-psychological principles, e.g. physical therapy. In addition long-term comparison of treatment and control groups was rendered difficult by the use of waiting list controls. Patients in these groups were commonly entered into an active treatment group or drop out of the trial.

When treatments were considered, across both comparisons (WLC and TC), there was considerable variability in quality and quantity of treatment as reported in the results. While some authors gave explicit accounts of the treatment procedures with reference to manualized interventions which were appropriately monitored, this was not universally so. It is possible that expediency and economy of reporting is a product of external pressures (e.g. editorial demands), but this does not account for what appear to be rather brief interventions delivered by relatively inexperienced therapists to chronically distressed patients for any realistic expectation of change to take place. In addition we note that the measurement of process variables, such as patients' expectations of change, adherence to treatment methods, and therapist competence, were generally lacking. In comparison with best practice in the psychotherapy outcome literature the design and implementation of psychological treatment trials for chronic pain has considerable scope for development (Kazdin, 1994; Lambert and Hill, 1994).

Our analysis of outcome measures revealed a lack in some domains which have economic importance and are of concern to health service managers, third party payers and patients themselves. Data were notably sparse on health service use, drug intake, uptake of additional treatment, and change in work and occupational status as a consequence of

treatment. The over reliance on self report measures is also notable. While many of these measures are psychometrically reliable the extent to which they are influenced by measurement reactivity is often unknown. We note that this feature is not confined to the field of pain (Smith et al., 1980) and while many psychological states can only be measured through self report the development of robust measures of direct observation or independent blind assessors would be beneficial. Kaplan (1990) has argued persuasively for the desirability of behavioural outcomes in health care trials.

4.3. Conclusion

Published randomized controlled trials provide good evidence for the effectiveness of cognitive behavioural therapy and behaviour therapy for chronic pain in adults. This systematic review raised methodological issues which should be considered in the design of future trials. Psychological treatment of chronic pain is complex, lengthy and variable, outcomes cannot be easily dichotomized, and it is rarely possible to blind patients and therapists to treatment conditions. We see the comments, criticisms and questions which arise from our review as a cause for optimism and we hope provide material for debate.

Acknowledgements

This work was partly funded by grants to Stephen Morley from the NHSE Northern and Yorkshire R and D, and by the NHS Health Technology Programme to Chris Eccleston and Amanda Williams. We thank Hannah Betts, Sarah Aldrich and Gill Hancock for support: Patrick McGrath, Steven Linton, Johan Vlaeyen and Frank Keefe for their encouragement and comments on this work. The following provided information on papers; G.D. Strauss, J.C. Parker and Karen Smarr. Finally we thank Henry McQuay and Andrew Moore.

References

Adams, C.E., Power, A., Frederick, K. and Lefebvre, C., An investigation of the adequacy of medline searches for randomized controlled trials (RCTs) of the effects of mental-health-care, Psychol. Med., 24 (1994) 741–748.

Altman, D.G., Better reporting of randomized controlled trials: the CONSORT statement, Br. Med. J., 313 (1996) 570–571.

Altmaier, E.M., Lehmann, T.R., Russell, D.W., Weinstein, J.N. and Kao, C.F., The effectiveness of psychological interventions for the rehabilitation of low-back pain: a randomized controlled trial evaluation, Pain, 49 (1992) 329–335.

Appelbaum, K.A., Blanchard, E.B., Hickling, E.J. and Alfonso, M., Cognitive behavioral treatment of a veteran population with moderate to severe rheumatoid arthritis, Behav. Ther., 19 (1988) 489–502.

Ashburn, M.A., Interdisciplinary pain management: center of excellence or home of the dinosaur?, Clin. J. Pain, 12 (1996) 258–259.

Beck, A.T., Rush, A.J., Shaw, B.F. and Emery, G., Cognitive Therapy of Depression, Guildford Press, New York, 1979.

Chalmers, T.C., Problems induced by meta-analyses, Stat. Med., 10 (1991) 971–980.

Bradley, L.A., Young, L.D., Anderson, K.O., Turner, R.A., Agudelo, C.A., McDaniel, L.K., Pisko, E.J., Semble, E.L. and Morgan, T.M., Effects of psychological therapy on pain behavior of rheumatoid arthritis patients, Treatment outcome and 6-month follow-up, Arth. Rheum., 30 (1987) 1105–1114.

Compas, B.E., Haaga, D.A.F., Keefe, F.J., Leitenberg, H. and Williams, D.A., A sampling of empirically supported treatments from health psychology: smoking, chronic pain, cancer and bulimia nervosa. J. Consult. Clin. Psychol., 66 (1999) 89–112.

Dickersin, K., Scherer, R. and Lefebvre, C., Identifying relevant studies for systematic reviews, Br. Med. J., 309 (1994) 1286–1291.

Egger, M. and Davey Smith, G., Meta-analysis: potentials and promise, Br. Med. J., 315 (1997) 1371–1374.

Egger, M., Davey Smith, G. and Phillips, A., Meta-analysis: principles and procedures, Br. Med. J., 315 (1997) 1533–1537.

Flor, H., Fydrich, T. and Turk, D.C., Efficacy of multidisciplinary pain treatment centers: a meta-analytic review, Pain, 49 (1992) 221–230.

Flor, H. and Birbaumer, N., Comparison of the efficacy of electromyographic biofeedback, cognitive-behavioral therapy, and conservative medical interventions in the treatment of chronic musculoskeletal pain, J. Consult. Clin. Psychol., 61 (1993) 653–658.

Fordyce, W.E., Fowler, R.S., Lehmann, J.F. and DeLateur, B.J., Some implications of learning on problems of chronic pain, J. Chronic Dis., 21 (1968) 179–190.

Fordyce, W.E., Fowler, R.S., Lehmann, J.F., Delateur, B.J., Sand, P.L. and Treischmann, R.B., Operant conditioning in the treatment of chronic pain, Arch. Phys. Med. Rehab., 54 (1973) 399–408.

Gatchel, R.J. and Turk, D.C., Psychological Approaches to Pain Management: A Practitioner's Handbook, Guilford Press, New York, 1996.

Gleser, L.J. and Olkin, I., Stochastically dependent effect sizes. In: H. Cooper and L.V. Hedges (Eds.), The Handbook of Research Synthesis. Russell Sage Foundation, New York, 1994, pp. 339–355.

Hedges, L.V. and Olkin, I. Statistical Methods for Meta-Analysis. Academic Press, Orlando, FL., 1985.

Hunter, J.E. and Schmidt, F.L. Methods of Meta-Analysis. Sage, Beverly Hills, CA., 1990.

Jadad, A.R., Carroll, D., Moore, A., McQuay, H., Developing a database of published reports of randomized clinical trials in pain research, Pain, 66 (1996a) 239–246.

Jadad, A.R., Moore, R.A., Carroll, D., Jenkinson, C., Reynolds, D.J.M., Gavaghan, D.J., McQuay, H.J., Assessing the quality of randomized clinical trials: is blinding necessary? Control. Clin. Trials, 17 (1996b) 1–12.

Jadad-Bechara, A.R., Meta-Analysis of Randomized Clinical Trials in Pain relief, Balliol College, University of Oxford, Oxford, 1994.

Kaplan, R.M., Behavior as the central outcome in health care, Am. Psychol., 45 (1990) 1211–1220.

Kazdin, A.E., Methodology design and evaluation in psychotherapy research. In: A.E. Bergin and S.L. Garfield (Eds.), Handbook of Psychotherapy and Behavior Change (4th edn.), Wiley, New York, 1994, pp. 19–71.

Keefe, F.J. and Block, A.R., Development of an observational method for assessing pain behavior in chronic low-back pain patients, Behav. Ther., 13 (1982) 363–375.

Keefe, F.J., Caldwell, D.S., Williams, D.A., Gil, K.M., Mitchell, D., Robertson, C., Martinez, S., Nunley, J., Beckham, J.C., Crisson, J.E. and Helms, M., Pain coping skills training in the management of osteoarthritic knee pain: a comparative study, Behav. Ther., 21 (1990a) 49–63.

Keefe, F.J., Caldwell, D.S., Williams, D.A., Gil, K.M., Mitchell, D., Robertson, C., Martinez, S., Nunley, J., Beckham, J.C., Crisson, J.E. and Helms, M., Pain coping skills training in the management of

osteoarthritic knee pain – II: follow-up results, Behav. Ther., 21 (1990b) 435–447.
Keefe, F.J., Caldwell, D.S., Baucom, D., Salley, A., Robinson, E., Timmons, K., Beaupre, P., Weisberg, J. and Helms, M., Spouse - assisted coping skills training in the management of osteoarthritic knee pain, Arth. Care Res., 9 (1996) 279–291.
Kerns, R.T., Turk, D.C. and Rudy, T.E., The West Haven–Yale multidimensional pain inventory, Pain, 23 (1985) 345–356.
Kerns, R.D., Turk, D.C., Holzman, A.D. and Rudy, T.E., Comparison of cognitive behavioral and behavioral approaches to the outpatient treatment of chronic pain, Clin. J. Pain, 1 (1986) 195–203.
Kraaimaat, F.W., Brons, M.R., Greenen, R. and Bijlsma, J.W.J., The effect of cognitive behavior therapy in patients with rheumatoid arthritis, Behav. Res. Ther., 33 (1995) 487–495.
Lambert, M.J. and Hill, C.E. Assessing psychotherapy outcomes and process. In: A.E. Bergin and S.L. Garfield (Eds.), Handbook of Psychotherapy and Behavior Change (4th edn.), Wiley, New York, 1994, pp. 72–113.
Linton, S.J. and Gotestam, K.G., A controlled study of the effects of applied relaxation plus operant procedures in the regulation of chronic pain, Br. J. Clin. Psychol., 23 (1984) 291–299.
Linton, S.J., Melin, L. and Stjernlof, K., The effects of applied relaxation and operant activity on chronic pain, Behav. Psychother., 13 (1985) 87–100.
Loeser, J.D., The role of chronic pain clinics in managing back pain. In: J.W. Frymoyer (Ed.), The Adult Spine: Principles and Practice, Raven Press, New York, 1991, pp. 221–229.
Lorig, K., Arthritis Self Management; Leader's Manual, Arthritis Foundation, Atlanta, GA., 1982.
Malone, M.D. and Strube, M.J., Meta-analysis of non-medical treatments for chronic pain, Pain, 34 (1988) 231–244.
McQuay, H., Nye, B.A., Carroll, D., Wiffen, P., Tramer, M. and Moore, A., A systematic review of antidepressants in neuropathic pain, Pain, 68 (1996) 217–227.
Moore, J.E. and Chaney, E.F., Cognitive behavioural treatment of chronic pain: effects of spouse involvement, J.Consult. Clin. Psychol., 53 (1985) 326–334.
Newton-John, T.R.O., Spence, S.H. and Schotte, D., Cognitive-behavioural therapy versus EMG biofeedback in the treatment of chronic low-back pain, Behav. Res. Ther., 33 (1995) 691–697.
NHSCRD, Undertaking Systematic Reviews of Research on Effectiveness: CDR Guidelines for Those Carrying Out or Commissioning Reviews, 4, York, 1996.
Nicholas, M.K., Wilson, P.H. and Goyen, J., Operant-behavioural and cognitive-behavioural treatment for chronic low-back pain, Behav. Res. Ther., 29 (1991) 238–255.
Nicholas, M.K., Wilson, P.H. and Goyen, J., Comparison of cognitive-behavioral group treatment and an alternative non-psychological treatment for chronic low-back pain, Pain, 48 (1992) 339–347.
O'Leary, K.D. and Borkovec, T.D., Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research, Am. Psychol., 33 (1978) 821–830.
O'Leary, A., Shoor, S., Lorig, K. and Holman, H.R., A cognitive behavioral treatment for rheumatoid arthritis, Health Psychol., 7 (1988) 537–544.
Parker, J.C., Frank, R.G., Beck, N.C., Smarr, K.L., Buesher, K.L., Smith, E.I., Anderson, S.K. and Walker, S.E., Pain management in rheumatoid arthritis patients: a cognitive behavioral approach, Arthritis Rheum., 31 (1988) 593–600.
Peters, J.L. and Large, R.G., A randomised control trial evaluating in- and outpatient pain management programmes, Pain, 41 (1990) 283–293.
Peters, J.L., Large, R.G. and Elkind, G., Follow-up results from a randomised controlled trial evaluating in- and outpatient pain management programmes, Pain, 50 (1992) 41–50.
Pilowsky, I., Spence, N., Rounsefell, B., Forsten, C. and Soda, J., Outpatient cognitive-behavioural therapy with amitriptyline for chronic non-malignant pain: a comparative study with 6-month follow-up, Pain, 60 (1995) 49–54.
Puder, R.S., Age analysis of cognitive-behavioral group therapy for chronic pain outpatients, Psychol. Aging, 3 (1988) 204–207.
Radojevic, V., Nicassio, P.M. and Weisman, M.H., Behavioral intervention with and without family support for rheumatoid arthritis, Behav. Ther., 23 (1992) 13–30.
Raudenbusch, S.W., Becker, B.J. and Kalaian, H., Modeling multivariate effect sizes, Psychol. Bull., 103 (1988) 111–120.
Rosenthal, R., Parametric measures of effect size. In: H. Cooper and L.V. Hedges (Eds.), The Handbook of Research Synthesis, Russell Sage Foundation, New York, 1994, pp. 231–244.
Schwartz, C.E., Chesney, M.A., Irvine, M.J. and Keefe, F.J., The control groups dilemma in clinical research: applications for psychosocial and behavioral medicine trials, Psychosomat. Res., 59 (1997) 362–371.
Shadish, W.R., Matt, G.E., Navarro, A.M., Siegle, G., Crits-Cristoph, P., Hazelrigg, M.D., Jorm, A.F., Lyons, L.C., Nietzel, M.T., Prout, H.T., Robinson, L., Smith, M.L., Svartberg, M. and Weiss, B., Evidence that therapy works in clinically representative conditions, J. Consult. Clin. Psychol., 65 (1997) 355–365.
Smith, M.L., Glass, G.V. and Miller, T.I., The Benefits of Psychotherapy. Johns Hopkins University Press, Baltimore, 1980.
Spence, S.H., Cognitive-behavior therapy in the management of chronic, occupational pain of the upper limbs, Behav. Res. Ther., 27 (1989) 435–466.
Spence, S.H., Cognitive-behavior therapy in the treatment of chronic, occupational pain of the upper limbs: a 2 year follow-up, Behav. Res. Ther., 29 (1991) 503–509.
Spence, S.H., Sharpe, L., Newton-John, T. and Champion, D., Effect of EMG biofeedback compared to applied relaxation training with chronic, upper extremity cumulative trauma disorders, Pain, 63 (1995) 199–203.
Stock, W.A., Systematic coding for research synthesis. In: H. Cooper and L.V. Hedges (Eds.), The Handbook of Research Synthesis, Russell Sage Foundation, New York, 1994, pp. 125–138.
Strauss, G.D., Spiegel, J.S., Daniels, M., Spiegel, T., Landsverk, J., Roy-Byrne, P., Edelstein, C., Ehlhardt, J., Falke, R., Hinden, L. and Zackler, L., Group therapies for rheumatoid arthritis: a controlled study of two approaches, Arthritis Rheum., 29 (1986) 1203–1209.
Thompson, S.G. and Pocock, S.J., Can meta-analyses be trusted?, Lancet, 338 (1991) 1127–1130.
Turk, D.C., Meichenbaum, D. and Genest, M., Pain and Behavioral Medicine. Guilford Press, New York, 1983.
Turk, D.C., Biopsychosocial perspective on chronic pain. In: R.J. Gatchel and D.C. Turk (Eds.), Psychological Approaches to Pain Management: A Practitioner's Handbook. Guilford Press, New York, 1996.
Turner, J.A., Comparison of group progressive-relaxation training and cognitive-behavioral group therapy for chronic low-back pain, J. Consult. Clin. Psychol., 50 (1982) 757–765.
Turner, J.A. and Clancy, S., Comparison of operant behavioral and cognitive behavioral group treatment for chronic low-back pain, J. Consult. Clin. Psychol., 56 (1988) 261–266.
Turner, J.A., Clancy, S., McQuade, K.J. and Cardenas, D.D., Effectiveness of behavioral therapy for chronic low-back pain: a component analysis, J. Consult. Clin. Psychol., 58 (1990) 573–579.
Turner, J.A. and Jensen, M.P., Efficacy of cognitive therapy for chronic low-back pain, Pain, 52 (1993) 169–177.
Turner, J.A., Deyo, R.A., Loeser, J.D., Von Korff, M. and Fordyce, W.E., The importance of placebo effects in pain treatment and research, J. Am. Med. Assoc., 271 (1994) 1609–1614.
Turner, J.A., Educational and behavioral interventions for back pain in primary care, Spine, 21 (1996) 2851–2859.
van Tulder, M.W., Assendelft, W.J.J., Koes, B.W. and Bouter, L.M., Spinal radiographic findings and non-specific low-back pain, Spine, 22 (1997) 427–434.
Vlaeyen, J.W.S., Haazen, I.W.C.J., Schuerman, J.A., Kole-Snijders, A.M.J. and van Eek, H., Behavioural rehabilitation of chronic low-back pain: comparison of an operant treatment, an operant cognitive

treatment and an operant-respondent treatment, Br. J. Clin. Psychol., 43 (1995) 95–118.

Vlaeyen, J.W.S., Teeken-Gruben, N.J.G., Goossens, M.E.J.B., Rutten-van Molken, M.M.P.H., Pelt, R.A.G.B., van Eek, H. and Heuts, P.H.T.G., Cognitive-educational treatment of fibromyalgia: a randomised clinical trial. I. Clinical effects, J. Rheumatol., 23 (1996) 1237–1245.

Williams, A.Cde.C., Richardson, P.H., Nicholas, M.K., Harding, V.R., Ridout, K.L., Ralphs, J.A., Richardson, I.H., Justins, D.M. and Chamberlain, J.H., Inpatient versus outpatient pain management: results of a randomised controlled trial, Pain, 66 (1996) 13–22.

J Physiol 595.13 (2017) pp 4141–4150

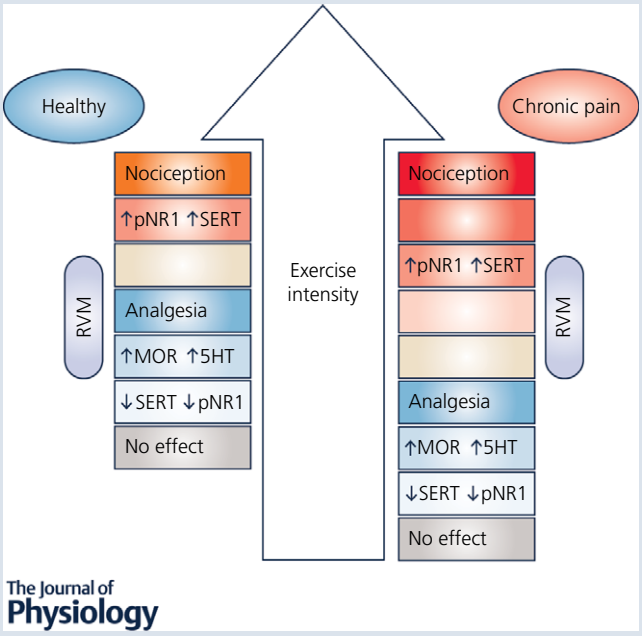
The Journal of Physiology

SYMPOSIUM REVIEW

Does exercise increase or decrease pain? Central mechanisms underlying these two phenomena

Lucas V. Lima, Thiago S. S. Abner and Kathleen A. Sluka

Department of Physical Therapy and Rehabilitation Science, Pain Research Program, University of Iowa, Iowa City, IA 52242, USA



Abstract Exercise is an integral part of the rehabilitation of patients suffering a variety of chronic musculoskeletal conditions, such as fibromyalgia, chronic low back pain and myofascial pain. Regular physical activity is recommended for treatment of chronic pain and its effectiveness has been established in clinical trials for people with a variety of pain conditions. However, exercise can also increase pain making participation in rehabilitation challenging for the person with pain. Animal models of exercise-induced pain have been developed and point to central mechanisms underlying this phenomena, such as increased activation of NMDA receptors in pain-modulating areas. Meanwhile, a variety of basic science studies testing different exercise protocols, show

Lucas Lima and **Thiago Abner** are graduate students in Physical Therapy at the Federal University of Sergipe, and did a research fellowship with Kathleen Sluka at the University of Iowa. **Kathleen Sluka** is a professor in the Department of Physical Therapy and Rehabilitation Science at the University of Iowa. She received a physical therapy degree from Georgia State University and a PhD in Anatomy from the University of Texas Medical Branch in Galveston. After a postdoctoral fellowship with Dr William D. Willis, she joined the faculty at the University of Iowa. Her research focuses on the neurobiology of musculoskeletal pain as well as the mechanisms and effectiveness of non-pharmacological pain treatments commonly used by physical therapists. She has published over 180 peer-reviewed manuscripts, numerous book chapters, and a textbook on Pain Mechanisms and Management for the Physical Therapist.

This review was presented at the symposium "Top-down control of pain", which took place at Physiology 2016, Dublin, Ireland, 29-31 July 2016.

DOI: 10.1113/JP273355

exercise-induced analgesia involves activation of central inhibitory pathways. Opioid, serotonin and NMDA mechanisms acting in rostral ventromedial medulla promote analgesia associated with exercise. This review explores and discusses current evidence on central mechanisms underlying exercised-induced pain and analgesia.

(Received 2 December 2016; accepted after revision 7 March 2017; first published online 29 March 2017)
Corresponding author K. A. Sluka: Department of Physical Therapy and Rehabilitation Science, 1-252 MEB, University of Iowa, Iowa City, IA 52242, USA. Email: Kathleen-sluka@uiowa.edu

Abstract figure legend Summary of the mechanisms in the rostral ventromedial medulla underlying the nociceptive and analgesic effects as exercise intensity increases. Differences between healthy and chronic pain patients is demonstrated by the lower intensity level of exercise necessary to produce both analgesia and nociception in the chronic pain patients.

Abbreviations 5-HT, serotonin; NMDA, *N*-methyl-D-aspartate; NRM, nucleus raphe magnus; NRO, nucleus raphe obscurus; NRP, nucleus raphe pallidus; PAG, periaqueductal grey; RVM, rostral ventromedial medulla.

Introduction

Exercise not only reduces pain perception, but also has effects on mental health, such as mood elevation and reduction of stress and depression, which are often associated with chronic pain conditions (Bement & Sluka, 2016). Exercise is a powerful tool in the management of those conditions, especially considering the Centers for Disease Control and Prevention's new opioid-prescribing guidelines, recommending a focus toward non-opioid and non-pharmacological treatments (Dowell *et al.* 2016). In healthy subjects, exercise increases thresholds for experimentally induced pain (Bement & Sluka, 2016). In clinical populations, exercise promotes analgesia in conditions such as low back pain, osteoarthritis, myofascial pain, chronic fatigue syndrome and fibromyalgia (Bement & Sluka, 2016). However, exercise has also been shown to increase pain in experimental and clinical settings, especially when a musculoskeletal pain condition is already established (Staud *et al.* 2005). Patients with fibromyalgia show greater increases in pain and perceived fatigue after performing a physically fatiguing task when compared to healthy subjects (Dailey *et al.* 2015). This increased pain to exercise in chronic pain patients is often a barrier to regular exercise, leading to a sedentary lifestyle that worsens the painful conditions and makes treatment even more difficult (Damsgard *et al.* 2010). Interestingly, contraction of painful muscles fails to activate pain inhibitory mechanisms in myalgia and fibromyalgia patients while it increases pressure pain thresholds in healthy subjects (Lannersten & Kosek, 2010). Exercise is, in most cases, one of the best approaches for managing chronic pain conditions, so understanding the mechanisms of both pain and analgesia induced by exercise is important to better define physical activity-related treatment protocols for people with pain.

Centrally, the rostral ventromedial medulla (RVM) is a key relay for pain modulation, playing a major role in exercise-induced pain and analgesia (Sluka & Rasmussen, 2010; Stagg *et al.* 2011; Sluka *et al.* 2012, 2013). Within the caudal brainstem, the nucleus raphe magnus (NRM), nucleus raphe obscurus (NRO) and nucleus raphe pallidus (NRP) are involved in modulation of both pain and motor outputs (Fields *et al.* 1995; Porreca *et al.* 2002; Zhuo *et al.* 2002; Da Silva *et al.* 2010*a*), making these nuclei potential links between physical activity and pain perception. Other pain-processing areas such as the periaqueductal grey (PAG) (Mathes & Kanarek, 2006; Stagg *et al.* 2011) and cortical areas (de Oliveira *et al.* 2010) have been implicated in exercise-induced pain and analgesia. *N*-Methyl-D-aspartate (NMDA) glutamate receptors in the RVM also play a key role in chronic muscle pain, including exercise-induced pain (Da Silva *et al.* 2010*a*; Sluka *et al.* 2012). Phosphorylation of the NR1 subunits of NMDA receptors in the caudal brainstem mediates the hyperalgesia in animal models of chronic musculoskeletal pain and exercise-induced pain (Sluka *et al.* 2012). On the other hand, opioidergic and serotonergic neurons are both expressed in the RVM (Basbaum & Fields, 1984) and there is recent evidence for the involvement of these systems in the analgesia induced by exercise (Stagg *et al.* 2011; Bobinski *et al.* 2015). Figure 1 illustrates the known mechanisms of exercise-induced pain and analgesia.

This review discusses animal studies that explore the underlying central mechanisms of both exercise induced pain and analgesia from different exercise protocols. We discuss the evidence with respect to type, duration, and frequency of exercise using different pain models.

Fatiguing exercise enhances pain

Pain and fatigue interactions. Clinically, physical fatigue is a common complaint in chronic musculoskeletal pain conditions, while chronic pain is common in chronic fatigue conditions (Vierck *et al.* 2001; Whiteside *et al.* 2004; Staud *et al.* 2005; Kadetoff & Kosek, 2007). The overlap

between muscle fatigue and pain syndromes suggests an interaction between fatigue and pain such that fatigue may enhance pain. Pain may be a factor in reducing adherence to regular exercise and rehabilitation, leading the patient to a sedentary life (Damsgard *et al.* 2010). It is proposed that muscle fatigue promotes changes in central nervous system function that cannot be explained only in the muscle itself (Davis & Bailey, 1997).

Fatiguing exercise-induced pain models. Several animal models were developed to better understand the interaction between muscle fatigue and pain. For example, when an acute bout of running wheel activity (2 h) was combined with intramuscular doses of saline of different pH (pH 4.0, 5.0, 6.0 or 7.2), enhanced hyperalgesia developed bilaterally when the pH 5.0 injections were combined with the fatigue task – no cutaneous hyperalgesia developed with pH 5.0 injections without fatigue (Yokoyama *et al.* 2007). In the initial studies, two 2 h runs prior to the first intramuscular pH 5.0 injection, and two 2 h runs prior to the second intramuscular pH 5.0 injection of acid saline produced an enhanced muscle hyperalgesia. Subsequently it was shown that a single 2 h or 30 min run prior to the subthreshold muscle insult produced the same widespread hyperalgesia (Sluka *et al.* 2012). Despite a 10% reduction in grip force after the 2 h fatiguing exercise, there were no changes in muscle P_{CO_2} ,P_{O_2}, lactate, creatinine kinase MB and phosphate suggesting minimal fatigue metabolites were released during the fatiguing task. These results show that muscle fatigue enhances the probability of the development of mechanical hyperalgesia in mice in response to intramuscular acid saline without muscle histological changes.

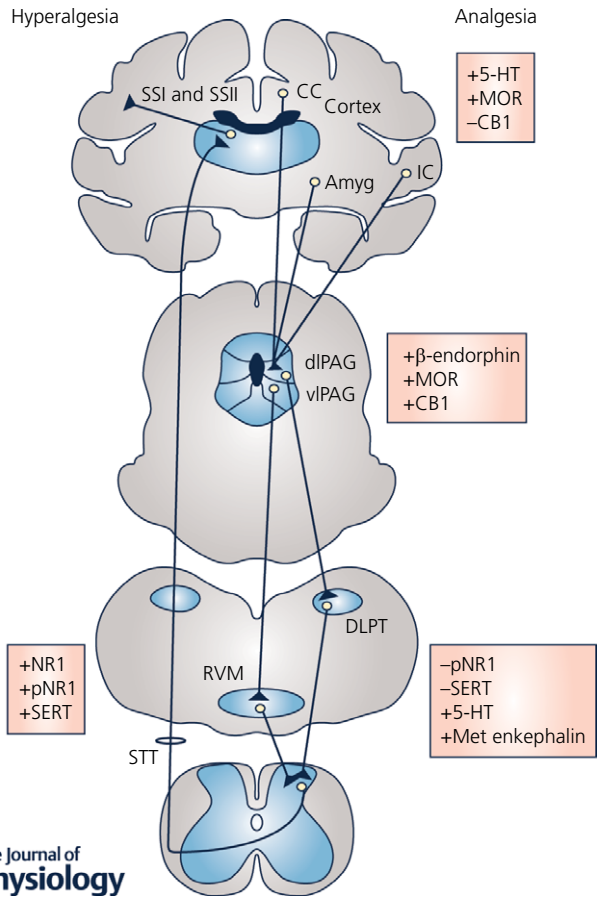


Figure 1. Overview of the underlying mechanisms of exercise-induced pain and analgesia
Known neurotransmitters and receptors that have been shown to be involved at different areas of the central nervous system are listed. The majority of studies have focused on the PAG and the RVM. Increases in serotonin and opioids, and activation of μ-opioid (MOR) and cannabinoid-1 (CB1) receptors are implicated in the exercise-induced analgesia. Further, the normally increased phosphorylation of the NR1 subunit of the NMDA receptor and the increased expression of serotonin transporter (SERT) that is increased by acute exercise are reduced by regular physical activity. +, increase; −, decrease; 5-HT, serotonin; CB1, cannabinoid receptor 1; DH, dorsal horn; MOR, μ-opioid receptor; PAG, periaqueductal grey; p-NR1, phosphorylated NR1; RVM, rostral ventromedial medulla.

Similarly, combining an acute bout of running wheel exercise with a low dose of intramuscular carrageenan injection (0.03%) produced widespread mechanical hyperalgesia. Interestingly, injection of carrageenan either 2 h before or 2 h after the fatigue task produced the same degree of mechanical hyperalgesia of the paw, but not the muscle (Sluka & Rasmussen, 2010). There was also an enhanced hyperalgesia in female mice that was eliminated by ovariectomy, suggesting oestradiol contributed to the development of exercise-induced hyperalgesia in this model.

To test if localized fatigue of the injected muscle was sufficient to induce the hyperalgesia, electrical stimulation of the muscle replaced the whole-body fatiguing task. When combining this electrically induced isometric contraction with pH 5.0 injections there was a significant hyperalgesia that developed in the ipsilateral muscle of male mice and bilaterally in the female mice (Gregory *et al.* 2013). Interestingly, the hyperalgesia was longer lasting and easier to induce in female mice. Hyperalgesia lasted for 2 weeks in males and over 1 month in females. Temporally separating the fatigue task and the muscle insult by 24 h resulted in bilateral hyperalgesia only in female mice, which suggests that the attenuation in response to muscle fatigue does not occur in females. Spatially separating the fatigue task and muscle insult by giving the fatigue task in the muscle contralateral to the injection also resulted in bilateral hyperalgesia only in female mice. In this case, ovariectomy had no effect on the sex differences suggesting oestradiol was not involved in the development of exercise-induced hyperalgesia in this model. It may be that the isometric fatiguing task favours a peripheral mechanism that results in release of fatigue metabolites

like acidic pH in muscle that subsequently activate acid sensing ion channels (ASICs). Indeed, we showed that blockade of ASIC3 prevents and ASIC3 knockout mice do not develop hyperalgesia in this localized fatigue-induced pain model (Gregory *et al.* 2016).

In summary, the studies described above show that fatiguing exercise can enhance hyperalgesia in both male and female mice, and this enhancement is greater in females. Interestingly, whole body exercise produced the female phenotype through oestradiol while the localized fatigue exercise task produces the enhancement in an oestradiol-independent manner. This highlights the complicated nature of nociceptive processing in males and females and suggests that there are task-dependent mechanisms involved in the enhancement of hyperalgesia by exercise.

Central mechanisms of fatiguing exercise-induced hyperalgesia. To examine potential brain sites that underlie exercise-induced hyperalgesia, C-fos immunostaining, as a marker of neuron activation, in the caudal brainstem was investigated. C-fos immunoreactivity showed an increase in the number of cells in the NRM, NRO and NRP after a 2 h running-wheel task, suggesting the caudal raphe might be involved in the development of exercise-induced hyperalgesia (Sluka *et al.* 2012). Since NMDA receptors in the RVM are involved in pain facilitation (Sluka & Rasmussen, 2010), NMDA receptors were blocked in the NRO/NRP during the fatiguing task when combined with 0.03% carrageenan. NMDA receptor blockade during the fatiguing task prevented the development of exercise-induced hyperalgesia. On the other hand, over-expression of the NR1 subunit of the NMDA receptor in the RVM, using a feline immunodeficiency virus expressing the complementary DNA to NR1, produced bilateral mechanical hyperalgesia of the paw and muscle (Da Silva *et al.* 2010*b*), supporting a role for NR1 in development of hyperalgesia. Since phosphorylation of NMDA receptors can enhance neuron excitability (Chen & Roche, 2007), the expression of the phosphorylated NR1 subunit was investigated. In the exercise-induced pain model induced by whole-body running wheel activity combined with 0.03% carrageenan or pH 5.0 injections, there was an increase in the number of cells stained for phosphorylated NR1 in the NRO, NRM, and NRP (Sluka *et al.* 2012; Lima *et al.* 2016). However, there were no differences in the number of p-NR1 labelled cells in the electrically stimulated fatigue task combined with two pH 5.0 injections (Gregory *et al.* 2013), suggesting different mechanisms in this model. Thus, NMDA receptor activation and phosphorylation of NMDA receptors underlies the development of hyperalgesia from a whole-body fatiguing task, but not from a localized fatigue task.

Exercise-induced analgesia

Mechanistic studies in human subjects. Exercise-induced analgesia and the underlying mechanisms have been investigated in several studies using healthy control human subjects and more recently in patient populations. Early studies show that high intensity running, or bicycle ergometry produced analgesia that was reversed by systemic naloxone, suggesting the involvement of opioids in exercise-induced analgesia (Janal *et al.* 1984; Olausson *et al.* 1986). Using a fatiguing isometric contraction, there were decreases in pain thresholds that were accompanied by a reduction in cortical excitability and motor evoked potentials assessed by transcranial magnetic stimulation (Bement *et al.* 2009). High levels of physical activity correlate with greater conditioned pain modulation, which is thought to measure central inhibition, in healthy controls (Geva & Defrin, 2013). Conditioned pain modulation is higher in athletes (Flood *et al.* 2017), and predicts exercise-induced analgesia in healthy subjects (Ellingson *et al.* 2014; Lemley *et al.* 2015; Stolzman & Bement, 2016). In people with osteoarthritis, there were significant increases in pressure pain thresholds in those with normal conditioned pain modulation, and decreases in pressure pain thresholds in those with reduced conditioned pain modulation, suggesting exercise and conditioned pain modulation use similar mechanisms (Fingleton *et al.* 2017). Further, both conditioned pain modulation and exercise-induced hypoalgesia predict greater pain relief 6 months after total knee replacement (Vaegter *et al.* 2017). Lastly, several studies show a reduction in temporal summation, a measure of central excitability, in healthy subjects and patient populations following aerobic and isometric exercise protocols (Koltyn *et al.* 2013; Henriksen *et al.* 2014; Naugle & Riley, 2014; Lemley *et al.* 2015; Stolzman & Bement, 2016; Vaegter *et al.* 2017). Thus, in human subjects there is evidence to support modulation of central nervous system function with enhanced inhibition and reduced excitation. A number of chronic pain conditions are associated with a loss of conditioned pain modulation and increased temporal summation, and thus lack of immediate effects of exercise, or even increases in pain with acute exercise, could be explained by this lack of inhibition and enhanced excitability. It is further likely that repeated regular exercise could restore the loss of conditioned pain modulation.

Animal models of exercise-induced analgesia. The first evidence of centrally mediated mechanisms came from animal studies using swimming as the exercise stimulus in healthy, non-injured rodents (Cooper & Carmody, 1982; Girardot & Holloway, 1984; Koltyn, 2000). Different protocols have been tested, testing different water temperatures and exercise durations (3–10 min). Although longer exercise protocols and colder water

temperatures seemed to produce a stronger analgesic effect (Cooper & Carmody, 1982; O'Connor & Chipkin, 1984), swimming interventions as short as 15 s and in warm water promoted increases in pain thresholds that were at least partially reversed by the opioid antagonist naloxone (Cooper & Carmody, 1982). These studies in healthy animals performed a single bout of the exercise task to produce analgesia. Similar results were found in the formalin model, where as little as 3 min of swimming with a single bout of exercise produced a reduction of pain behaviours that was reversed by naloxone (Carmody & Cooper, 1987; Kuphal *et al.* 2007). Since most studies showed that opioid antagonists only partially reversed exercise-induced analgesia, especially when lower temperatures and longer exercise times were used (Cooper & Carmody, 1982; Girardot & Holloway, 1984; Terman *et al.* 1986), it seems that other mechanisms could be involved, but also conditions other than exercise itself might have influenced the results, like changes in body temperature and stress (Koltyn, 2000).

Forced treadmill running in rodents has also been studied as an exercise stimulus and it excludes the temperature bias from swimming protocols. In a neuropathic pain model, 5 weeks of treadmill running with different frequencies (3 or 5 days week^{-1}) and intensities (10 or 16 m min^{-1} speeds) reversed the injury-induced hyperalgesia in an intensity- but not frequency-dependent manner (Stagg *et al.* 2011). A 5-day treadmill (15–30 min day^{-1}) protocol found similar results in a chronic muscle pain model, with reduction in bilateral mechanical hyperalgesia occurring as soon as immediately after the first session (Bement & Sluka, 2005). In both studies, the effects of exercise were reversed by administration of opioid antagonists, showing evidence of opioid mechanisms underlying the observed exercise-induced analgesia.

While treadmill running allows one to control the degree of physical activity each animal performs, it can produce a stress component (Contarteze *et al.* 2008), which itself could produce analgesia through activation of endogenous opioid and serotonergic systems (Yesilyurt *et al.* 2015), and thus confound interpretation of the results. One way to avoid this is by using running wheels placed in the animals' home cages. Rodents voluntarily exercise in running wheels in a consistent manner (Sherwin, 1998). Recent studies used running wheels to investigate exercise-induced analgesia (Smith & Yancey, 2003; Sluka *et al.* 2013; Grace *et al.* 2016; Leung *et al.* 2016) to isolate the effects of exercise from the influence of other stimuli. Different durations of running wheel activity, ranging from 5 consecutive days to 8 weeks and performed before or after the insult have been tested in different models, such as non-inflammatory chronic muscle pain (Sluka *et al.* 2013), exercise-induced pain (Sluka *et al.* 2013), acute inflammatory muscle pain (Sluka *et al.* 2013), neuropathic pain (Grace *et al.* 2016) and healthy control animals (Kanarek *et al.* 1998; Mathes & Kanarek, 2006). These studies showed the efficacy of running wheel activity in producing analgesia in healthy non-injured animals, but more importantly, in preventing and reversing hyperalgesia in different pain models. There is a duration-dependent effect. Importantly, in the studies investigating different pain models, the running wheels were removed from the cages at the time of induction of the model, and thus these studies compared physically active animals to physically inactive animals. Five days of wheel running prevents secondary, but not primary hyperalgesia, in the exercised-induced pain model and has no effect on hyperalgesia in a chronic non-inflammatory muscle pain model. On the other hand, 6–8 weeks of physical activity prevents both primary and secondary hyperalgesia in an exercise-induced pain model, a chronic non-inflammatory muscle pain model and a neuropathic pain model (Sluka *et al.* 2013; Grace *et al.* 2016), but not in an acute inflammatory pain model (Sluka *et al.* 2013). Further, 2 weeks of voluntary wheel running was unable to reverse hyperalgesia in mouse models of neuropathic pain and formalin-induced acute pain (Sheahan *et al.* 2015), but longer duration wheel running (6 weeks) successfully prevented and reversed hyperalgesia from a neuropathic pain model (Grace *et al.* 2016). Table 1 summarizes the exercise protocols used in animal studies. Thus, multiple different protocols have been used to produce analgesia in uninjured animals and in multiple pain models. These include swimming, treadmill exercise, and wheel running with a single bout of exercise producing analgesia to multiple days and weeks. The analgesic effects depend on duration (days or weeks), with longer training protocols producing more significant results. Further, while protocols applied after the injury can reverse the hyperalgesia, intriguingly making animals physically active prior to the insult prevents the development of the hyperalgesia in both neuropathic pain and muscle pain models.

Central mechanisms involved in exercise-induced analgesia. The RVM comprises, with the PAG and dorsal horn, a descending pain inhibitory system that both facilitates and inhibits noxious stimuli (Porreca *et al.* 2002). Within the RVM, NRM, NRO and NRP are nuclei known to be involved in pain modulation but are also involved in modulation of motor responses, making them potential key areas involved in exercise-induced analgesia mechanisms (Fields *et al.* 2006). Three types of cells exist in the RVM: ON-cells promote nociception when activated, OFF-cells inhibit nociception when activated, and neutral cells do not respond to noxious stimuli (Fields *et al.* 2006). We propose that a shift in the balance between ON- and OFF-cell activation defines hyperalgesia or analgesia from an exercise task. As discussed previously,

Table 1. Summary of studies examining exercise-induced analgesia

Pain model	Exercise intervention	Duration	Effect	Study
Healthy animals	Swimming (15 s to 7.5 min)	Single bout	Reduction in thermal hyperalgesia	Cooper & Carmody (1982)
	Cold water swim 3.5 min, 2°C	Single bout	Reduction in thermal hyperalgesia	Girardot & Holloway (1984)
	Swimming 3 min warm water, 2 min cold water	Single bout	Reduction in thermal hyperalgesia	O'Connor & Chipkin (1984)
	Voluntary wheel running	6 weeks	Reduction in thermal hyperalgesia	Smith & Yancey (2003)
		20 days	Increase in thermal hyperalgesia	Kanarek *et al.* (1998)
		24 h	No effect	
		3 weeks	Reduction in thermal hyperalgesia	Mathes & Kanarek (2006)
	Resistance exercise, 3 sets of 10 repetitions, 3 times week^{-1}	12 weeks	No effect	Galdino *et al.* (2010)
	Resistance exercise, 15 sets of 15 repetitions	1 day	Reduction in mechanical hyperalgesia	Galdino *et al.* (2014*a*)
	Treadmill running, 20 m min^{-1} speed, until fatigue (average of 49.06 $\pm$ 3 min)	Single bout	Reduction in thermal and mechanical hyperalgesia	Galdino *et al.* (2014*b*)
	Cold water swimming, 5 min, 1°C	Single bout	Reduction in thermal hyperalgesia	Vaswani *et al.* (1988)
Formalin test	Voluntary wheel running	2 weeks	No effect	Sheahan *et al.* (2015)
	Cold water swimming	Single bout	Reduction in pain scores and thermal hyperalgesia	Terman *et al.* (1986)
	Swimming 3 min warm water	Single bout	Reduction in pain scores	Carmody & Cooper (1987)
	Swimming 37°C water for 90 min day^{-1}	9 days	Reduction in pain scores	Kuphal *et al.* (2007)
Neuropathic pain	Swimming	18–25 days	Reduction in pain ratings, cold allodynia and thermal hyperalgesia	
	Treadmill running, 3 or 5 days week^{-1}, 16 m min^{-1} speed	5 weeks	Reduction in mechanical hyperalgesia	Stagg *et al.* (2011)
	Voluntary wheel running	6 weeks	Prevention of allodynia	Grace *et al.* (2016)
	Treadmill running, 30 min, 5 days week^{-1}, 10 m min^{-1} speed	2 weeks	Reduction in mechanical hyperalgesia	Bobinski *et al.* (2015)
	Voluntary wheel running	2 weeks	No effect	Sheahan *et al.* (2015)
	30 min day^{-1}, 5 days a week, 8–20 m min^{-1} speed	4 weeks	Reduction in thermal hyperalgesia and mechanical allodynia	Kim *et al.* (2015)
	Treadmill running, 20–60 min day^{-1}, 9 m min^{-1} speed	4 weeks	Reduction in mechanical allodynia	Korb *et al.* (2010)
Chronic muscle pain model	Treadmill running, 15–30 min day^{-1}, 6–10 m min^{-1} speed	5 days	Reduction in mechanical hyperalgesia	Bement & Sluka (2005)
	Voluntary wheel running	8 weeks	Prevention of primary and secondary hyperalgesia	Sluka *et al.* (2013)
		5 days	No effect	
		8 weeks	Prevention of primary and secondary hyperalgesia	Leung *et al.* (2016)
Exercise-enhanced pain	Voluntary wheel running	5 days	Prevention of secondary hyperalgesia	Lima *et al.* (2016)
		8 weeks	Prevention of secondary hyperalgesia	Sluka *et al.* (2013)
		5 days		
Acute muscle inflammation	Voluntary wheel running	8 weeks	No effect	Sluka *et al.* (2013)
		5 days		
Chemically induced nociception	Swimming, 30 min day^{-1}	5 days	Reduction in pain behaviours (abdominal constrictions)	Mazzardo-Martins *et al.* (2010)
	Swimming, 10–30 min day^{-1}	2 weeks	Reduction in pain behaviours (paw-licking) and mechanical hyperalgesia	Martins *et al.* (2017)

NMDA receptors in the RVM play a role in facilitation of nociception with an increase in phosphorylation of the NR1 subunit playing a critical role (Da Silva *et al.* 2010*a*, *b*; Sluka *et al.* 2012). Exercise-induced analgesia promotes the opposite response. Either 5 days or 8 weeks of wheel running prevented the increase in phosphorylation of NR1 in the RVM of mice induced with chronic non-inflammatory muscle pain or exercise-enhanced pain when compared to induced sedentary mice (Sluka *et al.* 2013). These data suggest that regular physical activity reduces facilitation in the caudal brainstem by modulating NMDA receptor function.

There is strong evidence that opioid mechanisms mediate exercise-induced analgesia in both human and animal studies (Koltyn, 2000). Several studies showed that the opioid antagonist naloxone, given systemically, blocks the analgesic effects of swimming and resistance exercise in healthy, uninjured animals (Cooper & Carmody, 1982; O'Connor & Chipkin, 1984; Galdino *et al.* 2010; Mazzardo-Martins *et al.* 2010; Martins *et al.* 2017), and treadmill running in chronic muscle pain (5-day running) and neuropathic pain models (5-weeks running) (Bement & Sluka, 2005; Stagg *et al.* 2011). Subsequent studies show that supraspinal naloxone blocks the analgesia produced by 5 weeks of treadmill running in a neuropathic pain model (Stagg *et al.* 2011). Further, there are increased concentrations of endogenous opioids systemically in both human subjects and in animals (Wildmann *et al.* 1986; Vaswani *et al.* 1988; Debruille *et al.* 1999; Stagg *et al.* 2011; Bidari *et al.* 2016), in the PAG and RVM in animals (Commons, 2003; Stagg *et al.* 2011; Kim *et al.* 2015), and increased μ-opioid receptor expression in the hippocampus of rats after both acute (7 days) and chronic (45 days) treadmill or wheel running (de Oliveira *et al.* 2010). Further, 4–6 weeks of voluntary wheel running produces cross-tolerance to μ-opioid agonists and physical dependence, effects similar to those resulting from chronic use of opioids (Kanarek *et al.* 1998; Smith & Yancey, 2003) and 3 weeks of wheel running attenuates the analgesia from morphine injected into the PAG of rats (Mathes & Kanarek, 2006). Thus, regular physical activity and exercise use central opioid receptors to produce analgesia.

Serotonin (5-HT) has also been implicated in exercise-induced analgesia. One hour of swimming increases 5-HT levels in the brainstem and hypothalamus, while 4 weeks of swimming extended this increase to the cerebral cortex (Dey *et al.* 1992). Similarly, 8 weeks of treadmill running showed increased levels of 5-HT in the midbrain and cortex (Brown *et al.* 1979), and 4 weeks of treadmill running increases 5-HT expression in the RVM (Korb *et al.* 2010). More recently, we extended these studies by examining the role of serotonin in a neuropathic pain model. We show that 2 weeks of low-intensity treadmill running in a neuropathic pain model increased 5-HT levels in the caudal brainstem, decreased expression of the serotonin transporter in the NRM, NRO and NRP, and altered serotonin receptor expression in the brainstem (Bobinski *et al.* 2015). Importantly, in neuropathic pain models there is an increase in serotonin transporter expression and a decrease in 5-HT in the brainstem; 2 weeks of treadmill running reversed these injury-induced changes. Further, systemic depletion of serotonin prevents the analgesia produced by treadmill running in neuropathic pain (Bobinski *et al.* 2015) and by high intensity swimming (30 min to 5 days) in the acetic acid writhing test (Mazzardo-Martins *et al.* 2010). Thus, there is emerging evidence that increases in supraspinal serotonin release, along with reductions in the serotonin transporter, play a significant role in the analgesia produced by regular exercise.

There are reasons to believe that the opioid and serotonergic mechanisms are not independently activated by exercise, but rather they interact to promote analgesia. Serotonergic neurons receive input from endogenous opioid peptides and both coexist in RVM neurons (Fields *et al.* 2006). Further evidence of this interaction is shown by blockade of analgesia from systemic or RVM-injected morphine following systemic depletion of serotonin, or blockade of serotonin receptors in the RVM (Schul & Frenk, 1991; Carruba *et al.* 1992). We recently tested this hypothesis by performing immunohistochemistry for serotonin transporter in μ-opioid receptor knockout mice induced with exercise-induced pain and comparing these to wild-type mice (Lima *et al.* 2016). μ-Opioid receptor knockout and wild-type mice were exposed to 5 days of wheel-running prior to induction the exercise-induced pain model, and compared to sedentary mice. Wheel running prevented the increase in the serotonin transporter in the RVM induced by the muscle insult in wild-type mice. However, in μ-opioid receptor knockout mice, wheel running had no effect on the increased serotonin transporter expression induced by muscle insult. Thus, these data suggest that μ-opioid receptor activation by exercise reduces expression of the serotonin transporter in the caudal brainstem to promote analgesia.

Endocannabinoids in the central nervous system also play a role in exercise-induced analgesia (Dietrich & McDaniel, 2004). Endocannabinoid receptors are present in pain-modulating areas of the brain and spinal cord (Herkenham *et al.* 1991) and activation of endocannabinoid receptors produces analgesia (Dietrich & McDaniel, 2004). Further, exercise increases circulating levels of the endocannabinoid *N*-arachidonylethanolamine in healthy human subjects (Koltyn *et al.* 2014). After both aerobic and resistance exercise tasks, there is an increased expression of the cannabinoid receptor CB_1 in the brain, including the PAG, in healthy uninjured animals. This effect is prevented by systemic and central blockade with cannabinoid

receptor antagonists (AM251 and AM630) (Galdino *et al.* 2014*a*, *b*). Since endocannabinoids have synergistic interactions with opioids to produce antinociception (Navarro *et al.* 1998), one could speculate that the same interaction occurs during exercise-induced analgesia. Thus, there is emerging evidence that endogenous endocannabinoids in the central nervous system contribute to the analgesia produced by regular exercise.

Conclusion

A single bout of fatiguing exercise in the presence of a chronic pain condition can exacerbate pain that is characterized by increased phosphorylation of NMDA receptors in the RVM, suggesting enhanced central facilitation. On the other hand, regular exercise promotes pain relief and is characterized by reduced NMDA receptor phosphorylation, suggesting reduced central facilitation. Further regular exercise reduces serotonin transporter expression, increases serotonin levels, and increases opioids in central inhibitory pathways including the PAG and RVM, suggesting exercise utilizes our endogenous inhibitory systems to reduce pain (Fig. 1). We propose that there is a balance between inhibition and excitation in the central nervous system that determines whether exercise will promote analgesia or promote pain. Several factors, such as fitness level, physical activity levels, and state of the injury or pain condition influence this balance. The great majority of the animal studies examining pain mechanisms are performed in physically inactive animals, and nearly all the exercise studies are focused on aerobic exercise. Further, there is no consistency regarding intensity, duration, frequency or exercise type making interpretation difficult. Understanding the mechanisms underlying different forms of exercise, as well as the different intensities and duration of exercise that produce analgesia, will be critically important to translate animal studies to human subjects, particularly those with acute and chronic pain.

References

Basbaum AI & Fields HL (1984). Endogenous pain control systems: brainstem spinal pathways and endorphin circuitry. *Ann Rev Neuroscience* **7**, 309–338.

Bement MKH & Sluka KA (2005). Low-intensity exercise reverses chronic muscle pain in the rat in a naloxone-dependent manner. *Arch Phys Med Rehabil* **86**, 1736–1740.

Bement MKH & Sluka KA (2016). Exercise-induced analgesia: an evidence-based review. In *Mechanisms and Management of Pain for the Physical Therapist*, 2nd edn, ed. Sluka KA, Ch. 10. pp. 177–201. Wolters Kuwer, IASP Press, Seattle.

Bement MKH, Weyer A, Hartley S, Yoon T & Hunter SK (2009). Fatiguing exercise attenuates pain-induced corticomotor excitability. *Neurosci Lett* **452**, 209–213.

Bidari A, Ghavidel-Parsa B, Rajabi S, Sanaei O & Toutounchi M (2016). The acute effect of maximal exercise on plasma beta-endorphin levels in fibromyalgia patients. *Korean J Pain* **29**, 249–254.

Bobinski F, Ferreira TA, Córdova MM, Dombrowski PA, da Cunha C, do Espírito Santo CC, Poli A, Pires RG, Martins-Silva C & Sluka KA (2015). Role of brainstem serotonin in analgesia produced by low-intensity exercise on neuropathic pain after sciatic nerve injury in mice. *Pain* **156**, 2595–2606.

Brown BS, Payne T, Kim C, Moore G, Krebs P & Martin W (1979). Chronic response of rat brain norepinephrine and serotonin levels to endurance training. *J Appl Physiol* **46**, 19–23.

Carmody J & Cooper K (1987). Swim stress reduces chronic pain in mice through an opioid mechanism. *Neurosci Lett* **74**, 358–363.

Carruba MO, Nisoli E, Garosi V, Sacerdote P, Panerai AE & Da Prada M (1992). Catecholamine and serotonin depletion from rat spinal cord: effects on morphine and footshock induced analgesia. *Pharmacol Res* **25**, 187–194.

Chen B-S & Roche KW (2007). Regulation of NMDA receptors by phosphorylation. *Neuropharmacology* **53**, 362–368.

Commons KG (2003). Translocation of presynaptic delta opioid receptors in the ventrolateral periaqueductal gray after swim stress. *J Comp Neurol* **464**, 197–207.

Contarteze RVL, Manchado FDB, Gobatto CA & De Mello MAR (2008). Stress biomarkers in rats submitted to swimming and treadmill running exercises. *Comp Biochem Physiol A Mol Integr Physiol* **151**, 415–422.

Cooper K & Carmody J (1982). The characteristics of the opioid-related analgesia induced by the stress of swimming in the mouse. *Neurosci Lett* **31**, 165–170.

Dailey DL, Keffala VJ & Sluka KA (2015). Do cognitive and physical fatigue tasks enhance pain, cognitive fatigue, and physical fatigue in people with fibromyalgia? *Arthritis Care Res* **67**, 288–296.

Damsgard E, Thrane G, Anke A, Fors T & Røe C (2010). Activity-related pain in patients with chronic musculoskeletal disorders. *Disabil Rehabil* **32**, 1428–1437.

Da Silva LF, DeSantana JM & Sluka KA (2010*a*). Activation of NMDA receptors in the brainstem, rostral ventromedial medulla, and nucleus reticularis gigantocellularis mediates mechanical hyperalgesia produced by repeated intramuscular injections of acidic saline in rats. *J Pain* **11**, 378–387.

Da Silva LFS, Walder RY, Davidson BL, Wilson SP & Sluka KA (2010*b*). Changes in expression of NMDA-NR1 receptor subunits in the rostral ventromedial medulla modulate pain behaviors. *Pain* **151**, 155–161.

Davis JM & Bailey SP (1997). Possible mechanisms of central nervous system fatigue during exercise. *Med Sci Sports Exerc* **29**, 45–57.

Debruille C, Luyckx M, Ballester L, Brunet C, Odou P, Dine T, Gressier B, Cayin M & Cayin J (1999). Serum opioid activity after physical exercise in rats. *Physiol Res* **48**, 129–134.

de Oliveira MSR, da Silva Fernandes MJ, Scorza FA, Persike DS, Scorza CA, da Ponte JB, de Albuquerque M, Cavalheiro EA & Arida RM (2010). Acute and chronic exercise modulates the expression of MOR opioid receptors in the hippocampal formation of rats. *Brain Res Bull* **83**, 278–283.

Dey S, Singh R & Dey P (1992). Exercise training: significance of regional alterations in serotonin metabolism of rat brain in relation to antidepressant effect of exercise. *Physiol Behav* **52**, 1095–1099.

Dietrich A & McDaniel W (2004). Endocannabinoids and exercise. *Br J Sports Med* **38**, 536–541.

Dowell D, Haegerich TM & Chou R (2016). CDC guideline for prescribing opioids for chronic pain—United States, 2016. *J Am Med Assoc* **315**, 1624–1645.

Ellingson LD, Koltyn KF, Kim JS & Cook DB (2014). Does exercise induce hypoalgesia through conditioned pain modulation? *Psychophysiology* **51**, 267–276.

Fields H, Malick A & Burstein R (1995). Dorsal horn projection targets of ON and OFF cells in the rostral ventromedial medulla. *J Neurophysiol* **74**, 1742–1759.

Fingleton C, Smart K, Doody C & Dip T (2017). Exercise-induced hypoalgesia in people with knee osteoarthritis with normal and abnormal conditioned pain modulation. *Clin J Pain* **33**, 395–404.

Flood A, Waddington G, Thompson K & Cathcart S (2017). Increased conditioned pain modulation in athletes. *J Sports Sci* **35**, 1066–1072.

Galdino G, Duarte I & Perez A (2010). Participation of endogenous opioids in the antinociception induced by resistance exercise in rats. *Braz J Med Biol Res* **43**, 906–909.

Galdino G, Romero T, da Silva JFP, Aguiar D, de Paula AM, Cruz J, Parrella C, Piscitelli F, Duarte I & Di Marzo V (2014*a*). Acute resistance exercise induces antinociception by activation of the endocannabinoid system in rats. *Anesth Analg* **119**, 702.

Galdino G, Romero TR, Silva JFP, Aguiar DC, de Paula AM, Cruz JS, Parrella C, Piscitelli F, Duarte ID & Di Marzo V (2014*b*). The endocannabinoid system mediates aerobic exercise-induced antinociception in rats. *Neuropharmacology* **77**, 313–324.

Geva N & Defrin R (2013). Enhanced pain modulation among triathletes: A possible explanation for their exceptional capabilities. *Pain* **154**, 2317–2323.

Girardot M-N & Holloway FA (1984). Cold water stress analgesia in rats: Differential effects of naltrexone. *Physiol Behav* **32**, 547–555.

Grace PM, Fabisiak TJ, Green-Fulgham SM, Anderson ND, Strand KA, Kwilasz AJ, Galer EL, Walker FR, Greenwood BN & Maier SF (2016). Prior voluntary wheel running attenuates neuropathic pain. *Pain* **157**, 2012–2023.

Gregory NS, Brito RG, Fusaro MC, Sluka KA (2016). ASIC3 is required for development of fatigue-induced hyperalgesia. *Mol Neurobiol* **53**, 1020–1030.

Gregory NS, Gibson-Corley K, Frey-Law L & Sluka KA (2013). Fatigue-enhanced hyperalgesia in response to muscle insult: induction and development occur in a sex-dependent manner. *Pain* **154**, 2668–2676.

Henriksen M, Klokker L, Graven-Nielsen T, Bartholdy C, Schjødt Jørgensen T, Bandak E, Danneskiold-Samsøe B, Christensen R & Bliddal H (2014). Association of exercise therapy and reduction of pain sensitivity in patients with knee osteoarthritis: a randomized controlled trial. *Arthritis Care Res* **66**, 1836–1843.

Herkenham M, Lynn AB, Johnson MR, Melvin LS, de Costa BR & Rice KC (1991). Characterization and localization of cannabinoid receptors in rat brain: a quantitative in vitro autoradiographic study. *J Neurosci* **11**, 563–583.

Janal MN, Colt EW, Clark WC & Glusman M (1984). Pain sensitivity, mood and plasma endocrine levels in man following long-distance running: effects of naloxone. *Pain* **19**, 13–25.

Kadetoff D & Kosek E (2007). The effects of static muscular contraction on blood pressure, heart rate, pain ratings and pressure pain thresholds in healthy individuals and patients with fibromyalgia. *Eur J Pain* **11**, 39–47.

Kanarek RB, Gerstein AV, Wildman RP, Mathes WF & D'Anci KE (1998). Chronic running-wheel activity decreases sensitivity to morphine-induced analgesia in male and female rats. *Pharmacol Biochem Behav* **61**, 19–27.

Kim Y-J, Byun J-H & Choi I-S (2015). Effect of exercise on μ-opioid receptor expression in the rostral ventromedial medulla in neuropathic pain rat model. *Ann Rehabil Med* **39**, 331–339.

Koltyn KF (2000). Analgesia following exercise. *Sports Med* **29**, 85–98.

Koltyn KF, Brellenthin AG, Cook DB, Sehgal N & Hillard C (2014). Mechanisms of exercise-induced hypoalgesia. *J Pain* **15**, 1294–1304.

Koltyn KF, Knauf MT & Brellenthin AG (2013). Temporal summation of heat pain modulated by isometric exercise. *Eur J Pain* **17**, 1005–1011.

Korb A, Bonetti LV, Da Silva SA, Marcuzzo S, Ilha J, Bertagnolli M, Partata WA & Faccioni-Heuser MC (2010). Effect of treadmill exercise on serotonin immunoreactivity in medullary raphe nuclei and spinal cord following sciatic nerve transection in rats. *Neurochem Res* **35**, 380–389.

Kuphal KE, Fibuch EE & Taylor BK (2007). Extended swimming exercise reduces inflammatory and peripheral neuropathic pain in rodents. *J Pain* **8**, 989–997.

Lannersten L & Kosek E (2010). Dysfunction of endogenous pain inhibition during exercise with painful muscles in patients with shoulder myalgia and fibromyalgia. *Pain* **151**, 77–86.

Lemley KJ, Hunter SK & Bement M (2015). Conditioned pain modulation predicts exercise-induced hypoalgesia in healthy adults. *Med Sci Sports Exerc* **47**, 176–184.

Leung A, Gregory NS, Allen L-AH & Sluka KA (2016). Regular physical activity prevents chronic pain by altering resident muscle macrophage phenotype and increasing interleukin-10 in mice. *Pain* **157**, 70–79.

Lima L, DeSantana J, Rasmussen L & Sluka K (2016). Short-duration physical activity prevents the development of exercise-enhanced hyperalgesia through opioid mechanisms. *J Pain* **17**, S95.

Martins DF, Siteneski A, Ludtke DD, Dal-Secco D & Santos AR (2017). High-intensity swimming exercise decreases glutamate-induced nociception by activation of G-protein-coupled receptors inhibiting phosphorylated protein kinase A. *Mol Neurobiol* (in press; https://doi.org/10.1007/s12035-016-0095-9).

Mathes WF & Kanarek RB (2006). Chronic running wheel activity attenuates the antinociceptive actions of morphine and morphine-6-glucouronide administration into the periaqueductal gray in rats. *Pharmacol Biochem Behav* **83**, 578–584.

Mazzardo-Martins L, Martins DF, Marcon R, dos Santos UD, Speckhann B, Gadotti VM, Sigwalt AR, Guglielmo LGA & Santos ARS (2010). High-intensity extended swimming exercise reduces pain-related behavior in mice: involvement of endogenous opioids and the serotonergic system. *J Pain* **11**, 1384–1393.

Naugle KM & Riley JL (2014). Self-reported physical activity predicts pain inhibitory and facilitatory function. *Med Sci Sports Exerc* **46**, 622–629.

Navarro M, Chowen J, Carrera MRA, del Arco I, Villanúa MA, Martin Y, Roberts AJ, Koob GF & de Fonseca FR (1998). CB1 cannabinoid receptor antagonist-induced opiate withdrawal in morphine-dependent rats. *Neuroreport* **9**, 3397–3402.

O'Connor P & Chipkin RE (1984). Comparisons between warm and cold water swim stress in mice. *Life Sci* **35**, 631–639.

Olausson B, Eriksson E, Ellmarker L, Rydenhag B, Shyu BC & Andersson S (1986). Effects of naloxone on dental pain threshold following muscle exercise and low frequency transcutaneous nerve stimulation: a comparative study in man. *Acta Physiol Scand* **126**, 299–305.

Porreca F, Ossipov MH & Gebhart G (2002). Chronic pain and medullary descending facilitation. *Trends Neurosci* **25**, 319–325.

Schul R & Frenk H (1991). The role of serotonin in analgesia elicited by morphine in the periaqueductal gray matter (PAG). *Brain Res* **553**, 353–357.

Sheahan TD, Copits BA, Golden JP & Gereau IV RW (2015). Voluntary exercise training: analysis of mice in uninjured, inflammatory, and nerve-injured pain states. *PLoS One* **10**, e0133191.

Sherwin C (1998). Voluntary wheel running: a review and novel interpretation. *Anim Behav* **56**, 11–27.

Sluka KA, Danielson J, Rasmussen L & Dasilva LF (2012). Exercise-induced pain requires NMDA receptor activation in the medullary raphe nuclei. *Med Sci Sports Exerc* **44**, 420–427.

Sluka KA, O'Donnell JM, Danielson J & Rasmussen LA (2013). Regular physical activity prevents development of chronic pain and activation of central neurons. *J Appl Physiol* **114**, 725–733.

Sluka KA & Rasmussen LA (2010). Fatiguing exercise enhances hyperalgesia to muscle inflammation. *Pain* **148**, 188–197.

Smith MA & Yancey DL (2003). Sensitivity to the effects of opioids in rats with free access to exercise wheels: μ-opioid tolerance and physical dependence. *Psychopharmacology (Berl)* **168**, 426–434.

Stagg NJ, Mata HP, Ibrahim MM, Henriksen EJ, Porreca F, Vanderah TW & Malan TP (2011). Regular exercise reverses sensory hypersensitivity in a rat neuropathic pain modelrole of endogenous opioids. *Anesthesiol* **114**, 940–948.

Staud R, Robinson ME & Price DD (2005). Isometric exercise has opposite effects on central pain mechanisms in fibromyalgia patients compared to normal controls. *Pain* **118**, 176–184.

Stolzman S & Bement MH (2016). Does exercise decrease pain via conditioned pain modulation in adolescents? *Pediatr Phys Ther* **28**, 470–473.

Terman GW, Morgan MJ & Liebeskind JC (1986). Opioid and non-opioid stress analgesia from cold water swim: importance of stress severity. *Brain Res* **372**, 167–171.

Vaegter HB, Handberg G, Emmeluth C & Graven-Nielsen T (2017). Preoperative hypoalgesia after cold pressor test and aerobic exercise is associated with pain relief six months after total knee replacement. *Clin J Pain* (in press; https://doi.org/10.1097/AJP.0000000000000428).

Vaswani KK, Richard CW & Tejwani GA (1988). Cold swim stress-induced changes in the levels of opioid peptides in the rat CNS and peripheral tissues. *Pharmacol Biochem Behav* **29**, 163–168.

Vierck CJ, Staud R, Price DD, Cannon RL, Mauderli AP & Martin AD (2001). The effect of maximal exercise on temporal summation of second pain (windup) in patients with fibromyalgia syndrome. *J Pain* **2**, 334–344.

Whiteside A, Hansen S & Chaudhuri A (2004). Exercise lowers pain threshold in chronic fatigue syndrome. *Pain* **109**, 497–499.

Wildmann J, Krüger A, Schmole M, Niemann J & Matthaei H (1986). Increase of circulating beta-endorphin-like immunoreactivity correlates with the change in feeling of pleasantness after running. *Life Sci* **38**, 997–1003.

Yesilyurt O, Seyrek M, Tasdemir S, Kahraman S, Deveci MS, Karakus E, Halici Z & Dogrul A (2015). The critical role of spinal 5-HT 7 receptors in opioid and non-opioid type stress-induced analgesia. *Eur J Pharmacol* **762**, 402–410.

Yokoyama T, Lisi TL, Moore SA & Sluka KA (2007). Muscle fatigue increases the probability of developing hyperalgesia in mice. *J Pain* **8**, 692–699.

Zhuo M, Sengupta J & Gebhart G (2002). Biphasic modulation of spinal visceral nociceptive transmission from the rostroventral medial medulla in the rat. *J Neurophysiol* **87**, 2225–2236.

Additional information

Competing interests

All authors declare no conflict of interest

Author contributions

L.V.L: designed, wrote and reviewed the manuscript; T.S.S.A.: wrote and reviewed the manuscript; K.A.S.: designed, wrote and reviewed the manuscript. All authors contributed to the writing of the manuscript and approved with the final version. All designated authors qualify for authorship, and all those who qualify for authorship are listed.

Funding

Funded by NIH grants R01 AR061371 and UM1 AR06338.

PLOS ONE

RESEARCH ARTICLE

Effective treatment options for musculoskeletal pain in primary care: A systematic overview of current evidence

Opeyemi O. Babatunde*, Joanne L. Jordan☯, Danielle A. Van der Windt☯, Jonathan C. Hill‡, Nadine E. Foster‡, Joanne Protheroe‡

Arthritis Research UK Primary Care Centre, Research Institute for Primary Care & Health Sciences, Keele University, Keele, United Kingdom

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* o.babatunde@keele.ac.uk

Citation: Babatunde OO, Jordan JL, Van der Windt DA, Hill JC, Foster NE, Protheroe J (2017) Effective treatment options for musculoskeletal pain in primary care: A systematic overview of current evidence. PLoS ONE 12(6): e0178621. https://doi.org/10.1371/journal.pone.0178621

Editor: Johannes Fleckenstein, University of Bern, SWITZERLAND

Received: October 20, 2016

Accepted: May 16, 2017

Published: June 22, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This manuscript presents independent research funded by the NIHR under its Programme Grants for Applied Research scheme (grant number: RP-PG-1211-20010) and a Centre of Excellence grant from the Arthritis Research UK (grant number: 20202). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute

Abstract

Background & aims

Musculoskeletal pain, the most common cause of disability globally, is most frequently managed in primary care. People with musculoskeletal pain in different body regions share similar characteristics, prognosis, and may respond to similar treatments. This overview aims to summarise current best evidence on currently available treatment options for the five most common musculoskeletal pain presentations (back, neck, shoulder, knee and multi-site pain) in primary care.

Methods

A systematic search was conducted. Initial searches identified clinical guidelines, clinical pathways and systematic reviews. Additional searches found recently published trials and those addressing gaps in the evidence base. Data on study populations, interventions, and outcomes of intervention on pain and function were extracted. Quality of systematic reviews was assessed using AMSTAR, and strength of evidence rated using a modified GRADE approach.

Results

Moderate to strong evidence suggests that exercise therapy and psychosocial interventions are effective for relieving pain and improving function for musculoskeletal pain. NSAIDs and opioids reduce pain in the short-term, but the effect size is modest and the potential for adverse effects need careful consideration. Corticosteroid injections were found to be beneficial for short-term pain relief among patients with knee and shoulder pain. However, current evidence remains equivocal on optimal dose, intensity and frequency, or mode of application for most treatment options.

for Health Research or the Department of Health. NF, an NIHR Senior Investigator, and JH are supported through an NIHR Research Professorship awarded to NF (NIHR-RP-011-015).

Competing interests: The authors have declared that no competing interests exist.

Conclusion

This review presents a comprehensive summary and critical assessment of current evidence for the treatment of pain presentations in primary care. The evidence synthesis of interventions for common musculoskeletal pain presentations shows moderate-strong evidence for exercise therapy and psychosocial interventions, with short-term benefits only from pharmacological treatments. Future research into optimal dose and application of the most promising treatments is needed.

Introduction

Pain as a result of musculoskeletal problems of the back, neck, shoulder, knee and multi-site pain is an increasing cause of diminished quality of life, and increased demands on healthcare [1–3]. Prognosis is often poor with many people reporting persistent symptoms 6 to 12 months after consulting their primary care practitioner [4, 5]. Furthermore, the likelihood of persistent or recurrent clinical symptoms may accentuate the physical, psychological, and socio-economic impacts of musculoskeletal pain.

Musculoskeletal pain is managed by a plethora of treatment options, most delivered in primary care by first contact clinicians such as general practitioners, physiotherapists, chiropractors and osteopaths. These include non-pharmacological treatments (e.g. self-management advice and education, exercise therapy, manual therapy and psychosocial interventions), complementary therapies (e.g. acupuncture), and pharmacological interventions (e.g. analgesics, non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroid injections). For those with refractory symptoms, surgical interventions (e.g. arthroscopic debridement, total knee replacements, and laminectomies) may be considered. However, for the overarching aim of reducing pain and improving function, recommendations are equivocal in respect to the effectiveness of various treatment options that are used across a range of common musculoskeletal pain presentations. For example, evidence for the effectiveness of corticosteroid injections for relief of shoulder or knee pain is inconsistent [6, 7]. Similarly, the efficacy and safety of simple analgesics and NSAIDs for reducing symptoms associated with osteoarthritis and back pain is uncertain [8–11]. In order to provide optimal care to patients with musculoskeletal pain and ensure the efficient use of healthcare resources, a comprehensive overview of the available evidence for the most effective treatment options for musculoskeletal pain presentations is essential.

Evidence from trials and systematic reviews indicate that most treatments for musculoskeletal pain provide small to moderate short-term benefits, with a lack of evidence for long-term effectiveness [12]. Also, there appears to be a wide heterogeneity in the response of patient symptoms to treatments, suggesting that some patients may benefit more from some treatments than others [12]. Due to an apparent lack of information on the comparative effectiveness of available treatment options, there is a need to summarise current evidence regarding the best treatments for musculoskeletal pain presentations.

Previous reviews and guidelines that describe the effectiveness of treatments for musculoskeletal pain specifically focus on single regional pain sites, such as shoulder pain [13, 14], knee pain [15, 16] or low back pain [17–20]. However, research evidence suggests that in the general population and those presenting to primary care, localised musculoskeletal pain frequently coexists in more than one body region [21, 22] and that those with different regional pains share similar underlying attributes, course of symptoms and prognostic factors [23, 24]. Nevertheless, for many patients, clinical decision-making regarding treatment is often

focussed on the specific body region without much recourse to the potential influence of prognostic factors or other co-existing pain problems. As a result, informed choices about which treatment might work best for which individual also remain a substantial clinical challenge. A more holistic view is perhaps difficult to obtain since trials and systematic reviews usually focus on a specific musculoskeletal pain site, comparing only two or three treatment options. To our knowledge there are no published reviews in which evidence regarding the comparative effectiveness of a wide range of treatments is systematically synthesised for the most common musculoskeletal pain presentations.

The aim of this study was to critically appraise current best evidence regarding the effectiveness of treatments to reduce pain and /or improve function for people with the five most common musculoskeletal pain presentations in primary care (i.e., back, neck, shoulder, knee and multi-site pain as indicated by Jordan et al.[25]). The specific objectives of this review were to:

1. identify effective treatment options for the five most common musculoskeletal pain presentations and
2. highlight gaps in evidence and priorities for policy or future research.

The review also identified, where available, evidence regarding patient subgroups most likely to respond to different treatment options.

Methods

Sources of data and search strategy

Integrated information from higher levels of evidence has been suggested as an "ideal source of evidence for clinical decision-making" by the Evidence Based Practice group (http://hsl.mcmaster.libguides.com/ebm). Therefore, using national clinical guidelines, policy documents, care pathways such as Map of Medicine (MoM), and clinical evidence summaries as a starting point, the search for evidence for this overview followed a pyramidal tract through a hierarchy of available evidence. Sources of evidence for the overview included: Clinical Knowledge Summaries, Map of Medicine, TRIP Database (systematic reviews and clinical guidelines), the Cochrane Library (including Cochrane database of systematic reviews, Database of abstracts of reviews of effects, Health technology assessment database), MEDLINE and EMBASE (using specific search filters to retrieve systematic reviews and clinical guidelines), reference lists of included systematic reviews and guidelines, research stakeholders and experts in the field of musculoskeletal research. Evidence sources were initially accessed in January 2014 and regularly checked for new updates at eight week intervals through to March 2015 whilst the review was ongoing. A Cochrane library search update was conducted in February and August 2016 in order to identify newly published Cochrane reviews.

All Cochrane reviews matching the inclusion criteria were included in the synthesis. Relevant non-Cochrane systematic reviews were added where there were no (up-to-date) Cochrane reviews summarising the effectiveness of a particular treatment. Additional searches of the bibliographic databases, MEDLINE and EMBASE (using narrow or specific search filters to retrieve systematic reviews and clinical guidelines) were carried out to identify and retrieve (1) relevant systematic reviews, and (2) more recently published relevant RCTs that had not yet been summarised in reviews or guidelines or where evidence gaps clearly existed. For the bibliographic database searches, retrieved search results were limited to published articles from 2000 until December 2014 initially, and then updated in August 2016. The search strategy and search terms for these additional searches are profiled in supplementary S1 File.

Relevant publications (guidelines, systematic reviews and meta-analyses of RCTs as well as recent RCTs which are yet to be summarised in reviews) were obtained and assessed against predefined eligibility criteria according to the study protocol by two reviewers.

Inclusion criteria.

- Study populations: Reviews/studies of adults (18 years and over) presenting with at least one of the five most common musculoskeletal pain presentations: back, neck, shoulder, knee and multi-site pain (the latter defined as musculoskeletal pain in more than one area of the body).
- Type of treatments: Reviews/studies of currently available treatments (including self-management advice and education, exercise therapy, manual therapy, pharmacological interventions (oral and topical analgesics, local injections), aids and devices, and other treatments (ultrasound, TENS, laser, acupuncture, ice / hot packs)) for musculoskeletal pain patients consulting in primary care were considered. Referral options for psychosocial interventions (such as cognitive-behavioural therapy and pain-coping skills) and surgery were also included. Comparison groups could include usual care, no intervention or other active interventions.
- Outcomes: Reviews/studies had to report outcomes of pain (e.g. intensity, widespreadness, bothersomeness, number of episodes, duration), and/or functional disability. These were considered primary outcomes for this review. Secondary outcomes such as psychological well-being / depression, catastrophising, quality of life (QOL), work related outcomes (e.g. sickness absence, return to work, days off work), and cost of treatment were highlighted, but were not required for inclusion in the review.

Exclusion criteria.

- Narrative reviews, letters, editorials, commentaries, and meeting abstracts were excluded, as were biomechanical, laboratory studies, animal studies as well as previous RCTs that were already summarised in included reviews, cohort, case-control, and cross-sectional studies.
- Reviews/studies published in other languages than English.
- Reviews/studies of musculoskeletal pain populations with suspected serious pathologies (e.g. suspected fracture, cancer, cauda equina syndrome), inflammatory arthritis, crystal disease, spondyloarthropathy, polymyalgia rheumatica, whiplash injuries, pregnancy-related pain problems, and vulnerable patients (e.g. experienced significant recent trauma, cognitive impairment, dementia, terminal illness).

Quality appraisal

In order to weigh the conclusion of reviews within our evidence summaries, the methodological quality of non-Cochrane systematic reviews was assessed using the 11-item 'assessment of multiple systematic reviews' (AMSTAR) checklist [26]. The guidelines and care pathways which were included in this evidence synthesis were not quality assessed as they all made use of published development processes based on explicit methodology.

Extraction of data

Data were extracted by one reviewer using a data collection form and independently checked for consistency and completeness by a second reviewer. Clarifications were sought where needed and disagreements between reviewers resolved by discussion. Data were extracted on

the effectiveness of non-pharmacological, pharmacological and surgical treatments for each musculoskeletal pain presentation separately, and where available, guidance or conclusions regarding patient subgroups mostly likely to respond to specific treatments. More specifically, data were extracted regarding:

- population characteristics (e.g. age, gender, symptom duration, musculoskeletal pain site and where possible musculoskeletal pain condition/diagnoses,
- treatments (type/intensity/dosage),
- primary and secondary outcome measures (as stated above),
- estimates of treatment effect (where pooled, and as presented in the systematic reviews),
- estimates of treatment effect for patient subgroups (where available),
- treatment setting (e.g. primary care), and
- sources of evidence.

Treatments were assessed for short-term (up to 3 months) and long-term (greater than 6 months) effectiveness based on the primary outcomes of pain and function.

Grading of evidence

Summaries of the overall evidence for the effectiveness of treatment options and strength of recommendations for each pain site were assessed based on (a modified) GRADE rating (http://www.gradeworkinggroup.org/). Summary evidence from all included reviews and guidelines were graded, taking into account the:

- Primary sources of data (e.g. guidelines, systematic reviews, RCTs): expert opinion or consensus in guidelines was rated as very weak evidence, while RCTs, systematic reviews and evidence-based guidelines were graded as higher level of evidence
- Quality of systematic reviews (Cochrane reviews or high methodological quality as assessed by AMSTAR checklist)
- Magnitude of effect where a standardised mean difference (SMD) of 0.2 was considered small, 0.5 as medium, and 0.8 as large according to Cohen [27], and for binary outcomes success rate, relative risk (RR) >2 was considered a medium to large effect size [28]
- Level of precision (confidence interval and level of significance; $p<0.05$)
- The consistency of results across systematic reviews or RCTs.

For each treatment option, evidence was graded as:

1. "Very weak evidence"—based solely on expert opinion or consensus in guidelines only or in the absence of systematic review evidence
2. "Limited evidence"—in the presence of little evidence from systematic reviews/evidence-based guidelines AND when there were small, inconsistent, or non-significant treatment effect sizes
3. "Moderate evidence"–in the presence of little evidence from systematic reviews/evidence-based guidelines (as in 2) but showing a medium to large treatment effect OR in the presence of strong evidence from high quality systematic reviews, but with small or inconsistent treatment effect sizes

4. "Strong evidence"—in the presence of strong evidence from high quality systematic reviews and evidence-based clinical guidelines AND medium or large effect sizes.

Each summary of evidence / analysis was graded using the adapted GRADE criteria as described above and a narrative synthesis was subsequently presented, indicating the strength of the evidence as very weak, limited, moderate, or strong.

Evidence synthesis

A narrative synthesis approach was undertaken. Given expected heterogeneity of sources of evidence, treatment settings, and the wide remit of this review (which covered currently available treatments in primary care and referral options for one or more musculoskeletal pain presentation), the evidence was summarised at a high level (using systematic reviews and guidelines where available), and therefore no new meta-analyses were conducted. However, pooled estimates of treatment effectiveness from systematic reviews, as well as comments on the consistency and magnitude of treatment effects were extracted and reported. Additional information from policy documents and guidelines on treatment recommendations and priorities, including the type of evidence from which it was generated (i.e. whether from RCTs, systematic reviews or expert opinion) was also noted. The gathered evidence was included in summary tables (S1–S7 Tables) to enable (indirect) comparisons to be made across pain sites for the various treatments. Gaps in the evidence were noted where no guidelines, systematic reviews or RCTs were found.

Results

Search results

A total of 3,588 unique citations (including Cochrane reviews) were retrieved from the electronic bibliographic databases. On assessing titles, abstracts and full texts against the inclusion criteria, 71 Cochrane systematic reviews met the selection criteria and were included. Non-Cochrane systematic reviews (n = 75) were only included where a gap not already covered by Cochrane reviews was identified, or if they represented new research that had not yet been considered within updated guidelines and care pathways. The remaining papers were excluded because they were not a systematic review (n = 798), focused on an area already covered by one of the included Cochrane systematic reviews (n = 234), were duplicate publications or did not fit the inclusion criteria (n = 2131). A summary of the review process outlining the selection of evidence is presented in Fig 1.

Quality appraisal

As Cochrane reviews followed a generic protocol specifying methods and review protocols go through a comprehensive peer review process prior to publication, the methodological quality of most Cochrane reviews included in this evidence synthesis was satisfactory (Fig 2). Cochrane reviews had flaws mainly associated with lack of searches for grey literature and/or no formal assessment of publication bias (Fig 2). As shown in Fig 3, methodological quality was less strong for non-Cochrane reviews, especially in terms of the comprehensiveness of the search strategy (including searches for grey literature), and listing of excluded studies (10%). Most reviews (81%) carried out some form of quality appraisal of included studies but study quality was not always incorporated into the evidence synthesis nor appropriately used to formulate conclusions (68%). Over half of the reviews (≈ 65%) minimised the risks of reviewer error and bias via duplicate processes for study selection and data extraction; and a very low proportion of reviews (16%) assessed the likelihood of publication bias.

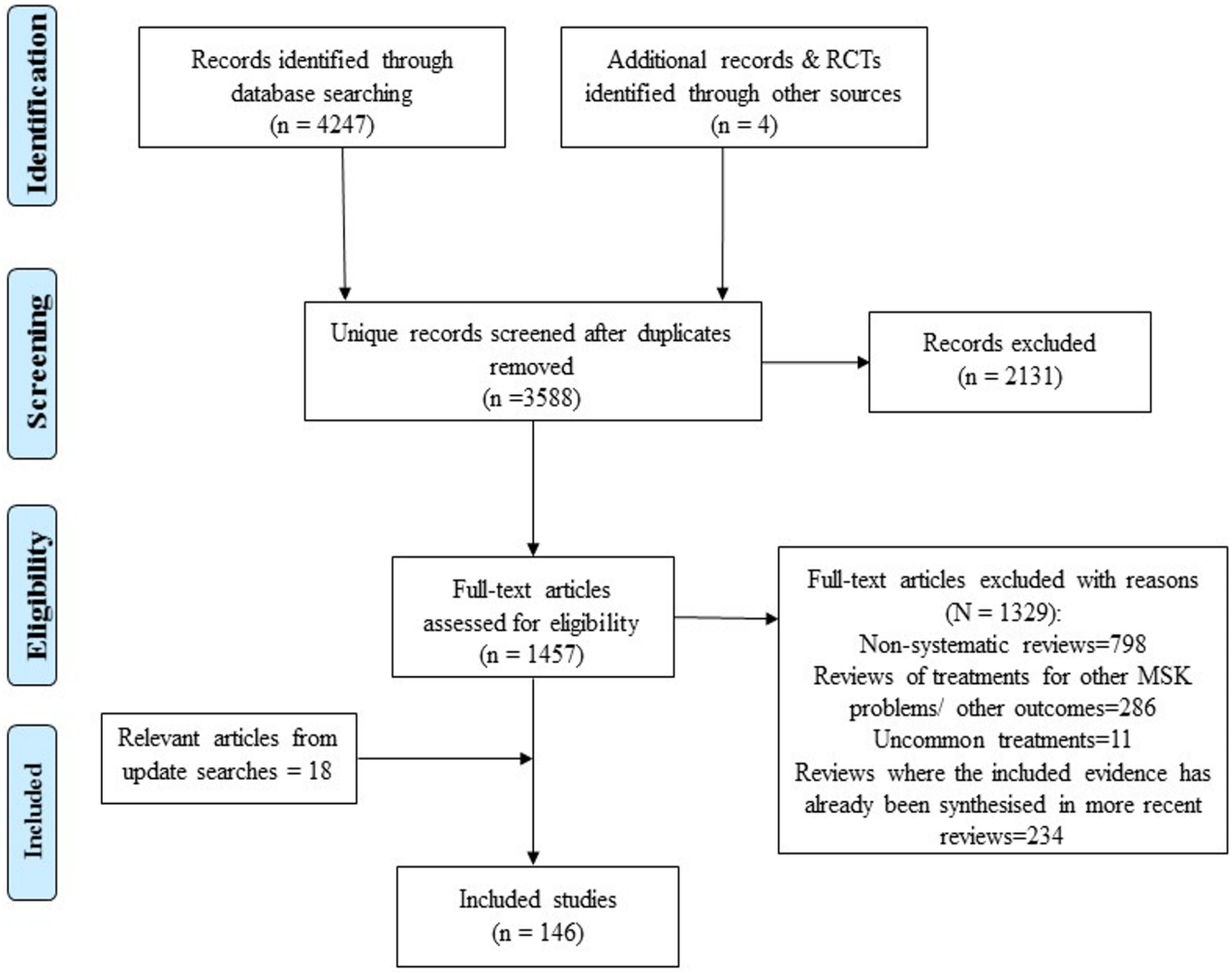


Fig 1. Review flow diagram (PRISMA).

https://doi.org/10.1371/journal.pone.0178621.g001

Included reviews and guidelines

Our searches identified reviews, guidelines and care pathways that covered a large range of treatment options for each musculoskeletal pain presentation. Based on specific review questions, authors of each review used particular criteria for identifying relevant trials in terms of setting, participants and interventions, resulting in variation across reviews in terms of the musculoskeletal condition, type and number of trials included, interventions, and reported outcomes of the reviews. A detailed description of the settings, populations, treatments and outcomes is provided in S1–S7 Tables. Each of the pain presentations (back, neck, shoulder, knee, multi-site pain) include several diagnostic categories, which are also summarised in the supplementary evidence tables.

Evidence synthesis

Effectiveness of available treatments for musculoskeletal pain was highlighted in the following order: self-management advice and education, exercise therapy, manual therapy, pharmacological

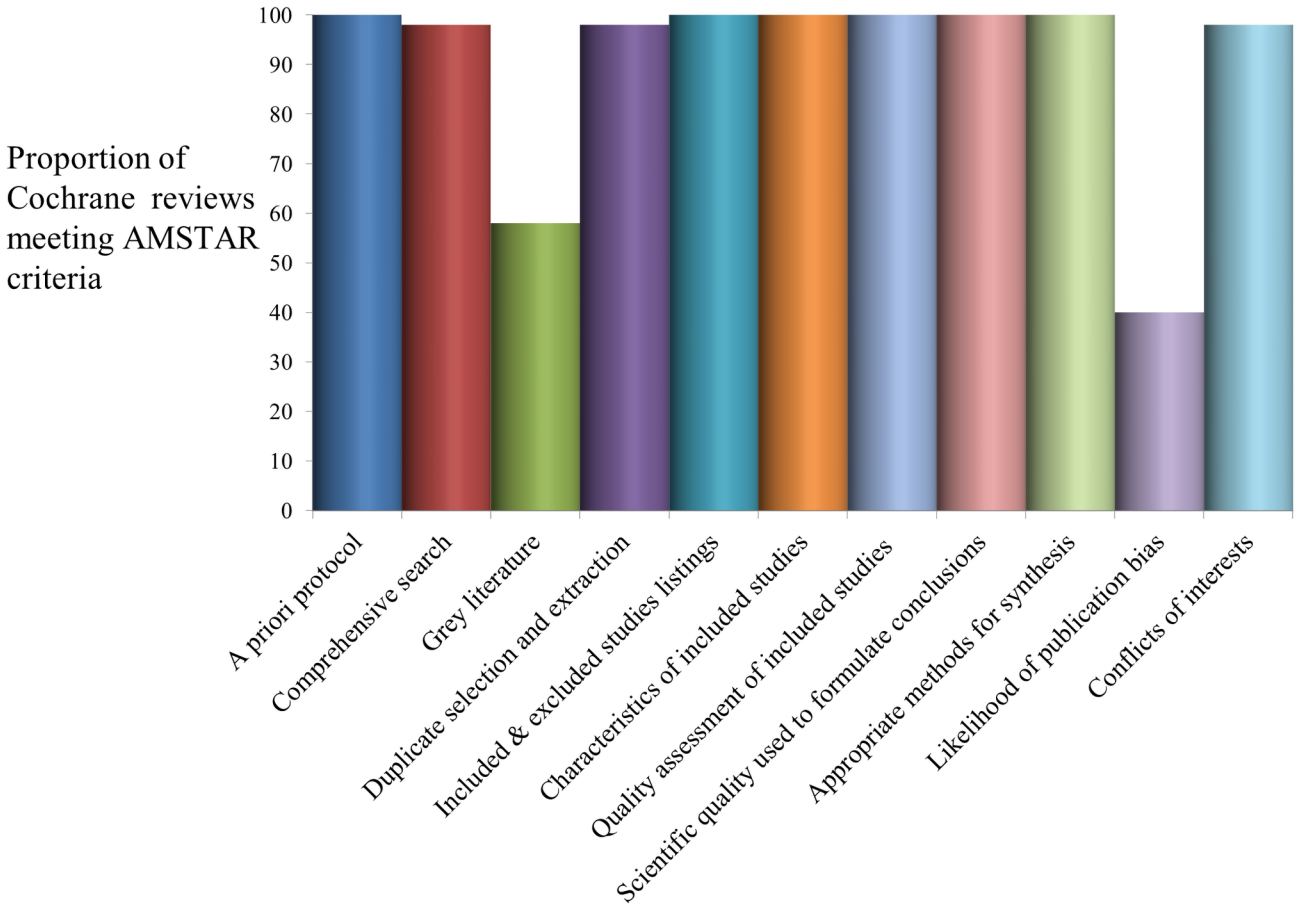


Fig 2. Quality assessment of contributing evidence from Cochrane reviews.

https://doi.org/10.1371/journal.pone.0178621.g002

interventions, aids and devices, other treatments (including ultrasound, TENS, laser, acupuncture, ice / hot packs), psychosocial interventions and surgery. A summary of the findings is presented in Table 1. The overall grade of evidence from included reviews and guidelines on effectiveness of treatment options was fairly consistent for each of the five musculoskeletal pain presentations. For instance, the strength of evidence in support of the beneficial effects of exercise therapy for all the five musculoskeletal pain sites ranged between moderate and strong whilst there was generally limited evidence for low to medium effectiveness of manual therapy across the pain sites. There was wide variability in terms of the application and mode of delivery of even the same treatments. Within guidelines, there was little evidence regarding specific patient subgroups and predictors of response to treatments. However, any information extracted, regarding patient subgroups most likely to respond to specific treatment options is summarised in the evidence tables (S1–S7 Tables).

Self-management advice and education. Evidence base: Evidence was extracted from two clinical guidelines, one clinical pathway and eight reviews about the effectiveness of self-management advice and education. As assessed by AMSTAR, the methodological quality of systematic reviews was moderate or high but the primary studies within those reviews were generally low or moderate in quality. Given mostly in the form of oral and / or written information, advice and education was directed at improving patients' understanding of their musculoskeletal pain, and self-management techniques, addressing patients' concerns about

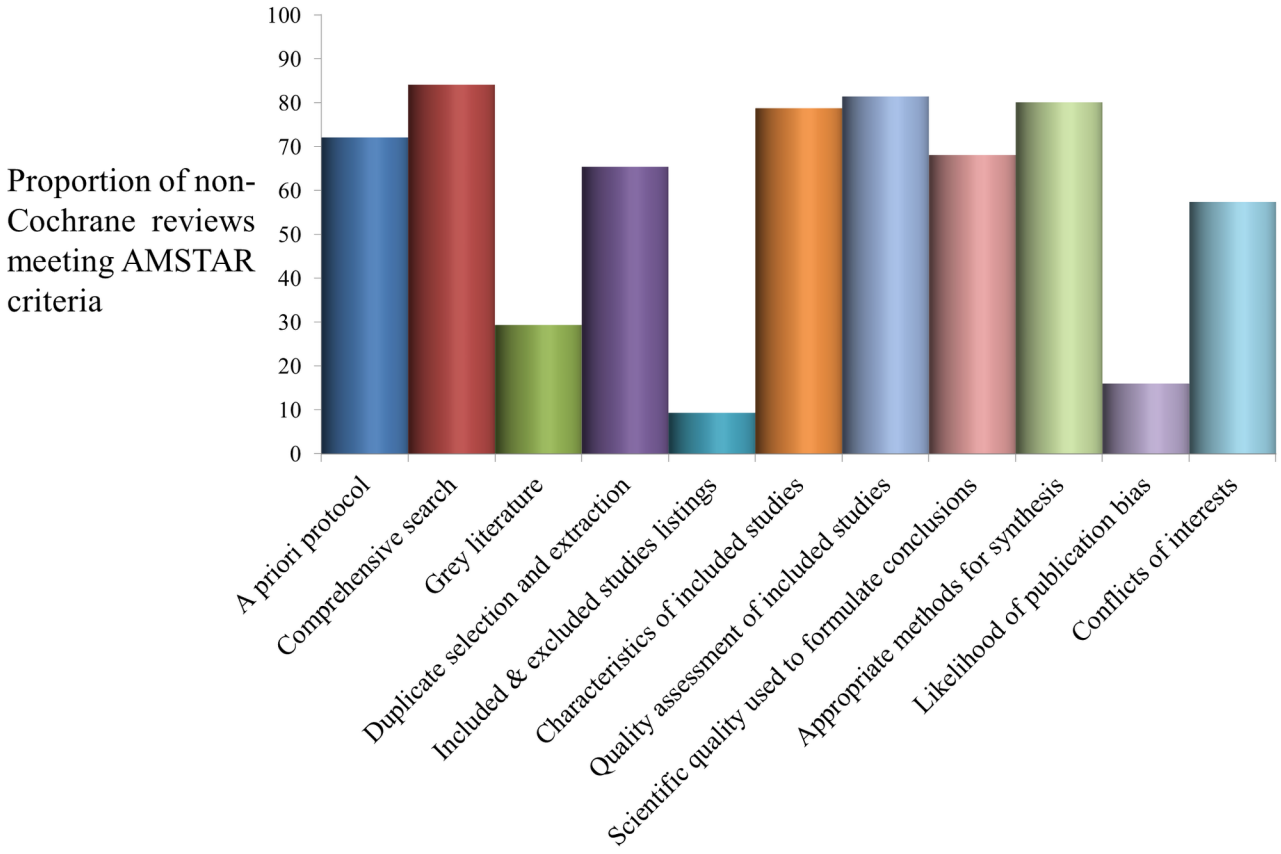


Fig 3. Quality assessment of contributing evidence from non-Cochrane systematic review & meta-analyses.

https://doi.org/10.1371/journal.pone.0178621.g003

serious causes and outcomes, supporting return to function, and minimising dependence on healthcare providers [20, 29–38].

Magnitude of effects: Self-management advice and education was typically provided to either individuals or patient groups, as part of an intervention programme and were not tested in isolation against a control treatment (S1 Table). Therefore, the evidence for the effectiveness of self-management advice and education alone on the outcomes of pain and function was difficult to assess. Where estimated, summary effect sizes were usually small and/or not statistically or clinically significant. For instance, for back pain patients who received self-management advice, Oliveira et al [37] reported a pooled Mean Difference (MD) at short-term (up to 3 months) follow-up for pain of -3.2 points on a 0–100 scale (95% CI, -5.1 to -1.3) and of -2.3 points (95% CI, -3.7 to -1.0) for function. There was no evidence regarding patient subgroups most likely to respond to self-management advice and education.

Strength of evidence: The evidence for self-management advice and education supporting expert opinion in clinical guidelines and consensus meetings as well as systematic reviews showed small effects on pain and function. Pooled results from meta-analyses tended to have wide confidence intervals although recommendations for the use of advice and education were consistent. Overall strength of evidence was graded as limited.

Bottom line: Despite the limited evidence-base, there were strong recommendations for the use of self-management advice and education as a first line treatment option for musculoskeletal pain.

Table 1. Summary of findings.

Evidence on treatment options across regional musculoskeletal pain presentations					
Treatment Options	**Evidence base**	**Regional pain**	**Outcomes** ***Pain Function Disability & other 2 0 Outcomes***	**Magnitude of Effects**	**Strength of evidence (Grade)**
Self-management advice & education	2 clinical guidelines, 1clinical pathway, 8 reviews.	Back, neck, shoulder, knee & multi-site pain.	Pain Function	Small effect sizes (e.g. -3.2 points (95% CI -5.1, -1.3) on a 0–100 scale for back pain, Oliveira et al. 2012). Beneficial effects not proven in the long term.	****Limited evidence**
Psychosocial interventions	2 guidelines, 1 clinical pathway, 10 reviews & 2 RCTs.	Back, neck, shoulder, knee & multi-site pain. Limited amount of evidence on shoulder & knee pain.	Pain Function Quality of life	Medium to large effect sizes (e.g. MD -5.18; 95% CI -9.79 to -0.57, Henschke et al. 2011) for pain on a scale of 1 to10). Beneficial effects demonstrated in short & long term.	***Moderate evidence**
Exercise Therapy	4 guidelines, 3 policy documents, 32 reviews, 1 RCT.	Back, neck, shoulder, knee & multi-site pain.	Pain Function Quality of life Work-related outcomes.	Medium to large summary effects sizes (e.g. SMD 0.65, 95% CI: -0.09 to 1.39 for multi-site pain, Busch et al 2007, & RR 7.74, 95% CI: 1.97 to 30.32 for shoulder pain, Green et al 2003) Beneficial effects in the short & long-term for all five pain presentations.	****Strong evidence**
Manual therapy	3 guidelines & 21 reviews.	Back, neck, & shoulder pain.	Pain Function	Small effect sizes (e.g. NNT 5, for neck pain, Gross et al. 2012, & MD: -4.16, 95% CI -6.97 to -1.36, on 0–100 point scale for back pain, Rubinstein et al. 2011). Short-term effect on chronic pain but no strong evidence of long-term effectiveness compared to other standard treatments.	****Limited evidence**
Pharmacological Treatments—(oral & topical analgesics)	3 guidelines, 1 clinical pathway & 30 reviews.	Back, neck, shoulder, knee & multi-site pain.	Pain Function Evidence on function less often reported.	Medium effect sizes (e.g. NNT 4.6 (95% CI 3.8 to 5.9 for NSAIDs compared to placebo, Mason et al. 2004). Cox-2 selective inhibitors and opioids reduce pain in the short-term but the risk of adverse effects such as gastrointestinal bleeding and opioids-induced hyperalgesia needs careful consideration.	***Moderate evidence**
Pharmacological Treatments–(Corticosteroid injections)	3 guidelines, 1 clinical pathway & 16 reviews.	Back, neck, shoulder, & knee pain. Limited effects on back and neck pain.	Pain	Medium to large effect sizes (e.g. RR: 3.11 (95% CI 1.61 to 6.01 using injections for relieving moderate to severe knee pain in the short term compared to placebo, Belamy et al 2006).	****Strong evidence**
Other treatments (Aids, Devices, complementary /alternative therapy)	5 guidelines, 1 clinical pathway, 1 policy document, & 20 reviews.	Back, neck, shoulder & knee pain.	Pain Function	Small, non-significant or inconsistent, summary effect sizes.	****Limited evidence**

(*Continued*)

Table 1. (*Continued*)

Evidence on treatment options across regional musculoskeletal pain presentations					
Treatment Options	**Evidence base**	**Regional pain**	**Outcomes** ***Pain Function Disability & other 2 0 Outcomes***	**Magnitude of Effects**	**Strength of evidence (Grade)**
Surgery	1 guideline, clinical pathway document, 17 reviews.	Back, neck, shoulder, knee & multi-site pain.	Pain Function Quality of life	Effect sizes (not often estimated). Beneficial effects on pain & function in the short term with little empirical evidence for sustained long-term improvement.	****Limited evidence**

*Very weak evidence: Expert opinions or consensus in guidelines only / Absence of evidence in a single systematic review.

** Limited evidence: Little empirical evidence from systematic reviews/evidence-based guidelines AND when there were small, inconsistent, or non-significant treatment effect sizes.

*** Moderate evidence: little empirical evidence from systematic reviews/evidence-based guidelines (as in limited evidence) but showing a medium to large treatment effect OR in the presence of strong empirical evidence from high quality systematic reviews, but with small or inconsistent treatment effect sizes across systematic reviews.

**** Strong evidence: Strong empirical evidence from high quality systematic reviews and evidence based clinical guidelines AND medium or large effect sizes.

https://doi.org/10.1371/journal.pone.0178621.t001

Exercise therapy. Evidence base: Synthesized evidence for the effectiveness of exercise on musculoskeletal pain included 10 Cochrane reviews [13, 18, 39–46], four guidelines [16, 20, 29, 47] and three policy documents [38, 48, 49]. Evidence from other reviews (n = 22), [36, 50–71] and one additional trial [72] were also considered. Quality of reviews ranged from moderate to high.

Magnitude of effects: Exercise therapy was determined to be beneficial for pain, function and quality of life in all five pain presentations [13, 16, 18, 20, 29, 36, 38, 47, 55, 56, 58–60, 63]. See supplementary S2 Table. Reviews and guidelines on exercise for neck pain [38, 58, 66] generally found exercises to be beneficial for function but no pooled estimates were provided. Exercise therapy led to clinically significant improvements in pain, function and quality of life for shoulder, knee, back and multi-site pain. In addition, medium to large summary effect sizes were reported in favour of exercise across the body of evidence, for example; RR 7.74; CI, 1.97 to 30.32 and RR 1.53; CI, 0.98 to 2.39 for improvement of shoulder pain and function respectively [13]; MD -1.46, CI -2.39 to -0.54 on a scale of 0 to 10) for pain as well as function (SMD 1.10, 95% CI, 0.58 to 1.63) for knee pain [46]; and MD 7.3 (95% CI, 3.7 to 10.9 points on a scale of 0–100) for low back pain [18]. With respect to multi-site pain[41], aerobic exercises was found to lead to improvement in global well-being (SMD 0.49, 95% CI, 0.23 to 0.75), physical function (SMD 0.66, 95% CI, 0.41 to 0.92) and pain (SMD 0.65, 95% CI, -0.09 to 1.39).

There appears to be little empirical evidence in favour of any particular exercise type, programme or mode of delivery, either as structured individual or group treatment for musculoskeletal pain [18, 39, 41, 48–50, 57, 65, 66, 68–70, 72, 73], although functional exercises (which adapt patients' exercises to their activities of daily living, and enables them to perform such activities more easily and without injuries) appear to be more beneficial than exercises not specifically targeting function. There was no evidence regarding patient subgroups most likely to respond to exercise therapy. While some contributing reviews included information on whether patient symptoms were acute and chronic, it was difficult to assess if any particular exercise therapy had better effects on acute or chronic symptoms.

Strength of evidence: On the basis of medium to large summary effects sizes from high quality reviews, and clinical guidelines, the strength of evidence for the effectiveness of exercise therapy for pain, function, and quality of life for patients with musculoskeletal pain, was graded as strong.

Bottom line: Current evidence shows significant positive effects in favour of exercise on pain, function, quality of life and work related outcomes in the short and long-term for all the musculoskeletal pain presentations (compared to no exercise or other control) but the evidence regarding optimal content or delivery of exercise in each case is inconclusive.

Manual therapy. Evidence base: Six Cochrane reviews [44, 45, 74–77], three guidelines [20, 38, 78], and 15 other systematic reviews [54, 58, 66, 79–90] contributed to the evidence synthesis on the effect of manual therapy for the five most common musculoskeletal pain presentations. The effects of manual therapy on pain and function were mostly examined in combination with other treatments and mostly for non-acute pain. Methodological quality of reviews was moderate or high, although as highlighted in many of the reviews, a number of the primary trials on which reviews were based were of low quality.

Magnitude of effects: Pooled estimates for the effectiveness of manual therapy for musculoskeletal pain were generally statistically significant, but variable in terms of size of the treatment effect S3 Table. Manipulation, mobilisation and massage (where indicated) were reported to be beneficial for immediate and or short-term (4–6 weeks) improvement in range of motion and function in both acute and chronic neck pain patients as well as those with whiplash [38, 58, 66, 78, 91]. For instance, thoracic manipulation was found to lead to significant pain reduction (number needed to treat (NNT) 7), and increased function (NNT 5) in acute neck pain patients whilst a single session of thoracic manipulation was reported to result in immediate pain reduction for chronic neck pain patients (NNT 5) compared to placebo [44]. In a recent Cochrane review of manual therapy for adhesive capsulitis, 46% of participants reported treatment success with manual therapy and exercise compared with 77% who had corticosteroid injections (summary RR 0.6, 95% CI, 0.44 to 0.83), with an absolute risk difference of 31% (13% to 48%). The number reporting adverse events did not differ (summary RR 1.07, 95% CI, 0.76 to 1.49) between groups [45]. As with neck pain, manual therapy offers some benefits for range of motion and function in shoulder pain presentations [54, 79, 80, 82, 83, 89].

For back pain, evidence suggests that manual therapy alone or in combination with other treatments may offer some benefit for pain and function [20, 81, 85, 88]. Most authors presented no pooled estimates of treatment effects due to large heterogeneity among included trials. Where presented, summary effect sizes were generally small compared to no manual therapy or other control (e.g. SMD -0.25 (95% CI, -0.46 to -0.04 for pain and SMD -0.22, (95% CI, -0.36 to -0.07 for function) with negative SMD indicating lower levels of pain or functional limitation for manual therapy) in the short term [75, 76]; and (MD -0.46 (95% CI, -1.18 to 0.26 on a scale of 0 to 10) for pain in the long term [76]. Compared with other treatments (e.g. general practitioner care, acupuncture, ultrasound, standard physiotherapy, analgesic therapy, exercise, or back school), manual therapy appears to confer little or no clinically important effect on pain intensity, functional status, global improvement or return to work among patients with acute, subacute or chronic back pain with or without sciatica [74–77, 92, 93]. Type and experience of professional delivering the therapy did not show any clinically significant effect of on musculoskeletal pain [73]. There was low quality evidence that the efficacy of manual therapy might differ for subgroups of patients, with manual therapy tending to be more effective for acute non-specific low back pain patients with mobility deficit [90].

Strength of evidence: Despite several high quality reviews examining the effects of manual therapy on pain and function for neck, shoulder and back pain, current evidence generally

shows small summary effect sizes or concludes no clinical effectiveness of manual therapy compared to sham or other active treatments. Overall strength of evidence was graded as limited.

Bottom line: Current evidence regarding manual therapy is beset by heterogeneity across clinical trials. Due to paucity of high quality evidence, it is uncertain if the efficacy of manual therapy might be different for different patient subgroups or influenced by the type and experience of professional delivering the therapy. On the whole, available evidence suggests that manual therapy may offer some beneficial effect on pain and function but it may not be superior to other non-pharmacological treatments (e.g. exercise) for patients with acute or chronic musculoskeletal pain.

Pharmacological treatments—Analgesics (oral & topical). Evidence base: Thirty systematic reviews of pharmacological interventions for musculoskeletal pain examined the effectiveness of analgesics (opioids and non-opioids) in the short and long-term as well as in acute and chronic pain presentations. Comparisons were against placebo [94, 95], other pharmacological agents [48, 49, 96–100], corticosteroid injections [101], and no treatment [102, 103]. A few comparisons were made with other treatments such as laser and acupuncture [7]. Over 60% of the reviews on oral and topical analgesics were of high methodological quality while the rest were moderate. Reviews highlighted that the quality of included primary trials ranged from low to high quality.

Magnitude of effects: Compared to placebo, acetaminophen (paracetamol) was not more effective (SMD 0.13, 95% CI, 0.04 to 0.22) for relieving knee and back pain [94, 100, 103]. NSAIDs and opioid analgesics (especially for acute pain) were generally found to be effective but beneficial effects were evident mostly in the short-term [7, 14, 16, 29, 38, 94, 104, 105]. Cyclooxygenase (Cox)-2 selective inhibitors (e.g. celecoxib), were found to be effective for musculoskeletal pain relief. However, these were more likely to be associated with higher risks of adverse cardiovascular and gastrointestinal events (hazard ratio 2.18, 95% CI 1.82, 2.61), compared to non-selective NSAIDs [48, 49]. In the long-term and for more chronic pain presentations, stepwise analgesia according to the WHO analgesic ladder (mostly based on expert opinion) may be recommended [20, 29, 106–109]. Medium effect sizes were commonly reported S4 Table. For instance, topical NSAIDs were found to be more beneficial compared to placebo with summary RR of 1.9 (95% CI, 1.7 to 2.2) and a NNT of 4.6 (95% CI, 3.8 to 5.9) in the short-term [98, 99, 110, 111]. Furthermore, duloxetine, commonly used for multi-site pain may be carefully considered where there has been inadequate clinical response to initial pharmacologic treatments [48]. The effects of analgesics for improving function were less often reported in included reviews and guidelines.

Strength of evidence: With consistent medium summary effect sizes reported across moderate to high quality systematic reviews and clinical guidelines, there is moderate evidence that pharmacological therapies are beneficial for the short-term relief of musculoskeletal pain. Overall strength of evidence was graded as moderate.

Bottom line: NSAIDs, Cox-2 selective inhibitors and opioids reduce pain in the short-term, but the effect size is modest and the potential for adverse effects such as gastrointestinal bleeding and opioids-induced hyperalgesia need careful consideration.

Pharmacological interventions–injections. Evidence base: The evidence base for the effectiveness of injections for musculoskeletal pain involved the synthesis of three clinical guidelines [16, 112, 113] and one care pathway document [38], six Cochrane reviews [104, 114–118] and 13 other systematic reviews [7, 64, 66, 101, 119–127]. The systematic reviews were mostly high in methodological quality.

Magnitude of effects: A care pathway document [38], one guideline [49] and seven systematic reviews [64, 101, 116, 117, 119, 120, 125] supported evidence for the short-term (<4

weeks) benefits of corticosteroid injections for relieving moderate to severe shoulder pain (summary RR 1.43 (95% CI, 0.95 to 2.16) for corticosteroid injection compared with NSAIDs [119]). Likewise for knee pain, corticosteroid injections were found to be effective in the short-term for relieving moderate to severe pain compared to placebo ((summary RR: 3.11 (95% CI, 1.61 to 6.01); NNT 3 to 4) [115, 128]. Though corticosteroid injections were found to relieve pain, there was a lack of evidence for clinically significant effects on function [115]. For knee pain, viscosupplements such as intra-articular hyaluronate injections were found to be better than placebo (SMD 0.60, 95% CI 0.37 to 0.83) for reducing pain and improving function (SMD 0.61, 95% CI 0.35 to 0.87) in the short term (1–4 weeks). However, high clinical and statistical heterogeneity, evidence of publication bias and low quality trials preclude definitive recommendations about routine use in clinical practice [49].

Furthermore, the available evidence did not suggest injections are effective for the management of neck pain [66, 113, 121–124] or back pain [38, 104, 118]. Overall, there was no strong evidence for the use of epidural spinal injections with or without steroids, as benefits (immediate reductions in pain) were small and not sustained [114, 126, 127]. It appears the short-term pain relief offered by epidural spinal injections are hampered by significant heterogeneity, and that the severity and subtype of pathology may affect outcome [114, 126, 127].

Generally, in the long-term, injections may be no more effective than non-pharmacological interventions such as exercise [7, 64, 113, 116, 117, 119, 120, 125]. Evidence also suggests that the addition of corticosteroid injections to local anaesthetic does not confer improved symptom relief in the long-term [121, 122] however, expert opinion and guideline recommendations support its use prior to, or alongside, exercise and self-management advice [38, 64, 101, 112, 113, 119]. Although injections were often offered for acute pain relief and to enable patients to tolerate exercise therapy, there was no evidence regarding patient subgroups most likely to respond to injections.

Strength of evidence: Supported by high quality reviews, and clinical guidelines, medium to strong effect sizes across the various sources of evidence, injections offer clinically significant benefits for relieving moderate to severe shoulder and knee pain but in the short-term (up to 3 months) only. Overall, the strength of evidence was graded as strong.

Bottom line: The evidence indicates that injections offer short-term pain relief for shoulder and knee pain but effectiveness for back and neck pain is uncertain. Across the musculoskeletal pain presentations for which pharmacological injections may be given for pain relief, current evidence is equivocal on the optimal procedure (e.g. guided vs. unguided), frequency, dose and active component of the injections (though corticosteroid injections are more often reported in literature).

Aids & devices—Orthotics, tapes, braces, cervical collars and other support devices. Evidence base: The evidence for the effectiveness of aids and devices for pain and function included five guidelines [16, 20, 47, 112, 129], one clinical pathway [38], four Cochrane reviews [66, 77, 130–132], two best evidence syntheses [58, 133] and a meta-analysis [134]. The quality of reviews was moderate.

Magnitude of effects: Either as stand-alone treatment or mostly in combination with other treatments, aids and devices for musculoskeletal pain have generally shown small effects (see supplementary S5 Table) on pain, function or work outcomes [16, 20, 38, 58, 66, 77, 131–133]. Routine use of collars has not been found to confer any clinically significant benefits for neck pain [38, 58, 66, 133]. This may be attributed to marginal pain relief (in the short-term), and inclination to induce rest and inactivity hence prolonging disability. Patellar taping has been shown to have some beneficial effects (in the short-term) on pain and function in patients with patellofemoral pain [16, 20, 47, 112, 129]. Warden et al.[134] reported significantly less pain on a 100-mm scale (weighted mean difference (WMD) = -20.1, 95% CI, -26.0 to -14.3, $p <$

.001) for patellofemoral pain patients treated with medially directed tape compared to patients treated with no tape or patients treated with sham tape (WMD = -13.3, 95% CI, -18.1 to -8.4, $p < .001$). There is very weak empirical evidence for the beneficial effects of knee braces but in grade II and III collateral ligament injuries, short-term (4–6 weeks) application of a hinged brace may be considered as part of rehabilitation [38, 47]. Empirical evidence suggests lumbar supports are not effective for improving pain and function in back pain patients [20]. This review did not find any evidence regarding specific patient subgroups for which aids and devices might be most beneficial.

Strength of evidence: Supported mostly by expert opinion or consensus in guidelines as well as small, inconsistent, or non-significant treatment summary effect sizes from systematic reviews, overall evidence for the use of aids and devices in the management of musculoskeletal pain is graded as limited.

Bottom line: For neck, shoulder, back and knee pain presentations, available evidence does not justify routine use of aids and devices for effective improvement of pain, function, and / or work outcomes.

Other treatments: Acupuncture, ultrasound, TENS, laser, ice / hot packs. Evidence base: Contributing evidence on the effectiveness of acupuncture, therapeutic ultrasound, TENS, laser, and superficial ice / hot packs for pain and function included five guidelines [16, 20, 47, 112, 129], one policy and one clinical pathway document [38, 48], 14 Cochrane reviews [13, 48, 76, 135–148] and 18 systematic reviews [36, 55, 59, 149–159]. The quality of reviews was mostly moderate with some reviews having high methodological quality. However, within the reviews and clinical documents, there was large heterogeneity and significant publication bias in primary studies.

Magnitude of effects: Compared to treatments such as analgesia, and exercise, these interventions have been less frequently evaluated, and the quality of RCTs is generally low. Also, for many of these treatments (i.e., therapeutic ultrasound, laser, and superficial ice / hot packs), reports of high clinical and methodological heterogeneity within the trials contributing to reviews preclude statistical pooling of effect estimates. There was also no evidence regarding specific patient subgroups which might benefit most from these treatments.

For acupuncture, available evidence from a good quality individual patient data meta-analysis suggests that acupuncture may be effective for short-term relief of back pain and knee pain with medium summary effect sizes (SMD 0.55 (95% CI, 0.51 to 0.58) and (SMD 0.42 (95% CI, 0.37 to 0.46)) respectively compared with usual care or no acupuncture [158]. However, effects on function were reported to be minimal and not maintained at longer-term follow-up [20, 139, 149, 152, 158]. Similarly for neck and shoulder pain, acupuncture was only found to be effective for short-term (immediately post-treatment and at short-term follow-up) symptom relief (SMD -0.37 (95% CI, -0.61 to -0.12)) and (WMD 3.53 (95% CI, 0.74 to 6.32 on a scale of 1–100)) compared to placebo [140, 148].

TENS was no more effective for reducing pain than placebo in chronic back pain [136, 141, 160, 161], neck pain [142], shoulder pain [145], knee [147] and chronic musculoskeletal pain [144, 150]. Ultrasound and shockwave therapy do not appear to significantly improve clinical outcomes for acute and chronic low back pain [162]. Also, for those with shoulder and/or neck pain, evidence suggests ultrasound does not confer significant or added benefit over placebo or other treatments [47, 55, 101, 140, 153, 157]. The evidence on effectiveness of laser therapy for shoulder pain [59, 159], or acute or chronic neck pain was inconclusive [151]. With regards to knee pain, other treatments including ultrasound, electromagnetic fields, low level laser therapy, TENS, biofeedback, neuromuscular electrical stimulation may confer added benefits to exercise and / or surgical treatment but empirical and clinical effect sizes are

small and only supported by weak evidence [16, 47, 48, 112, 129, 143, 146, 147]. (Please refer to supplementary S5 Table for more details regarding other treatments).

Strength of evidence: There was little empirical evidence for the effectiveness of other treatments including ultrasound, TENS, laser, and superficial ice / hot packs. Presented summary effect sizes and estimates were often small, inconsistent, and non-significant. Although medium short-term effects were found for the effects of acupuncture on back and knee pain, overall strength of evidence was graded as limited.

Bottom line: The evidence for the clinical effectiveness of most of these other treatment options was not substantiated by strong evidence. Either as stand alone or in combination with other treatments, the often small effect sizes as a result of these treatments for improving musculoskeletal pain and function was mostly not clinically significant.

Psychosocial interventions. Evidence base: Evidence base for the effectiveness of psychosocial interventions (referred to various interventions such as cognitive-behavioural therapy and pain-coping skills, used to support people for overcoming challenges and maintenance of good health) included one guideline [20] and an overview of guidelines [163], one care pathway [38], four Cochrane reviews [164–168] and seven systematic reviews [17, 133, 169–173]. The quality of reviews ranged from moderate to high. Due to gaps in available systematic reviews of shoulder pain regarding psychosocial interventions, additional evidence from RCTs [174, 175] was extracted for shoulder pain.

Magnitude of effects: Reviews of psychosocial treatments for the management of musculoskeletal pain included a wide range of approaches that aimed to achieve increased self-management, behavioural and/or cognitive changes alongside biomedical management of pain S6 Table. Interventions were often multimodal and involved multidisciplinary treatment. At long-term follow-up, medium summary effect sizes (e.g. SMD 0.23; (95% CI, 0.43 to 0.040) compared to usual care and SMD 0.48 (95% CI, 0.93 to 0.04) compared to other active treatments [172]) were reported for pain, function and/ or other psychosocial related-outcome measures such as quality of life. With the exception of a few studies in back pain and neck pain, where patient recruitment and outcome reporting were based on targeted groups of patients receiving a psychosocial intervention according to baseline complexity of patients' pain presentations [38, 164, 173], there was wide variability in the characteristics of patients included in trials. The effectiveness of psychosocial interventions for the management of shoulder, knee, and neck pain presentations was less well researched compared to those of back pain. Psychosocial interventions in combination with other treatment options appear to provide additional benefit for all musculoskeletal pain presentations. However, there was no consensus on specific treatment components, providers and settings for optimal outcomes [20, 38, 163–167, 170, 172–175]. Furthermore, methodological issues regarding primary studies reported by the systematic reviews, such as high attrition rates, incomplete outcome reporting, mixed treatment regimens and generally low sample numbers and patient heterogeneity made conclusions tentative.

Strength of evidence: Except for shoulder and knee pain, where the strength of evidence was limited, current evidence for the beneficial effects of psychosocial interventions for neck, back and multi-site pain is supported by moderate to high quality reviews, medium effect sizes with precise confidence intervals and this is consistent across sources of evidence. Overall, the strength of evidence was graded as moderate.

Bottom line: Available evidence suggests beneficial effects of psychosocial interventions, particularly for patients identified as having a poor prognosis prior to treatment. Also, outcome of psychosocial treatment appears to be influenced by other factors such as patient prognosis, the healthcare professional providing treatment, the settings for treatment delivery and the components of treatment.

Surgery. Evidence base: Evidence for the effectiveness of surgery for the musculoskeletal pain presentations (excluding multi-site pain) was synthesised from one guideline and a care pathway document [38, 47], nine Cochrane reviews [176–184] and eight systematic reviews [66, 185–191]. Reviews were mostly high in methodological quality.

Magnitude of effects: Most guidelines specify that surgical treatments are indicated in a small proportion of patients (as low as 8%) for neck, shoulder, back and knee pain presentations [38, 47, 181]. Within the body of synthesised evidence (supplementary S1 and S7 Tables), the presence of serious pathology, substantial pain and disability or symptoms which are refractory to conservative treatment were prominent indications for surgery [38, 47, 180, 191], but the roles of such factors in determining the long-term clinical outcome of treatment was equivocal [38, 180]. Based on clinical judgement and expert opinion, current evidence suggests early surgical intervention may be considered on a case by case basis [38, 47]. Generally for neck, shoulder, knee and back pain, when indicated, there is moderate evidence that surgical intervention does provide benefits for pain, and function compared to waiting list controls or conservative treatments including analgesia and exercise in the short-term [38, 66, 176, 178–180, 187]. In specific cases, such as arthroscopic debridement and joint lavage of the knee, available evidence indicates no clinically important benefit (SMD -0.11, 95% CI, to 0.42 to 0.21) for pain or function compared to control (SMD -0.10, 95% CI, -0.30 to 0.11) at three months [182]. Available evidence suggests there are no long-term benefits of surgical procedures for clinical outcomes compared with conservative treatment [177, 184–188, 190]. Neither was there strong evidence for a significant difference in favour of any particular surgical technique for any of the pain sites [182, 183, 189, 191].

Strength of evidence: Though reviews were mostly high in methodological quality, summary effect sizes were small. Overall strength of evidence of long-term effectiveness of surgery is limited except where directly indicated by specific serious pathology such as end-stage degenerative knee joint disease, persistent pain and functional limitation which are refractory to conservative treatments.

Bottom line: The effectiveness of surgery as a first line treatment option is not established in current literature. The current evidence base is limited in terms of quantity, especially comparing surgical versus conservative interventions but there is moderate evidence from guidelines, Cochrane reviews and other systematic reviews to support short-term efficacy of surgical interventions for pain and function for specific neck, shoulder, knee and back pain presentations. Available evidence also suggests that surgery is not superior to conservative treatment options in the long-term.

Discussion

This review has systematically identified, synthesised and graded a large body of evidence on the effectiveness of treatment for musculoskeletal pain presentations. For most pain presentations, non-pharmacological treatments especially exercise therapy as well as psychosocial interventions, produced medium to large effects on pain and function, with corticosteroid injections potentially offering short-term benefit in those with knee and shoulder pain. NSAIDs and opioids (where appropriate) also offer short-term benefit for musculoskeletal pain, but the potential for adverse effects need careful consideration.

The effectiveness of exercise therapy, psychosocial interventions and corticosteroid injections was consistently supported by empirical evidence of mostly medium effect sizes provided by meta-analyses of RCTs, by guidelines, and expert opinion for musculoskeletal pain. With regard to intensity, and modes of applications of most treatments, the amount of clinical contact, the type of provider, setting, and delivery modes/techniques for effective treatment varied

widely and, as yet, there is limited evidence to support choices regarding optimal delivery of these treatments. Therefore, further research to investigate the optimal dose and application of these treatment options is needed.

In this review, there was little information within the evidence base in relation to patient subgroups most likely to respond to different treatment options. Where available, for each treatment option, evidence regarding patient characteristics such as baseline pain severity and function, duration of pain, and previous pain episodes have been documented S1–S7 Tables. For most treatment options apart from manual therapy (due to low quality evidence for differential effects of manual therapy across patient subgroups), and psychosocial interventions (where moderate evidence supports targeting patient subgroups to psychosocial intervention according to baseline complexity), it is not certain if clinical outcomes for most treatment options may be improved by targeting patient subgroups. Given that there are many factors (including patient characteristics and risk of poor outcome) which may influence outcome of treatment, it is likely that, an optimal approach to management of musculoskeletal pain may involve strategic selection of treatments best suited for different patients. Future trials should be designed to bridge this gap in evidence for the management of musculoskeletal pain.

It is worth noting that in many of the reviews, guidelines and trials contributing to this evidence-base, individual treatments were rarely used in isolation. Therefore, the evidence for the isolated effectiveness of treatments in some reviews was difficult to assess. For instance, self-management advice and education was typically provided as part of intervention package rather than tested in isolation against a control treatment. Consequently, there was little empirical evidence about its effectiveness despite consistent support of the beneficial effects (by expert opinion and consensus in guidelines). This could impact the quality and level of evidence for the beneficial effects of otherwise promising treatments.

Overall completeness and applicability of evidence

As expected, given the breadth of this review there was wide heterogeneity in study populations, outcomes, and statistical methods for estimating summary effect measures in the included systematic reviews. Interpreting findings within this overview was also complicated by variability in both the intervention and control conditions (placebo, no treatment, active treatments) examined within the reviews, making it difficult to summarise evidence regarding the magnitude of treatment effects. Furthermore, in this overview, the treatments provided in individual studies could not be described in detail; settings, exact content, intensity or dose of interventions may have varied; many interventions may have required specialist staff (e.g. injection, acupuncture, manual therapy, surgery) and the training and skills of providers are likely to have varied over time and locations. Control conditions were frequently not described in reviews and trials, and the definition of terms "routine care", "standard care" or "no intervention" may vary depending on setting and country. In the conduct of this study however, concerted efforts were made to capture and report available contextual information when summarizing evidence regarding treatment effectiveness.

Strengths and limitations of the review

Where possible, given the wide remit of this review, a number of steps were taken to ensure methodological rigour. The focus was on publications providing high quality evidence or recommendations, including Cochrane and other high quality reviews, well-developed clinical guidelines that met specific quality assurance criteria, and evidence-based multidisciplinary care plans as outlined in care pathways. For Cochrane reviews, all reviews used protocols that aimed to minimise bias whilst for non-Cochrane reviews, evidence of using systematic

methods was a pre-requisite for inclusion in this study. In addition, separate structured and systematic searches of bibliographic databases were conducted to identify additional trials not covered in previous reviews, where gaps concerning the effectiveness of specific treatment options were identified.

This review provides evidence summaries regarding the effectiveness of a wide range of treatments for the five most common musculoskeletal pain presentations in primary care, drawing together findings from a large evidence base. To facilitate this rapid evidence summary, the methodology evolved as a rapid application of systematic review methods to synthesising evidence. Efforts were made to capture, appraise and synthesise the best available evidence in a systematic yet rapid fashion. Definitive elements of typical systematic review methods such as a comprehensive and systematic search of best available evidence, pre-specified inclusion and exclusion criteria, quality appraisal and synthesis have all been preserved. Further strengths of this review included independent assessment of eligibility for inclusion and data extraction for contributing reviews, data checks, appraisal of the quality of systematic reviews, and a standardised approach to synthesising evidence.

There are several limitations to this review. First, there was no independent assessment of the methodological quality of primary trials that were included in the reviews. As this is an overview of current best available evidence, methodological quality assessment of included primary studies depended largely on the ratings of systematic review authors rather than our own assessment of the details presented in the individual studies. The overview of evidence incorporating reviews of multiple interventions across many musculoskeletal pain conditions therefore may not follow strictly the process generally applied in a single systematic review of one intervention on a single target population. However, much care has been taken to ensure that our approaches to searching for evidence, quality appraisal and grading of available evidence, and synthesis (as highlighted in the methods section) were as rigorous and as transparent as possible.

In this overview, evidence on effectiveness of treatment options for musculoskeletal pain has been presented based on pain presentations at different body regions rather than on specific clinical diagnoses given available evidence of similarity of patient characteristics, prognosis and clinical course of musculoskeletal pain presentations irrespective of specific clinical diagnoses [23, 24]. However, information on the specific clinical diagnosis for which evidence was derived is indicated in the supplementary evidence S1–S7 Tables.

Practice implications

Across health systems globally, there is wide variation in clinical management of musculoskeletal pain patients whereby the most effective treatment options are not consistently used, leading to inefficient care, unnecessary costs and in some cases harm [3, 192]. In a clinical field with so many treatment options, this summary of evidence provides patients, clinicians, managers, policy-makers, and researchers with a helpful "one stop" overview of the evidence for treatments.

In this review, despite an extensive search for evidence, there was a paucity of evidence on treatment for those with multi-site pain. This musculoskeletal pain presentation, often managed as chronic widespread pain and / or fibromyalgia has been less examined in the literature because effectiveness of most treatment options has traditionally been compared on a pairwise basis and according to individual regional pain presentations. However, regional pains are known to co-exist in individual patients [84]. Patients included in most of the studies addressing management of single site pain are likely to have pain in other sites as well. Hence future research needs to investigate interventions that address these multiple sites of pain, in order to better inform clinical practice.

The lack of information regarding patient subgroups most likely to respond to specific treatment options, equivocal recommendations on the optimal mode of treatments, as well as the obvious focus of treatment approaches on single pain sites rather than the individual with multi-site musculoskeletal pain are key specific gaps in the current body of knowledge identified in this review.

Conclusions

Effective healthcare depends on high quality evidence. Best available evidence shows that patients with musculoskeletal pain problems in primary care can be managed effectively with non-pharmacological treatments such as self-management advice, exercise therapy, and psychosocial interventions. Pharmacological interventions such as corticosteroid injections (for knee and shoulder pain) were shown to be effective treatment options for the short-term relief of musculoskeletal pain and may be used in addition to non-pharmacological treatments. NSAIDs and opioids also offer short-term benefit for musculoskeletal pain, but the potential for adverse effects must be considered. Furthermore, the optimal treatment intensity, methods of application, amount of clinical contact, and type of provider or setting, are unclear for most treatment options.

Supporting information

S1 File. MSK systematic review search strategy.
(DOCX)

S2 File. PRISMA checklist.
(DOC)

S1 Table. MSK Self management advice & education.
(DOCX)

S2 Table. MSK exercise therapy.
(DOCX)

S3 Table. MSK manual therapy.
(DOCX)

S4 Table. MSK pharmacological therapy.
(DOCX)

S5 Table. MSK_other therapies.
(DOCX)

S6 Table. MSK psychosocial Rx.
(DOCX)

S7 Table. MSK surgery.
(DOCX)

Author Contributions

Data curation: OOB JJ DAVdW.

Formal analysis: OOB JJ DAVdW.

Funding acquisition: NEF.

Investigation: OOB JJ DAVdW JCH NEF JP.

Methodology: OOB JJ DAVdW.

Project administration: DAVdW JP.

Resources: OOB JJ DAVdW JCH NEF JP.

Supervision: DAVdW JP.

Validation: OOB JJ DAVdW JCH NEF JP.

Visualization: OOB JJ DAVdW JCH NEF JP.

Writing – original draft: OOB.

Writing – review & editing: OOB JJ DAVdW JCH NEF JP.

References

1. Arthritis and Osteoporosis Victoria & Deloitte Access Economics. A problem worth solving: the rising cost of musculoskeletal conditions in Australia Elsternwick [Victoria], 2013. http://www.arthritisvic.org.au/Research/AOV-Funded-Research/Completed/A-Problem-Worth-Solving/APWS.aspx. Accessed 24th April 2015.
2. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 2013, 380(9859):2163–2196.
3. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bulletin of the World Health Organization 2003, 81(9):646–656. PMID: 14710506
4. Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. BMJ 2008, 337:a171. https://doi.org/10.1136/bmj.a171 PMID: 18614473
5. Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. Pain 2005, 117(3):304–313. https://doi.org/10.1016/j.pain.2005.06.017 PMID: 16153772
6. Strunce JB, Boyles RE: Can we shoot for better for patients with shoulder pain? Physical Therapy Reviews 2012, 17(2):132–133.
7. van der Sande R, Rinkel WD, Gebremariam L, Hay EM, Koes BW, Huisstede BM. Subacromial Impingement Syndrome: Effectiveness of Pharmaceutical Interventions–Nonsteroidal Anti-Inflammatory Drugs, Corticosteroid, or Other Injections: A Systematic Review. Archives of physical medicine and rehabilitation 2013, 94(5):961–976. https://doi.org/10.1016/j.apmr.2012.11.041 PMID: 23246416
8. Mallen C, Hay E. Managing back pain and osteoarthritis without paracetamol. Bmj-Brit Med J 2015, 350.
9. Machado GC, Maher CG, Ferreira PH, Pinheiro MB, Lin CW, Day RO, et al. Efficacy and safety of paracetamol for spinal pain and osteoarthritis: systematic review and meta-analysis of randomised placebo controlled trials. BMJ 2015, 350:h1225. https://doi.org/10.1136/bmj.h1225 PMID: 25828856
10. Malanga G, Wolff E: Evidence-informed management of chronic low back pain with nonsteroidal anti-inflammatory drugs, muscle relaxants, and simple analgesics. Spine J 2008, 8(1):173–184. https://doi.org/10.1016/j.spinee.2007.10.013 PMID: 18164465
11. Enthoven WTM, Roelofs PDDM, Deyo RA, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for chronic low back pain. Cochrane Database Syst Rev 2016 (2): CD012087. https://doi.org/10.1002/14651858.CD012087 PMID: 26863524
12. Foster NE, Dziedzic KS, van der Windt DA, Fritz JM, Hay EM. Research priorities for non-pharmacological therapies for common musculoskeletal problems: nationally and internationally agreed recommendations. BMC Musculoskelet Disord 2009, 10:3.
13. Green S, Buchbinder R, Hetrick S. Physiotherapy interventions for shoulder pain. Cochrane Database Syst Rev 2003(2):CD004258. https://doi.org/10.1002/14651858.CD004258 PMID: 12804509
14. Mitchell C, Adebajo A, Hay E, Carr A. Shoulder pain: diagnosis and management in primary care. BMJ 2005, 331(7525):1124–1128. https://doi.org/10.1136/bmj.331.7525.1124 PMID: 16282408
15. Bennell KL, Hinman RS. A review of the clinical evidence for exercise in osteoarthritis of the hip and knee. Journal of Science and Medicine in Sport 2011, 14(1):4–9. https://doi.org/10.1016/j.jsams.2010.08.002 PMID: 20851051

16. National Institute for Health and Care Excellence (NICE) 2014. Osteoarthritis: care and management in adults. https://www.nice.org.uk/guidance/cg177?unlid=38189754201622015349. Accessed 30th June 2014.
17. Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary rehabilitation for chronic low back pain: systematic review. BMJ 2001, 322(7301):1511–1516. PMID: 11420271
18. Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev 2005(3):CD000335. https://doi.org/10.1002/14651858.CD000335.pub2 PMID: 16034851
19. Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005, 142(9):776–785. PMID: 15867410
20. National Institute for Health and Care Excellence (NICE) 2009. Low back pain: early management of persistent non-specific low back pain. https://www.nice.org.uk/guidance/cg88?unlid=562639842201511301453 5. Accessed 18th July 2014.
21. Hartvigsen J, Davidsen M, Hestbaek L, Sogaard K, Roos EM. Patterns of musculoskeletal pain in the population: a latent class analysis using a nationally representative interviewer-based survey of 4817 Danes. Eur J Pain 2013, 17(3):452–460. https://doi.org/10.1002/j.1532-2149.2012.00225.x PMID: 23042697
22. Kamaleri Y, Natvig B, Ihlebaek CM, Bruusgaard D. Localized or widespread musculoskeletal pain: does it matter? Pain 2008, 138(1):41–46. https://doi.org/10.1016/j.pain.2007.11.002 PMID: 18077092
23. Mallen CD, Peat G, Thomas E, Dunn KM, Croft PR. Prognostic factors for musculoskeletal pain in primary care: a systematic review. Br J Gen Pract 2007, 57(541):655–661. PMID: 17688762
24. Henschke N, Ostelo RW, Terwee CB, van der Windt DA. Identifying generic predictors of outcome in patients presenting to primary care with nonspinal musculoskeletal pain. Arthritis care & research 2012, 64(8):1217–1224.
25. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord 2010, 11:144. https://doi.org/10.1186/1471-2474-11-144 PMID: 20598124
26. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007, 7:10. https://doi.org/10.1186/1471-2288-7-10 PMID: 17302989
27. Cohen J. Statistical power analysis for the behavioral sciences 2nd edn. New Jersey Lawrence Erlbaum; 1988.
28. McGough JJ, Faraone SV. Estimating the size of treatment effects: moving beyond p values. Psychiatry (Edgmont) 2009, 6(10):21–29.
29. British Pain Society (BPS) 2013. Guidelines for Pain Management Programmes for adults. https://www.britishpainsociety.org/static/uploads/resources/files/pmp2013_main_FINAL_v6.pdf. Accessed 15th July 2016.
30. Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. Cochrane Database Syst Rev 2010(6):CD007612. https://doi.org/10.1002/14651858.CD007612.pub2 PMID: 20556780
31. Engers A, Jellema P, Wensing M, van der Windt DA, Grol R, van Tulder MW. Individual patient education for low back pain. Cochrane Database Syst Rev 2008(1):CD004057. https://doi.org/10.1002/14651858.CD004057.pub3 PMID: 18254037
32. Hagen KB, Hilde G, Jamtvedt G, Winnem M. Bed rest for acute low-back pain and sciatica. Cochrane Database Syst Rev 2004(4):CD001254. https://doi.org/10.1002/14651858.CD001254.pub2 PMID: 15495012
33. Poquet N, Lin CWC, Heymans MW, van Tulder MW, Esmail R, Koes BW, Maher CG. Back schools for acute and subacute non-specific low-back pain. Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD008325.
34. Heymans MW, van Tulder MW, Esmail R, Bombardier C, Koes BW. Back schools for nonspecific low back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine (Phila Pa 1976) 2005, 30(19):2153–2163.
35. Liddle SD, Gracey JH, Baxter GD. Advice for the management of low back pain: a systematic review of randomised controlled trials. Man Ther 2007, 12(4):310–327. https://doi.org/10.1016/j.math.2006.12.009 PMID: 17395522
36. Mannerkorpi K, Henriksson C. Non-pharmacological treatment of chronic widespread musculoskeletal pain. Best Pract Res Clin Rheumatol 2007, 21(3):513–534. https://doi.org/10.1016/j.berh.2007.04.001 PMID: 17602997

37. Oliveira VC, Ferreira PH, Maher CG, Pinto RZ, Refshauge KM, Ferreira ML. Effectiveness of self-management of low back pain: systematic review with meta-analysis. Arthritis Care Res (Hoboken) 2012, 64(11):1739–1748.

38. Map of Medicine. Care pathways 2014. http://mapofmedicine.com/access-map/. Accessed 18th July 2016.

39. Bartels EM, Juhl CB, Christensen R, Hagen KB, Danneskiold-Samsøe B, Dagfinrud H, et al. Aquatic exercise for the treatment of knee and hip osteoarthritis. Cochrane Database Syst Rev 2016 (3): CD005523. https://doi.org/10.1002/14651858.CD005523.pub3 PMID: 27007113

40. Bidonde J, Busch AJ, Webber SC, Schachter CL, Danyliw A, Overend TJ, et al. Aquatic exercise training for fibromyalgia. Cochrane Database Syst Rev 2014, 10:CD011336.

41. Busch AJ, Barber KA, Overend TJ, Peloso PM, Schachter CL. Exercise for treating fibromyalgia syndrome. Cochrane Database Syst Rev 2007(4):CD003786. https://doi.org/10.1002/14651858.CD003786.pub2 PMID: 17943797

42. Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. Cochrane Database Syst Rev 2013, 12:CD010884.

43. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. Cochrane Database of Syst Rev 2015(1): CD004376. https://doi.org/10.1002/14651858.CD004376.pub3 PMID: 25569281

44. Gross A, Kay TM, Paquin JP, Blanchette S, Lalonde P, Christie T, et al. Exercises for mechanical neck disorders. Cochrane Database Syst Rev 2015, 1:CD004250. https://doi.org/10.1002/14651858.CD004250.pub5 PMID: 25629215

45. Page MJ, Green S, Kramer S, Johnston RV, McBain B, Chau M, Buchbinder R. Manual therapy and exercise for adhesive capsulitis (frozen shoulder). Cochrane Database Syst Rev 2014, 8:CD011275.

46. van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SM, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. Cochrane Database Syst Rev 2015, 1:CD010387. https://doi.org/10.1002/14651858.CD010387.pub2 PMID: 25603546

47. New Zealand Guidelines Group (NZGG). The diagnosis and management of soft tissue shoulder injuries and related disorders. 2003. New Zealand. http://www.acc.co.nz/PRD_EXT_CSMP/groups/external_communications/documents/guide/wcm001684.pdf. Accessed 12th December 2013.

48. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012, 64(4):465–474.

49. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010, 18(4):476–499. https://doi.org/10.1016/j.joca.2010.01.013 PMID: 20170770

50. Aladro-Gonzalvo AR, Araya-Vargas GA, Machado-Diaz M, Salazar-Rojas W. Pilates-based exercise for persistent, non-specific low back pain and associated functional disability: a meta-analysis with meta-regression. J Body Mov Ther 2013, 17(1):125–136.

51. Bidonde J, Busch AJ, Bath B, Milosavljevic S. Exercise for adults with fibromyalgia: an umbrella systematic review with synthesis of best evidence. Curr Rheumatol Rev 2014, 10(1):45–79. PMID: 25229499

52. Choi BKL, Verbeek JH, Tam WWS, Jiang JY. Exercises for prevention of recurrences of low-back pain. Occupational and Environmental Medicine 2010, 67(11):795–796. https://doi.org/10.1136/oem.2010.059873 PMID: 20959397

53. Cleland J, Schulte C, Durall C. The role of therapeutic exercise in treating instability-related lumbar spine pain: A systematic review. J Back Musculoskelet Rehabil 2002, 16(2):105–115. PMID: 22387407

54. Desmeules F, Cote CH, Fremont P. Therapeutic exercise and orthopedic manual therapy for impingement syndrome: a systematic review. Clin J Sport Med 2003, 13(3):176–182. PMID: 12792213

55. Grant HJ, Arthur A, Pichora DR. Evaluation of interventions for rotator cuff pathology: a systematic review. J Hand Ther 2004, 17(2):274–299. https://doi.org/10.1197/j.jht.2004.02.013 PMID: 15162111

56. Hanratty CE, McVeigh JG, Kerr DP, Basford JR, Finch MB, Pendleton A, et al. The effectiveness of physiotherapy exercises in subacromial impingement syndrome: a systematic review and meta-analysis. Semin Arthritis Rheum 2012, 42(3):297–316. https://doi.org/10.1016/j.semarthrit.2012.03.015 PMID: 22607807

57. Hauser W, Klose P, Langhorst J, Moradi B, Steinbach M, Schiltenwolf M, et al. Efficacy of different types of aerobic exercise in fibromyalgia syndrome: a systematic review and meta-analysis of

randomised controlled trials. Arthritis Res Ther 2010, 12(3):R79. https://doi.org/10.1186/ar3002 PMID: 20459730

58. Hurwitz EL, Carragee EJ, van der Velde G, Carroll LJ, Nordin M, Guzman J, et al. Treatment of neck pain: noninvasive interventions: results of the Bone and Joint Decade 2000–2010 Task Force on Neck Pain and Its Associated Disorders. Spine (Phila Pa 1976) 2008, 33(4 Suppl):S123–152.

59. Kromer TO, Tautenhahn UG, de Bie RA, Staal JB, Bastiaenen CHG. Effects of Physiotherapy in Patients with Shoulder Impingement Syndrome: A Systematic Review of the Literature. Journal of Rehabilitation Medicine 2009, 41(11):870–880. https://doi.org/10.2340/16501977-0453 PMID: 19841837

60. Kuhn JE. Exercise in the treatment of rotator cuff impingement: a systematic review and a synthesized evidence-based rehabilitation protocol. J Shoulder Elbow Surg 2009, 18(1):138–160. https://doi.org/10.1016/j.jse.2008.06.004 PMID: 18835532

61. Littlewood C, Ashton J, Chance-Larsen K, May S, Sturrock B. Exercise for rotator cuff tendinopathy: a systematic review. Physiotherapy 2012, 98(2):101–109. https://doi.org/10.1016/j.physio.2011.08.002 PMID: 22507359

62. Macedo LG, Maher CG, Latimer J, McAuley JH. Motor control exercise for persistent, nonspecific low back pain: a systematic review. Physical therapy 2009, 89(1):9–25. https://doi.org/10.2522/ptj.20080103 PMID: 19056854

63. Marinko LN, Chacko JM, Dalton D, Chacko CC. The effectiveness of therapeutic exercise for painful shoulder conditions: a meta-analysis. J Shoulder Elbow Surg 2011, 20(8):1351–1359. https://doi.org/10.1016/j.jse.2011.05.013 PMID: 21889366

64. Murphy RJ, Carr AJ. Shoulder pain. BMJ Clin Evid 2010, 2010.

65. Pereira LM, Obara K, Dias JM, Menacho MO, Guariglia DA, Schiavoni D, et al. Comparing the Pilates method with no exercise or lumbar stabilization for pain and functionality in patients with chronic low back pain: systematic review and meta-analysis. Clinical Rehabilitation 2012, 26(1):10–20. https://doi.org/10.1177/0269215511411113 PMID: 21856719

66. Teasell RW, McClure JA, Walton D, Pretty J, Salter K, Meyer M, et al. A research synthesis of therapeutic interventions for whiplash-associated disorder: part 1—overview and summary. Pain Res Manag 2010, 15(5):287–294. PMID: 21038007

67. Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, et al. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. BMJ 2013, 347:f5555. https://doi.org/10.1136/bmj.f5555 PMID: 24055922

68. Waller B, Lambeck J, Daly D. Therapeutic aquatic exercise in the treatment of low back pain: a systematic review. Clinical Rehabilitation 2009, 23(1):3–14. https://doi.org/10.1177/0269215508097856 PMID: 19114433

69. Wang XQ, Zheng JJ, Yu ZW, Bi X, Lou SJ, Liu J, et al. A meta-analysis of core stability exercise versus general exercise for chronic low back pain. PLoS One 2012, 7(12):e52082. https://doi.org/10.1371/journal.pone.0052082 PMID: 23284879

70. Wells C, Kolt GS, Marshall P, Hill B, Bialocerkowski A. Effectiveness of Pilates exercise in treating people with chronic low back pain: a systematic review of systematic reviews. BMC Med Res Methodol 2013, 13:7.

71. Steffens D, Maher C, Pereira L, Stevens M, Oliveira V, Chapple M, et al. Prevention of Low Back Pain: A Systematic Review and Meta-analysis. JAMA internal medicine 2016:1–10.

72. Rasmussen-Barr E, Ang B, Arvidsson I, Nilsson-Wikmar L. Graded exercise for recurrent low-back pain: a randomized, controlled trial with 6-, 12-, and 36-month follow-ups. Spine (Phila Pa 1976) 2009, 34(3):221–228.

73. Kuczynski JJ, Schwieterman B, Columber K, Knupp D, Shaub L, Cook CE. Effectiveness of physical therapist administered spinal manipulation for the treatment of low back pain: a systematic review of the literature. Int J Sports Phys Ther 2012, 7(6):647–662. PMID: 23316428

74. Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal manipulative therapy for low-back pain. Cochrane Database Syst Rev 2013, 1:CD000447.

75. Rubinstein SM, Terwee CB, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low-back pain. Cochrane Database Syst Rev 2012, 9:CD008880.

76. Walker BF, French SD, Grant W, Green S. Combined chiropractic interventions for low-back pain. Cochrane Database Syst Rev 2010(4):CD005427. https://doi.org/10.1002/14651858.CD005427.pub2 PMID: 20393942

77. Wegner I, Widyahening IS, van Tulder MW, Blomberg SE, de Vet HC, Bronfort G, et al. Traction for low-back pain with or without sciatica. Cochrane Database Syst Rev 2013, 8:CD003010.

78. Brosseau L, Wells GA, Poitras S, Tugwell P, Casimiro L, Novikov M, et al. Ottawa Panel evidence-based clinical practice guidelines on therapeutic massage for low back pain. Journal of bodywork and movement therapies 2012, 16(4):424–455. https://doi.org/10.1016/j.jbmt.2012.04.002 PMID: 23036876

79. Bokarius AV, Bokarius V. Evidence-based review of manual therapy efficacy in treatment of chronic musculoskeletal pain. Pain Pract 2010, 10(5):451–458. https://doi.org/10.1111/j.1533-2500.2010.00377.x PMID: 20412502

80. Brantingham JW, Cassa TK, Bonnefin D, Jensen M, Globe G, Hicks M, et al. Manipulative therapy for shoulder pain and disorders: expansion of a systematic review. J Manipulative Physiol Ther 2011, 34 (5):314–346. https://doi.org/10.1016/j.jmpt.2011.04.002 PMID: 21640255

81. Bronfort G, Haas M, Evans R, Leininger B, Triano J. Effectiveness of manual therapies: the UK evidence report. Chiropr Osteopat 2010, 18:3.

82. Brudvig TJ, Kulkarni H, Shah S. The effect of therapeutic exercise and mobilization on patients with shoulder dysfunction: a systematic review with meta-analysis. J Orthop Sports Phys Ther 2011, 41 (10):734–748. https://doi.org/10.2519/jospt.2011.3440 PMID: 21891875

83. Camarinos J, Marinko L. Effectiveness of manual physical therapy for painful shoulder conditions: a systematic review. J Man Manip Ther 2009, 17(4):206–215. https://doi.org/10.1179/106698109791352076 PMID: 20140151

84. Carnes D, Mars TS, Mullinger B, Froud R, Underwood M. Adverse events and manual therapy: a systematic review. Man Ther 2010, 15(4):355–363. https://doi.org/10.1016/j.math.2009.12.006 PMID: 20097115

85. Ernst E, Harkness E. Spinal manipulation: a systematic review of sham-controlled, double-blind, randomized clinical trials. J Pain Symptom Manage 2001, 22(4):879–889. PMID: 11576805

86. Furlan AD, Giraldo M, Baskwill A, Irvin E, Imamura M. Massage for low-back pain. Cochrane Database Syst Rev 2015, 9:CD001929.

87. Gebremariam L, Hay EM, van der Sande R, Rinkel WD, Koes BW, Huisstede BM. Subacromial impingement syndrome—effectiveness of physiotherapy and manual therapy. Br J Sports Med 2014, 48(16):1202–1208. https://doi.org/10.1136/bjsports-2012-091802 PMID: 24217037

88. Goertz CM, Pohlman KA, Vining RD, Brantingham JW, Long CR. Patient-centered outcomes of high-velocity, low-amplitude spinal manipulation for low back pain: a systematic review. J Electromyogr Kinesiol 2012, 22(5):670–691. https://doi.org/10.1016/j.jelekin.2012.03.006 PMID: 22534288

89. Ho CY, Sole G, Munn J. The effectiveness of manual therapy in the management of musculoskeletal disorders of the shoulder: a systematic review. Man Ther 2009, 14(5):463–474. https://doi.org/10.1016/j.math.2009.03.008 PMID: 19467911

90. Slater SL, Ford JJ, Richards MC, Taylor NF, Surkitt LD, Hahne AJ. The effectiveness of sub-group specific manual therapy for low back pain: a systematic review. Man Ther 2012, 17(3):201–212. https://doi.org/10.1016/j.math.2012.01.006 PMID: 22386046

91. Patel KC, Gross A, Graham N, Goldsmith CH, Ezzo J, Morien A, et al. Massage for mechanical neck disorders. Cochrane Database Syst Rev 2012, 9:CD004871.

92. Rubinstein SM, Terwee CB, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low back pain: an update of the cochrane review. Spine (Phila Pa 1976) 2013, 38(3):E158–177.

93. Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain: an update of a Cochrane review. Spine (Phila Pa 1976) 2011, 36 (13):E825–846.

94. Moore RA, Derry S, Wiffen PJ, Straube S, Aldington DJ. Overview review: Comparative efficacy of oral ibuprofen and paracetamol (acetaminophen) across acute and chronic pain conditions. Eur J Pain 2015, 19(9):1213–1223. https://doi.org/10.1002/ejp.649 PMID: 25530283

95. Siler AC, Gardner H, Yanit K, Cushman T, McDonagh M. Systematic review of the comparative effectiveness of antiepileptic drugs for fibromyalgia. J Pain 2011, 12(4):407–415. https://doi.org/10.1016/j.jpain.2010.09.007 PMID: 21146466

96. Cepeda MS, Camargo F, Zea C, Valencia L. Tramadol for osteoarthritis. Cochrane Database Syst Rev 2006(3):CD005522. https://doi.org/10.1002/14651858.CD005522.pub2 PMID: 16856101

97. Cording M, Derry S, Phillips T, Moore RA, Wiffen PJ. Milnacipran for pain in fibromyalgia in adults. Cochrane Database Syst Rev 2015 (10): CD008244. https://doi.org/10.1002/14651858.CD008244.pub3 PMID: 26482422

98. Derry S, Matthews PR, Wiffen PJ, Moore RA. Salicylate-containing rubefacients for acute and chronic musculoskeletal pain in adults. Cochrane Database Syst Rev 2014, 11:CD007403.

99. Derry S, Moore RA, Gaskell H, McIntyre M, Wiffen PJ. Topical NSAIDs for acute musculoskeletal pain in adults. Cochrane Database Syst Rev 2015, 6:CD007402.

100. Towheed TE, Maxwell L, Judd MG, Catton M, Hochberg MC, Wells G. Acetaminophen for osteoarthritis. Cochrane Database Syst Rev 2006(1):CD004257. https://doi.org/10.1002/14651858.CD004257.pub2 PMID: 16437479

101. Johansson K, Oberg B, Adolfsson L, Foldevi M. A combination of systematic review and clinicians' beliefs in interventions for subacromial pain. Br J Gen Pract 2002, 52(475):145–152. PMID: 11885825

102. da Costa BR, Nuesch E, Reichenbach S, Juni P, Rutjes AW. Doxycycline for osteoarthritis of the knee or hip. Cochrane Database Syst Rev 2012, 11:CD007323. https://doi.org/10.1002/14651858.CD007323.pub3 PMID: 23152242

103. Davies RA, Maher CG, Hancock MJ. A systematic review of paracetamol for non-specific low back pain. Eur Spine J 2008, 17(11):1423–1430. https://doi.org/10.1007/s00586-008-0783-x PMID: 18797937

104. Heintjes E, Berger MY, Bierma-Zeinstra SM, Bernsen RM, Verhaar JA, Koes BW. Pharmacotherapy for patellofemoral pain syndrome. Cochrane Database Syst Rev 2004(3):CD003470. https://doi.org/10.1002/14651858.CD003470.pub2 PMID: 15266488

105. Kuijpers T, van Middelkoop M, Rubinstein SM, Ostelo R, Verhagen A, Koes BW, et al. A systematic review on the effectiveness of pharmacological interventions for chronic non-specific low-back pain. Eur Spine J 2011, 20(1):40–50. https://doi.org/10.1007/s00586-010-1541-4 PMID: 20680369

106. Chaparro LE, Furlan AD, Deshpande A, Mailis-Gagnon A, Atlas S, Turk DC. Opioids compared to placebo or other treatments for chronic low-back pain. Cochrane Database Syst Rev 2013, 8:CD004959.

107. Deshpande A, Furlan A, Mailis-Gagnon A, Atlas S, Turk D. Opioids for chronic low-back pain. Cochrane Database Syst Rev 2007(3):CD004959. https://doi.org/10.1002/14651858.CD004959.pub3 PMID: 17636781

108. Martell BA, O'Connor PG, Kerns RD, Becker WC, Morales KH, Kosten TR, et al. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. Ann Intern Med 2007, 146(2):116–127. PMID: 17227935

109. Stanos SP, Galluzzi KE. Topical therapies in the management of chronic pain. Postgrad Med 2013, 125(4 Suppl 1):25–33.

110. Mason L, Moore RA, Edwards JE, Derry S, McQuay HJ. Topical NSAIDs for chronic musculoskeletal pain: systematic review and meta-analysis. BMC Musculoskelet Disord 2004, 5:28. https://doi.org/10.1186/1471-2474-5-28 PMID: 15317652

111. Mason L, Moore RA, Edwards JE, Derry S, McQuay HJ. Topical NSAIDs for acute pain: a meta-analysis. BMC Fam Pract 2004, 5:10. https://doi.org/10.1186/1471-2296-5-10 PMID: 15147585

112. American Academy of Orthopaedic Surgeons (AAOS). Treatment of Osteoarthritis of the Knee: Evidence-Based Guideline. 2013.

113. Institute for Clinical Systems Improvement (ICSI). Assessment and Management of Chronic Pain. In. Bloomington,MN, US: ICSI; 2011.

114. Staal JB, de Bie R, De Vet H, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low-back pain. Cochrane Database Syst Rev 2008, 3(3).

115. Bellamy N, Campbell J, Robinson V, Gee T, Bourne R, Wells G. Intraarticular corticosteroid for treatment of osteoarthritis of the knee. Cochrane Database Syst Rev 2006(2):CD005328. https://doi.org/10.1002/14651858.CD005328.pub2 PMID: 16625636

116. Buchbinder R, Green S, Youd JM. Corticosteroid injections for shoulder pain. Cochrane Database Syst Rev 2003(1):CD004016. https://doi.org/10.1002/14651858.CD004016 PMID: 12535501

117. Buchbinder R, Green S, Youd JM, Johnston RV. Oral steroids for adhesive capsulitis. Cochrane Database Syst Rev 2006(4):CD006189. https://doi.org/10.1002/14651858.CD006189 PMID: 17054278

118. Waseem Z, Boulias C, Gordon A, Ismail F, Sheean G, Furlan AD. Botulinum toxin injections for low-back pain and sciatica. Cochrane Database Syst Rev 2011(1):CD008257. https://doi.org/10.1002/14651858.CD008257.pub2 PMID: 21249702

119. Arroll B, Goodyear-Smith F. Corticosteroid injections for painful shoulder: a meta-analysis. British Journal of General Practice 2005, 55(512):224–228. PMID: 15808040

120. Blanchard V, Barr S, Cerisola FL. The effectiveness of corticosteroid injections compared with physiotherapeutic interventions for adhesive capsulitis: A systematic review. Physiotherapy 2010, 96(2):95–107. https://doi.org/10.1016/j.physio.2009.09.003 PMID: 20420956

121. Bono CM, Ghiselli G, Gilbert TJ, Kreiner DS, Reitman C, Summers JT, et al. An evidence-based clinical guideline for the diagnosis and treatment of cervical radiculopathy from degenerative disorders. Spine Journal 2011, 11(1):64–72. https://doi.org/10.1016/j.spinee.2010.10.023 PMID: 21168100

122. Diwan SA, Manchikanti L, Benyamin RM, Bryce DA, Geffert S, Hameed H, et al. Effectiveness of Cervical Epidural Injections in the Management of Chronic Neck and Upper Extremity Pain. Pain Physician 2012, 15(4):E405–E434. PMID: 22828692

123. Langevin P, Peloso PMJ, Lowcock J, Nolan M, Weber J, Gross A, et al. Botulinum toxin for subacute/chronic neck pain. Cochrane Database of Systematic Reviews 2011(7).

124. Peloso PM, Khan M, Gross AR, Carlesso L, Santaguida L, Lowcock J, et al. Pharmacological Interventions Including Medical Injections for Neck Pain: An Overview as Part of the ICON Project. Open Orthop J 2013, 7:473–493. https://doi.org/10.2174/1874325001307010473 PMID: 24155805

125. Shah N, Lewis M. Shoulder adhesive capsulitis: systematic review of randomised trials using multiple corticosteroid injections. Br J Gen Pract 2007, 57(541):662–667. PMID: 17688763

126. Benyamin RM, Manchikanti L, Parr AT, Diwan S, Singh V, Falco F, et al. The effectiveness of lumbar interlaminar epidural injections in managing chronic low back and lower extremity pain. Pain Physician 2012, 15(4):E363–404. PMID: 22828691

127. Cohen SP, Bicket MC, Jamison D, Wilkinson I, Rathmell JP. Epidural steroids: a comprehensive, evidence-based review. Regional anesthesia and pain medicine 2013, 38(3):175–200. https://doi.org/10.1097/AAP.0b013e31828ea086 PMID: 23598728

128. Juni P, Hari R, Rutjes AW, Fischer R, Silletta MG, Reichenbach S, et al. Intra-articular corticosteroid for knee osteoarthritis. Cochrane Database Syst Rev 2015, 10:CD005328.

129. National Health and Medical Research Council (NHMRC). Guideline for the Non-surgical Management of Hip and Knee Osteoarthritis In: Guideline for the Non-surgical Management of Hip and Knee Osteoarthritis Edited by The Royal Australian College of General Practitioners; 2009.

130. Duivenvoorden T, Brouwer RW, van Raaij TM, Verhagen AP, Verhaar JAN, Bierma-Zeinstra SMA. Braces and orthoses for treating osteoarthritis of the knee. Cochrane Database Syst Rev 2015 (3): CD004020. https://doi.org/10.1002/14651858.CD004020.pub3 PMID: 25773267

131. D'Hondt N E, Aufdemkampe G, Kerkhoffs GM, Struijs PA, Verheul C, van Dijk CN. Orthotic devices for treating patellofemoral pain syndrome. Cochrane Database Syst Rev 2009(1):CD002267. https://doi.org/10.1002/14651858.CD002267.pub2 PMID: 19160208

132. Hossain M, Alexander P, Burls A, Jobanputra P. Foot orthoses for patellofemoral pain in adults. Cochrane Database Syst Rev 2011(1):CD008402. https://doi.org/10.1002/14651858.CD008402.pub2 PMID: 21249707

133. Carroll LJ, Cassidy JD, Peloso PM, Giles-Smith L, Cheng CS, Greenhalgh SW, et al. Methods for the best evidence synthesis on neck pain and its associated disorders: the Bone and Joint Decade 2000–2010 Task Force on Neck Pain and Its Associated Disorders. J Manipulative Physiol Ther 2009, 32(2 Suppl):S39–45.

134. Warden SJ, Hinman RS, Watson MA Jr., Avin KG, Bialocerkowski AE, Crossley KM. Patellar taping and bracing for the treatment of chronic knee pain: a systematic review and meta-analysis. Arthritis Rheum 2008, 59(1):73–83. https://doi.org/10.1002/art.23242 PMID: 18163413

135. Cameron M, Chrubasik S. Topical herbal therapies for treating osteoarthritis. Cochrane Database Syst Rev 2013, 5:CD010538.

136. Brosseau L, Welch V, Wells G, deBie R, Gam A, Harman K, et al. Low level laser therapy (Classes I, II and III) for treating osteoarthritis. Cochrane Database Syst Rev 2003(2):CD002046. https://doi.org/10.1002/14651858.CD002046 PMID: 12804422

137. Brosseau L, Yonge KA, Robinson V, Marchand S, Judd M, Wells G, et al. Thermotherapy for treatment of osteoarthritis. Cochrane Database Syst Rev 2003(4):CD004522. https://doi.org/10.1002/14651858.CD004522 PMID: 14584019

138. French SD, Cameron M, Walker BF, Reggars JW, Esterman AJ. Superficial heat or cold for low back pain. Cochrane Database Syst Rev 2006(1):CD004750. https://doi.org/10.1002/14651858.CD004750.pub2 PMID: 16437495

139. Furlan AD, van Tulder MW, Cherkin DC, Tsukayama H, Lao L, Koes BW, et al. Acupuncture and dry-needling for low back pain. Cochrane Database Syst Rev 2005(1):CD001351. https://doi.org/10.1002/14651858.CD001351.pub2 PMID: 15674876

140. Green S, Buchbinder R, Hetrick S. Acupuncture for shoulder pain. Cochrane Database Syst Rev 2005 (2):CD005319. https://doi.org/10.1002/14651858.CD005319 PMID: 15846753

141. Khadilkar A, Odebiyi DO, Brosseau L, Wells GA. Transcutaneous electrical nerve stimulation (TENS) versus placebo for chronic low-back pain. Cochrane Database Syst Rev 2008(4):CD003008. https://doi.org/10.1002/14651858.CD003008.pub3 PMID: 18843638

142. Kroeling P, Gross A, Graham N, Burnie SJ, Szeto G, Goldsmith CH, et al. Electrotherapy for neck pain. Cochrane Database Syst Rev 2013, 8:CD004251.

143. Manheimer E, Cheng K, Linde K, Lao L, Yoo J, Wieland S, et al. Acupuncture for peripheral joint osteoarthritis. Cochrane Database Syst Rev 2010(1):CD001977. https://doi.org/10.1002/14651858.CD001977.pub2 PMID: 20091527

144. Nnoaham KE, Kumbang J. Transcutaneous electrical nerve stimulation (TENS) for chronic pain. Cochrane Database Syst Rev 2008(3):CD003222. https://doi.org/10.1002/14651858.CD003222.pub2 PMID: 18646088

145. Page MJ, Green S, Kramer S, Johnston RV, McBain B, Buchbinder R. Electrotherapy modalities for adhesive capsulitis (frozen shoulder). Cochrane Database Syst Rev 2014, 10:CD011324.

146. Rutjes AW, Nuesch E, Sterchi R, Juni P. Therapeutic ultrasound for osteoarthritis of the knee or hip. Cochrane Database Syst Rev 2010(1):CD003132. https://doi.org/10.1002/14651858.CD003132.pub2 PMID: 20091539

147. Rutjes AW, Nuesch E, Sterchi R, Kalichman L, Hendriks E, Osiri M, et al. Transcutaneous electrostimulation for osteoarthritis of the knee. Cochrane Database Syst Rev 2009(4):CD002823. https://doi.org/10.1002/14651858.CD002823.pub2 PMID: 19821296

148. Trinh K, Graham N, Irnich D, Cameron ID, Forget M. Acupuncture for neck disorders. Cochrane Database Syst Rev 2016 (5): CD004870. https://doi.org/10.1002/14651858.CD004870.pub4 PMID: 27145001

149. Hutchinson AJ, Ball S, Andrews JC, Jones GG. The effectiveness of acupuncture in treating chronic non-specific low back pain: a systematic review of the literature. J Orthop Surg Res 2012, 7:36. https://doi.org/10.1186/1749-799X-7-36 PMID: 23111099

150. Johnson M, Martinson M. Efficacy of electrical nerve stimulation for chronic musculoskeletal pain: a meta-analysis of randomized controlled trials. Pain 2007, 130(1–2):157–165. https://doi.org/10.1016/j.pain.2007.02.007 PMID: 17383095

151. Kadhim-Saleh A, Maganti H, Ghert M, Singh S, Farrokhyar F. Is low-level laser therapy in relieving neck pain effective? Systematic review and meta-analysis. Rheumatol Int 2013, 33(10):2493–2501. https://doi.org/10.1007/s00296-013-2742-z PMID: 23579335

152. Lam M, Galvin R, Curry P. Effectiveness of acupuncture for nonspecific chronic low back pain: a systematic review and meta-analysis. Spine (Phila Pa 1976) 2013, 38(24):2124–2138.

153. Maund E, Craig D, Suekarran S, Neilson A, Wright K, Brealey S, et al. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. Health Technol Assess 2012, 16(11):1–264. https://doi.org/10.3310/hta16110 PMID: 22405512

154. Orrock PJ, Myers SP. Osteopathic intervention in chronic non-specific low back pain: a systematic review. BMC Musculoskelet Disord 2013, 14:129. https://doi.org/10.1186/1471-2474-14-129 PMID: 23570655

155. Parkinson L, Sibbritt D, Bolton P, van Rotterdam J, Villadsen I. Well-being outcomes of chiropractic intervention for lower back pain: a systematic review. Clin Rheumatol 2013, 32(2):167–180. https://doi.org/10.1007/s10067-012-2116-z PMID: 23149906

156. Posadzki P, Ernst E. Yoga for low back pain: a systematic review of randomized clinical trials. Clin Rheumatol 2011, 30(9):1257–1262. https://doi.org/10.1007/s10067-011-1764-8 PMID: 21590293

157. Robertson VJ, Baker KG. A review of therapeutic ultrasound: effectiveness studies. Phys Ther 2001, 81(7):1339–1350. PMID: 11444997

158. Vickers AJ, Cronin AM, Maschino AC, Lewith G, MacPherson H, Foster NE, et al. Acupuncture for chronic pain: individual patient data meta-analysis. Arch Intern Med 2012, 172(19):1444–1453. https://doi.org/10.1001/archinternmed.2012.3654 PMID: 22965186

159. Faber E, Kuiper JI, Burdorf A, Miedema HS, Verhaar JA. Treatment of impingement syndrome: a systematic review of the effects on functional limitations and return to work. J Occup Rehabil 2006, 16(1):7–25. https://doi.org/10.1007/s10926-005-9003-2 PMID: 16705497

160. Dubinsky RM, Miyasaki J. Assessment: efficacy of transcutaneous electric nerve stimulation in the treatment of pain in neurologic disorders (an evidence-based review): report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Neurology 2010, 74(2):173–176. https://doi.org/10.1212/WNL.0b013e3181c918fc PMID: 20042705

161. van Middelkoop M, Rubinstein SM, Kuijpers T, Verhagen AP, Ostelo R, Koes BW, et al. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. Eur Spine J 2011, 20(1):19–39. https://doi.org/10.1007/s00586-010-1518-3 PMID: 20640863

162. Seco J, Kovacs FM, Urrutia G. The efficacy, safety, effectiveness, and cost-effectiveness of ultrasound and shock wave therapies for low back pain: a systematic review. Spine J 2011, 11(10):966–977. https://doi.org/10.1016/j.spinee.2011.02.002 PMID: 21482199

163. Koes BW, van Tulder M, Lin C-WC, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. European Spine Journal 2010, 19(12):2075–2094. https://doi.org/10.1007/s00586-010-1502-y PMID: 20602122

164. Henschke N, Ostelo RWJG, van Tulder MW, Vlaeyen JWS, Morley S, Assendelft WJJ, et al. Behavioural treatment for chronic low-back pain. Cochrane Database of Systematic Reviews 2010(7).

165. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RW, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2014, 9:CD000963.

166. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults. Cochrane Database Syst Rev 2003(2):CD002194. https://doi.org/10.1002/14651858.CD002194 PMID: 12804428

167. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary biopsychosocial rehabilitation for subacute low back pain among working age adults. Cochrane Database Syst Rev 2003(2):CD002193. https://doi.org/10.1002/14651858.CD002193 PMID: 12804427

168. Ostelo RW, van Tulder MW, Vlaeyen JW, Linton SJ, Morley SJ, Assendelft WJ. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005(1):CD002014. https://doi.org/10.1002/14651858.CD002014.pub2 PMID: 15674889

169. Bunzli S, Gillham D, Esterman A. Physiotherapy-provided operant conditioning in the management of low back pain disability: A systematic review. Physiother Res Int 2011, 16(1):4–19. https://doi.org/10.1002/pri.465 PMID: 20310071

170. Dixon KE, Keefe FJ, Scipio CD, Perri LM, Abernethy AP. Psychological interventions for arthritis pain management in adults: a meta-analysis. Health Psychol 2007, 26(3):241–250. https://doi.org/10.1037/0278-6133.26.3.241 PMID: 17500610

171. Hoffman BM, Papas RK, Chatkoff DK, Kerns RD. Meta-analysis of psychological interventions for chronic low back pain. Health Psychol 2007, 26(1):1–9. https://doi.org/10.1037/0278-6133.26.1.1 PMID: 17209691

172. Richmond H, Hall AM, Copsey B, Hansen Z, Williamson E, Hoxey-Thomas N, Cooper Z, Lamb SE. The Effectiveness of Cognitive Behavioural Treatment for Non-Specific Low Back Pain: A Systematic Review and Meta-Analysis. Plos One 2015, 10(8).

173. Sveinsdottir V, Eriksen HR, Reme SE. Assessing the role of cognitive behavioral therapy in the management of chronic nonspecific back pain. J Pain Res 2012, 5:371–380. https://doi.org/10.2147/JPR.S25330 PMID: 23091394

174. De Bruijn C, Goossens M, De Bie R, Ament A, Geraets J, Dinant G-J. Cost-effectiveness of an education and activation program for patients with acute and subacute shoulder complaints compared to usual care. International journal of technology assessment in health care 2007, 23(01):80–88.

175. Geraets JJ, Goossens ME, de Groot IJ, de Bruijn CP, de Bie RA, Dinant GJ, et al. Effectiveness of a graded exercise therapy program for patients with chronic shoulder complaints. Aust J Physiother 2005, 51(2):87–94. PMID: 15924511

176. Buchbinder R, Green S, Youd JM, Johnston RV, Cumpston M. Arthrographic distension for adhesive capsulitis (frozen shoulder). Cochrane Database Syst Rev 2008(1):CD007005. https://doi.org/10.1002/14651858.CD007005 PMID: 18254123

177. Coghlan JA, Buchbinder R, Green S, Johnston RV, Bell SN. Surgery for rotator cuff disease. Cochrane Database Syst Rev 2008(1):CD005619. https://doi.org/10.1002/14651858.CD005619.pub2 PMID: 18254085

178. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis. Cochrane Database Syst Rev 2005(4):CD001352. https://doi.org/10.1002/14651858.CD001352.pub3 PMID: 16235281

179. Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse. Cochrane Database Syst Rev 2007(2):CD001350. https://doi.org/10.1002/14651858.CD001350.pub4 PMID: 17443505

180. Jacobs W, Van der Gaag NA, Tuschel A, de Kleuver M, Peul W, Verbout AJ, et al. Total disc replacement for chronic back pain in the presence of disc degeneration. Cochrane Database Syst Rev 2012, 9:CD008326.

181. Nikolaidis I, Fouyas IP, Sandercock PA, Statham PF. Surgery for cervical radiculopathy or myelopathy. Cochrane Database Syst Rev 2010(1):CD001466. https://doi.org/10.1002/14651858.CD001466.pub3 PMID: 20091520

182. Reichenbach S, Rutjes AW, Nuesch E, Trelle S, Juni P. Joint lavage for osteoarthritis of the knee. Cochrane Database Syst Rev 2010(5):CD007320. https://doi.org/10.1002/14651858.CD007320.pub2 PMID: 20464751

183. Singh JA, Sperling J, Buchbinder R, McMaken K. Surgery for shoulder osteoarthritis. Cochrane Database Syst Rev 2010(10):CD008089. https://doi.org/10.1002/14651858.CD008089.pub2 PMID: 20927773

184. Tamaoki MJ, Belloti JC, Lenza M, Matsumoto MH, Gomes Dos Santos JB, Faloppa F. Surgical versus conservative interventions for treating acromioclavicular dislocation of the shoulder in adults. Cochrane Database Syst Rev 2010(8):CD007429. https://doi.org/10.1002/14651858.CD007429.pub2 PMID: 20687087

185. Dorrestijn O, Stevens M, Winters JC, van der Meer K, Diercks RL. Conservative or surgical treatment for subacromial impingement syndrome? A systematic review. J Shoulder Elbow Surg 2009, 18 (4):652–660. https://doi.org/10.1016/j.jse.2009.01.010 PMID: 19286397

186. Gebremariam L, Hay EM, Koes BW, Huisstede BM. Effectiveness of surgical and postsurgical interventions for the subacromial impingement syndrome: a systematic review. Arch Phys Med Rehabil 2011, 92(11):1900–1913. https://doi.org/10.1016/j.apmr.2011.06.006 PMID: 22032225

187. Grant JA, Schroeder N, Miller BS, Carpenter JE. Comparison of manipulation and arthroscopic capsular release for adhesive capsulitis: a systematic review. J Shoulder Elbow Surg 2013, 22(8):1135–1145. https://doi.org/10.1016/j.jse.2013.01.010 PMID: 23510748

188. Huisstede BM, Koes BW, Gebremariam L, Keijsers E, Verhaar JA. Current evidence for effectiveness of interventions to treat rotator cuff tears. Man Ther 2011, 16(3):217–230. https://doi.org/10.1016/j.math.2010.10.012 PMID: 21146445

189. Laupattarakasem P, Laopaiboon M, Kosuwon W, Laupattarakasem W. Meta-analysis comparing bioabsorbable versus metal interference screw for adverse and clinical outcomes in anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 2014, 22(1):142–153. https://doi.org/10.1007/s00167-012-2340-8 PMID: 23238925

190. Maffulli N, Longo UG, Loppini M, Berton A, Spiezia F, Denaro V. Tissue engineering for rotator cuff repair: an evidence-based systematic review. Stem Cells Int 2012, 2012:418086. https://doi.org/10.1155/2012/418086 PMID: 25098365

191. Paxton ES, Stock MV, Brophy RH. Meniscal repair versus partial meniscectomy: a systematic review comparing reoperation rates and clinical outcomes. Arthroscopy 2011, 27(9):1275–1288. https://doi.org/10.1016/j.arthro.2011.03.088 PMID: 21820843

192. Organisation for Economic Co-operation and Development (OECD). Health at a Glance 2013: OECD Indicators. OECD Publishing 2013. https://www.oecd.org/els/health-systems/Health-at-a-Glance-2013.pdf. Accessed 20th June 2016.