Cochrane Library

Cochrane Database of Systematic Reviews

Self-management education programmes for osteoarthritis (Review)

Kroon FPB, van der Burg LRA, Buchbinder R, Osborne RH, Johnston RV, Pitt V

Kroon FPB, van der Burg LRA, Buchbinder R, Osborne RH, Johnston RV, Pitt V.
Self-management education programmes for osteoarthritis.
Cochrane Database of Systematic Reviews 2014, Issue 1. Art. No.: CD008963.
DOI: 10.1002/14651858.CD008963.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Self-management education programmes for osteoarthritis

Féline PB Kroon[1a], Lennart RA van der Burg[1b], Rachelle Buchbinder[2], Richard H Osborne[3], Renea V Johnston[2], Veronica Pitt[4]

[1]Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands. [2]Monash Department of Clinical Epidemiology, Cabrini Hospital, Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Malvern, Australia. [3]Director, Public Health Innovation, Deakin, Deakin University, Burwood, Australia. [4]National Trauma Research Institute, The Alfred Hospital, Monash University, Melbourne, Australia

[a]Joint first author. [b]Joint first author

Contact: Rachelle Buchbinder, Monash Department of Clinical Epidemiology, Cabrini Hospital, Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Suite 41, Cabrini Medical Centre, 183 Wattletree Road, Malvern, Victoria, 3144, Australia. rachelle.buchbinder@monash.edu.

Editorial group: Cochrane Musculoskeletal Group.
Publication status and date: New, published in Issue 1, 2014.

Citation: Kroon FPB, van der Burg LRA, Buchbinder R, Osborne RH, Johnston RV, Pitt V. Self-management education programmes for osteoarthritis. *Cochrane Database of Systematic Reviews* 2014, Issue 1. Art. No.: CD008963. DOI: 10.1002/14651858.CD008963.pub2.

ABSTRACT

Background

Self-management education programmes are complex interventions specifically targeted at patient education and behaviour modification. They are designed to encourage people with chronic disease to take an active self-management role to supplement medical care and improve outcomes.

Objectives

To assess the effectiveness of self-management education programmes for people with osteoarthritis.

Search methods

The Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, PyscINFO, SCOPUS and the World Health Organization (WHO) International Clinical Trial Registry Platform were searched, without language restriction, on 17 January 2013. We checked references of reviews and included trials to identify additional studies.

Selection criteria

Randomised controlled trials of self-management education programmes in people with osteoarthritis were included. Studies with participants receiving passive recipients of care and studies comparing one type of programme versus another were excluded.

Data collection and analysis

In addition to standard methods we extracted components of the self-management interventions using the eight domains of the Health Education Impact Questionnaire (heiQ), and contextual and participant characteristics using PROGRESS-Plus and the Health Literacy Questionnaire (HLQ). Outcomes included self-management of osteoarthritis, participant's positive and active engagement in life, pain, global symptom score, self-reported function, quality of life and withdrawals (including dropouts and those lost to follow-up). We assessed the quality of the body of evidence for these outcomes using the GRADE approach.

Main results

We included twenty-nine studies (6,753 participants) that compared self-management education programmes to attention control (five studies), usual care (17 studies), information alone (four studies) or another intervention (seven studies). Although heterogeneous, most interventions included elements of skill and technique acquisition (94%), health-directed activity (85%) and self-monitoring and

insight (79%); social integration and support were addressed in only 12%. Most studies did not provide enough information to assess all PROGRESS-Plus items. Eight studies included predominantly Caucasian, educated female participants, and only four provided any information on participants' health literacy. All studies were at high risk of performance and detection bias for self-reported outcomes; 20 studies were at high risk of selection bias, 16 were at high risk of attrition bias, two were at high risk of reporting bias and 12 were at risk of other biases. We deemed attention control as the most appropriate and thus the main comparator.

Compared with attention control, self-management programmes may not result in significant benefits at 12 months. Low-quality evidence from one study (344 people) indicates that self-management skills were similar in active and control groups: 5.8 points on a 10-point self-efficacy scale in the control group, and the mean difference (MD) between groups was 0.4 points (95% confidence interval (CI) -0.39 to 1.19). Low-quality evidence from four studies (575 people) indicates that self-management programmes may lead to a small but clinically unimportant reduction in pain: the standardised mean difference (SMD) between groups was -0.26 (95% CI -0.44 to -0.09); pain was 6 points on a 0 to 10 visual analogue scale (VAS) in the control group, treatment resulted in a mean reduction of 0.8 points (95% CI -0.14 to -0.3) on a 10-point scale, with number needed to treat for an additional beneficial outcome (NNTB) of 8 (95% CI 5 to 23). Low-quality evidence from one study (251 people) indicates that the mean global osteoarthritis score was 4.2 on a 0 to 10-point symptom scale (lower better) in the control group, and treatment reduced symptoms by a mean of 0.14 points (95% CI -0.54 to 0.26). This result does not exclude the possibility of a clinically important benefit in some people (0.5 point reduction included in 95% CI). Low-quality evidence from three studies (574 people) showed no signficant difference in function between groups (SMD -0.19, 95% CI -0.5 to 0.11); mean function was 1.29 points on a 0 to 3-point scale in the control group, and treatment resulted in a mean improvement of 0.04 points with self-management (95% CI -0.10 to 0.02). Low-quality evidence from one study (165 people) showed no between-group difference in quality of life (MD -0.01, 95% CI -0.03 to 0.01) from a control group mean of 0.57 units on 0 to 1 well-being scale. Moderate-quality evidence from five studies (937 people) shows similar withdrawal rates between self-management (13%) and control groups (12%): RR 1.11 (95% CI 0.78 to 1.57). Positive and active engagement in life was not measured.

Compared with usual care, moderate-quality evidence from 11 studies (up to 1,706 participants) indicates that self-management programmes probably provide small benefits up to 21 months, in terms of self-management skills, pain, osteoarthritis symptoms and function, although these are of doubtful clinical importance, and no improvement in positive and active engagement in life or quality of life. Withdrawal rates were similar. Low to moderate quality evidence indicates no important differences in self-management , pain, symptoms, function, quality of life or withdrawal rates between self-management programmes and information alone or other interventions (exercise, physiotherapy, social support or acupuncture).

Authors' conclusions

Low to moderate quality evidence indicates that self-management education programmes result in no or small benefits in people with osteoarthritis but are unlikely to cause harm.

Compared with attention control, these programmes probably do not improve self-management skills, pain, osteoarthritis symptoms, function or quality of life, and have unknown effects on positive and active engagement in life. Compared with usual care, they may slightly improve self-management skills, pain, function and symptoms, although these benefits are of unlikely clinical importance.

Further studies investigating the effects of self-management education programmes, as delivered in the trials in this review, are unlikely to change our conclusions substantially, as confounding from biases across studies would have likely favoured self-management. However, trials assessing other models of self-management education programme delivery may be warranted. These should adequately describe the intervention they deliver and consider the expanded PROGRESS-Plus framework and health literacy, to explore issues of health equity for recipients.

PLAIN LANGUAGE SUMMARY

Self-management education programmes for osteoarthritis

This review shows that in people with osteoarthritis:

Self-management education programmes may not improve self-management skills, osteoarthritis symptoms, function, quality of life and dropout rates but may reduce pain modestly compared with attention control. Active and positive engagement in life was not reported.

Self-management education programmes may slightly improve self-management skills, pain and function but may not improve active and positive engagement in life, osteoarthritis symptoms, quality of life and dropout rates compared with usual care.

Self-management education programmes probably do not improve outcomes compared with provision of information alone or compared with other interventions (exercise, physiotherapy, social support or acupuncture).

What is osteoarthritis and what are self-management education programmes?

Osteoarthritis (OA) is a disease of the joints, such as your knee or hip, or the joints in your hands. The joint cartilage that lines the joint gradually thins, narrowing the joint space. In severe cases, no cartilage remains between the bones, and the bones rub together when the joint is moved, making the joint painful and sometimes unstable.

Self-management education programmes are behavioural interventions designed to encourage people with chronic disease to take an active role in the management of their own condition. These programmes aim to improve outcomes for patients by supporting, not replacing, medical care. The content used to educate patients about their condition and to explain how they can best manage their symptoms varies between programmes.

Best estimate of what happens to people with osteoarthritis who undergo self-management programmes:

People who completed a self-management programme rated their self-management skills to be 0.4 points better (0.4 points worse to 1.2 points better) on a scale of 1 to 10 (higher score means better self-management) after 12 months (4% absolute improvement; 4% worse to 12% better).

- People who completed a self-management programme rated their self-management skills as 6.2 points on a scale of 1 to 10.

- People who received attention control rated their self-management skills as 5.8 points on a scale of 1 to 10.

People who completed a self-management programme rated their pain to be 0.8 points lower (0.3 to 0.14 points lower) on a scale of 0 to 10 (lower score means less pain) after 12 months (8% absolute improvement).

- People who completed a self-management programme rated their pain as 5 points on a scale of 0 to 10.

- People who received attention control rated their pain as 5.8 points on a scale of 0 to 10.

People who completed a self-management programme rated their osteoarthritis symptoms to be 0.14 points lower (0.54 points lower to 0.26 points higher) on a scale of 0 to 10 (lower score means fewer symptoms) after 12 months (1% absolute improvement).

- People who completed a self-management programme rated their symptoms as 4.1 points on a scale of 0 to 10.

- People who received attention control rated their symptoms as 4.2 points on a scale of 0 to 10.

People who completed a self-management programme rated their function to be 0.04 points lower (0.02 points lower to 0.10 points higher) on a scale of 0 to 3 (lower score means better function) after 12 months (4% absolute improvement).

- People who completed a self-management programme rated their function as 1.25 points on a scale of 0 to 3.

- People who received attention control rated their function as 1.29 points on a scale of 0 to 3.

People who completed a self-management programme rated their quality of life to be 0.01 points lower (0.03 points lower to 0.01 points higher) on a scale of 0 to 1 (higher score means better quality of life) after 12 months (1% absolute worsening).

- People who completed a self-management programme rated their quality of life as 0.56 points on a scale of 0 to 1.

- People who received attention control rated their quality of life as 0.57 points on a scale of 0 to 1.

One more person out of 100 dropped out of self-management programmes (1% absolute improvement).

- 13 out of 100 people who received a self-management programme dropped out.

- 12 out of 100 people who received attention control dropped out.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. SMP compared to Attention control for osteoarthritis

SMP compared with attention control for osteoarthritis

Patient or population: patients with osteoarthritis
Settings: primary care, or outpatient
Intervention: SMP
Comparison: attention control

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Attention control	SMP				
Self-management of OA Arthritis self-efficacy scale (ASES). Scale from 1 to 10, higher better Follow-up: 12 months	Mean self-management of osteoarthritis in the control groups was **5.8 points**	Mean self-management of osteoarthritis in the intervention groups was **0.4 points higher** (0.4 lower to 1.2 higher)		344 (one study)	⊕⊕⊖⊖ **low** [1,2]	MD 0.4 (-0.39 to 1.19) Absolute mean improvement 4% (4% worse to 12% improved) Relative improvement 7% (7% worse to 21% improved)[3]
Positive and active engagement in life —not measured	See comment	See comment	Not estimable	-	See comment	No studies measured this outcome
Pain Multiple tools[4]. Scale from: 0 to 10, 0 = no pain Follow-up: six to 12 months	Mean pain ranged across control groups from **5.67 to 6.19 points**	Mean pain in the intervention groups was **0.8 points lower** (0.3 to 0.14 lower)		575 (three studies)	⊕⊕⊖⊖ **low** [1,2]	SMD -0.26 (-0.44 to -0.09) Absolute reduction in pain 8% (3% to 14% reduction). Relative reduction in pain 13% (5% to 22% reduction). NNTB = 8 (5 to 23)[5]
Global OA scores AIMS2 (average of physical, affect, and pain subscales). Scale from 0 to 10, lower better Follow-up: nine months	Mean global osteoarthritis symptom score in the control group was **4.22 points**	Mean global osteoarthritis symptom score in the intervention group was **0.14 points lower** (0.54 lower to 0.26 higher)		251 (one study)	⊕⊕⊖⊖ **low** [2,6]	Absolute reduction 1.4% (5.4% reduction to 2.6% increase). Relative reduction 3% (11% reduction to 5% increase)
Self-reported function	Mean self-reported function in the control groups was	Mean self-reported function in the intervention groups was		574 (three studies)	⊕⊕⊖⊖ **low** [1,2]	SMD -0.19 (-0.5 to 0.11)

Multiple tools.[7] Lower score better Follow-up: 12 months	**1.29 points on 0 to 3 scale** [8]	**0.04 points lower** (0.02 lower to 0.10 higher)				Absolute improvement in function 4% (2% reduction to 11% improvement). Relative improvement 11% (6% reduction to 30% improvement)[8]
Quality of life Quality of well-being scale. Scale from 0 to 1, higher better Follow-up: 12 months	Mean quality of life in the control groups was **0.57 units**	Mean quality of life in the intervention groups was **0.01 lower** (0.03 lower to 0.01 higher)		165 (one study)	⊕⊕⊖⊖ **low** [2,9]	Absolute mean reduction 1% (95% CI 3% lower to 1% higher) Relative reduction 2% (95% CI 5% lower to 1% higher)
Withdrawals Follow-up: six to 12 months	**117 per 1,000**	**130 per 1,000** (91 to 183)	**RR 1.11** (0.78 to 1.57)	937 (five studies)	⊕⊕⊕⊖ **moderate** [2]	Absolute risk difference 1% increase (95% CI 3% decrease to 5% increase). Relative percentage change 11% increase (95% CI 22% decrease to 57% increase)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) 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; **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]One large study was conducted in Veteran population that was mostly men (92.7%), limiting applicability of findings.
[2]Design flaws, including participants were not blind to group allocation in all trials, other trials had unclear randomisation method or concealment of allocation and unbalanced withdrawals across treatment groups, render the evidence susceptible to bias.
[3]Estimated relative changes based on mean (SD) ASES score in attention control group at baseline 5.8 (2.0) from Allen 2010.
[4]Pain VAS, pain on walking VAS and pain subscale of the arthritis impact measurement scale (AIMS).
[5]Estimated using mean (SD) for control group VAS pain on walking at baseline 6.28 (3.18) from Mazzuca 1997, and an assumed minimal clinically important difference of 1.5 points in 10-point pain scale.
[6]Approximately half of total study population in trial had rheumatoid arthritis (data not included); data are presented for the OA subgroup; small sample size and wide CIs reduce precision.
[7]AIMS physical disability subscale, AIMS2 function subscale and HAQ disability subscale.
[8]Assumed risk from Mazzuca 1997 control group at 12 months mean HAQ disability scale:1.29 (SD 0.70); 0 to 3 scale, lower score better. Absolute risk difference estimated from control group SD at baseline from the same study (SD 0.66); and relative percent change using mean control group HAQ score at baseline (1.13).
[9]Potential imprecision due to data available only from a single study (n = 165).

Summary of findings 2. SMP compared with usual care for osteoarthritis

SMP compared with usual care

Patient or population: patients with osteoarthritis
Settings: community, outpatient, primary care
Intervention: SMP
Comparison: usual care or no treatment or wait list control

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/No treatment/Wait list	SMP				
Self-management of OA Multiple tools[1]. Scale from 1 to 10, higher better Follow-up: three to 21 months	Mean self-management of osteoarthritis in the control groups, using ASES 1 to 10-point scale (10 is better), was **3.7 points** [2]	Mean self-management of osteoarthritis in the intervention groups was **0.13 points higher** (0.02 to 0.23 higher)		1,706 (11 studies)	⊕⊕⊕⊖ **moderate** [3]	SMD 0.16 (0.03 to 0.29) Absolute mean improvement 1.3% (0.2% to 2.3% improvement). Relative improvement 3.5% (0.65% to 6.3% improvement). NNTB 13 (7 to 69)[2]
Positive and active engagement in life Multiple tools[4] Follow-up: six to 12 months	Mean positive and active engagement in life in the control groups, based on SF-36 subscale for role emotional, 0 to 100 scale (100 best), was **57 points** [5]	Mean positive and active engagement in life in the intervention groups was **0.4 points higher** (8 lower to 8.4 higher)		357 (three studies)	⊕⊕⊕⊖ **moderate** [3]	SMD 0.01 (-0.2 to 0.21) Absolute mean improvement 0.4% (8% worsening to 8.8% improvement). Relative improvement 0.5% (10% worsening to 10% improvement)
Pain Multiple tools[6] Follow-up: three to 21 months	Mean pain in the control groups, based on 0 to 10 VAS scale (0 is no pain), was **3.5 points** [7]	Mean pain in the intervention groups was **0.5 points lower** (0.23 to 0.1 lower)		2,083 (14 studies)	⊕⊕⊕⊖ **moderate** [3]	SMD -0.19 (-0.28 to -0.10) Absolute mean reduction 5.5% (8% to 3% reduction). Relative reduction 16% (23% to 8% reduction). NNTB 11 (7 to 21)
Global OA scores Multiple tools[8] Follow-up: six to 21 months	Mean global osteoarthritis scores in the control groups, based on 0 to 96 point WOMAC scale (lower is better), was **35 points** [9]	Mean global osteoarthritis score in the intervention groups was **5.0 points lower** (7.6 to 2.7 lower)[9]		1,957 (seven studies)	⊕⊕⊕⊖ **moderate** [3]	SMD -0.25 (-0.37 to -0.13) Absolute mean improvement 5% (3% to 8% improvement). Relative improvement 13% (7% to 19% improvement). NNTB 10 (7 to 19)[9]

Function—Self-reported Scale from 0 to 68. Lower score is better Follow-up: six to 21 months	Mean function self-reported in the control groups, based on 0 to 68 WOMAC subscale (lower is better), was **25 points** [10]	Mean function self-reported in the intervention groups was **2.6 points lower** (3.9 to 1.3 lower)[10]		2,254 (13 studies)	⊕⊕⊕⊖ **moderate** [3]	SMD -0.18 (-0.27 to -0.09) Absolute improvement 4% (2% to 6% improvement). Relative improvement 10% (5% to 15% improvement). NNTB 14 (9 to 27)[10]
Quality of life Multiple tools[11] Follow-up: six to 21 months	Mean quality of life in the control groups, based on -0.11 to 1.0 EQ-5D scale (higher score is better), was **0.66 points** [12]	Mean quality of life in the intervention groups was **0.006 points higher** (0.03 lower to 0.04 higher)		1,383 (eight studies)	⊕⊕⊕⊖ **moderate** [3]	SMD 0.02 (-0.09 to 0.13) Absolute improvement 0.6% (2.7% worsening to 3.9% improvement). Relative improvement 1% (4.5% worsening to 6.5% improvement)
Withdrawals Losses to follow-up Follow-up: three to 21 months	**172 per 1,000**	**171 per 1,000** (128 to 229)	**RR 0.99** (0.74 to 1.33)	3,738 (16 studies)	⊕⊕⊖⊖ **low** [3,13]	Absolute risk difference 0% (3% lower to 4% higher). Relative difference 10% lower (26% lower to 33% higher)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) 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; **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]Arthritis self-efficacy scale (ASES), ASES subscale for pain, arthritis helplessness index (AHI) and the health education impact questionnaire (heiQ).
[2]Self-management measured using ASES 1 to 10 scale; 10 is best score, taken from Heuts 2005: mean (SD) baseline ASES score in control group was 3.7 (0.8), and mean (SD) ASES final score in control group was 3.7 (0.9); for number needed to treat for an additional beneficial outcome (NNTB) calculation, the minimal clinically important difference not known, assumed as 0.5.
[3]Participants and study personnel were not blind to group allocations; other issues included unclear randomisation and concealment of allocation; thus trials were at risk of selection, performance and detection biases. One trial, Victor 2005, had inconsistent results compared with the other 10 trials, possibly related to a high risk of bias in that trial, but we do not believe it was significant enough to downgrade the evidence further.
[4]Quality of life short form 36 (SF-36) subscale for role emotional and heiQ subscale for positive and active engagement in life.
[5]Positive and active engagement in life calculated from Victor 2005, using SF-36, 0 to 100 scale (100 is highest score): mean (SD) baseline score in control group was x (y); and mean final score in control group was 57 points.
[6]Arthritis impact measurement scale (AIMS), visual analogue scale (VAS) and Western Ontario McMaster Universities Arthritis Index (WOMAC).
[7]Pain calculated from Heuts 2005, using VAS 0 to 10 scale (0 is no pain): Mean (SD) baseline hip pain score in control group was 3.5 (2.9); and mean final score in control group was 3.5 (2.7).

[8]Western Ontario McMasters University Arthritis Index (WOMAC), Arthritis Impact Measurement Scale (AIMS2) and self-rated global health questionnaires.

[9]Global disease scores taken from Hurley 2007, using WOMAC 0 to 96 point scale (lower score better): Mean (SD) baseline score in the control group was 38.4 (19.82); and mean final score in control group was 35 points. For number needed to treat for an additional beneficial outcome (NNTB) calculation, the minimal clinically important difference not known, assumed as 0.5.

[10]Self-reported function based on Hurley 2007, using WOMAC function 0 to 68 point scale (lower score better): Mean (SD) baseline score in the control group was 27.2 (14.6); and mean final score in the control group was 25 points. For number needed to treat for an additional beneficial outcome (NNTB) calculation, the minimal clinically important difference not known, assumed as 0.5.

[11]Quality of life short form 36 (SF-36), quality of life, quality of well-being scale and EQ-5D.

[12]Quality of life taken based on Hurley EQ-5D (0 to 1 scale; higher score better): Control group mean (SD) at baseline was 0.6 (0.3) and at follow-up was 0.66 (0.3).

[13]Inconsistency across studies regarding whether greater number of withdrawals in the self-management group or control group.

Summary of findings 3. SMP compared with information only for osteoarthritis

SMP compared with information only for osteoarthritis

Patient or population: patients with osteoarthritis
Settings: community, outpatient, primary care
Intervention: SMP
Comparison: information only

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Information only	SMP				
Self-management of OA Multiple tools[1]. Scale from 5 to 35 points (higher better) Follow-up: 12 months	Mean self-management of osteoarthritis in the control groups, based on 5 to 35 point ASES self-efficacy pain scale, was **19.2 points** [2]	Mean self-management of osteoarthritis in the intervention groups was **1.3 points higher** (0.26 lower to 2.82 higher)		760 (three studies)	⊕⊕⊕⊖ **moderate** [3]	SMD 0.20 (-0.04 to 0.44) Absolute mean improvement 4% (1% reduction to 9% improvement). Relative improvement 7% (1% reduction to 15% improvement)[2]
Positive and active engagement in life heiQ subscale for positive and active engagement in life. Scale from 1 to 6 (higher better) Follow-up: 12 months	Mean positive and active engagement in life in the control group, on 1 to 6 point scale (higher better), was **4.76 points** [4]	Mean positive and active engagement in life in the intervention group was **0.2 points lower** (0.59 lower to 0.18 higher)		93 (one study)	⊕⊕⊖⊖ **low** [3,5]	Absolute mean worsening 3% (10% worse to 3% improved). Relative mean worsening 4% (12% worse to 4% improved)[4]
Pain	Mean pain in the control group, based on 0 to 20	Mean pain in the intervention groups was		751 (three studies)	⊕⊕⊕⊖ **moderate** [3]	SMD -0.07 (-0.21 to 0.08)

WOMAC subscale for pain . Scale from 0 to 20 (lower better) Follow-up: 12 months	WOMAC pain subscale (lower is better), was **8.5 points** [2]	**0.3 points lower** (0.8 lower to 0.3 higher)				Absolute mean reduction in pain 1.3% (4.0% reduction to 1.5% increase). Relative mean reduction 3% (9% reduction to 3% increase)
Global OA scores Multiple tools[6]. Scale from 0 to 96 (lower score better) Follow-up: 12 months	Mean global OA scores in the control group, based on 0 to 96 total WOMAC score (lower better), was **41.1 points** [2]	Mean global OA score in the intervention group was **0.8 points lower** (3.7 lower to 2.1 higher)		751 (three studies)	⊕⊕⊕⊖ **moderate** [3]	SMD -0.06 (-0.28 to 0.16) Absolute mean improvement 0.8% (2% worse to 4% improved). Relative mean improvement 2% (5% worse to 9% improvement)[2]
Self-reported function Multiple tools[7]. Scale from 0 to 68 (lower better) Follow-up: six to 12 months	Mean self-reported function in the control group, based on 0 to 68 WOMAC function scale (lower better), was **28.9 points** [2]	Mean self-reported function in the intervention groups was **1.1 points lower** (2.7 lower to 0.6 higher)		854 (four studies)	⊕⊕⊕⊖ **moderate** [3]	SMD -0.09 (-0.22 to 0.05) Absolute mean improvement 2% (4% improved to 1% worse). Relative mean improvement 4% (10% improved to 2% worse)[2]
Quality of life Multiple tools[8]. Scale from 0 to 100 (higher better) Follow-up: 12 months	Mean quality of life in the control group, based on 0 to 100 point scale (higher better), was **55.9 points** [2]	Mean quality of life in the intervention group was **0.5 points higher** (1 lower to 2 higher)		648 (two studies)	⊕⊕⊕⊖ **moderate** [3]	SMD 0.05 (-0.1 to 0.21) Absolute mean improvement 0.5% (1% worsening to 2% improvement). Relative mean improvement 1% (2% worsening to 4% improvement)[2]
Withdrawals Losses to follow-up Follow-up: six to 12 months	**243 per 1,000**	**389 per 1,000** (182 to 827)	**RR 1.6** (0.75 to 3.4)	1,251 (four studies)	⊕⊕⊖⊖ **low** [3,9]	Absolute difference 6% higher withdrawals (8% lower to 19% higher). Relative increase 60% (25% decrease to 240% increase)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) 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; **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]ASES subscale for pain and heiQ subscale for self-monitoring and insight.

[2]Control group baseline and final values taken from Buszewicz 2006, used to estimate the mean difference between groups and absolute and relative changes: mean (SD) self-management score, based on arthritis self-efficacy pain 5 to 35 point scale (higher score better), in the control group at baseline was 19.2 (6.4), and at follow-up was 18.8 (6.5) points; mean (SD) pain score, based on 0 to 20 point WOMAC pain subscale (lower better) in the control group at baseline, was 8.7 (3.7) points, and at follow-up was 8.5 (3.9) points; mean (SD) global OA score, based on WOMAC 0 to 96 point scale (lower better), in the control group at baseline was 41.6 (13.32) and at follow-up was 41.4 points; mean (SD) function, based on WOMAC 0 to 68 (lower better), in the control group at baseline was 29.1 (12.7) and at follow-up was 28.9 points; mean (SD) quality of life, based on SF-36 mental component score 0 to 100 (higher better), in the control group at baseline was 50.6 (10.6) and at follow-up was 55.9 points.

[3]Design flaws, including participants were not blind to group allocation in all trials, some trials had unclear randomisation method or concealment of allocation and unbalanced withdrawals across treatment groups; thus the evidence is susceptible to selection, performance, detection or attrition biases.

[4]Baseline and final value control group heiQ scores (1 to 6 point scale, higher better) from Ackerman 2012: Mean (SD) at baseline was 4.8 (0.8); mean at follow-up was 4.76 points.

[5]Findings based on a single study.

[6]WOMAC and the hip and knee multi-attribute priority tool (MAPT).

[7]Function subscales of WOMAC and the Dutch AIMS-SF.

[8]Mental health component of the short form 36 (SF-36) and assessment of quality of life (AQoL).

[9]Inconsistency across studies regarding whether greater number of withdrawals in the self-management group or the control group.

Summary of findings 4. SMP compared with non-SMP intervention for osteoarthritis

SMP compared with non-SMP intervention for osteoarthritis

Patient or population: patients with osteoarthritis
Settings: community, outpatient or physiotherapy clinic, age care facility
Intervention: SMP
Comparison: non-SMP intervention

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Non-SMP intervention	SMP				
Self-management of OA Arthritis self-efficacy scale. Scale from 30 to 300. Follow-up: one to 12 months	Mean self-management of osteoarthritis in the control groups, based on 30 to 300 ASES scale (higher score better), was **220.46 points** [1]	Mean self-management of osteoarthritis in the intervention groups was **12 points higher** (0 to 24 higher)[1]		175 (three studies)	⊕⊕⊕⊖ **moderate** [2]	SMD 0.33 (0 to 0.66) Absolute mean improvement 4% (0% to 9% improvement). Relative improvement 5.7% (0% to 11.3% improvement)
Positive and active engagement in life Not measured	See comment	See comment	Not estimable	0 (0)	See comment	No studies measured this outcome

Pain Multiple tools[3]. Scale from 0 to 20. Lower score is better. Follow-up: one to 12 months	Mean pain in the control groups, based on 0 to 20 WOMAC pain subscale (lower better), was **4.2 points** [4]	Mean pain in the intervention groups was **0.3 points lower** (1.2 lower to 0.5 higher)[4]		321 (five studies)	⊕⊕⊕⊖ **moderate** [2]	SMD -0.09 (-0.36 to 0.17) Absolute mean reduction 1.4% (-5.8 to 2.7%). Relative reduction 6% (-20% 9.5%)[4]
Global OA scores WOMAC. Scale from 0 to 240. Lower score is better. Follow-up: 12 weeks	Mean global osteoarthritis scores in the control group, based on 0 to 240 WOMAC scale (lower better), was **66.8 points** [5]	Mean global osteoarthritis score in the intervention group was **11.6 points higher** (5.6 lower to 28.7 higher)		98 (one study)	⊕⊕⊕⊖ **moderate** [2]	SMD 0.27 (-0.13 to 0.67) Absolute mean worsening 4.8% (12% worsening to 2.3% improvement). Relative worsening 12.7% (31.4% worsening to 6.1% improvement)[5]
Function—Self-reported Multiple tools[3]. Scale from 0 to 68. Lower score is better. Follow-up: one to 12 months	Mean function self-reported in the control groups, based on 0 to 68 point WOMAC scale (lower better), was **12.2 points** [4]	Mean function self-reported in the intervention groups was **0.04 standard deviations higher** (0.34 lower to 0.42 higher)		216 (three studies)	⊕⊕⊕⊖ **moderate** [2]	SMD 0.04 (-0.34 to 0.42) Absolute worsening 0.6% (6.4% worsening to 5.2% improvement). Relative worsening 2.6% (27.4% worsening to 22.2% improvement)
Quality of life Multiple tools[6]. Scale from 0 to 1. Higher score is better. Follow-up: 12 to 36 months	Mean quality of life in the control groups was **0.73** [4]	Mean quality of life in the intervention groups was **0.06 standard deviations lower** (0.49 lower to 0.36 higher)		226 (two studies)	⊕⊕⊕⊖ **moderate** [2]	SMD -0.06 (95% CI -0.49 to 0.36) Absolute worsening 0.54% (4.4% worsening to 3.2% improvement). Relative worsening 0.7% (5.8% worsening to 4.3% improvement)[4]
Withdrawals Losses to follow-up. Follow-up: one to 36 months	**243 per 1,000**	**209 per 1,000** (168 to 265)	**RR 0.86** (0.69 to 1.09)	919 (seven studies)	⊕⊕⊕⊖ **moderate** [2]	Absolute difference of 2% fewer withdrawals (7% fewer to 12% more). Relative percentage change of 14% fewer withdrawals (31% fewer to 9% more)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) 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; **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]Baseline and final values for comparison group taken from Keefe 2004, ASES 30 to 300 self-efficacy scale (higher better): Control group baseline mean (SD) was 215 (36.95); control group follow-up score was 220.46.

[2]Trials had design flaws, making results susceptible to bias, including participants and study personnel were not blind to group allocations; unclear randomisation and concealment of allocation; and selective reporting.

[3]Arthritis impact measurement scale (AIMS) and Western Ontario McMasters University Arthritis Index (WOMAC).

[4]Control group baseline and final values taken from Jessep 2009. WOMAC 0 to 20 point pain subscale (lower better): Control group mean (SD) pain at baseline was 5.7 (3.2) and at 12-month follow-up was 4.2 (4.0). Control group WOMAC 0 to 68 point function subscale (lower better): Control group mean (SD) at baseline was 15.9 (10.4) and at follow-up was 12.2. Control group quality of life 0 to 1 point EQ-5D scale (higher is better): Control group mean (SD) at baseline was 0.76 (0.09) and at follow up was 0.73.

[5]Baseline and final control group means from Maurer 1999: Control group mean at baseline was 91.5 points on 0 to 240 WOMAC scale (0 is better) and at 12 weeks was 66.8 points; control group SD taken from Wolfe 1999 (42.9).

[6]Quality of well-being scale and EQ-5D.

BACKGROUND

Description of the condition

Osteoarthritis (OA) is a degenerative joint condition that affects primarily the weight-bearing joints such as the hips, knees, ankles and spine but may also affect the hands. Worldwide, OA affects approximately 10% of men and 18% of women 60 years of age or older (Woolf 2003), and incidence increases with age.

The impact of OA includes pain, difficulty in performing activities of daily living, dependency on family and friends for assistance, reduced quality of life, lost productivity and personal economic impact associated with ongoing care and management (ABS 2004; March 2004). The major costs of OA within the health sector are related to joint replacement, visits to general practitioners (GPs) and specialists, prescription and over-the-counter medications and allied health care. As the condition is not reversible, the growing number of people with OA will result in a greater burden of disease; current estimates indicate that OA will be the fourth leading cause of disability by the year 2020 (Woolf 2003).

Description of the intervention

Support is available for coordinated delivery of patient self-management education programmes to improve healthcare outcomes for people with OA (Osborne 2004). Self-management education programmes are distinct from simple patient education or skills training, as they encourage people with chronic disease to take an active role in the management of their own condition. Self-management education programmes aim to improve outcomes for patients by supporting, not replacing, medical care (Walker 2003).

Self-management education programmes are complex behavioural interventions comprising a package of interventions specifically targeted at patient education and behaviour modification. Programmes vary in the content used to educate patients about their condition and to explain how they can best manage their symptoms. Some programmes specifically focus on managing the chronic condition itself, whereas other programmes may take a more holistic approach to managing the overall general well-being of the individual. Substantial variation exists in the delivery of self-management education programmes, such as the mode (face-to-face, Internet, telephone), the audience (group, individual), the duration (single session, several months, ongoing), the frequency (once a week, once every two months) and the personnel (healthcare professionals, lay leaders).

How the intervention might work

The pathology associated with joints affected by OA is typically irreversible. Aside from joint replacement, interventions are usually targeted at maintaining or improving life *with* the condition rather than improving the condition itself. A variety of terms in the literature describe self-management, including self-care, self-monitoring, self-help and social support (Walker 2003). Several models of self-management are known (Osborne 2004); however, the core steps involve (1) engaging in activities that promote health and prevent adverse sequelae; (2) interacting with healthcare providers; (3) performing improved self-monitoring of physical and emotional status; and (4) managing the effects of illness on a person's ability to function in important roles and on emotions, self-esteem and relationships with others (Von Korff 1997). The skills required for these tasks include problem solving, decision making, finding and utilising resources, forming partnerships with healthcare workers and taking action (Lorig 2003).

Assessing the characteristics and impact of self-management education programmes

Studies of self-management education programmes have varied widely in their attempts to quantify the potential impact of these programmes on participant health and well-being. This has resulted in significant heterogeneity in outcome assessment across studies and has contributed to inconsistencies in reported effectiveness of programmes. Understanding which outcomes are most relevant to assessment of the effectiveness of self-management education programmes is required, so that programmes can be assessed systematically on the basis of outcomes that we know are important to participants.

The Arthritis Self-Efficacy Scale (ASES) was the first arthritis-specific instrument developed to measure the effects of arthritis self-management programmes (Lorig 1989). It consists of three subscales (pain, function and other symptoms) and includes efficacy expectation items that ask individuals how certain they are that they can perform a specific activity, for example, walking 100 feet on flat ground in seven seconds; as well as performance attainment items, for example, how certain they are that they can control their fatigue or deal with the frustration of arthritis. Although these items capture an individual's ability to self-manage and therefore are useful in measuring outcomes of self-management education programmes, the validity of the ASES as a true self-efficacy measure has been questioned (Brady 1997; Brady 2011). Although the ASES includes items pertaining to efficacy expectations, it does not ask about an individual's confidence that different behaviours will produce the desired outcome (outcome expectations)—an integral component of Bandura's theory of self-efficacy (Bandura 1977). In addition, the function subscale items appear to capture perceived physical function rather than self-efficacy belief.

Recent research has been undertaken to identify key indicators of effective self-management interventions from the patient perspective (Osborne 2007). Development of the Health Education Impact Questionnaire (heiQ) involved extensive engagement and consultation with consumers and healthcare professionals regarding the outcomes they consider to be valuable and direct benefits of self-management programmes. Eight independent domains were described and form the basis of the constructs of the questionnaire. Domains identified as key indicators of effective self-management programmes include health-directed behaviour; positive and active engagement in life; emotional well-being; self-monitoring and insight; constructive attitudes and approaches; skill and technique acquisition; social integration and support; and health service navigation. The constructs used in the heiQ have been shown as robust across a range of settings (Nolte 2007; Nolte 2009).

Why it is important to do this review

Several previous systematic reviews have summarised the effects of self-management programmes (Chodosh 2005; Devos-Comby 2006; Warsi 2003) or arthritis patient education interventions (Hirano 1994; Lorig 1987; Superio-Cabuslay 1996) for people with OA or for mixed populations (including people with chronic diseases such as OA). However, some of these previous reviews

have (1) combined data from studies that included both people with OA and individuals with rheumatoid arthritis (Warsi 2003); (2) restricted inclusion of studies to those in which participants have OA affecting the knee only (Devos-Comby 2006); (3) restricted inclusion of studies to those in which the comparison intervention was a non-steroidal anti-inflammatory drug treatment only (Superio-Cabuslay 1996); or (4) did not employ rigorous and systematic methods of searching, appraising and synthesising the evidence as necessary to produce reliable systematic reviews on the effects of healthcare interventions (Hirano 1994; Lorig 1987).

Existing clinical practice guidelines uniformly recommend self-management for OA of the hip and knee (March 2010; Zhang 2007). If broad implementation of self-management education programmes for OA are to be considered, a strong evidence base of support is needed. Evidence must show that self-management education programmes improve functional, psychological and/or social outcomes for people with OA. Furthermore, it is important to identify any variables that may significantly affect outcomes of the intervention (e.g. age, stage of disease, comorbidities), as well as contextual issues around content, settings and methods of delivery of self-management education programmes. A systematic review of all randomised controlled trials (RCTs) to date would determine whether self-management education programmes are effective in improving outcomes for people with OA.

OBJECTIVES

To assess the effectiveness of self-management education programmes for people with osteoarthritis.

METHODS

Criteria for considering studies for this review

Types of studies

We included RCTs or quasi-randomised trials in which group assignments were determined through methods other than true randomisation (i.e. alternate assignment).

Types of participants

We included studies of people of all age groups diagnosed with OA as defined in the included trials. Studies involving people with conditions other than OA (i.e. mixed populations) were included only if outcomes for people with OA were presented separately, or if a high proportion of participants (90% or greater) had OA.. We planned to contact trial authors to obtain separate data for people with OA if these had not been reported.

Types of interventions

We included structured self-management education programmes that were judged as being primarily educational and that addressed self-management of OA, arthritis in general or living with chronic disease. Programme components that directly address self-management may include fostering skills in managing OA, such as problem solving, goal setting, decision making, self-monitoring and coping with the condition, as well as providing interventions to manage pain or improve physical and psychological functioning.

Structured programmes delivered by healthcare professionals, lay leaders or both were included, irrespective of whether the programme was delivered to a group of participants or on an individual basis. Studies describing interventions for self-management that are not delivered within a structured format or by some form of organised content delivery were excluded. All modes of delivery, such as face-to-face or interventions delivered by post, Internet, or telephone, were included, provided they incorporated an iterative process of interaction between the participant and programme facilitators. Studies were not excluded on the basis of duration or location of self-management education programmes. We excluded interventions that were judged as treating participants as passive recipients of care (e.g. provision of information alone). Studies that focused solely on exercise were not included.

Programmes that incorporate education for carers or relatives were included, provided the intervention was primarily intended for the person with OA.

Studies that compared outcomes of people with OA assigned to a self-management education programme versus those of individuals who did not receive a self-management educational intervention (i.e. information only, no treatment, usual care, waiting list control, or alternative interventions not considered self-management education programmes) were eligible for inclusion. Studies comparing different self-management education programmes without an appropriate comparison group (i.e. no programme) were excluded. Co-interventions were considered, provided the comparison group received the same co-intervention.

Studies that compared one type of self-management programme versus another were excluded.

Types of outcome measures

No studies were excluded on the basis of outcomes reported. All immediate (up to six weeks from the start of the intervention), intermediate (up to and including one year after the intervention) and longer-term outcomes (longer than one year after the intervention) are reported.

Main outcomes

The following outcomes were selected as the most relevant and are included in the Summary of findings tables.

- Self-management of OA (participant's self-monitoring and insight into living with OA).
- Participant's positive and active engagement in life (including return to work, fulfilling his or her role within the family).
- Pain,
- Global OA scores (e.g. Western Ontario and McMaster Universities Arthritis Index (WOMAC), Lequesne Osteoarthritis Index global score).
- Self-reported function (e.g. function as measured on the WOMAC function subscale).
- Quality of life (including participant-reported general health status).
- Withdrawals (including withdrawals related or unrelated to the study intervention (dropouts) and individuals lost to follow-up). Originally, we planned to assess adverse events or withdrawals (when reported reasons for withdrawal are related to the intervention); however, none of the included trials reported this outcome.

Other outcomes

We also included the following outcomes, which are relevant to the impact of self-management.

- Performance measures (e.g. the six-minute walking distance test, the timed up-and-go test).
- Emotional distress (including depression, anxiety, stress).
- Health-directed activity (including adherence, exercise, diet, relaxation).
- Social integration and support (including social participation, social network, social input).
- Health service navigation (visits to healthcare professionals, emergency room visits, hospital admissions, length of stay).
- Skill and technique acquisition (including knowledge about the condition and how symptoms can be managed).
- Constructive attitude and approaches (including changes in perceived impact of OA on participant's life).
- Participant satisfaction.

In the light of the common use of arthritis-specific self-efficacy measures, particularly the ASES (Lorig 1989), as outcome measures in trials of self-management education programmes, we considered whether to include self-efficacy as a separate outcome. However, as these tools may not comprehensively capture all aspects of self-efficacy theory and may include items measuring performance attainment, we considered an important or meaningful distinction between the outcome of 'self-management' and 'self-efficacy' to be insufficient to justify treating these as separate outcomes in the review. We therefore included arthritis self-efficacy scales within the primary outcome of self-management.

Search methods for identification of studies

We searched the following electronic databases for primary studies, up to January 17, 2013.

- The Cochrane Central Register of Controlled Trials (CENTRAL via *The Cochane Library,* Issue 1, 2013).
- MEDLINE (Ovid 2005 to January 17, 2013).
- EMBASE (Ovid 2010 to January 2013).
- CINAHL (EBSCOHost).
- PsycINFO (1806 to January 2013).
- Dissertation Abstracts (ProQuest January 2013).
- SCOPUS (January 2013).

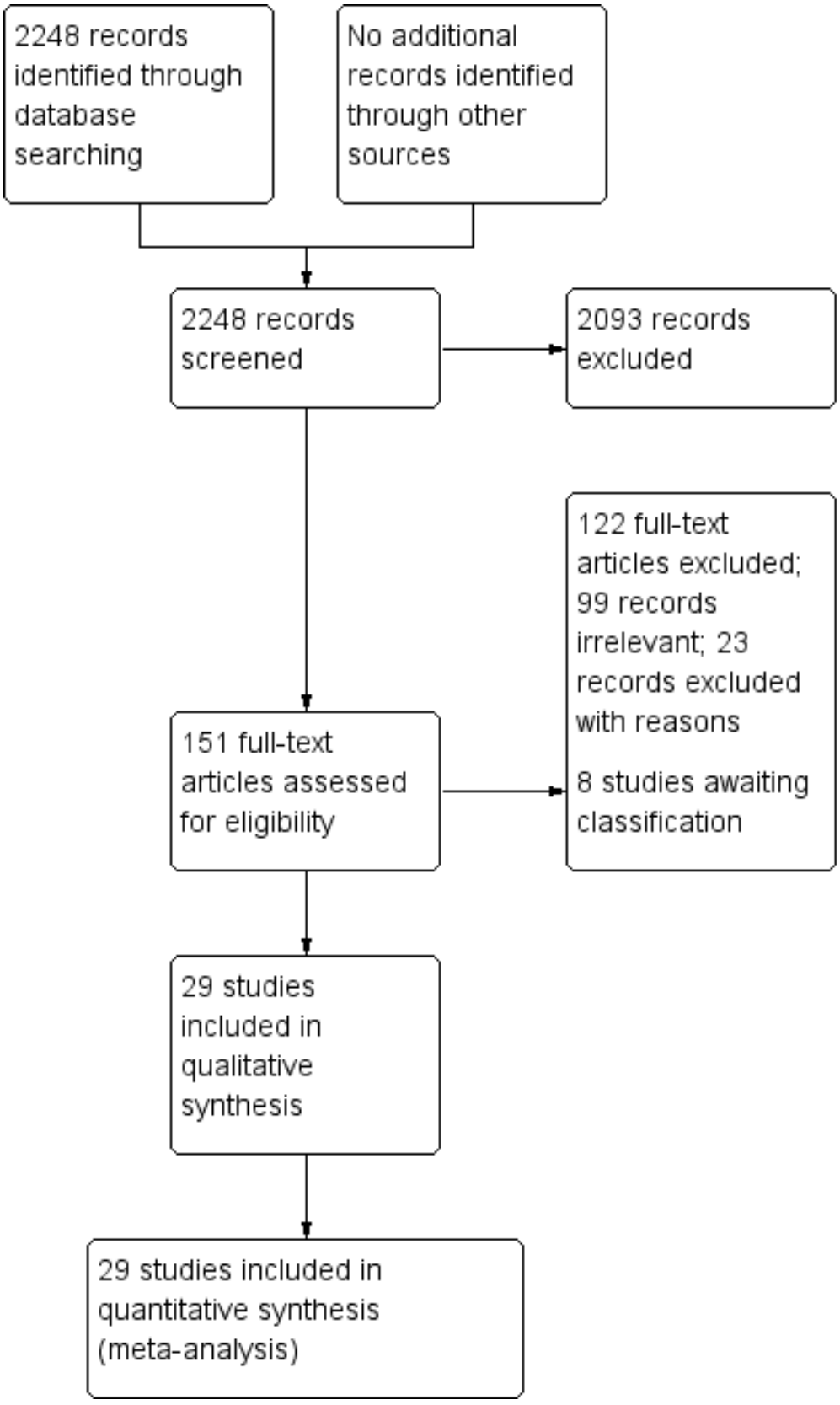
The Database of Abstracts of Reviews of Effects (DARE) via *The Cochane Library* was searched to identify potentially relevant reviews and the reference lists screened to identify primary studies. Reference lists of relevant studies were also screened to identify potential studies for inclusion in the review.

We also searched the World Health Organization International Clinical Trial Registry Platform (ICTRP) (www.who.int/ictrp/en) to identify trials in progress.

Search strategies and time periods for each database are listed in the appendices (Appendix 1; Appendix 2; Appendix 3; Appendix 4; Appendix 5; Appendix 6; Appendix 7; Appendix 8). The search strategy combined text words and controlled vocabulary for describing arthritis/OA and self-management programmes using the Cochrane highly sensitive methodological filter for controlled trials (Lefebvre 2011). We applied no restrictions based on language or publication status. The MEDLINE and EMBASE searches were limited to 2005 to 2013 and 2010 to 2013, respectively, to reduce duplication, as RCTs from both of these databases are included in CENTRAL for the years before these time periods.

A flow diagram for the search results and for selection of studies is presented in Figure 1.

Figure 1. Flowchart.



Data collection and analysis

Selection of studies

Two review authors (VP and FPBK) assessed the titles and available abstracts of all studies identified by the initial search and excluded any clearly irrelevant studies. Studies were not excluded on the basis of the language of published articles. We included both published and unpublished reports. Two review authors (FPBK and LRAB) independently applied the selection criteria to full-text reports of potentially eligible studies. The review authors resolved any disagreements through discussion until consensus was reached or by arbitration by a third review author (VP, RB or RHO) when required. Reasons for exclusion are provided in cases where studies could be considered plausible for inclusion but were excluded from the review.

Data extraction and management

Two review authors (FPBK and LRAB) independently extracted data from the included trials, including information about the study population, interventions, analyses, outcomes and sources of funding, using a standardised data extraction form specifically designed and piloted for this review.

Contextual factors and characteristics of the population relevant for addressing potential issues in health equity were extracted using the PROGRESS-Plus concept (place of residence; race, ethnicity and culture; occupation; sex; religion; education; socioeconomic status; social capital; age; disability; and sexual orientation) (Bambas 2004; Borkhoff 2011).

Health literacy of the population may be another potentially important issue of relevance to health equity that is not currently captured in the PROGRESS-Plus framework. The World Health Organization describes health literacy as “the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health” (World Health Organization 1998). In both developing and developed countries, health and social policies are emerging that highlight health literacy as a key determinant of a person’s ability to optimally manage his or her health and ensure equitable access to and use of services (Committee on Health Literacy 2004; Commonwealth of Aus 2009; United Nations Economic and Social Council 2010). We therefore also extracted information regarding health literacy of the study population, if available, using the nine domains of the Health Literacy Questionnaire (HLQ) (Osborne 2013).

- Feeling understood and supported by healthcare providers.
- Having sufficient information to manage my health.
- Actively managing my health.
- Social support for health.
- Appraisal of health information.
- Ability to actively engage with healthcare providers.
- Navigating the healthcare system.
- Ability to find good health information.
- Understand health information well enough to know what to do.

The following information was collected for each self-management education programme.

- Intended audience (people with OA, arthritis or chronic disease).
- Mode (delivered on a one-to-one basis or to groups of participants).
- Personnel (led by healthcare professionals or by trained facilitators).
- Delivery method (face-to-face, written, audio, video, phone, Internet).
- Language (English or other languages).
- Format (tailored to the individual's needs or delivered in standard format).
- Location (hospital, GP clinic, community setting, home).
- Duration (number and frequency of sessions, hours per session, total duration of programme).

We also extracted information about the components of each self-management education intervention using the eight domains described in the Health Education Impact Questionnaire (heiQ) (Osborne 2007). Each of these domains have been identified as an independent outcome indicator of effective self-management interventions and has been found to be robust across settings (Osborne 2007). We assessed whether interventions were developed on the basis of an explicit theoretical framework (e.g. models of behavioural theory) or a set of principles (e.g. principles of adult education), and whether each of the following components was addressed within each programme.

- Health-directed activity.
- Positive and active engagement in life.
- Emotional distress.
- Self-monitoring and insight.
- Constructive attitudes and approaches.
- Skill and technique acquisition.
- Social integration and support.
- Health service navigation.

To assess the effects of an intervention, we extracted raw data for outcomes of interest (means and standard deviations for continuous outcomes and number of events for dichotomous outcomes) when available in the published reports.

We contacted the authors of all studies to obtain more information as needed.

If a study reported multiple time points within immediate, intermediate or longer-term outcomes, only the longest time point was extracted.

Assessment of risk of bias in included studies

Two review authors (FPBK and LRAB) independently assessed the risk of bias in each included study against key criteria: random sequence generation; allocation concealment; blinding of outcomes; incomplete outcome data; and selective outcome reporting. We also considered other sources of bias, such as whether the intervention was delivered as intended, whether groups were comparable at baseline and whether contamination between groups was possible. Assessments were conducted in accordance with methods recommended by The Cochrane Collaboration (Higgins 2011).

Selective outcome reporting was judged on the basis of whether all outcomes assessed in a trial have been reported. When possible,

we obtained trial protocols for comparison of planned outcome assessment versus the outcome data available from each trial. For trials published after July 1, 2005, we searched for trial protocols using the Clinical Trial Register at the International Clinical Trials Registry Platform of the World Health Organization (http://apps.who.int/trialssearch; DeAngelis 2004). We also constructed outcome matrices listing primary outcomes reported across the included studies to identify inconsistencies in outcome reporting to indicate possible selective outcome reporting. We planned to use the Outcome Reporting Bias In Trials (ORBIT) classification system to describe whether selective outcome reporting was suspected and the potential reason for it (Kirkham 2010).

Each of the domains assessed for risk of bias is explicitly judged as follows: Yes (low risk of bias); No (high risk of bias); or Unclear (lack of information or uncertainty over the potential for bias). Review authors resolved disagreements through discussion until consensus was reached, or consulted a third review author (RB) to resolve disagreements if necessary.

Measures of treatment effect

Point estimates and 95% confidence intervals (CIs) were calculated for outcomes of individual RCTs whenever possible. Point estimates for dichotomous outcomes are expressed as risk ratios (RRs). For continuous outcomes, results are summarised as mean difference (MDs) if the same tool has been used to measure the same outcome across separate studies. Alternatively, we summarised treatment effects using the standardised mean difference (SMD) when studies measured the same outcome but employed different tools. If results could not be summarised as point estimates with 95% CIs, we tabulated results for each outcome.

Studies included in forest plots were listed in order of the weight (from lowest to highest weight) that each individual study contributed to the overall summary estimate.

Unit of analysis issues

When appropriate, we incorporated results of cluster-randomised trials into meta-analyses using the generic inverse variance method in RevMan (Deeks 2011). Effect estimates (e.g. RR and 95% CI) for relevant outcomes were extracted from cluster trials that had appropriately accounted for the cluster design. When trials had not appropriately accounted for the design effect, we corrected the standard errors of effect estimates by using an intraclass correlation coefficient that was obtained from the trial report or estimated from similar studies.

Dealing with missing data

We contacted trial authors if the type of arthritis had not been specified or when studies involved mixed populations, in an attempt to obtain separate data for people with OA. Clarification was also sought for descriptions of interventions (e.g. setting, mode of delivery, format, duration), trial conduct (e.g. method of random sequence generation, method of allocating participants to treatment groups, blinding of trial personnel) and availability of unpublished data for outcomes that were measured.

For outcomes assessed using standard scales (e.g. WOMAC, quality of life scales), we attempted to present overall scores when possible. If results were presented only for separate subscales, we used results of subscales considered to be most relevant to the outcome of interest and recorded instances where this applied in the Notes section of the Characteristics of included studies.

When the number of people assessed for an outcome was unclear, we imputed this on the basis of the number of people originally randomly assigned to the study groups. For continuous measures, missing standard deviation (SD) values were estimated from other measures such as standard error (SE), P values or confidence intervals whenever possible, or they were imputed on the basis of SD values in similar trials (Higgins 2011). For dichotomous outcomes, percentages were used to estimate the number of events or the number of people assessed for an outcome. All data imputations are recorded in the Notes section of the Characteristics of included studies.

Assessment of heterogeneity

Before a meta-analysis was conducted, studies were assessed for similarities with respect to characteristics of the self-management education programmes, comparison groups and outcomes. Studies judged by the review authors as being too different from each other were not combined in the analysis but were described separately in the text of the review.

For studies judged as sufficiently similar, statistical heterogeneity was assessed visually by looking at the scatter of effect estimates on the forest plots and by determining the I^2 statistic (Higgins 2003). The I^2 statistic was used as an indication of the proportion of heterogeneity, with higher values indicating a higher proportion of heterogeneity, using the following as a rough guide for interpretation: 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 considerable heterogeneity (Deeks 2011). In cases of considerable heterogeneity (defined as $I^2 \geq 75\%$), we explored the data further by comparing characteristics of individual studies and any subgroup analyses and reported any differences when interpreting the results of this review, or we reported I^2 values whenever unexplained statistical heterogeneity was present.

Assessment of reporting biases

Selective outcome reporting was assessed using the approach described previously in this protocol (see Assessment of risk of bias in included studies). The potential impact of selective outcome reporting on the overall results of the review was discussed in terms of suspected reasons for the missing outcome data and the size, strength and direction of the effect.

To assess for potential small-study effects in meta-analyses (i.e. the intervention effect is more beneficial in smaller studies), we compared effect estimates derived from a random-effects model and from a fixed-effect model of meta-analysis. In the presence of small-study effects, the random-effects model will give a more beneficial estimate of the intervention than the fixed-effect estimate (Sterne 2011).

It was planned that the potential for small-study effects in the main outcomes of the review would be further explored using funnel plots if at least 10 studies were included in a meta-analysis.

Data synthesis

Included studies were grouped and assessed according to whether they compared self-management education programmes versus:

- an attention control group (i.e. participants received the same contact hours with programme providers, but the content delivered was unrelated to self-management of their condition);
- a group that received no treatment or usual care or were placed on a waiting list to attend the self-management programme at a later date;
- an information-only group (i.e. educational materials, programme handbook);
- a group that received an alternate intervention that was not a self-management education programme (i.e. exercise or diet plan); or
- a group that received acupuncture. This intervention was considered separately from other studies comparing self-management programmes versus alternative interventions (i.e. comparison described in the previous bullet) because, unlike alternative interventions in the other trials, this comparison is not a behavioural intervention.

We considered the first two comparisons to be the most important for addressing the objectives of this review. Evidence from physical therapy trials suggests that the quality of the therapeutic relationship influences clinical outcomes such as pain and physical function (Hall 2010; Pinto 2012). This effect may also apply to self-management education programme providers. An attention control may control for any effect of contact time with programme providers, and we considered this to be the comparator with the lowest risk of bias in determining the true effect of self-management education programmes. We considered usual care to be an important comparator as well, as this reflects routine care. However, comparisons versus usual care are generally unblinded (analogous to no treatment), while an attention control allows for blinding of participants (of utmost importance when outcomes are participant assessed) so is closer to a sham/placebo control.

When studies were considered to be sufficiently similar in terms of the self-management education programme delivered and the comparison intervention provided, we pooled outcomes in a meta-analysis using the random-effects method (Deeks 2011). Forest plots display individual study results sorted by weight in ascending order.

To minimise outcome reporting bias, if data from more than one self-efficacy scale were reported for a trial, we extracted data according to the following hierarchy.

- Self-efficacy on a visual analogue scale.
- Arthritis self-efficacy scale mean score.
- Arthritis self-efficacy subscale (pain or other symptoms).
- Self-efficacy on other scales (i.e. Arthritis Helplessness Index, heiQ self-monitoring and insight).

If data on more than one pain scale were provided for a trial, we referred to a previously described hierarchy of pain-related outcomes (Jüni 2006) and presented data on the pain scale that was highest on this list.

- Global pain.
- Pain on walking.
- WOMAC pain subscore.
- Composite pain scores other than WOMAC.
- Pain on activities other than walking.
- Rest pain or pain during the night.

If data on more than one global OA scale were provided for a trial, we extracted data according to the following hierarchy.

- Patient global assessment score.
- Self-rated global health.
- MAPT (Multi-attribute Arthritis Prioritisation Tool).
- WOMAC total score.
- AIMS (Arthritis Impact Measurement Scales) total score.

Similarly, if data on more than one self-reported function scale were provided for a trial, we extracted data according to the hierarchy presented below (Rutjes 2009).

- Global disability score.
- Walking disability.
- WOMAC disability subscore.
- Composite disability scores other than WOMAC.
- Disability other than walking.

If data on more than one quality of life scale were provided for a trial, we extracted data according to the following hierarchy.

- Short Form (SF)-36.
- SF-12.
- EuroQoL.
- SIP (Sickness Impact Profile).
- NHP (Nottingham Health Profile).
- Other validated quality of life scores.

If data on a quality of life scale were provided in both a multi-question format and a visual analogue scale format, we chose the first, as we judged this would provide a more accurate measure of quality of life and patient-reported global health status.

If data for both anxiety and depression were presented, we chose to extract only the data on depression for the outcome of emotional distress, as we judged that depression overall was more consistently reported as a measure of emotional distress when compared with anxiety.

Summary of findings

We presented the main outcomes of the review in Summary of findings tables (self-management, positive and active engagement in life, withdrawals, pain, global OA scores, self-reported function and quality of life) to provide key information concerning the quality of evidence, the magnitude of effect of the interventions examined and the sum of available data on the main outcomes, as recommended by The Cochrane Collaboration (Schünemann 2011a). The 'Summary of findings' tables provide an overall grading of the evidence related to each of the main outcomes based on the GRADE approach (Schünemann 2011b).

Overall outcome data presented in the Summary of findings tables are based on the longest time points measured in each study.

Separate analyses (forest plots not shown) were performed to include all possible studies per outcome, as we did not have a prespecified primary time point, and there did not seem to be an effect of time. Outcomes pooled using SMDs were re-expressed as a mean difference by multiplying the SMD by a representative control group baseline standard deviation from a trial, using a familiar instrument.

In the comments column, we calculated the absolute percentage change and the relative percentage change; and, for outcomes with statistically significant differences between intervention groups, we calculated the number needed to treat for an additional beneficial outcome (NNTB).

For dichotomous outcomes, the absolute risk difference was calculated using the risk difference statistic in RevMan and the result expressed as a percentage; the relative percentage change was calculated as the risk ratio -1 and was expressed as a percentage; and the NNT from the control group event rate and the risk ratio were determined using the Visual Rx NNT calculator (Cates 2008).

For continuous outcomes, the absolute risk difference was calculated as the mean difference between intervention and control groups in the original measurement units (divided by the scale), expressed as a percentage; the relative difference was calculated as the absolute change (or mean difference) divided by the baseline mean of the control group from a representative trial. We used the Wells calculator to obtain the NNTB for continuous measures (available at the Cochrane Musculoskeletal Group (CMSG) Editorial office; http://musculoskeletal.cochrane.org/). The minimal clinically important difference (MCID) for each outcome was determined for input into the calculator. We assumed an MCID of 1.5 points on a 10-point pain scale, 0.5 points on the 10-point ASES (self-management) scale and 0.5 points on the WOMAC function 0 to 68-point scale.

Subgroup analysis and investigation of heterogeneity

We undertook an exploratory analysis of studies assessing the effects of self-management education programmes and issues of health equity. We compared effect sizes for the major outcomes of the review, as well as self-reported pain across studies comparing self-management education programmes versus a control group. Studies were classified according to whether the study population consisted mainly of Caucasian, educated, older females or populations drawn from minority groups within the community (e.g. culturally and linguistically diverse populations).

Sensitivity analysis

A sensitivity analysis was conducted and was based on whether participants were randomly allocated and group assignments had been adequately concealed.

RESULTS

Description of studies

Results of the search

The initial database search identified 2,248 records (see Figure 1). We assessed 151 possibly eligible papers in full text. Of these, 29 studies published between 1990 and 2012, involving 6,753 participants (range 32 to 570), met the inclusion criteria for this review. A further eight trials were identified in an updated search performed on 17 January 2013, and, as they were unlikely to alter the conclusions of the review, these trials will be assessed when the review is updated (see Studies awaiting classification).

Included studies

A full description of all included studies is provided in the Characteristics of included studies table.

Design

Of the 29 included studies, three were cluster-RCTs (Hurley 2007; Mazzuca 2004; Victor 2005), and the remaining 26 studies were RCTs. All studies were published in English.

Participants

A description of participants using the PROGRESS-Plus framework is shown in Table 1. Studies were most commonly conducted in the US (17 studies, 58.6%) followed by the UK (four studies, 13.8%), The Netherlands (three studies, 10.3%) and Australia (two studies, 6.9%). Single studies were performed in Spain, Sweden and Hong Kong (China). Most participants (68%) were female, and the average age of participants was 64.8 years. Race and ethnicity were reported in 15 studies (Allen 2010; Berman 2004; Blixen 2004; Buszewicz 2006; Calfas 1992; Cronan 1997; Hughes 2004; Lorig 2008; Maisiak 1996; Mazzuca 1997; Mazzuca 2004; McKnight 2010; Murphy 2008; Victor 2005; Yip 2007), and 70.2% of participants were described as white or Caucasian. PROGRESS-Plus domains that were least described across studies included occupation (nine studies) and socioeconomic status (six studies). Only four trials provided any information related to health literacy, and this was limited to the domains of social support for health (three trials), navigating the healthcare system (two trials), actively managing my health (two trials) and ability to actively engage with healthcare providers (one trial).

Other characteristics of participants (location and duration of OA, body mass index (BMI)) are listed in Table 2. Location of osteoarthritis was reported in 20 of 29 studies (69%), and the predominant location was the knee. BMI was reported in only nine of 29 studies (31%). The duration of OA was reported in 15 studies and ranged from a few months to longer than 20 years.

Intervention

Although self-management education programmes differed in mode (individual or group), personnel (healthcare professionals or trained facilitators), delivery method (face-to-face, telephone, Internet) and duration, all were considered to include an element of self-management (see Table 3). Across 29 studies, 34 self-management education interventions were assessed. Twenty-two of these interventions were group sessions, eight were individual sessions and four involved a combination of group and individual sessions. Most programmes were delivered face-to-face (25 interventions), two were provided over the telephone and one was delivered over the Internet. The remaining six interventions delivered a combination of face-to-face and telephone sessions. The total duration of the programmes ranged from four weeks to 12 months, although most programmes lasted six weeks. The frequency of sessions ranged from four per week to one per month, and most programmes were delivered on a weekly basis.

A description of the components of the programmes based on the heiQ framework is shown in Table 4. The mean number of heiQ items included in individual programmes was 4.4 (out of a possible eight components). Most programmes included skill and technique acquisition (32 programmes, 94%), health-directed activity (29 programmes, 85%) and self-monitoring and insight (27 programmes, 79%). The least included component was social integration and support (four programmes, 12%).

Comparator

Self-management education programmes were compared with attention control (five studies), usual care (17 studies), information-only (four studies) or alternative interventions that did not include self-management (seven studies incorporating interventions such as exercise, physiotherapy or social support). All of these studies considered the self-management education programme to be an active intervention. In contrast, one additional study included self-management education as an inactive control group and compared it with both acupuncture as the active intervention and sham acupuncture (Berman 2004).

Outcomes

A limited number of studies comparing self-management programmes versus usual care or information only reported the main outcomes prespecified in the protocol for this review. The main outcome of positive and active engagement in life was not reported in any trials comparing self-management versus an attention control or alternative intervention.

For one study, the only data that could be extracted consisted of information regarding withdrawals (Calfas 1992).

Excluded studies

One hundred twenty-two studies assessed in full text were excluded. Ninety-nine were judged irrelevant, and 23 studies were excluded according to reasons provided in Characteristics of excluded studies. Studies were excluded if they involved mixed arthritis populations without subgroup data available for people with OA (12 studies; Barlow 2000; Ehrlich-Jones 2001; Goeppinger 1989; Laforest 2008; Laforest 2008a; Lindroth 1989; Lorig 1985; Lorig 1999a; Lorig 1999b; Lorig 2005; Nour 2006; Solomon 2002), if they compared two different self-management education programmes (six studies; Coleman 2010; Hoogeboom 2010; Lorig 1998; Martire 2003a; Martire 2008; Murphy 2010) or if the intervention was judged as failing to fulfil our inclusion criteria for a self-management programme (five studies; Bezalel 2010; Ettinger 1997; Fernandes 2009; Fernandes 2010; Focht 2005).

Risk of bias in included studies

The risk of bias was assessed for each study (see Characteristics of included studies), and the results are summarised in Figure 2 and Figure 3.

Figure 2. Summary of the risk of bias across all included studies.

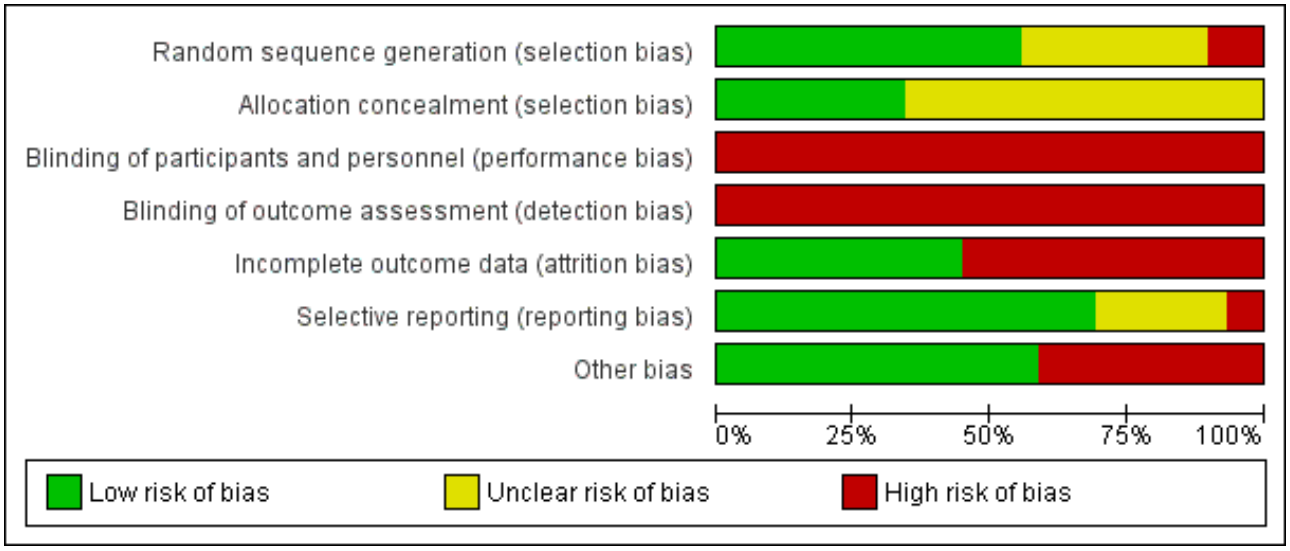


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

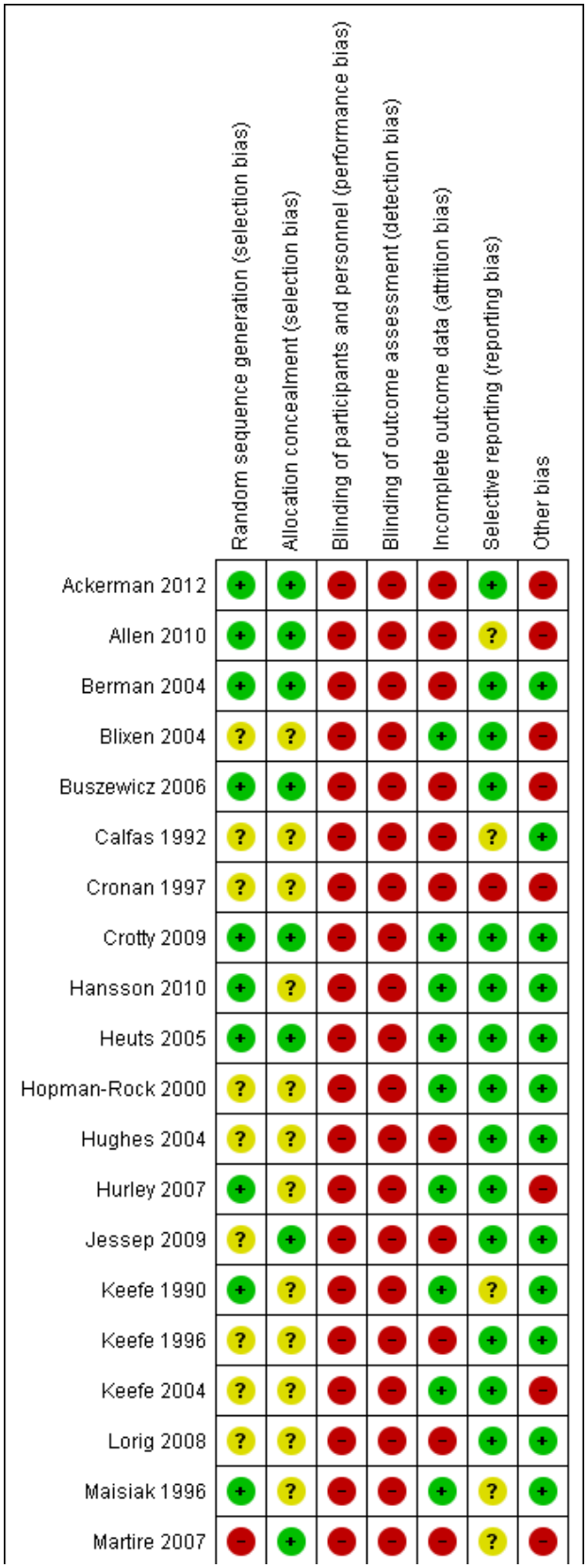
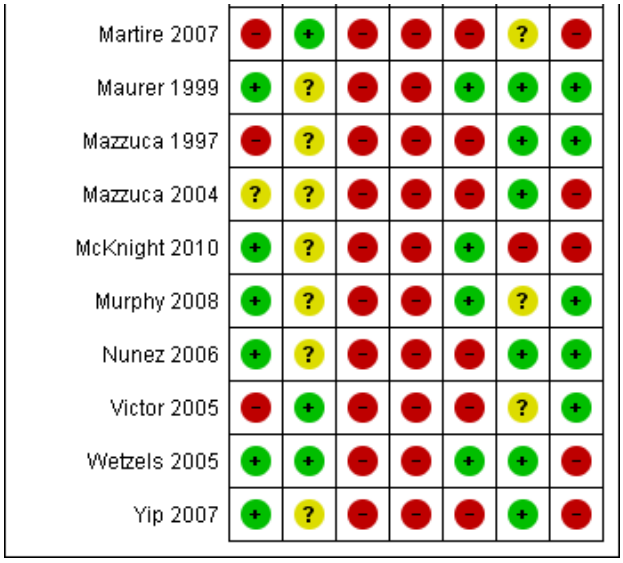


Figure 3. (Continued)



Allocation

Seven studies described adequate sequence generation and allocation concealment and were assessed as being at low risk of selection bias (Ackerman 2012; Allen 2010; Berman 2004; Buszewicz 2006; Crotty 2009; Heuts 2005; Wetzels 2005). Nine trials had adequate random sequence generation, but the risk of bias in the concealment of allocation was unclear (Hansson 2010; Hurley 2007; Keefe 1990; Maisiak 1996; Maurer 1999; McKnight 2010; Murphy 2008; Nunez 2006; Yip 2007). Three trials were assessed as adequately concealing allocation but as having high risk (Martire 2007; Victor 2005) or unclear risk (Jessep 2009) of bias in random sequence generation. Mazzuca 1997 was assessed as having inadequate random sequence generation and unclear risk of bias in allocation concealment. The remaining nine trials were assessed as having unclear risk of selection bias, based on both random sequence generation and allocation concealment.

Blinding

It is not possible to blind personnel delivering self-management programmes, and all participants were aware of the treatment they were receiving; therefore all studies were at high risk of performance bias and dectection bias (for self-reported outcomes). Of 14 studies that included assessor-reported outcomes, eight were assessed as low risk, as the outcome assessors were blinded (Berman 2004; Hansson 2010; Hopman-Rock 2000; Jessep 2009; Maurer 1999; Murphy 2008; Victor 2005; Wetzels 2005), four were assessed as unclear risk (Cronan 1997; Hurley 2007; Keefe 2004; Yip 2007) and two were assessed as high risk for detection bias (Hughes 2004; McKnight 2010).

Incomplete outcome data

Thirteen trials were assessed as low risk for attrition bias (Blixen 2004; Crotty 2009; Hansson 2010; Heuts 2005; Hopman-Rock 2000; Hurley 2007; Keefe 1990; Keefe 2004; Maisiak 1996; Maurer 1999; McKnight 2010; Murphy 2008; Wetzels 2005). The remaining 16 studies had unexplained incomplete outcome data and were judged as being at high risk of attrition bias.

Selective reporting

Twenty studies were assessed as being at low risk for reporting bias. Seven trials were assessed as having an unclear risk (Allen 2010; Calfas 1992; Keefe 1990; Maisiak 1996; Martire 2007; Murphy 2008; Victor 2005), as the impact of minor outcomes that were not reported and/or the impact of not reporting outcomes at all time points was unclear. The remaining two studies (Cronan 1997; McKnight 2010) were assessed as being at high risk for reporting bias because they failed to report the primary outcome for this review.

Other potential sources of bias

All three cluster-RCTs accounted for the effects of clustering in the analysis of their results (Hurley 2007; Mazzuca 2004; Victor 2005).

In 12 trials (Ackerman 2012; Allen 2010; Blixen 2004; Buszewicz 2006; Cronan 1997; Hurley 2007; Keefe 2004; Martire 2007; Mazzuca 2004; McKnight 2010; Wetzels 2005; Yip 2007), another potential source of bias was identified. These included baseline differences between groups (Keefe 2004; Martire 2007), low adherence to the intervention (Buszewicz 2006; Cronan 1997; McKnight 2010), differences in adherence between groups (Hurley 2007), risk of contamination between groups (Mazzuca 2004) and cultural bias in self-reported measures (Yip 2007).

Effects of interventions

See: **Summary of findings for the main comparison** SMP compared to Attention control for osteoarthritis; **Summary of findings 2** SMP compared with usual care for osteoarthritis; **Summary of findings 3** SMP compared with information only for osteoarthritis; **Summary of findings 4** SMP compared with non-SMP intervention for osteoarthritis

Comparison 1. Self-management programmes versus attention control

Main outcomes

One trial with 344 participants (Allen 2010) found no difference between a self-management education programme and attention

control in terms of improving self-management skills after 12 months (mean difference (MD) 0.40, 95% confidence interval (CI) -0.39 to 1.19) (Analysis 1.1).

Positive and active engagement in life was not reported by any of the trials comparing self-management versus attention control.

One trial (Keefe 1990) with 68 participants found greater pain reduction post-treatment in the self-management intervention group than in the attention control group (SMD -0.62, 95% CI -1.11 to -0.13)—equivalent to a mean difference of 1.55 cm (95% CI 0.33 to 2.78 cm) on a 10-cm VAS (Analysis 1.2). Three trials with 577 participants (Allen 2010; Keefe 1990; Mazzuca 1997) assessed pain at longer time points (six to 12 months); pooled analysis also found greater pain reduction in the self-management intervention group than in the attention control group (SMD -0.26, 95% CI -0.44 to -0.09) (Analysis 1.2). This is equivalent to a mean difference of 0.65 cm (95% CI 0.23 to 1.10 cm) on a 10-cm VAS—a finding that is unlikely to be of clinical significance.

No overall between-group differences were noted in the other main efficacy outcomes: global OA scores (one trial, SMD -0.14, 95% CI -0.54 to 0.26), self-reported function in the short term (one trial, SMD -0.13, 95% CI -0.49 to 0.23) and in the intermediate term (three trials, SMD -0.19, 95% CI -0.50 to 0.11) and quality of life (one trial, SMD -0.01, 95% CI -0.03 to 0.01) (Analysis 1.3; Analysis 1.4; Analysis 1.5).

In five studies with 937 participants (Allen 2010; Calfas 1992; Keefe 1990; Maisiak 1996; Mazzuca 1997), no difference was observed in the rate of withdrawals between self-management (61/446, 14%) and attention control (55/471, 12%) groups (RR 1.11, 95% CI 0.78 to 1.57) (Analysis 1.6).

Other outcomes

No overall between-group differences in emotional distress were noted in the short term (one trial, SMD -0.37, 95% CI -0.85 to 0.11) and in the intermediate term (two trials, SMD 0.02, 95% CI -0.18 to 0.21) (Analysis 1.7).

Functional performance, health-directed activity, social integration and support, health service navigation, skill and technique acquisition and constructive attitudes and approaches were not reported by any of the trials in this comparison.

Comparison 2. Self-management programmes versus usual care

Main outcomes

Five trials with 721 participants (Hopman-Rock 2000; Keefe 2004; Martire 2007; Victor 2005; Yip 2007) reported no difference between a self-management programme and usual care in terms of self-management skills within one month post-treatment (SMD 0.22, 95% CI -0.00 to 0.45) (Analysis 2.1); this is equivalent to a mean difference of 0.11 points (95% CI -0.00 to 0.23) on the 1 to 6-point (higher score is better) heiQ subscale self-monitoring and insight. However, moderate statistical heterogeneity was present (I^2 = 48%) because findings of one trial were not consistent (Victor 2005) with findings of the other four trials. Victor 2005 was a cluster-RCT with more withdrawals in the intervention than in the control group (40% vs 27%) and with high risk of selection, reporting and attrition biases.

Results of 10 trials with 1,647 participants (Allen 2010; Blixen 2004; Crotty 2009; Hansson 2010; Heuts 2005; Hopman-Rock 2000; Lorig 2008; Martire 2007; Victor 2005; Yip 2007) show no difference between a self-management programme and usual care in improving self-management skills from three to 12 months (SMD 0.14, 95% CI 0.00 to 0.27) (Analysis 2.1); this is equivalent to a mean difference of 0.07 points (95% CI 0.00 to 0.14) on the heiQ subscale self-monitoring and insight (1 to 6 scale, higher score is better). Moderate statistical heterogeneity was present (I^2 = 41%) because the results of Victor 2005 were not consistent with the findings of the other nine trials. One trial with 195 participants (Heuts 2005) assessed self-management skills over 21 months and found no difference between participants in a self-management programme and those receiving usual care (SMD 0.23, 95% CI -0.05 to 0.51) (Analysis 2.1); this finding is equivalent to a mean difference of 0.12 (95% CI -0.03 to 0.26) on the heiQ subscale self-monitoring and insight (1 to 6 scale, higher score is better). The impact of the failure of Cronan 1997 to report self-management skills is likely to be small.

One trial with 143 participants (Victor 2005) found no difference between a self-management programme and usual care in terms of participants' positive and active engagement in life after one month (SMD -0.23, 95% CI -0.57 to 0.11) (Analysis 2.2); this is equivalent to a mean difference of -0.18 (95% CI -0.46 to 0.09) on the heiQ subscale for positive and active engagement in life (1 to 6 scale, higher score is better). Three trials with 357 participants (Crotty 2009; Nunez 2006; Victor 2005) also showed no difference at six to 12 months (SMD 0.01, 95% CI -0.20 to 0.21); this is equivalent to a mean difference of 0.01 (95% CI -0.16 to 0.17) on the heiQ subscale positive and active engagement in life (1 to 6 scale, higher score is better).

Findings of six trials with 766 participants (Hopman-Rock 2000; Keefe 1990; Keefe 2004; Martire 2007; Victor 2005; Yip 2007) indicate that pain improved significantly in the self-management intervention group compared with the usual care group in the time frame post-treatment to one month (SMD -0.26, 95% CI -0.41 to -0.10) (Analysis 2.3); this is equivalent to a mean difference of 0.65 cm (95% CI 0.25 to 1.00 cm) on a 10-cm VAS, which is unlikely to be of clinical significance. In 13 trials with 7,447 participants (Allen 2010; Blixen 2004; Crotty 2009; Heuts 2005; Hopman-Rock 2000; Hurley 2007; Keefe 1990; Lorig 2008; Martire 2007; Mazzuca 2004; Nunez 2006; Victor 2005; Yip 2007), pain improved significantly in the self-management group in the time frame three to 12 months (SMD -0.17, 95% CI -0.26 to -0.08); this is equivalent to a mean difference of 0.43 cm (95% CI 0.2 to 0.65 cm) on a 10-cm VAS—a finding also unlikely to be of clinical significance. One trial with 213 participants (Heuts 2005) found no difference in pain between the self-management programme and usual care after 21 months (SMD -0.18, 95% CI -0.45 to 0.09).

The findings of two studies with 319 participants (Martire 2007; Yip 2007) show that global OA scores improved significantly in the self-management intervention compared with the usual care group up to one week post-treatment (SMD -0.34, 95% CI -0.59 to -0.09). This result could be clinically meaningful (MD 0.71 points, 95% CI 0.19 to 1.23 points on the WOMAC) (Analysis 2.4). In seven studies with 1,351 participants (Heuts 2005; Hurley 2007; Lorig 2008; Maisiak 1996; Martire 2007; Nunez 2006; Yip 2007), global OA scores improved in the self-management education programme compared with usual care between three and 12 months (SMD -0.28, 95% CI -0.39 to -0.17) (Analysis 2.4). However, this finding is

unlikely to be of clinical significance (MD 0.59 points, 95% CI 0.36 to 0.82 points, on the WOMAC; 0 to 96-point scale, lower score is better). One study with 197 participants (Heuts 2005) reported that global OA scores improved significantly in the self-management education programme compared with the usual care group after 21 months (SMD -0.29, 95% CI -0.56 to -0.02) (Analysis 2.4), but this finding is unlikely to be of clinical significance (MD 0.61, 95% CI 0.04 to -1.17 points on the 0 to 96-point WOMAC).

Findings of five studies with 714 participants (Hopman-Rock 2000; Keefe 1990; Martire 2007; Victor 2005; Yip 2007) show no difference in self-reported function between participants in self-management programmes and those provided with usual care up to one month post-treatment (SMD -0.01, 95% -0.19 to 0.18) (Analysis 2.5). Thirteen studies with 2,176 participants (Allen 2010; Blixen 2004; Crotty 2009; Heuts 2005; Hopman-Rock 2000; Hurley 2007; Keefe 1990; Lorig 2008; Martire 2007; Mazzuca 2004; Nunez 2006; Victor 2005; Yip 2007) showed that self-reported function improved significantly in self-management programmes compared with usual care between three and 12 months (SMD -0.16, 95% CI -0.25 to -0.08). However, this finding is unlikely to be of clinical significance (MD 0.35, 95% CI 0.17 to -0.55 points on the 0 to 68-point WOMAC function subscale) (Analysis 2.5). One study with 199 participants (Heuts 2005) showed no difference in self-reported function between participants in self-management education programmes and usual care after 21 months (SMD -0.27, 95% CI -0.55 to 0.01) (Analysis 2.5). Similarly, no between-group differences were seen in performance measures of function in the short term (one trial, SMD 0.33, 95% CI -0.07 to 0.73) and in the intermediate term (two trials, SMD 0.06, 95% CI -0.24 to 0.36) (Analysis 2.6).

No overall between-group difference in quality of life was noted at any time point (two trials assessed up to six weeks post-treatment: SMD 0.14, 95% CI -0.47 to 0.75; eight trials assessed up to one year: SMD 0.03, 95% CI -0.08 to 0.14; and two studies assessed beyond one year: SMD 0.10, 95% CI -0.10 to 0.31) (Analysis 2.7).

Findings of 16 trials with 3,738 participants (Allen 2010; Blixen 2004; Cronan 1997; Hansson 2010; Heuts 2005; Hopman-Rock 2000; Hurley 2007; Keefe 1990; Keefe 2004; Lorig 2008; Maisiak 1996; Martire 2007; Mazzuca 2004; Nunez 2006; Victor 2005; Yip 2007) show that no differences in numbers of withdrawals were found between groups (RR 0.99, 95% CI 0.74 to 1.33) (Analysis 2.8).

Other outcomes

Three studies with 262 participants (Keefe 1990; Keefe 2004; Victor 2005) reported no difference in emotional distress between self-management programmes and usual care in the time frame post-treatment and one month (SMD 0.01, 95% CI -0.44 to 0.45), although considerable statistical heterogeneity (I^2 = 64%) was observed (Analysis 2.9). Eight studies with 1,427 participants (Allen 2010; Blixen 2004; Crotty 2009; Hurley 2007; Keefe 1990; Lorig 2008; Nunez 2006; Victor 2005) reported no difference in terms of emotional distress between self-management programmes and usual care between six and 12 months (SMD 0.11, 95% CI -0.06, 0.28), although moderate statistical heterogeneity was present (I^2 = 48%).

One study with 182 participants (Yip 2007) indicated that self-management programmes improved health-directed activity compared with usual care after one week (SMD 0.67, 95% CI 0.37 to 0.97) (Analysis 2.10). This result could be clinically meaningful (MD 0.64, 95% CI 0.35-0.92 points on the heiQ 1 to 6-point subscale health-directed activities). Findings of three studies with 626 participants (Crotty 2009; Lorig 2008; Yip 2007) indicate that self-management education programmes improved health-directed activity compared with usual care between six and 12 months (SMD 0.25, 95% CI 0.05 to 0.46). However, this finding is unlikely to be of clinical importance (MD 0.24, 95% CI 0.05 to 0.44 points on the heiQ 1 to 6-point subscale health-directed activities).

One study with 152 participants (Crotty 2009) reported that skill and technique acquisition improved significantly in participants in the self-management intervention compared with usual care after six months (MD 0.26, 95% CI 0.01 to 0.51) (Analysis 2.11). However, this finding is unlikely to be of clinical importance (MD 0.22, 95% CI 0.01 to 0.43 points on the heiQ 1 to 6-point subscale skill and technique acquisition).

Another study with 56 participants (Keefe 2004) reported that constructive attitudes and approaches improved significantly post-treatment in participants in the self-management intervention compared with usual care (SMD 1.04, 95% CI 0.44 to 1.63) (Analysis 2.12). This result could be clinically important (MD 0.83, 95% CI 0.35 to 1.30 points on the 1 to 6-point heiQ subscale constructive attitudes and approaches). One study with 152 participants (Crotty 2009) showed no difference in terms of constructive attitudes and approaches between self-management programmes and usual care after six months (SMD 0.11, 95% CI -0.20 to 0.43).

No between-group differences were described for any of the other reported outcomes, including social integration and support in the short term (one trial, SMD -0.19, 95% CI -0.52 to 0.15) and in the intermediate term (three trials, SMD -0.08, 95% CI -0.30 to 0.14) (Analysis 2.13) and health service navigation (two trials, SMD 0.15, 95% CI -0.03 to 0.34) (Analysis 2.14).

Comparison 3. Self-management programmes versus information only

Main outcomes

One study with 90 participants (Ackerman 2012) found no difference between a self-management education programmes and information only in improving self-management skills over six weeks (SMD 0.06, 95% CI -0.36 to 0.47) (Analysis 3.1); this is equivalent to a mean improvement of 0.03 (95% CI -0.18 to 0.24 cm) points with self-management on the 1 to 6-point heiQ subscale self-monitoring and insight. Three trials with 760 participants (Ackerman 2012; Buszewicz 2006; Hughes 2004) showed no difference between a self-management programme and information only in terms of improving self-management skills over 12 months (SMD 0.20, 95% CI -0.04 to 0.44); this is equivalent to a mean improvement of 0.10 (95% CI -0.02 to 0.22) points on the 1 to 6-point heiQ subscale self-monitoring and insight, although moderate statistical heterogeneity was present (I^2 = 41%). Hughes 2004 differed from the other two trials in that it had a higher withdrawal rate in the control group than in the intervention group (68% vs 50%), and the study incorporated a minimal number of components considered important for self-management programmes (2/8 heiQ components compared with 6/8 (Ackerman 2012) and 7/8 (Buszewicz 2006)).

One study with 90 participants (Ackerman 2012) showed no difference between a self-management programme and information only in improving participants' positive and active

engagement in life after six weeks (MD -0.08, 95% CI -0.41 to 0.26) or 12 months (MD -0.20, 95% CI -0.59 to 0.18) on the 1 to 6-point heiQ subscale positive and active engagement in life) (Analysis 3.2).

No overall between-group differences were noted in the other main efficacy outcomes: pain (intermediate term, three trials, SMD -0.07, 95% CI -0.21 to 0.08), global OA (short term, one trial, SMD 0.09, 95% CI -0.33 to 0.50; and intermediate term, three trials, SMD -0.06, 95% CI -0.28 to 0.16), self-reported function (four trials up to 12 months, SMD -0.09, 95% CI -0.22 to 0.05) and quality of life (intermediate term, two trials, SMD 0.05, 95% CI -0.10 to 0.21) (Analysis 3.3; Analysis 3.4; Analysis 3.5; Analysis 3.7).

Findings of four trials with 1,251 participants (Ackerman 2012; Buszewicz 2006; Hughes 2004; Wetzels 2005) showed no between-group differences with respect to numbers of withdrawals (RR 1.60, 95% CI 0.75, 3.40) (Analysis 3.8). However, considerable statistical heterogeneity (I^2 = 89%) was found to be due to larger differences in favour of the intervention in two small trials and small but statistically significant differences in opposite directions in the remaining two large trials.

Other outcomes

No differences between groups were noted for any of the other reported outcomes, including emotional distress (three trials up to 12 months, SMD 0.00, 95% CI -0.30 to 0.30) (Analysis 3.9); health-directed activity (short term, one trial, MD 0.24, 95% CI -0.09 to 0.57; and intermediate term, one trial, MD 0.21, 95% CI -0.14 to 0.56) (Analysis 3.10); social integration and support (short term, one trial SMD -0.02, 95% CI -0.44 to 0.40; and intermediate term, two trials, SMD -0.02, 95% CI -0.39 to 0.35) (Analysis 3.11); health service navigation (short term, one trial, MD -0.04, 95% CI -0.33 to 0.26; and intermediate term, one trial, MD 0.14, 95% CI -0.13 to 0.41) (Analysis 3.12); skill and technique acquisition (short term, one trial, MD 0.15, 95% CI -0.13 to 0.44; and intermediate term, one trial, MD -0.06, 95% CI -0.35 to 0.23) (Analysis 3.13); and constructive attitudes and approaches (short term, one trial, MD 0.16, 95% CI -0.12 to 0.44; and intermediate term, one trial, MD -0.20, 95% CI -0.50 to 0.10) (Analysis 3.14).

Comparison 4. Self-management programmes versus alternate interventions

Main outcomes

Findings of three trials with 186 participants (Keefe 1996; Keefe 2004; Murphy 2008) showed no difference between a self-management programme and an alternate intervention in terms of improving self-management skills immediately post-treatment (SMD 0.42, 95% CI -0.05 to 0.89); this is equivalent to a mean difference of 0.21 (95% CI -0.03 to 0.45) on the heiQ subscale self-monitoring and insight (Analysis 4.1), although moderate statistical heterogeneity was present (I^2 = 56%). One trial with 81 participants (Keefe 1996) indicated that self-management skills improved at 12 months for those attending a self-management programme compared with arthritis education (SMD 0.54, 95% CI 0.05 to 1.04) (Analysis 4.1); this is equivalent to a mean improvement of 0.27 (95% CI 0.03 to 0.52) points on the 1 to 6-point heiQ subscale self-monitoring and insight. This result could be of clinical significance. The failure of McKnight 2010 to report self-management skills is likely to have had little impact on these results.

Positive and active engagement in life was not reported by any of the trials comparing self-management versus alternate interventions.

No overall between-group differences were reported in terms of pain (short term, five trials, SMD 0.03, 95% CI -0.29 to 0.36; and intermediate term, two trials, SMD -0.18, 95% CI -0.56 to 0.19) (Analysis 4.2); global OA scores (one trial, SMD 0.27, 95% CI -0.13 to 0.67) (Analysis 4.3); self-reported function (short term, three trials, SMD 0.23, 95% CI -0.03 to 0.48; and intermediate term, two trials, SMD -0.17, 95% CI -0.54 to 0.20) (Analysis 4.4); functional performance (one trial, SMD -0.09, 95% CI -0.64 to 0.46) (Analysis 4.5); and quality of life in the short term (one trial, SMD 0.24, 95% CI -0.28 to 0.76), the intermediate term (two trials, SMD -0.01, 95% CI -0.28 to 0.26) and the long term (one trial, SMD -0.23, 95% CI -0.55 to 0.10) (Analysis 4.6).

Seven trials with 880 participants (Cronan 1997; Jessep 2009; Keefe 1996; Keefe 2004; Maurer 1999; McKnight 2010; Murphy 2008) showed no differences in the numbers of withdrawals between groups (RR 0.86, 95% CI 0.69 to 1.09) (Analysis 4.7).

Other outcomes

No between-group differences were noted for emotional distress in the short term (three trials, SMD 0.13, 95% CI -0.28 to 0.55) or in the intermediate term (two trials, SMD 0.18, 95% CI -0.19 to 0.55) (Analysis 4.8).

In two trials with 135 participants (Keefe 1996; Keefe 2004), constructive attitudes and approaches increased in participants in the self-management intervention compared with an alternate intervention (arthritis education and exercise) immediately post-treatment (SMD 0.92, 95% CI 0.49 to 1.34) (Analysis 4.9); this is equivalent to a mean improvement of 0.74 (95% CI 0.39 to 1.07) on the 1 to 6-point heiQ subscale constructive attitudes and approaches, which could be of clinical significance. One trial with 81 participants (Keefe 1996) reported that constructive attitudes and approaches improved significantly in participants in the self-management intervention group compared with those in an alternate intervention group after 12 months (SMD 0.62, 95% CI 0.12, 1.12); this is equivalent to a mean improvement of 0.50 (95% CI 0.10 to 0.90) points on the 1 to 6-point heiQ subscale constructive attitudes and approaches. This result could be of clinical significance.

The other outcomes in this review, including health-directed activity, social integration and support, health service navigation and skill and technique acquisition, were not reported by any of the trials comparing self-management programmes versus alternative interventions.

Comparison 5. Self-management programmes versus acupuncture

One study (Berman 2004) with 570 participants compared a self-management programme versus acupuncture. This study was considered separately from the other studies comparing self-management programmes versus alternative interventions (Comparison 4) because, unlike in the other trials, the comparison was not a behavioural intervention. This study was judged to be at high risk of bias because of lack of blinding and high attrition (see Characteristics of included studies), increasing the risk of bias in favour of acupuncture.

Main outcomes

Self-management of OA, self-monitoring and insight, positive and active engagement in life and quality of life were not reported in this trial.

No difference was found between treatment groups in global OA scores after four weeks (SMD 0.05, 95% CI -0.15 to 0.25) and after 26 weeks (SMD -0.10, 95% CI -0.32 to 0.12) (Analysis 5.1); these differences are equivalent to a decrease of -0.10 points (95% CI -0.52 to 0.31) on the WOMAC 96-point scale after four weeks and -0.21 points (95% CI -0.68 to 0.25) after 26 weeks.

Pain improved more in the acupuncture group than in the self-management group after four weeks (SMD 0.95, 95% CI 0.74 to 1.16) and after 26 weeks (SMD 1.37, 95% CI 1.13 to 1.61) (Analysis 5.2); this is equivalent to a mean reduction in pain of 2.38 cm (95% CI 1.85 to 2.90 cm) on a 10-cm VAS after four weeks and a mean reduction of 3.43 cm (95% CI 2.83 to 4.03 cm) after 26 weeks and is likely to represent a clinically meaningful difference.

Self-reported function increased in the acupuncture group compared with the self-management education programme group after four weeks (SMD 1.22, 95% CI 1.00 to 1.44) and after 26 weeks (SMD 1.53, 95% CI 1.29 to 1.77) (Analysis 5.3); this is equivalent to a mean improvement in function of 2.67 (95% CI 2.19 to 3.15) points on the 68-point WOMAC function scale after four weeks and a mean improvement of 3.35 (95% CI 2.83 to 3.88) points after 26 weeks —differences that are likely to be clinically meaningful. Functional performance increased in the acupuncture group compared with the self-management group after 26 weeks (SMD -0.30, 95% CI -0.52 to -0.08) (Analysis 5.4); this is equivalent to a mean improvement of 97.8 (95% CI 26.1 to 169.5) feet on the six-minute walk distance test and could show clinical significance.

Significantly more withdrawals were reported in the self-management group than in the acupuncture group (RR 1.96, 95% CI 1.58, 2.42) (Analysis 5.5).

Other outcomes

The other outcomes in this review, including emotional distress, health-directed activity, social integration and support, health service navigation, skill and technique acquisition and constructive attitudes and approaches, were not reported in this trial.

Sensitivity analysis

The results were robust to excluding trials that did not randomly allocate participants (results not shown). No trials were judged to be at high risk of bias for treatment allocation concealment. However, the results were robust to excluding trials with unclear treatment allocation concealment (results not shown).

Subgroup analyses

We had planned a subgroup analysis based on whether the study population consisted mainly of Caucasian, educated, older females. However, as the mean age of participants was 64.8 years across all studies, the subgroup analysis considered only whether the study population consisted mainly of Caucasian, educated females. Eight studies included predominantly Caucasian, educated females (Berman 2004; Cronan 1997; Hughes 2004; Lorig 2008; Maisiak 1996; Mazzuca 2004; McKnight 2010; Murphy 2008), and five studies did not (Allen 2010; Blixen 2004; Mazzuca 1997; Victor 2005; Yip 2007). One study (Berman 2004) was judged too different to be included in this analysis, and not enough information was available for review authors to determine subgroup status for the remaining 16 studies.

For participants' positive and active engagement in life, no subgroup analysis was possible, as not enough trials had assessed this outcome. Results for the other subgroup analyses varied by outcome. For self-management in OA and self-reported function, self-management programmes appeared to be more beneficial in trials that primarily included Caucasian, educated females compared with trials that did not primarily include this subgroup (self-management in OA: SMD 0.29, 95% CI 0.07 to 0.50 vs SMD 0.03, 95% CI -0.29 to 0.36; Analysis 6.1; self-reported function: SMD -0.20, 95% CI -0.37 to -0.02 vs SMD -0.06, 95% CI -0.21, 0.08; Analysis 6.2). On the other hand, for self-reported pain, self-management programmes appeared more beneficial in trials that did not primarily include Caucasian, educated females (SMD -0.11, 95%CI -0.30 to 0.07 vs SMD -0.20, 95% CI -0.35 to -0.05) (Analysis 6.3). No difference in withdrawals was noted between subgroups (Analysis 6.4).

Because no strong treatment effects were indicated for self-management programmes in the Results of this review, we did not conduct other subgroup analyses as described in the protocol (see Differences between protocol and review).

Funnel plots

The potential for small-study effects was explored using funnel plots in outcomes with more than 10 studies; this included self-management of OA of Analysis 3.1 (Figure 4), withdrawals of Analysis 3.3 (Figure 5) and pain of Analysis 3.4 (Figure 6). Based on the appearance of these funnel plots, we judged that the pooled results for these outcomes were not biased by small-study effects.

Figure 4. Funnel plot of comparison: 3 SMP versus Usual care/No treatment/Wait list, outcome: 3.1 Self-management of OA.

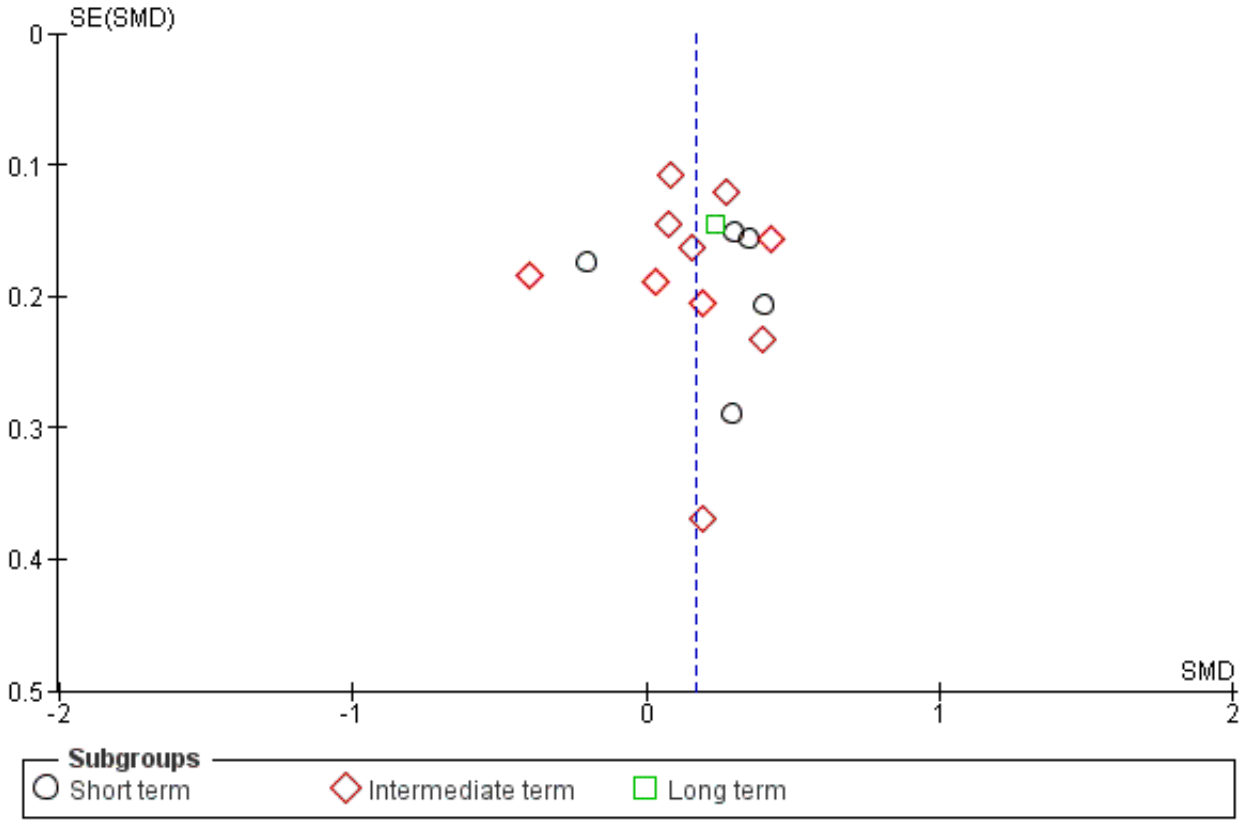


Figure 5. Funnel plot of comparison: 3 SMP versus Usual care/No treatment/Wait list, outcome: 3.3 Dropouts.

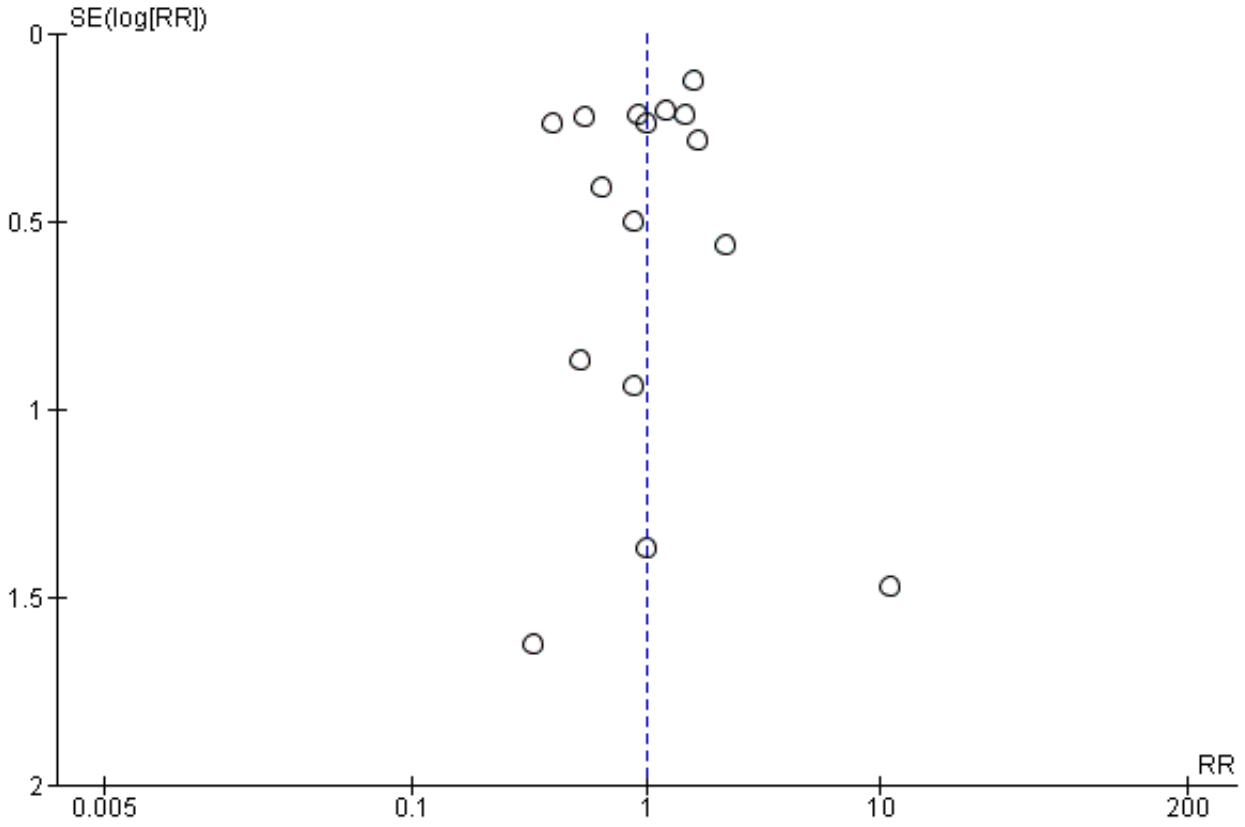
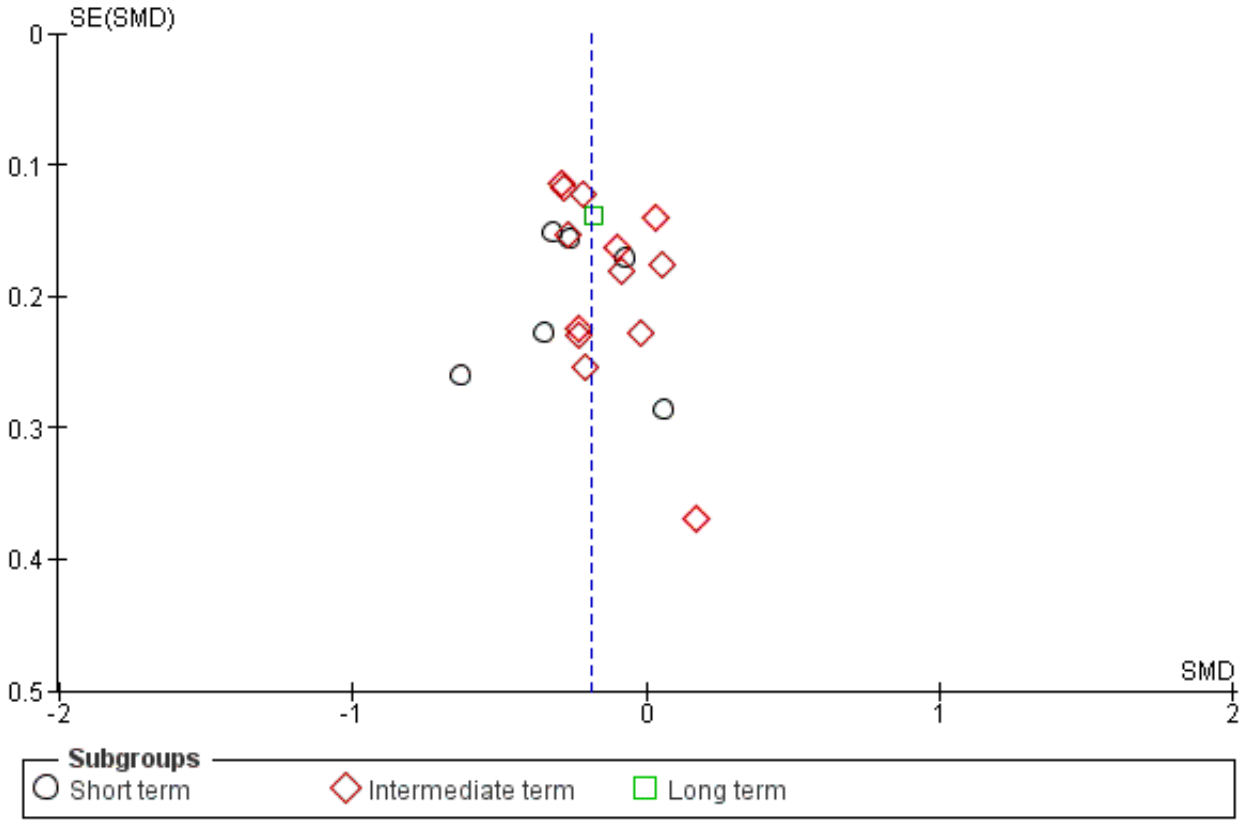


Figure 6. Funnel plot of comparison: 3 SMP versus Usual care/No treatment/Wait list, outcome: 3.4 Pain.



DISCUSSION

Summary of main results

This systematic review assessed the effects of self-management education programmes for OA by employing rigorous and systematic methods of searching, appraising and synthesising the evidence. As overall outcome data presented in the Summary of findings tables are based on the longest time points measured in each study, some differences in effect estimates from results presented will be noted according to immediate, intermediate and longer-term outcomes. Evidence of low to moderate quality indicated lack of benefit of self-management education programmes in comparison with attention control, information-only or alternate interventions. Evidence of low to moderate quality showed a small, statistically significant benefit of self-management education programmes compared with usual care; however, these findings were unlikely to be of clinical importance.

Compared with attention control, evidence of low to moderate quality indicated that self-management education programmes may not result in significant benefit at 12 months (Summary of findings for the main comparison). Although a small difference in pain favoured self management education programmes (low-quality evidence, three trials, 575 participants; SMD -0.26, 95% CI -0.44 to -0.09), this is unlikely to be of clinical importance, and no between-group differences were observed for any of the other measured main outcomes (e.g. self-management skills) (low-quality evidence, one trial, N = 344; MD 0.4 points, 95% CI -0.39 to 1.19) or for withdrawal rates (moderate-quality evidence, five trials, 937 participants; RR 1.11, 95% CI 0.78 to 1.57). No trial measured positive and active engagement in life.

Compared with usual care, moderate-quality evidence (11 trials, N = 1706) suggested small but clinically unimportant benefits favouring self-management education programmes up to 21 months (Summary of findings 2). Differences favoured self-management education programmes for self-management skills (absolute improvement (SMD 3.78%, 95% CI 0.71% to 6.85%), pain (SMD -0.19, 95% CI -0.28 to -0.1), function (SMD 0.08, 95% CI 0.09 to 0.27) and global osteoarthritis symptoms (SMD -0.28, 95% CI -0.39 to -0.17), but no between-group differences in quality of life (SMD 0.02, 95% CI -0.09 to 0.13) or in positive and active engagement in life (SMD 0.01, 95% CI -0.2 to 0.21) were noted. Evidence of low quality (16 trials, N = 3,738) suggested similar withdrawal rates (RR 0.99, 95% CI 0.74 to 1.33).

Low to moderate evidence from four studies (up to 1,251 participants in total) comparing self-management education programmes with provision of information alone showed no differences in outcome in terms of self-management of OA, positive and active engagement in life, pain, global OA scores, function, quality of life or study withdrawals (Summary of findings 3).

Evidence of low to moderate quality obtained from seven studies (up to 919 participants in total) showed no differences in outcome with respect to self-management of OA, positive and active engagement in life, pain, global OA scores, function, quality of life or study withdrawals for self-management education

programmes compared with alternative interventions (e.g. exercise, physiotherapy, social support, acupuncture) (Summary of findings 4).

This review found several modest beneficial effects of self-management education programmes although none of these was considered to be of clinical importance. Although we do not have data on patient expectations for self-management education programs in OA, a recent survey of patients with chronic low back pain has shown that, on average, recipients of care need to see effect sizes (i.e. additional improvement to natural recovery) ranging from 20% to 30% on pain and disability with interventions such as exercise and non-steroidal anti-inflammatory drugs to consider them worthwhile (Ferreira 2011). It is therefore likely that individual patients would not consider any of the identified small beneficial effects of the studied self-management education programmes worthwhile.

It is important to note that all of the efficacy outcomes that we considered were continuous outcome measures. However for chronic pain, patients may report a very good or a very poor response to treatment, making interpretation of average changes in continuous pain measures difficult (Moore 2010). Although average effect sizes in our review were small, we cannot exclude a bimodal (rather than Gaussian) distribution of response, whereby a proportion of participants may have derived a large, clinically relevant benefit of the self-management intervention, although, by extension, another group would have had to derive a large, clinically relevant harm from the intervention —a scenario that is unlikely given that the intervention of interest was self-management education programmes. However, although no significant difference in withdrawal rates was noted between self-management education programmes and control groups, and no differences were evident between trials of short or longer duration of follow-up, a high withdrawal rate across studies suggests that participant adherence to self-management education programmes, as delivered in the trials included in this review, may be less than optimal.

Overall completeness and applicability of evidence

This review was limited to evidence provided from RCTs or quasi-RCTs undertaken to assess the effects of self-management education programmes compared with receipt of information only, no treatment, usual care, waiting list control or alternative interventions that are not considered self-management education programmes.

Several trials that were potentially eligible for inclusion in this review were not included because they involved mixed populations with chronic disease or unspecified 'arthritis' (Barlow 2000; Ehrlich-Jones 2001; Goeppinger 1989; Laforest 2008; Laforest 2008a; Lindroth 1989; Lorig 1985; Lorig 1999a; Lorig 1999b; Lorig 2005; Nour 2006; Solomon 2002). These trials failed to meet our prespecified criterion that a high proportion of included participants must have OA (90% or greater), or that outcomes for people with OA must be reported separately. Although this is an arbitrary threshold, the rationale for excluding studies based on this criterion was the presumed limited applicability of the results of these studies to people with OA.

This review was also limited to studies that compared self-management education programmes versus no self-management or an alternative intervention. We did not assess the superiority of one type of self-management programme compared with another; this resulted in exclusion of other randomised trials (Coleman 2010; Hoogeboom 2010; Lorig 1998; Martire 2003a; Martire 2008; Murphy 2010).

The RCTs included in this review were generally of short duration, and limited outcome assessments were available at longer time points (i.e. longer than a year after completion of the programme); however, the lack of short-term effects of self-management programmes shown in the analysis of this review would indicate that any longer-term effects are unlikely.

We expected that apart from pain, function, quality of life and global OA score, self-management or self-efficacy would be a main efficacy outcome in studies included in this review, as improvement in this outcome should be a major goal of a self-management education programme. However, only 13 of 29 (45%) studies included in our review specified this as an outcome. Similarly, another main efficacy outcome, positive and active engagement in life, was not reported in any trials that compared self-management versus attention control or an alternative intervention.

In addition, several main outcomes in the five comparisons that we undertook were reported by only one study (Analysis 1.1; Analysis 1.3; Analysis 1.5; Analysis 3.2; Analysis 3.6; Analysis 3.10; Analysis 3.12; Analysis 3.13; Analysis 3.14; Analysis 2.11; Analysis 4.3; Analysis 4.5). Although two studies were assessed as being at high risk for reporting bias because they purportedly measured but failed to report a primary outcome for this review (Cronan 1997; McKnight 2010), and an additional seven trials were assessed as having unclear risk of reporting bias, it is unlikely that this would have appreciably altered our results, as the direction of bias is likely to have been towards the null.

Only studies conducted by Keefe et al (Keefe 1990; Keefe 1996; Keefe 2004) used the CSQ (Coping Strategies Questionnaire) (see Analysis 2.12; Analysis 4.9). Although the validity of this questionnaire has been extensively tested for the French (Irachabal 2008) and German (Verra 2006) versions, validation studies for the original (American) version of the CSQ are scarce. The study performed by Rosentiel and Keefe (Rosentiel 1983) considered that the CSQ had good internal reliability; however, other studies have questioned its construct validity (Robinson 1997; Steward 2001), and Geisser 1994 concluded that individual CSQ subscales may have greater utility in terms of examining coping, appraisal and pain adjustment compared with the composite scores.

When the characteristics of study populations were assessed using the PROGRESS-Plus framework, a predominance of older, Caucasian, educated females was seen in the included studies. This potentially raises questions regarding the applicability of results from this review to other groups. In investigating whether this combination of factors moderate outcome, we performed subgroup analyses that yielded conflicting results. Self-management programmes appeared more beneficial for Caucasian, educated female participants with respect to self-management of OA and self-reported function, but for self-reported pain, self-management programmes appeared more beneficial in trials that did not primarily include this subgroup. Our data contrast with the findings of a systematic review of RCTs of self-management interventions for chronic musculoskeletal pain performed to identify predictors, moderators and mediators of

outcome (Miles 2011). Upon examining 16 RCTs involving 4,047 participants that included appropriate analyses of moderators and/or mediators, Miles 2011 found only tentative evidence to support age and gender as moderators of outcomes resulting from self-management interventions because the data were insufficient. However, those review authors found strong evidence that self-efficacy, depression, pain catastrophising and physical activity are all important influences on participant outcome, irrespective of treatment, although evidence suggested that pain catastrophising and physical activity can mediate outcomes resulting from self-management.

Many studies in our review did not provide enough information to permit assessment of all items in the PROGRESS-Plus framework. On average, information was available for 5.6 of the nine items. This lack of information means that we were unable to fully assess whether the results of this review are applicable to all individuals with OA. As pointed out by Furler 2011, disadvantaged people might be less able to access support for self-management from healthcare professionals, and the quality of services may be lower, potentially increasing rather than decreasing health inequities. In addition to the PROGRESS-Plus framework, we assessed health literacy in each study population, as it is likely to be a key determinant of a person's ability to optimally manage his or her health and to ensure equitable access to and use of services. Only 4 of 29 (14%) trials provided any information on the health literacy of their study populations.

Quality of the evidence

For the main comparison, self-management education programmes versus attention control (five studies, N = 937), the overall quality of evidence was graded as low to moderate (Summary of findings for the main comparison). Evidence was downgraded because of methodological limitations of the studies, including lack of participant blinding, inadequate randomisation or concealment of allocation and greater numbers of withdrawals in the intervention group.

The overall quality of evidence for self-management education programmes compared with usual care was low to moderate (16 studies, N = 3,738) (Summary of findings 2). Evidence was downgraded, as participants and study personnel were not blind to group allocations, and some studies had inadequate randomisation or concealment of allocation.

The quality of evidence comparing self-management education programmes versus the provision of information alone (4 studies, N = 340) was graded as low to moderate using the GRADE approach (Summary of findings 3). Evidence was downgraded because of lack of participant blinding in all studies and because of additional methodological issues, including inadequate randomisation and unbalanced losses to follow-up across groups.

For studies comparing self-management education programmes versus alternative interventions (7 studies, N = 919), the overall quality of evidence was moderate (Summary of findings 4). Evidence was downgraded because of lack of blinding of participants and study personnel and unclear randomisation and concealment of allocation in some studies.

Although the evidence was downgraded, we are not convinced that further well-designed trials would substantially change the estimate of the effects, or the direction of the effects. Confounding from biases across studies would have likely favoured self-management education programmes; thus it is unlikely that correcting for these biases would overturn the direction of the results; it may, in fact, drive some of the small but clinically unimportant improvements seen with self-management over usual care towards the null (i.e. no significant differences between groups).

As well as grading the overall quality of the evidence, which relies on an assessment of risk of bias of individual trials, a novel aspect of our review included an attempt to assess the quality of the self-management interventions that were being evaluated in the trials on the basis of the eight domains described in the heiQ (Table 4). At least some evidence indicates that elements of skill and technique acquisition were addressed in 94% of the self-management education programmes, health-directed activity was addressed in 85% and self-monitoring and insight in 79%. However, social integration and support were addressed in only 12% of the self-management education programmes. On average, interventions resulted in evidence of delivery of about four of the eight components (range two to seven), and limited evidence suggested that they were delivered in a high-quality fashion. It is important to note that although we used inclusion of a component as a proxy for quality, we were unable to ascertain how well the included components were delivered. Furthermore, it could be argued that although we based our assessment on the heiQ, key therapeutic components of self-management interventions remain unknown.

Potential biases in the review process

We believe that all relevant published studies were identified for inclusion in this review. A thorough search strategy was devised, and all major databases were searched for relevant studies with no language restrictions applied. Two review authors assessed the trials for inclusion in the review, and a third review author adjudicated any discussions or discrepancies.

Apart from the risk of bias of the included trials, the biggest limitation of the review process was that the self-management education programmes in the trials differed in mode, personnel, delivery method and duration. Moreover, the trials varied in their outcome measures. For several outcomes, such as pain and function, we elected to pool different measures that may not necessarily be measuring the exact same concept. All post hoc decisions regarding choice of outcome data for inclusion in analyses of this review were recorded in the Notes section of Characteristics of included studies.

Agreements and disagreements with other studies or reviews

Previous systematic reviews that have investigated the effects of self-management interventions for OA have reported broadly similar findings, although their inclusion criteria have all varied in some way from ours (e.g. inclusion of trials of interventions that provided only information, inclusion of mixed populations of arthritis or musculoskeletal pain in general, inclusion of trials in older participants only, inclusion of trials that included a combination of exercise and self-management education) (Chodosh 2005; Devos-Comby 2006; Du 2011; Smith 2009; Walsh 2006; Warsi 2003).

Chodosh 2005 included 14 studies for OA in a meta-analysis that assessed chronic disease self-management programmes in older adults (age criterion for inclusion not specified). Although minimal overlap of included trials with our review was noted (only four of their included trials were included in our review, and an additional three studies included in their review were excluded from ours because they included a mixed study population, and data for people with OA were not reported separately), they drew similar conclusions—pooled effects of self-management interventions were statistically significant but clinically trivial for pain and function outcomes. Warsi 2003 included 15 trials that investigated self-management education in people with OA or mixed populations; these review authors found similarly small benefits in terms of pain and function and noted that the overall high dropout rate of 19% raises concerns about the validity of these findings.

Devos-Comby 2006 included 16 studies investigating exercise and/or self-management interventions for participants with knee OA. Compared with control interventions, these review authors found a significant but modest benefit of self-management interventions in terms of psychological but not physical well-being. In contrast, they found that exercise regimens led to improvement in physical health (by self-report and direct measures) and in overall impact of OA, while perceived psychological health remained unchanged. Walsh 2006 included 10 randomised controlled trials in a systematic review of combined exercise and self-management programmes for people with OA knee or hip. Although most trials reported significant benefits in terms of pain (seven of 10 trials) and function (eight of 10 trials), these review authors identified many methodological weaknesses, and clinical heterogeneity precluded meta-analysis. Smith 2009 included 13 trials in a systematic review of self-management education and/or exercise interventions for knee OA and concluded that no evidence indicated that self-management reduced pain or improved function or quality of life compared with a wait list or no treatment group, and limited evidence from one trial suggested that self-management was more effective than standard care in reducing pain and improving function. Similar to Devos-Comby 2006, these review authors reported that the exercise component appeared to provide a small but significant benefit in terms of reducing pain, improving function and improving aspects of quality of life. Taken together, these reviews suggest that programmes that include an exercise component, as well as components directed at improving psychological outcomes, may be worthwhile, but lack of high-quality evidence precludes confident extrapolation of these research findings into clinical practice.

Du 2011 performed a systematic review to determine the effectiveness of self-management programmes for pain and disability in chronic musculoskeletal pain conditions. Trials were included only if interventions focused primarily on managing pain and minimising disability, and pain and disability were the primary outcomes of interest. These review authors included 16 trials of mixed populations of arthritis (only four of their included trials were included in our review, and an additional four studies included in their review were excluded from ours because they included a mixed study population, and data for people with OA were not reported separately). These review authors found that self-management programmes resulted in only small to moderate effects in terms of improving pain and disability in the long term (and no improvements in disability were seen in the medium term).

In keeping with the findings of our review, Jüni 2006 devised a league table of selected interventions for osteoarthritis in terms of effect sizes typically found in large-scale randomised controlled trials (comprising at least 100 participants) and found that formal participant education interventions have a minimal effect on pain (effect size -0.10, 95% CI -0.19 to 0.01— for an approximate difference of -0.25 on a 10-cm visual analogue pain scale).

Reviews on self-management education programmes in other chronic conditions have generally found small positive results. One Cochrane review examining self-management education programmes in chronic obstructive pulmonary disease found that self-management education was associated with a reduction in hospital admissions without detrimental effects on other outcome parameters (Effing 2007). Another Cochrane review found that education in asthma self-management, which involves self-monitoring by peak expiratory flow or by symptoms, coupled with regular medical review and a written action plan, has been found to improve health outcomes in adults with asthma (Gibson 2002).

A Cochrane review of self-management education programmes in type 2 diabetes found that group-based training in self-management strategies is effective in improving fasting blood glucose levels, glycate haemoglobin and diabetes knowledge and in reducing systolic blood pressure levels, body weight and the requirement for diabetes medication (Deakin 2005). Finally, in a review that assessed the effects of self-management education programmes in chronic conditions, lay-led self-management education programmes were found to result in small, short-term improvements in participants' self-management, self-rated health, cognitive symptom management and healthcare use (Foster 2009). It may be that the nature of the disease influences the outcomes that can be achieved through self-management education programmes. Unlike chronic obstructive pulmonary disease, for asthma and diabetes, in which clear demonstrable complications or deterioration in the condition can occur as a result of poor management, the management of OA is largely concerned with managing the symptoms of the persistent underlying pathological condition and impact on quality of life as measured by subjective tools. Further, other chronic conditions tend to be evaluated on the basis of specific clinical outcomes relevant to the condition using measurement tools or tests that can measure the outcomes with high precision.

We found that data were insufficient to allow review authors to perform a preplanned subgroup analysis to explore whether a relationship could be discerned between any of the component domains addressed in the self-management education programmes and participant outcomes (Pitt 2011). However, Carnes 2012 performed a systematic review to specifically uncover the evidence for effectiveness of different self-management course characteristics and components for chronic musculoskeletal pain. Upon review of 46 RCTs involving 8,539 participants, these review authors reported more beneficial effects in group-delivered courses that included healthcare professional input and slightly more consistent beneficial effects for courses with a psychological component.

AUTHORS' CONCLUSIONS

Implications for practice

We found low to moderate evidence suggesting that self-management education programmes, as delivered in the studies included in this review, result in no or small benefits in patients with OA. Compared with attention control, the comparator with the least risk of bias, as it controls for any effect of contact time with programme providers, self-management education programmes probably do not improve self-management skills, pain, osteoarthritis symptoms, function or quality of life in people with OA, and their effects on positive and active engagement in life are unknown. Compared with usual care, self-management programmes do not result in more positive and active engagement in life but may improve self-management skills, pain, osteoarthritis symptoms and function. However, all apparent benefits were small and unlikely to be of clinical importance. Compared with provision of information alone or alternative interventions (e.g. exercise, physiotherapy, social support, acupuncture), self-management education programmes do not improve self-management of OA, positive and active engagement in life, pain, global OA scores, function or quality of life. We found no evidence that self-management programmes cause harm.

Implications for research

Although we downgraded the evidence to moderate or low for most outcomes, we believe that further studies investigating the effects of self-management education programmes, as delivered in the trials in this review, are not likely to substantially change the conclusions of this review. Confounding from biases across studies would have likely favoured self-management; thus it is unlikely that correcting for these biases would overturn the direction of the results and may, in fact, drive some of the small but clinically unimportant improvements seen with self-management over usual care, towards the null (i.e. no significant between-group differences).

However, it is possible that other models of self-management education programmes that differ in mode of delivery, type of audience, duration and frequency of sessions, personnel used to teach self-management skills, etc., may enhance self-management of OA. Further trials of different self-management education programmes, particularly those that are tailored to the individual, may therefore be warranted.

We found a mismatch between the aims of self-management education programmes and the outcomes used to measure success. Studies often sought to measure pain as a main outcome. Although this is certainly a desirable outcome, OA is a chronic condition, and conservative interventions, including self-management education programmes, are not intended to 'cure' pain but rather to enable people to have a reasonable life despite their ongoing pain. The stated aim of self-management education programmes for people with OA is to educate people about their condition and teach them how best to manage their symptoms. It seems reasonable therefore that indicators of knowledge, self-management skills and self-efficacy should be included as key outcomes in studies assessing the effects of a self-management education programme in OA. We suggest that a measure of fatigue should be considered as an outcome in the next updated version of this review. Most studies (27/29 studies) did not include fatigue as an outcome, although it has been reported to be a major problem in patients suffering from chronic diseases such as OA (Snijders 2011).

Comprehensive evaluation of self-management education programmes using robust and well-validated tools will improve this field. The heiQ was developed from patients' and clinicians' perspectives on what are valued outcomes of self-management programmes (Osborne 2007), and it has been a useful evaluation tool in many settings (e.g. Cadhilac 2011; Francis 2009; Greenhalgh 2009; Osborne 2011; Packer 2012; Wanitkun 2011). Although self-efficacy scales are often applied, the interpretation, validity and value of the derived scores have been debated; these require further rigorous testing in relevant clinical settings in well-defined populations (Brady 1997; Brady 2011).

Further trials should adequately describe the self-management education programme that they deliver to enable better assessment of the therapeutic quality of the programme and to enable an assessment of the likely impact of intended outcomes. This will also allow between-study comparisons. The methods we employed to describe the self-management education programmes included in this review may serve as a good starting point. We suggest that trials include a more detailed description of study participants, taking into account the expanded PROGRESS-Plus framework and health literacy, as well as other potential predictors and moderators of treatment outcome. We detected differences between subgroups in several outcomes, suggesting that it might be important for future studies to explore issues of health equity.

ACKNOWLEDGEMENTS

We wish to thank Denise O'Connor and Simon French, who contributed to writing the original protocol; and Louise Falzon for designing the search strategy and undertaking the searches.

We wish to thank Alina Roman who assisted in initial study selection.

The review authors wish to acknowledge Dr K Bertocci for sending a copy of her thesis, and Dr N Foster for sending a copy of her article. The review authors also wish to acknowledge Drs K Allen, B Berman, M Buszewicz, S Coleman, E Hansson, M Hurley, R Maisiak, L Martire, S Murphy, I Ackerman, R Schumacher, K Lorig and J Sit for providing additional information on request.

REFERENCES

References to studies included in this review

Ackerman 2012 *{published and unpublished data}*

* Ackerman IN, Buchbinder R, Osborne RH. Challenges in evaluating an Arthritis Self-Management Program for people with hip and knee osteoarthritis in real-world clinical settings. *Journal of Rheumatology* 2012;**39**(5):1047-55.

Osborne RH, Buchbinder R, Ackerman IN. Can a disease-specific education program augment self-management skills and improve Health-Related Quality of Life in people with hip or knee osteoarthritis?. *BMC Musculoskeletal Disorders* 2006;**7**:90. [DOI: Cn-00574027; PUBMED: 17134516]

Allen 2010 *{published and unpublished data}*

Allen KD, Oddone EZ, Coffman CJ, Datta SK, Juntilla KA, Lindquist JH, et al. Telephone-based self-management of osteoarthritis: a randomized trial. *Annals of Internal Medicine* 2010;**153**(9):570-9.

Allen KD, Oddone EZ, Stock JL, Coffman CJ, Lindquist JH, Juntilla KA, et al. The self-management of osteoarthritis in veterans (SeMOA) study: design and methodology. *Contemporary Clinical Trials* 2008;**29**(4):596-607.

Berman 2004 *{published and unpublished data}*

Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. *Annals of Internal Medicine* 2004;**141**(12):901-10. [DOI: Cn-00505978; PUBMED: 15611487]

Blixen 2004 *{published data only}*

Blixen CE, Bramstedt KA, Hammel JP, Tiley BC. A pilot study of health education via a nurse-run telephone self-management programme for elderly people with osteoarthritis. *Journal of Telemedicine and Telecare* 2004;**10**(1):44-9. [DOI: Cn-00470756; PUBMED: 15006216]

Buszewicz 2006 *{published and unpublished data}*

Buszewicz M, Rait G, Griffin M, Nazareth I, Patel A, Atkinson A, et al. Self management of arthritis in primary care: randomised controlled trial. *BMJ* 2006;**333**(7574):879. [DOI: Cn-00568416; PUBMED: 17040926]

Patel A, Buszewicz M, Beecham J, Griffin M, Rait G, Nazareth I, et al. Economic evaluation of arthritis self management in primary care. *BMJ* 2009;**339**:b3532. [DOI: Cn-00721961; PUBMED: 19773329]

Calfas 1992 *{published data only}*

Calfas KJ. The effects of a cognitive-behavioral intervention upon function and depression in osteoarthritis. PhD/ Dissertation Abstracts International 1991.

Calfas KJ, Kaplan RM, Ingram RE. One-year evaluation of cognitive-behavioral intervention in osteoarthritis. *Arthritis Care and Research* 1992;**5**(4):202-9. [DOI: Cn-00090424; PUBMED: 1489766]

Cronan 1997 *{published data only}*

Cronan TA, Groessl E, Kaplan RM. The effects of social support and education interventions on health care costs. *Arthritis and Research* 1997;**10**(2):99-110. [DOI: Cn-00255687; PUBMED: 9313398]

Cronan TA, Hay M, Groessl E, Bigatti S, Gallagher R, Tomita M. The effects of social support and education on health care costs after three years. *Arthritis Care and Research* 1998;**11**(5):326-34. [DOI: Cn-00157316; PUBMED: 9830877]

Gallagher RA, Miller C, Cronan TA, Groessl E. Gender differences in participation and responsiveness to a health intervention for older Americans. *Women & Health* 1997;**25**(3):63-81. [DOI: Cn-00142992; PUBMED: 9273984]

Groessl E, Cronan TA. A cost analysis of self-management programs for people with chronic illness. *American Journal of Community Psychology* 2000;**28**(4):455-80. [DOI: Cn-00330467; PUBMED: 10965386]

Groessl EJ. Participation and empowerment: factors related to positive outcomes in a health intervention with older people with osteoarthritis. PhD/University of California, San Diego 1999.

Crotty 2009 *{published data only}*

Crotty M, Prendergast J, Battersby MW, Rowett D, Graves SE, Leach G, et al. Self-management and peer support among people with arthritis on a hospital joint replacement waiting list: a randomised controlled trial. *Osteoarthritis and Cartilage* 2009;**17**(11):1428-33. [DOI: Cn-00743584; PUBMED: 19486959]

Hansson 2010 *{published and unpublished data}*

Hansson EE, Jonsson-Lundgren M, Ronnheden AM, Sorensson E, Bjarnung A, Dahlberg LE. Effect of an education programme for patients with osteoarthritis in primary care—A randomized controlled trial. *BMC Musculoskeletal Disorders* 2010;**11**:244.

Heuts 2005 *{published data only}*

Heuts PH, de Bie R, Drietelaar M, Aretz K, Hopman-Rock M, Bastiaenen CH, et al. Self-management in osteoarthritis of hip or knee: a randomized clinical trial in a primary healthcare setting. *The Journal of Rheumatology* 2005;**32**(3):543-9. [DOI: Cn-00512972; PUBMED: 15742451]

Hopman-Rock 2000 *{published data only}*

Hopman-Rock M, Westhoff MH. The effects of a health educational and exercise program for older adults with osteoarthritis for the hip or knee. *Journal of Rheumatology* 2000;**27**(8):1947-54. [DOI: Cn-00331216; PUBMED: 10955337]

Hughes 2004 *{published data only}*

Hughes SL, Seymour RB, Campbell R, Pollak N, Huber G, Sharma L. Impact of the fit and strong intervention on older adults with osteoarthritis. *The Gerontologist* 2004;**44**(2):217-28. [DOI: Cn-00469463; PUBMED: 15075418]

Hughes SL, Seymour RB, Campbell RT, Huber G, Pollak N, Sharma L, et al. Long-term impact of Fit and Strong! on older adults with osteoarthritis. *The Gerontologist* 2006;**46**(6):801-14. [DOI: Cn-00617581; PUBMED: 17169935]

Hurley 2007 *{published data only}*

Hurley M, Jessop S, Walsh N, Ratcliffe J. A community-based integrated rehabilitation programme for chronic knee pain. *Rheumatology* 2009;**48**:i4.

Hurley M, Walsh N, Mitchell H, Nicholas J. Improvements in physical function were sustained for 2 1/2 years following escape-knee pain: an integrated rehabilitation programme for chronic knee pain. *Rheumatology* 2010;**49**:i16-17.

Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E, et al. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. *Arthritis and Rheumatism* 2007;**57**(7):1211-19. [DOI: Cn-00619769; PUBMED: 17907147]

Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Williamson E, Jones RH, et al. Economic evaluation of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain. *Arthritis and Rheumatism* 2007;**57**(7):1220-9. [DOI: Cn-00619770; PUBMED: 17907207]

Jessep 2009 *{published data only}*

Jessep SA, Walsh NE, Ratcliffe J, Hurley MV. Long-term clinical benefits and costs of an integrated rehabilitation programme compared with outpatient physiotherapy for chronic knee pain. *Physiotherapy* 2009;**95**(2):94-102. [DOI: Cn-00718549; PUBMED: 19627690]

Keefe 1990 *{published data only}*

Keefe FJ, Caldwell DS, Williams DA, Gil KM, Mitchell D, Robertson C, Martinez S, et al. Pain coping skills training in the management of osteoarthritic knee pain: II. Follow-up results. *Behavior Therapy* 1990;**21**(4):435-47. [DOI: Cn-00364121]

Keefe FJ, Caldwell DS, Williams DA, Gil KM, Mitchell D, Robertson C, Martinez S, et al. Pain coping skills training in the management of osteoarthritic knee pain: a comparative study. *Behavior Therapy* 1990;**21**(1):49-62. [DOI: Cn-00388097]

Keefe 1996 *{published data only}*

Keefe FJ, Caldwell DS, Baucom D, Salley A, Robinson E, Timmons K, et al. Spouse-assisted coping skills training in the management of knee pain in osteoarthritis: long-term followup results. *Arthritis Care and Research* 1999;**12**(2):101-11. [DOI: Cn-00168154; PUBMED: 10513498]

Keefe FJ, Caldwell DS, Baucom D, Salley A, Robinson E, Timmons K, et al. Spouse-assisted coping skills training in the management of osteoarthritic knee pain. *Arthritis Care and Research* 1996;**9**(4):279-91. [DOI: Cn-00135816; PUBMED: 8997917]

Keefe 2004 *{published data only}*

Keefe FJ, Blumenthal J, Baucom D, Affleck G, Waugh R, Caldwell DS, et al. Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. *Pain* 2004;**110**(3):539-49. [DOI: Cn-00490497; PUBMED: 15288394]

Lorig 2008 *{published data only}*

Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. *Arthritis and Rheumatism* 2008;**59**(7):1009-17. [DOI: Cn-00640424; PUBMED: 18576310]

Maisiak 1996 *{published and unpublished data}*

Maisiak R, Austin J, Heck L. Health outcomes of two telephone interventions for patients with rheumatoid arthritis or osteoarthritis. *Arthritis and Rheumatism* 1996;**39**(8):1391-9. [DOI: Cn-00127841; PUBMED: 8702449]

Martire 2007 *{published and unpublished data}*

Martire LM, Schulz R, Keefe FJ, Rudy TE, Starz TW. Couple-oriented education and support intervention: effects on individuals with osteoarthritis and their spouses. *Rehabilitation Psychology* 2007;**52**(2):121-32. [DOI: Cn-00642217]

Maurer 1999 *{published and unpublished data}*

Maurer BT, Stern AG, Kinossian B, Cook KD, Schumacher HR. Osteoarthritis of the knee: isokinetic quadriceps exercise versus an educational intervention. *Archives of Physical Medicine and Rehabilitation* 1999;**80**(10):1293-9. [DOI: Cn-00168433; PUBMED: 10527090]

Mazzuca 1997 *{published data only}*

Mazzuca SA, Brandt KD, Katz BP, Chambers M, Byrd D, Hanna M. Effects of self-care education on the health status of inner-city patients with osteoarthritis of the knee. *Arthritis and Rheumatism* 1997;**40**(8):1466-74. [DOI: Cn-00142539; PUBMED: 9259427]

Mazzuca SA, Brandt KD, Katz BP, Hanna MP, Melfi CA. Reduced utilization and cost of primary care clinic visits resulting from self-care education for patients with osteoarthritis of the knee. *Arthritis and Rheumatism* 1999;**42**(6):1267-73. [DOI: Cn-00164125; PUBMED: 10366121]

Mazzuca 2004 *{published data only}*

Mazzuca SA, Brandt KD, Katz BP, Ragozzino LR, G'Sell PM. Can a nurse-directed intervention reduce the exposure of patients with knee osteoarthritis to nonsteroidal antiinflammatory drugs?. *Journal of Clinical Rheumatology* 2004;**10**(6):315-22. [DOI: Cn-00516982]

McKnight 2010 *{published data only}*

McKnight PE, Afram A, Kashdan TB, Kasle S, Zautra A. Coping self-efficacy as a mediator between catastrophizing and physical functioning: treatment target selection in an osteoarthritis sample. *Journal of Behavioral Medicine* 2010;**33**(3):239-49. [DOI: Cn-00750770; PUBMED: 20177766]

McKnight PE, Kasle S, Going S, Villanueva I, Cornett M, Farr J, et al. A comparison of strength training, self-management, and the combination for early osteoarthritis of the knee. *Arthritis Care and Research* 2010;**62**(1):45-53. [DOI: Cn-00734981; PUBMED: 20191490]

Murphy 2008 *{published and unpublished data}*

Murphy SL, Strasburg DM, Lyden AK, Smith DM, Koliba JF, Dadabhoy DP, et al. Effects of activity strategy training on pain and physical activity in older adults with knee or hip osteoarthritis: a pilot study. *Arthritis and Rheumatism* 2008;**59**(10):1480-7.

Nunez 2006 *{published data only}*

Nunez M, Nunez E, Segur JM, Macule F, Quinto L, Hernandez MV, et al. The effect of an educational program to improve health-related quality of life in patients with osteoarthritis on waiting list for total knee replacement: a randomized study. *Osteoarthritis and Cartilage* 2006;**14**(3):279-85.

Victor 2005 *{published data only}*

Victor CR, Ross F, Axford J. Capturing lay perspectives in a randomized control trial of a health promotion intervention for people with osteoarthritis of the knee. *Journal of Evaluation in Clinical Practice* 2004;**10**(1):63-70. [DOI: Cn-00469710; PUBMED: 14731152]

Victor CR, Triggs E, Ross F, Lord J, Axford JS. Lack of benefit of a primary care-based nurse-led education programme for people with osteoarthritis of the knee. *Clinical Rheumatology* 2005;**24**(4):358-64. [DOI: Cn-00528413; PUBMED: 15937632]

Wetzels 2005 *{published data only}*

Wetzels R, van Weel C, Grol R, Wensing M. Family practice nurses supporting self-management in older patients with mild osteoarthritis: a randomized trial. *BMC Family Practice* 2008;**9**:7. [DOI: Cn-00622367; PUBMED: 18226255]

Yip 2007 *{published and unpublished data}*

Yip Y, Sit JW, Wong DYS, Chong SYC, Chung L. A 1-year follow-up of an experimental study of a self-management arthritis programme with an added exercise component of clients with osteoarthritis of the knee. *Psychology, Health and Medicine* 2008;**13**(4):402-14.

Yip YB, Sit JW, Fung KK, Wong DY, Chong SY, Chung LH, et al. Effects of a self-management arthritis programme with an added exercise component for osteoarthritic knee: randomized controlled trial. *Journal of Advanced Nursing* 2007;**59**(1):20-8. [DOI: Cn-00618227; PUBMED: 17559610]

Yip YB, Sit JW, Fung KK, Wong DY, Chong SY, Chung LH, et al. Impact of an arthritis self-management programme with an added exercise component for osteoarthritic knee sufferers on improving pain, functional outcomes, and use of health care services: an experimental study. *Patient Education and Counseling* 2007;**65**(1):113-21. [DOI: Cn-00572535; PUBMED: 17010554]

References to studies excluded from this review

Barlow 2000 *{published data only}*

Barlow JH, Turner AP, Wright CC. A randomized controlled study of the Arthritis Self-Management Programme in the UK. *Health Education Research* 2000;**15**:665-80.

Bezalel 2010 *{published data only}*

Bezalel T, Carmeli E, Katz-Leurer M. The effect of a group education programme on pain and function through knowledge acquisition and home-based exercise among patients with knee osteoarthritis: a parallel randomised single-blind clinical trial. *Physiotherapy* 2010;**96**:137-43.

Coleman 2010 *{published data only}*

Coleman S. To compare the effectiveness of OAK self-management program, delivered by health professionals, with the Arthritis Self-Management Program (ASMP), delivered by trained lay leaders, as determined by improvements in pain, quality of life and physical function. World Health Organization International Clinical Trials Registry Platform.

Coleman S, McQuade J, Rose J Inderjeeth C, Carroll G, Briffa NK. Self-management for osteoarthritis of the knee: does mode of delivery influence outcome?. *BMC Musculoskeletal Disorders* 2010;**11**:56.

Ehrlich-Jones 2001 *{unpublished data only}*

Ehrlich-Jones LS. Self-care behavior of African-American elderly with arthritis. PhD/University of Illinois at Chicago, 2001.

Ettinger 1997 *{published data only}*

Ettinger WH, Burns R, Messier SP, Applegate W, Rejeski WJ, Morgan T, et al. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis: The Fitness Arthritis and Seniors Trial (FAST). *JAMA* 1997;**277**:25-31.

Fernandes 2009 *{published data only}*

Fernandes L, Storheim K, Nordsletten L, et al. Effect of patient education and supervised exercise in patients with hip osteoarthritis: a randomized clinical trial. *Osteoarthritis and Cartilage* 2009;**17**:S170.

Fernandes 2010 *{published data only}*

Fernandes L, Storheim K, Sandvik L, Nordsletten L, Risberg MA. Efficacy of patient education and supervised exercise vs patient education alone in patients with hip osteoarthritis: a single blind randomized clinical trial. *Osteoarthritis and Cartilage* 2010;**18**:1237-43.

Focht 2005 *{published data only}*

Focht BC, Rejeski W, Ambrosius WT, Katula JA, Messier SP. Exercise, self-efficacy, and mobility performance in overweight and obese older adults with knee osteoarthritis. *Arthritis Care and Research* 2005;**53**:659-65.

Goeppinger 1989 *{published data only}*

Goeppinger J, Arthur MW, Baglioni AJ Jr, Brunk SE, Brunner CM. A reexamination of the effectiveness of persons with arthritis of self-care education. *Arthritis and Rheumatism* 1989;**32**(6):705-16.

Hoogeboom 2010 *{published data only}*

Hoogeboom TJ, Stukstette MJPM, De Bie RA, Cornelissen J, den Broeder AA, van den Ende CH. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. *BMC Musculoskeletal Disorders* 2010;**11**:142.

Laforest 2008 *{published data only}*

Laforest S, Nour K, Gignac M, Gauvin L, Parisien M, Poirier MC. Short-term effects of a self-management intervention on health status of housebound older adults with arthritis. *Journal of Applied Gerontology* 2008;**27**:539-67.

Laforest 2008a *{published data only}*

Laforest S, Nour K, Parisien M, Poirier MC, Gignac M, Lankoande H. "I'm taking charge of my arthritis": designing a targeted self-management program for frail seniors. *Physical and Occupational Therapy in Geriatrics* 2008;**26**:45-66.

Lindroth 1989 *{published data only}*

Lindroth Y, Bauman A, Barnes C, McCredie M, Brooks PM. A controlled evaluation of arthritis education. *British Journal of Rheumatology* 1989;**28**:7-12.

Lorig 1985 *{published data only}*

Lorig K, Lubeck D, Kraines RG, Seleznick M, Holman HR. Outcomes of self-help education for patients with arthritis. *Arthritis and Rheumatism* 1985;**28**(6):680-5.

Lorig 1998 *{published data only}*

Lorig K, Gonzalez VM, Laurent DD, Morgan L, Laris BA. Arthritis self-management program variations: three studies. *Arthritis Care and Research* 1998;**11**:448-54.

Lorig 1999a *{published data only}*

Lorig KR, Sobel DS, Stewart AL, Brown BW Jr, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization. *Medical Care* 1999;**37**:5-14.

Lorig 1999b *{published data only}*

Lorig K, Gonzalez VM, Ritter P. Community-based Spanish language arthritis education program: a randomized trial. *Medical Care* 1999;**37**(9):957-63.

Lorig 2005 *{published data only}*

Lorig K, Ritter PL, Plant K. A disease-specific self-help program compared with a generalized chronic disease self-help program for arthritis patients. *Arthritis and Rheumatism* 2005;**53**(6):950–7.

Martire 2003a *{published data only}*

Martire LM, Schulz R, Keefe FJ, Starz TW, Osial TA Jr, Dew MA, et al. Feasibility of a dyadic intervention for management of osteoarthritis: a pilot study with older patients and their spousal caregivers. *Aging and Mental Health* 2003;**7**:53-60.

Martire 2008 *{published data only}*

Martire LM, Schulz R, Keefe FJ, Rudy TE, Starz TW. Couple-oriented education and support intervention for osteoarthritis: effects on spouses' support and responses to patient pain. *Families, Systems and Health: The Journal of Collaborative Family HealthCare* 2008;**26**:185-95.

Murphy 2010 *{published data only}*

Murphy SL, Lyden AK, Smith DM, Dong Q, Koliba JF. Effects of a tailored activity pacing intervention on pain and fatigue for adults with osteoarthritis. *American Journal of Occupational Therapy* 2010;**64**:869-76.

Nour 2006 *{published data only}*

Nour K, Laforest S, Gauvin L, Gignac M. Behavior change following a self-management intervention for housebound older adults with arthritis: an experimental study. *International Journal of Behavioral Nutrition and Physical Activity* 2006;**3**:12.

Solomon 2002 *{published data only}*

Solomon DH, Warsi A, Brown-Stevenson T, Farrell M, Guathier S, Mikels D, et al. Does self-management education benefit all populations with arthritis? A randomized controlled trial in a primary care physician network. *The Journal of Rheumatology* 2002;**29**:362-8.

References to studies awaiting assessment

Allen 2011 *{published data only}*

Allen KD, Bosworth HB, Coffman C, Lindquist JH, Sperber NR, Weinberger M, et al. Effects of a telephone based osteoarthritis self-management program on communication with health care providers. *Arthritis and Rheumatism* 2011;**63**(Suppl 10):1582.

Coleman 2012 *{published data only}*

Coleman S, Briffa NK, Carroll G, Inderjeeth C, Cook N, McQuade J. A randomised controlled trial of a self-management education program for osteoarthritis of the knee delivered by health care professionals. *Arthritis Research and Therapy* 2012;**14**(1):R21.

Hurley 2012 *{published data only}*

Hurley MV, Walsh NE, Mitchell H, Nicholas J, Patel A. Long-term outcomes and costs of an integrated rehabilitation program for chronic knee pain: a pragmatic, cluster randomized, controlled trial. *Arthritis Care and Research* 2012;**64**(2):238-47.

Schlenk 2011 *{published data only}*

Schlenk EA, Lias JL, Sereika SM, Dunbar-Jacob J, Kwoh CK. 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.

Somers 2012 *{published data only}*

Huebner JL, Landerman LR, Somers TJ, Kraus VB, Guilak F, Blumenthal JA, et al. Inflammatory biomarkers of OA, IL-6 and leptin are modifiable in overweight/obese OA patients with a protocol that combines training in pain coping skills and weight management. *Osteoarthritis and Cartilage* 2012;**20**:S38.

Somers TJ, Blumenthal JA, Guilak F, Kraus VB, Schmitt DO, Babyak MA, et al. Pain coping skills training and lifestyle behavioral weight management in patients with knee osteoarthritis: a randomized controlled study. *Pain* 2012;**153**(6):1199-209.

Stukstette 2011 *{published data only}*

Stukstette MJ, Dekker J, Den Broeder AA, Van der Laan WN, Bijlsma JWJ, Van den Ende CHM. A multidisciplinary and multidimensional program for hand osteoarthritis is not effective: results of a randomized controlled study. *Arthritis and Rheumatism* 2011;**63**(Suppl 10):1570.

Von Korff 2012 *{published data only}*

Von Korf M, Vitiello MV, McCurry SM, Balderson BH, Moore AL, Baker LD, et al. Group interventions for co-morbid insomnia and osteoarthritis pain in primary care: the lifestyles cluster randomized trial design. *Contemporary Clinical Trials* 2012;**33**(4):759-68.

Wu 2011 *{published data only}*

Kao MJ, Wu MP, Tsai MW, Chang WW, Wu SF. The effectiveness of a self-management program on quality of life for knee osteoarthritis (OA) patients. *Archives of Gerontology and Geriatrics* 2012;**54**(2):317-24.

Wu SF, Kao MJ, Wu MP, Tsai MW, Chang WW. Effects of an osteoarthritis self-management programme. *Journal of Advanced Nursing* 2011;**67**(7):1491-501.

References to ongoing studies

Allen 2012 *{published data only}*

Allen KD, Bosworth HB, Brock DS, Chapman JG, Chatterjee R, Coffman CJ, et al. Patient and provider interventions for managing osteoarthritis in primary care: protocols for two randomized controlled trials. *BMC Musculoskeletal Disorders* 2012;**13**:60.

Bennell 2012 *{published data only}*

Bennell KL, Ahamed Y, Bryant C, Jull G, Hunt MA, Kenardy J, et al. A physiotherapist-delivered integrated exercise and pain coping skills training intervention for individuals with knee osteoarthritis: a randomised controlled trial protocol. *BMC Musculoskeletal Disorders* 2012;**13**:129.

Additional references

ABS 2004

Australian Bureau of Statistics. Musculoskeletal conditions in Australia: a snapshot. Commonwealth Government of Australia, 2004.

Bambas 2004

Bambas L, Mahmood Q. Global Equity Gauge Alliance. Course reader: Health equity? Research to action. Cape Town: School of Public Health, University of the Western Cape, 2004. www.gega.org.za/download/ResearchtoAction04.pdf (accessed 21 January 2006).

Bandura 1977

Bandura A. Self-efficacy: toward a unifying theory of behaviour change. *Psychology Review* 1977;**2**:191-215.

Borkhoff 2011

Borkhoff CM, Wieland ML, Myasoedova E, Ahmad Z, Welch V, Hawker GA, et al. Reaching those most in need: a scoping review of interventions to improve health care quality for disadvantaged populations with osteoarthritis. *Arthritis Care and Research* 2011;**63**(1):39-52. [DOI: 10.1002/acr.20349]

Brady 1997

Brady TJ. Do common arthritis self-efficacy measures really measure self-efficacy?. *Arthritis Care and Research* 1997;**10**(1):1-8.

Brady 2011

Brady, TJ. Measures of self-efficacy: Arthritis Self-Efficacy Scale (ASES), Arthritis Self-Efficacy Scale-8 Item (ASES-8), Children's Arthritis Self-Efficacy Scale (CASE), Chronic Disease Self-Efficacy Scale (CDSES), Parent's Arthritis Self-Efficacy Scale (PASE), and Rheumatoid Arthritis Self-Efficacy Scale (RASE). *Arthritis Care and Research* 2011;**63**(S11):S473-85.

Cadhilac 2011

Cadilhac DA, Hoffmann S, Kilkenny M, Lindley R, Lalor E, Osborne RH, Batterbsy M. A phase II multicentered, single-blind, randomized, controlled trial of the stroke self-management program. *Stroke* 2011;**42**(6):1673-9.

Carnes 2012

Carnes D, Homer KE, Miles CL, Pincus T, Underwood M, Rahman A, Taylor SJC. Effective delivery styles and content for self-management interventions for chronic musculoskeletal pain: a systematic literature review. *Clinical Journal of Pain* 2012;**28**:344–54.

Cates 2008 [Computer program]

Cates C. Visual Rx 2.0 NNT Calculator, 2008. www.nntonline.net, Accessed on 1st of February 2013.

Chodosh 2005

Chodosh J, Morton SC, Mojica W, Maglione M, Suttorp MJ, Hilton L, et al. Meta-analysis: chronic disease self-management programs for older adults. Annals of Internal Medicine 2005; Vol. 143, issue 6:427-38.

Committee on Health Literacy 2004

Committee on Health Literacy. In: Nielson-Bohlman L, Panzer A, Hamlin B, Kindig D editor(s). Health Literacy: A Prescription to End Confusion. Washington, DC: The National Academies Press, 2004.

Commonwealth of Aus 2009

Commonwealth of Australia. A Healthier Future for All Australians—Final Report of the National Health and Hospitals Reform Commission. ACT: Commonwealth of Australia, 2009.

Deakin 2005

Deakin TA, McShane CE, Cade JE, Williams R. Group-based training for self-management strategies in people with type 2 diabetes mellitus. *Cochrane Database of Systematic Reviews* 2005, Issue 2. [DOI: 10.1002/14651858.CD003417.pub2]

DeAngelis 2004

DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. JAMA. 2004/09/10 2004; Vol. 292, issue 11:1363-4. [1538-3598: (Electronic)]

Deeks 2011

Deeks JJ, Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. 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.

Devos-Comby 2006

Devos-Comby L, Cronan T, Roesch SC. Do exercise and self-management interventions benefit patients with osteoarthritis of the knee? A metaanalytic review. *Journal of Rheumatology* 2006;**33**(4):744-56.

Du 2011

Du S, Yuan C, Xiao X, Chu J, Qiu Y, Qian H. Self-management programs for chronic musculoskeletal pain conditions: a systematic review and meta-analysis. *Patient Education and Counseling* 2011;**85**:e299-e310.

Effing 2007

Effing T, Monninkhof EEM, van der Valk PP, Zielhuis GGA, Walters EH, van der Palen JJ, et al. Self-management education for patients with chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD002990.pub2]

Ferreira 2011

Ferreira ML, Ferreira PH, Latimer J, Barrett B, Ostelo R, Grotle M, et al. Consumers' perceptions of the smallest worthwhile effect of interventions for non-specific low back pain. *World Confederation of Physical Therapists* 2011;**97**:eS479.

Foster 2009

Foster G, Taylor SJC, Eldridge S, Ramsay J, Griffiths CJ. Self-management education programmes by lay leaders for people with chronic conditions. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD005108.pub2]

Francis 2009

Francis KL, Matthews BL, Van Mechelen W, Bennell KL, Osborne RH. Effectiveness of a community-based osteoporosis education and self-management course: a wait list controlled trial. *Osteoporosis International* 2009;**20**(9):1563-70.

Furler 2011

Furler J, Harris M, Rogers A. Equity and long-term condition self-management. *Chronic Illness* 2011;**7**:3-5. [DOI: 10.1177/1742395310386978]

Geisser 1994

Geisser ME, Robinson ME, Henson CD. The coping strategies questionnaire and chronic pain adjustment: a conceptual and empirical reanalysis. *Clinical Journal of Pain* 1994;**10**:98-106.

Gibson 2002

Gibson PG, Powell H, Wilson A, Abramson MJ, Haywood P, Bauman A, et al. Self-management education and regular practitioner review for adults with asthma. *Cochrane Database of Systematic Reviews* 2002, Issue 3. [DOI: 10.1002/14651858.CD001117]

Greenhalgh 2009

Greenhalgh T. Chronic illness: beyond the expert patient. *British Medical Journal* 2009;**338**(7695):629-31.

Hall 2010

Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. *Physical Therapy* 2010;**90**:1099-110.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003/09/06 2003; Vol. 327, issue 7414:557-60. [1468-5833: (Electronic)]

Higgins 2011

Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. 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.

Hirano 1994

Hirano PC, Laurent DD, Lorig K. Arthritis patient education studies, 1987-1991: a review of the literature. *Patient Education and Counseling* 1994;**24**(1):9-54.

Irachabal 2008

Irachabal S, et al. Pain coping strategies: French adaptation of the coping strategies questionnaire (CSQ-F). *Encephale* 2008;**34**:47-53.

Jüni 2006

Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best Practice and Research Clinical Rheumatology 2006; Vol. 20, issue 4:721-40.

Kirkham 2010

Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, Williamson PR. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ. 2010/02/17 2010; Vol. 340:c365. [1468-5833: (Electronic)]

Lefebvre 2011

Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.cochrane-handbook.org.

Lorig 1987

Lorig K, Konkol L, Gonzalez V. Arthritis patient education: a review of the literature. *Patient Education and Counseling* 1987;**10**(3):207-52.

Lorig 1989

Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. *Arthritis and Rheumatism* 1989;**32**(1):37-44.

Lorig 2003

Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. *Annals of Behavioral Medicine* 2003;**26**(1):1-7.

March 2004

March LM, Bagga H. Epidemiology of osteoarthritis in Australia. *Medical Journal of Australia* 2004;**180**(5 Suppl):6-10.

March 2010

March L, Amatya B, Osborne RH, Brand C. Developing a minimum standard of care for treating people with osteoarthritis of the hip and knee. Best Practice and Research Clinical Rheumatology 2010; Vol. 24, issue 1:121-45.

Miles 2011

Miles CL, Pincus T, Carnes D, Homer KE, Taylor SJC, Bremner SA, 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. *European Journal of Pain* 2011;**15**(8):775.e1–1.

Moore 2010

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.

Nolte 2007

Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. The extent and breadth of benefits from participating in chronic disease self-management courses: a national patient-reported outcomes survey. Patient Education and Counseling. 2006/10/10 2007; Vol. 65, issue 3:351-60.

Nolte 2009

Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. Tests of measurement invariance failed to support the application of the "then-test". Journal of Clinical Epidemiology. 2009/07/15 2009; Vol. 62, issue 11:1173-80.

Osborne 2004

Osborne RH, Spinks JM, Wicks IP. Patient education and self-management programs in arthritis. *Medical Journal of Australia* 2004;**180**(5 Suppl):23-6.

Osborne 2007

Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient Education and Counseling* 2007;**66**(2):192-201.

Osborne 2011

Osborne RH, Batterham R, Livingston J. The quality, impact and implementation of chronic disease self management across settings: the international experience of the Health Education Impact Questionnaire (heiQ) quality monitoring system. *Nursing Clinics of North America* 2011;**46**(2):255-70.

Osborne 2013

Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded theory, psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013;**13**:658. [DOI: 10.1186/1471-2458-13-658]

Packer 2012

Packer TL, Boldy D, Ghahari S, Melling L, Parsons R, Osborne RH. Self-management programs conducted within a practice setting: who participates, who benefits and what can be learned?. *Patient Education and Counseling* 2012;**87**(1):93-100.

Pinto 2012

Pinto RZ, Ferreira ML, Oliveira VC, Franco MR, Adams R, Maher CG, Ferreira PH. Patient-centred communication is associated with positive therapeutic alliance: a systematic review. *Journal of Physiotherapy* 2012;**58**:77-87.

Pitt 2011

Pitt V, O'Connor D, French SD, Osborne RH, Buchbinder R. Self-management education programmes for osteoarthritis. *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.CD008963]

Robinson 1997

Robinson ME, Myers CD, Sadler IJ, Riley JL 3rd, Kvaal SA, Geisser ME. Bias effects in three common self-report pain assessment measures. *Clinical Journal of Pain* 1997;**13**(1):74-81.

Rosentiel 1983

Rosentiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. *Pain* 1983;**17**(1):33-44.

Rutjes 2009

Rutjes AWS, Nüesch E, Sterchi R, Kalichman L, Hendriks E, Osiri M, et al. Transcutaneous electrostimulation for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2009, Issue 4. [DOI: 10.1002/14651858.CD002823.pub2]

Schünemann 2011a

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.

Schünemann 2011b

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, et al. Chapter 12: Interpreting results and drawing conclusions. 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.

Smith 2009

Smith C, Kumar S, Pelling N. The effectiveness of self-management educational interventions for osteoarthritis of the knee. *JBI Library of Systematic Reviews* 2009;**7**:1089-116.

Snijders 2011

Snijders GF. Fatigue in knee and hip osteoarthritis: the role of pain and physical function. *Rheumatology (Oxford)* 2011;**50**:1894-900.

Sterne 2011

Sterne JAC, Egger M, Moher D (editors). Chapter 10: Addressing reporting biases. In: Higgins J, 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.

Steward 2001

Steward MW, Harvey ST, Evans IM. Coping and catastrophizing in chronic pain: a psychometric analysis and comparison of two measures. *Journal of Clinical Psychology* 2001;**57**:131-8.

Superio-Cabuslay 1996

Superio-Cabuslay E, Ward MM, Lorig KR. Patient education interventions in osteoarthritis and rheumatoid arthritis: a meta-analytic comparison with nonsteroidal antiinflammatory drug treatment. *Arthritis Care and Research* 1996;**9**(4):292-301.

United Nations Economic and Social Council 2010

United Nations Economic and Social Council (ECOSOC). Health literacy and the Millennium Development Goals: United Nations Economic and Social Council (ECOSOC) regional meeting background paper (abstracted). *Journal of Health Communication in Healthcare* 2010;**15(Suppl 2)**:211-23.

Verra 2006

Verra ML, Angst F, Lehmann S, Aeschlimann A. Translation, cross-cultural adaptation, reliability, and validity of the German version of the coping strategies questionnaire (CSQ-D). *Journal of Pain* 2006;**7**:327-36.

Von Korff 1997

Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. *Annals of Internal Medicine* 1997;**127**(12):1097-102.

Walker 2003

Walker C, Swerissen H, Belfrage J. Self-management: its place in the management of chronic illnesses. *Australian Health Review* 2003;**26**(2):34-42.

Walsh 2006

Walsh NE, Mitchell HL, Reeves BC, Hurley MV. Integrated exercise and self-management programmes in osteoarthritis of the hip and knee: a systematic review of effectiveness. *Physical Therapy Reviews* 2006;**11**(4):289-97.

Wanitkun 2011

Wanitkun N, Batterham R, Vichathai C, Leetongin G, Osborne RH. Building equity in chronic disease management in Thailand: a whole-system provincial trial of systematic, pro-active chronic illness care. *Chronic Illness* 2011;**7**(1):31-44.

Warsi 2003

Warsi A, LaValley MP, Wang PS, Avorn J, Solomon DH. Arthritis self-management education programs: a meta-analysis of the effect on pain and disability. *Arthritis and Rheumatism* 2003;**48**(8):2207-13.

Wolfe 1999

Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. *Rheumatology* 1999;**38**:355-61.

Woolf 2003

Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bulletin of the World Health Organization* 2003;**81**:646-56.

World Health Organization 1998

World Health Organization. Health Promotion Glossary. *Health Promotion International* 1998;**13**(4):349-64.

Zhang 2007

Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part I: Critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis and Cartilage 2007; Vol. 15, issue 9:981-1000.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Ackerman 2012

Methods	**Study design:** RCT (blocks four to six, stratified by site), multi-centre, two arms, non-blinded

Ackerman 2012 *(Continued)*

Country in which trial was carried out: Australia

Method of recruitment of participants:referral by orthopaedic surgeons or rheumatologists

Setting:outpatients (secondary and tertiary care)

Was the sample size justified with a priori calculation of effect size/power?Yes

Length of follow-up:12 months

Dropouts:14 (24%) dropped out from the intervention group (11 were unable to be allocated to a course or had their course cancelled, one did not attend the scheduled course, one was scheduled for joint replacement, one died); two (3%) dropped out from the control group (no reasons provided)

Participants

Inclusion criteria

- Hip or knee OA diagnosis from radiology reports or able to be classified according to ACR criteria
- Aged 18 years or over
- Referred to an orthopaedic surgeon or rheumatologist
- Sufficient English language skills and vision to self-complete questionnaires
- A reasonable expectation of attending six sessions of the ASMP

Exclusion criteria

- Cognitive dysfunction
- Previous participation in an ASMP or similar education programme
- Placement on an orthopaedic waiting list for joint replacement surgery
- Scheduled joint replacement surgery

Baseline characteristics

Baseline characteristics were similar among all treatment groups

Intervention group: ASMP (N = 58 randomly assigned, 39 analysed after six weeks, 40 analysed after three months, 38 analysed after 12 months of follow-up (FU))

Location of OA: 36% hip, 57% knee, 7% hip and knee

BMI (mean (IQR)): 30 (24 to 35) kg/m^2

PROGRESS-Plus

Place of residence: Australia

Race, ethnicity and culture: 69% Australian-born

Occupation: 29% paid employment, 59% retired, 7% not working because of OA or medical condition, 5% unemployed

Sex: 38% male, 62% female

Education: 12% primary school or less, 47% years seven to 10, 14% years 11 and 12, 11% trade/technical education, 16% university

Social capital: 62% married or living with partner

Age (mean (SD)), years: 63.5 (10.8)

Control group: ASMP booklet (N = 62 randomly assigned, 51 analysed after six weeks, 55 analysed after three months, 56 analysed after 12 months of FU)

Location of OA: 26% hip, 68% knee, 6% hip and knee

BMI (mean (IQR)): 29 (26 to 35) kg/m^2

Ackerman 2012 *(Continued)*

	PROGRESS-Plus *Place of residence:* Australia *Race, ethnicity and culture:* 68% Australian-born *Occupation:* 23% paid employment, 65% retired, 10% not working because of OA or medical condition, 2% unemployed *Sex:* 42% male, 58% female *Education:* 12% primary school or less, 45% years seven to 10, 17% years 11 and 12, 18% trade/technical education, 8% university *Social capital:* 65% married or living with partner *Age (mean (SD)), years:* 66.6 (10.9)
Interventions	***Intervention: ASMP*** *Description:* The Stanford ASMP covers management of pain and fatigue, physical activity, managing emotions, health-related problem solving and communication with doctors. Participants also received a copy of the arthritis self-help book *Intended audience:* people with OA of the knee or hip *Mode:* group sessions *Personnel:* one peer leader, one healthcare professional *Delivery method:* face-to-face *Language:* English *Format:* standard format *Location:* in community-based and hospital locations *Duration:* one session per week, lasting 2.5 hours, for six weeks *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* no *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* yes *Skill and technique acquisition:* yes *Social integration and support:* No *Health service navigation:* yes ***Comparator*** *Type:* education only *Description:* Participants were mailed a copy of the arthritis self-help book. No advice was given regarding use of the book

Ackerman 2012 *(Continued)*

	Additional treatment during trial: unclear
Outcomes	*Outcome assessed at:* baseline, six weeks and three and 12 months of follow-up **Primary outcomes of study** • Quality of life (assessment of quality of life (AQoL), -0.04 to 1.00, higher score is better) **Secondary outcomes of study** • Health-directed activities (heiQ health-directed activities subscale, 1 to 6, higher score is better) • Positive and active engagement in life (heiQ positive and active engagement in life subscale, 1 to 6, higher score is better) • Skill and technique acquisition (heiQ skill and technique acquisition subscale, 1 to 6, higher score is better) • Constructive attitudes and approaches (heiQ constructive attitudes and approaches subscale, 1 to 6, higher score is better) • Self-management (heiQ self-monitoring and insight subscale, 1 to 6, higher score is better) • Health service navigation (heiQ health service navigation subscale, 1 to 6, higher score is better) • Social integration and support (heiQ social integration and support subscale, 1 to 6, higher score is better) • Emotional distress (heiQ emotional distress subscale, 1 to 6, higher score is better) • Global OA scores (WOMAC subscales pain 0 to 20, stiffness 0 to 8, function 0 to 68, lower scores are better) • Emotional distress (Kessler Psychological Distress Scale (K10), 10 to 50, lower score is better) • Global OA scores (disease severity and need for surgery on the hip and knee multi-attribute priority tool (MAPT), 0 to 100, lower score is better)
Notes	We extracted the following outcomes at six weeks (short term) and at 12 months (intermediate term) for the analyses in this review: self-management (heiQ subscale self-monitoring and insight), engagement in life (heiQ subscale positive and active engagement in life), pain (WOMAC subscale pain), global OA scores (MAPT), function self-reported (WOMAC subscale function), quality of life (AQoL), emotional distress (K10), health-directed activity (heiQ subscale health-directed activity), social integration and support (heiQ subscale social integration and support), health service navigation (heiQ subscale health service navigation), skill and technique acquisition (heiQ subscale skill and technique acquisition), constructive attitudes and approaches (heiQ subscale constructive attitudes and approaches) and dropouts (proportion of missing participants) The study was funded by the National Health and Medical Research Council of Australia (Project Grant number 400210); Dr Ackerman was supported in part by an Australian National Health and Medical Research Council Public Health (Australia) Training Fellowship (#520004); Prof Buchbinder was supported in part by Australian National Health and Medical Research Council Practitioner Fellowships (#334010 and #606429); Prof Osborne was supported in part by an Australian National Health and Medical Research Council Population Health Career Development Award (#400391) The author (I Ackerman) provided the unpublished manuscript and additional information about the trial on request The prior calculated sample size was not reached; therefore, a type II error has possibly occurred

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “For each site, group allocation was assigned using a computer-generated random list in permuted blocks of 4 or 6” Comment: This method of random sequence generation has low risk of introducing bias

Ackerman 2012 *(Continued)*

Allocation concealment (selection bias)	Low risk	Quote: "Group allocation was concealed using opaque sealed envelopes, with individual envelopes opened at the coordinating centre by a research assistant not associated with the study and verified by an independent observer" Comment: This method of allocation concealment has low risk of introducing bias
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Participants and investigators were not blinded" Comment: Blinding of participants to the intervention was not feasible, with high risk of biasing results. Although personnel who conducted the intervention were not blinded either, risk of bias remains low, as the control intervention did not come in contact with study personnel
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: No information provided on blinding of outcome assessment; however, most outcomes are subjective, and participants are not blinded to group allocation introducing a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "Statistical analysis was performed using all randomised participants who provided at least one post-baseline assessment. Intervention group participants who did not receive the allocated intervention were not included in post-baseline analyses" Comment: The number of dropouts in both groups differs substantially (3% in the control vs 24% in the intervention group), with a reasonably high number of dropouts in the intervention group. An intention-to-treat analysis was performed
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in methods or in the protocol have been reported in the results
Other bias	High risk	Quote: "Of the intervention group (n = 58), 44 (76%) participants received the intervention as allocated. Of those who commenced the ASMP, only 21 participants (47%) attended all sessions. The median (IQR) number of sessions attended was 5 (4-6)" Comment: Participants who commenced the ASMP had quite high compliance. However, a large number of the intervention group did not receive the allocated intervention (24%), thereby introducing a risk of bias

Allen 2010

Methods	**Study design:**RCT (block randomisation, stratified by race), single-centre, three arms, outcome assessment blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**DVAMC's electronic medical record system. **Setting:** primary care (Veterans Affairs Medical Center) **Was the sample size justified with a priori calculation of effect size/power?**yes **Length of follow-up:** 12 months **Dropouts:** 28 (16%) dropped out from the self-management education programme (seven excluded, nine lost to follow-up, 12 withdrew), 17 (10%) dropped out from health education (four excluded, nine lost to follow-up, four withdrew) and 17 (10%) dropped out from usual care (six excluded, six lost to follow-up, five withdrew)

Allen 2010 *(Continued)*

Participants

Criteria for defining the condition being treated (OA)

- Physician diagnosis of hip or knee
- Radiographic evidence of knee/hip OA

Inclusion criteria

- A physician's diagnosis of hip or knee OA with radiographic evidence
- Enrolled in primary care at the Durham VAMC
- Current and persistent joint symptoms
- A physician visit during the study period

Exclusion criteria

- Diagnosis of other systemic rheumatic disease (e.g. RA, FM)
- Hospitalised for cardiovascular disease (e.g. stroke, MI)
- Diagnosis of metastatic cancer in past three months
- Active diagnosis of psychosis or diagnosis of dementia
- Any other serious health condition that limits participation
- On waiting list for arthroplasty
- Resident in a nursing home
- Severely impaired hearing/speech
- No access to a telephone
- Participation in another interventional study

Baseline characteristics

Baseline characteristics were similar in all treatment groups

Intervention group: osteoarthritis self-management Intervention (N = 174 randomly assigned, 172 analysed)

Location of OA: 82% knee, 12% hip, 6% knee and hip

BMI (mean (SD)): 32.0 (7.0) kg/m^2

Duration of OA (mean (SD)): 16.5 (12.7) years

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 55% white, 45% non-white, 1% Hispanic

Occupation: 38% employed

Sex: 91% male, 9% female

Education: 33% high school education or less

Socioeconomic status: 28% inadequate income

Social capital: 72% married

Age (mean (SD)), years: 60.3 (10.3)

Disability: 30% fair or poor health

Control group: health education intervention (N = 175 randomly assigned, 172 analysed)

Location of OA: 79% knee, 16% hip, 5% knee and hip

BMI (mean (SD)): 31.6 (6.5) kg/m^2

Allen 2010 *(Continued)*

Duration of OA (mean (SD)): 15.8 (12.0) years

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 53% white, 47% non-white, 3% Hispanic

Occupation: 38% employed

Sex: 93% male, 7% female

Education: 34% high school education or less

Socioeconomic status: 27% inadequate income

Social capital: 65% married

Age (mean (SD)), years: 60.3 (10.8)

Disability: 37% fair or poor health

Control group: usual care (N = 174 randomly assigned, 171 analysed)

Location of OA: 79% knee, 17% hip, 4% knee and hip

BMI (mean (SD)): 31.8 (6.5) kg/m^2

Duration of OA (mean (SD)): 15.9 (11.9) years

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 54% white, 46% non-white, 2% Hispanic

Occupation: 39% employed

Sex: 94% male, 6% female

Education: 33% high school education or less

Socioeconomic status: 22% inadequate income

Social capital: 71% married

Age (mean (SD)), years: 59.7 (10.1)

Disability: 30% fair or poor health

Note: Several participants were excluded from analysis because they did not meet eligibility criteria after subsequent medical record review (two in the intervention group, three in the health education group and three in the usual care group)

Interventions	**Intervention: osteoarthritis self-management intervention** *Description:* The intervention included two main components: providing education related to managing OA symptoms, and helping participants develop goals and action plans related to OA management. Participants were asked to identify and write down one or more goals related to their OA symptoms and management, as well as weekly action plans for achieving these goals *Intended audience:* people with OA of the knee and/or hip *Mode:* individual *Personnel:* health professionals

Allen 2010 *(Continued)*

Delivery method: telephone, written, audio and video material

Language: English

Format: tailored to individual's needs

Location: home

Duration: monthly telephone calls for 12 months, mean duration of calls 9.0 minutes

Additional treatment during trial: usual care for OA (incl analgesic and anti-inflammatory medication)

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: yes

Comparator: health education intervention

Type: attention control

Description: Participants received written and audio materials regarding common health problems, as well as related screening recommendations. The health educator called participants to review key points from the modules, to assess whether participants were being appropriately screened and to make suggestions for screening as needed

Frequency: monthly telephone calls for 12 months, mean duration of calls 4.9 minutes

Additional treatment during trial: usual care for OA (incl analgesic and anti-inflammatory medication)

Comparator: usual care

Type: usual care

Description: Participants received their usual care for OA

Outcomes	Outcome assessed at: baseline and 12 months of follow-up **Primary outcomes of study** • Pain (Arthritis Impact Measurement Scales-2 (AIMS-2) subscale pain, 0 to 10, lower score is better) **Secondary outcomes of study** • Function self-reported (AIMS-2 physical function subscale, 0 to 10, lower score is better) • Emotional distress (AIMS-2 affect subscale, 0 to 10, lower score is better) • Self-management (Arthritis Self-Efficacy Scale (ASES), 1 to 10, higher score is better) • Pain (VAS, 0 to 10, lower score is better)
Notes	We extracted the following outcomes at 12 months (intermediate term) for the analyses in this review: self-management (ASES), pain (VAS), function self-reported (AIMS-2), emotional distress (AIMS-2) and dropouts (proportion of missing participants)

Allen 2010 *(Continued)*

	Funding was provided by the US Department of Veterans Affairs Health Services Research and Development Service Author (K Allen) provided additional information about the trial on request The study sample consisted mainly of male veterans, which is not a representative sample of general primary care Participants were reimbursed $10 after baseline assessment and $10 after follow-up assessment

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomisation is computer generated, stratified by race (white vs. non-white) using block sized of 12 (…)." Comment: correct method of randomisation with low risk of bias
Allocation concealment (selection bias)	Low risk	Quote: "Randomisation was (…) maintained separately from participant enrolment (…)." Comment: The randomisation sequence was adequately concealed during participant enrolment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "(…) study coordinator informs participants of their group assignment (…)." Comment: Participants were informed about their allocated treatment before the start of the intervention. Personnel were not adequately blinded either and administered both interventions
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "(…) research assistants conducting the assessments are blinded to participants' group" Comment: Personnel assessing study outcomes were blinded; however, most outcomes are subjective and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "(…) we fit linear mixed models by using an intention-to-treat approach" Comment: An intention-to-treat-analysis without imputation techniques was used. The number of withdrawals is more than twofold higher in the SMP intervention group compared with the health education or usual care group (15.1% vs 8.1% vs 8.2%). Reasons for dropout and postintervention exclusions are not reported
Selective reporting (reporting bias)	Unclear risk	Comment: The authors did not report the self-management behaviours, analgesic/anti-inflammatory medication use and intervention adherence/intensity that were mentioned in the previously published protocol. All prespecified outcomes that we considered important for the review were reported
Other bias	High risk	Quote: "We assumed a common baseline value among treatment groups" Comment: Incorrect assumption that baseline values are similar after randomisation in all treatment groups Comment: The same health educator conducts the telephone calls in the SMP and the health education group. Authors attempted to minimise the risk of contamination by using standardised scripts

Berman 2004

Methods	**Study design:**RCT, multi-centre, three arms, outcome assessment and two treatment arms blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**through print and radio advertisements **Setting:**general population **Was the sample size justified with a priori calculation of effect size/power?**yes **Length of follow-up:**26 weeks (6.5 months) **Dropouts:** 50 (26%) in the true acupuncture group dropped out (17 were disqualified for medical reasons, 33 withdrew), 52 (27%) in the sham acupuncture group dropped out (27 were disqualified for medical reasons, 25 withdrew) and 99 (52%) dropped out from the educational control group (29 were disqualified for medical reasons, 70 withdrew)
Participants	**Inclusion criteria** • OA of the knee with radiographic evidence of at least one osteophyte at the tibiofemoral joint (Kellgren-Lawrence grade 2 or higher) • Aged 50 years or older • Moderate or greater clinically significant knee pain on most days during the past month • Willingness to be randomly assigned **Exclusion criteria** • Presence of serious medical conditions that precluded participation in the study • Bleeding disorders that might contraindicate acupuncture • Intra-articular corticosteroid or hyaluronate injections during the past six months • Knee surgeries during the past six months • Concomitant use of topical capsaicin cream during the past six months • Previous experience with acupuncture • Any planned events (including total knee replacement) that would interfere with participation in the study during the following 26 weeks **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: education group (N = 189 randomly assigned, 174 analysed at baseline, 124 analysed after four weeks of FU, 125 analysed after eight weeks of FU, 113 analysed after 14 weeks of FU, 108 analysed after 26 weeks of FU)*** *Location of OA:* 100% knee *Duration of OA:* 44.3% less than five years, 24.3% six to 10 years, 31.4% longer than 10 years PROGRESS-Plus *Place of residence:* US, around Baltimore or Towson (both in Maryland) or around New York City (New York) *Race, ethnicity and culture:* 66.7% white, 31.7% African American, 1.6% other *Sex:* 32.8% male, 67.2% female *Education:* 35.1% no college, 64.9% some college *Age (mean (SD)), years:* 65.1 (8.8)

Berman 2004 *(Continued)*

Disability: 74.7% one target knee, 25.3% two target knees; 73.4% moderate or lesser pain when walking on flat surface, 26.6% severe or extreme pain when walking on flat surface

Control group: true acupuncture (N = 190 randomly assigned, 186 analysed at baseline, 173 analysed after four weeks of FU, 169 analysed after eight weeks of FU, 158 analysed after 14 weeks of FU, 142 analysed after 26 weeks of FU)

Location of OA: 100% knee

Duration of OA: 53.8% less than five years, 19.9% six to 10 years, 25.8% longer than 10 years

PROGRESS-Plus

Place of residence: US, around Baltimore or Towson (both in Maryland) or around New York City (New York)

Race, ethnicity and culture: 70.0% white, 27.4% African American, 2.6% other

Sex: 36.8% male, 63.2% female

Education: 32.8% no college, 67.2% some college

Age (mean (SD)), years: 65.2 (8.4)

Disability: 75.0% one target knee, 25.0% two target knees; 76.5% moderate or lesser pain when walking on flat surface, 23.5% severe or extreme pain when walking on flat surface

Control group: sham acupuncture (N = 191 randomly assigned, 183 analysed at baseline, 163 analysed after four weeks of FU, 161 analysed after eight weeks of FU, 157 analysed after 14 weeks of FU, 141 analysed after 26 weeks of FU)

Location of OA: 100% knee

Duration of OA: 53.0% less than five years, 18.0% six to 10 years, 29.0% longer than 10 years

PROGRESS-Plus

Place of residence: US, around Baltimore or Towson (both in Maryland) or around New York City (New York)

Race, ethnicity and culture: 70.7% white, 26.7% African American, 2.6% other

Sex: 38.2% male, 61.8% female

Education: 25.4% no college, 74.6% some college

Age (mean (SD)), years: 66.2 (8.7)

Disability: 71.1% one target knee, 28.9% two target knees; 75.5% moderate or lesser pain when walking on flat surface, 24.5% severe or extreme pain when walking on flat surface

Interventions	***Intervention: education***

Description: The education-attention control intervention consisted of six two-hour group sessions based on the Arthritis Self-Management Program and taught by an experienced Arthritis Foundation–trained patient education specialist. In addition, educational materials were mailed to the education group periodically in an attempt to equalise the amount of experimental contact in all groups. Many topics were taught to the education control group, including types of arthritis, various treatments for arthritis, self-management of arthritis, creating an action plan to manage arthritis, body mechanics for home and office, products that make mobility and daily routine activities easier and safer, physical fitness and flexibility exercises for arthritis, pain management, depression and acceptance

*Intended audience: p*eople with arthritis of the knee

Mode: group sessions (mean number of participants: seven)

Berman 2004 *(Continued)*

Personnel: trained patient education specialist

Delivery method: face-to-face

Language: English

Format: standard format

Location: -

Duration: six sessions, each of two hours, delivered every other week for a total duration of 12 weeks

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: unclear

Emotional well-being: yes

Self-monitoring and insight: no

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: yes

Comparator: true acupuncture

Type: alternate intervention

Description: The true acupuncture group underwent 26 weeks of gradually tapering treatment. The acupuncture point selection was based on Traditional Chinese Medicine meridian theory to treat knee joint pain

Frequency: eight weeks of two treatments per week, two weeks of one treatment per week, four weeks of one treatment every other week, 12 weeks of one treatment per month

Additional treatment during trial: unclear

Comparator: sham acupuncture

Type: attention control

Description: For the sham treatment, a combined insertion and non-insertion procedure was modified. The sham acupuncture procedure was given on the same schedule as in the experimental group and used the same active needle placements, except that actual insertion did not occur at the nine critical points

Frequency: eight weeks of two treatments per week, two weeks of one treatment per week, four weeks of one treatment every other week, 12 weeks of one treatment per month

Additional treatment during trial: unclear

Outcomes	*Outcome assessed at:* baseline and four, eight, 14 and 26 weeks of follow-up **Primary outcomes of study** • Pain (WOMAC pain subscale, 0 to 20, lower score is better) • Function self-reported (WOMAC function subscale, 0 to 68, lower score is better)

Berman 2004 *(Continued)*

Secondary outcomes of study

- Quality of life (SF-36 physical component score, 0 to 10, higher is better)
- How does OA affect you? (patient global assessment score)
- Functional performance (six-minute walk test, as many feet as possible, higher is better)
- Adverse events
- Which treatment participants were believed to receive

Notes	We extracted the following outcomes at week four (short term) and week 26 (intermediate term) for the analyses in this review: pain (WOMAC), function self-reported (WOMAC), quality of life (SF-36), functional performance (six-minute walk test) and dropouts (proportion of missing participants) Funding was provided by the National Centre for Complementary and Alternative Medicine (National Institutes of Health Cooperative Agreement I01 AT-00171), with advice and encouragement by the National Institute of Arthritis and Musculoskeletal and Skin Diseases Author (B Berman) sent additional information about the trial on request Data analysis: Change scores were combined with end point scores using generic inverse variance

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "(…) each cohort at each site was randomly assigned to 1 of 3 groups by a computer-generated process using randomly selected blocks of 3, 6 and 9" Comment: This method of random sequence generation has low risk of selection bias
Allocation concealment (selection bias)	Low risk	Quote: "We assured allocation concealment by using disguised letter codes that were generated and sent to the site coordinators by a central statistical core" Comment: Allocation was sufficiently concealed to minimise risk of bias
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "(…) the participants themselves (in the true acupuncture and sham acupuncture groups) (…) were blinded to group assignment" Comment: Although the acupuncture groups were blinded, it was not possible to blind the education group, which may have introduced a risk of bias. It was not possible to blind personnel to treatment allocation either
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The research assistants who collected assessments from participants, (…) and the statistician were blinded to group assignment" Comment: Outcome assessors were blinded to treatment allocation; however, most outcomes are subjective, and participants in the control group are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "At 26 weeks, 43% of the participants in the education group and 25% in each of the true and sham acupuncture groups were not available for analysis" Quote: "Thus, we present only the results from the analyses that used all available data" Comment: Dropout rate was very high and differed among treatment groups, introducing high risk of attrition bias. Also, dropouts from the education and

Berman 2004 *(Continued)*

		true acupuncture group reported significantly more pain at baseline than was reported by completers. No intention-to-treat analysis was presented
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in the Methods section are reported in the Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Blixen 2004

Methods	
Participants	*Criteria for defining the condition being treated (OA)* • Documented diagnosis of OA in arthritis/rheumatology clinic **Inclusion criteria** • Documented diagnosis of OA in an arthritis/rheumatology clinic • Aged 60 years or older • Visit to department of rheumatology in previous six months **Exclusion criteria:** none **Baseline characteristics** Income and marital status differed significantly between groups (P < 0.05) at baseline ***Intervention group: telephone health education strategy (N = 16 randomly assigned, 15 analysed)*** *Duration of OA (mean (IQR)):* 10 (2 to 30) years PROGRESS-Plus *Place of residence:* US, urban *Race, ethnicity and culture:* 69% Caucasian, 31% African American *Sex:* 56% male, 44% female *Education:* 6% grade 7 to 9, 25% high school, 44% 1 to 4 years college, 6% college graduate, 19% professional/graduate school *Socioeconomic status:* 19% < $10,000 per year, 7% $10,000 to 19,999 per year, 7% $20,000 to 29,999 per year, 29% $30,000 to 39,999 per year, 14% $40,000 to 49,999 per year, 29% > $50,000 per year *Social capital:* 75% married, 25% divorced *Age (mean (SD)), years:* 71.7 (6.3) *Disability:* 13% had joint replacement surgery, mean (SD) number of comorbid conditions 2.2 (1.0) ***Control group: usual care (N = 16 randomly assigned, 15 analysed)*** *Duration of OA (mean (IQR)):* 6.5 (2 to 18) years PROGRESS-Plus *Place of residence:* US, urban *Race, ethnicity and culture:* 75% Caucasian, 25% African American

Blixen 2004 *(Continued)*

	Sex: 69% male, 31% female *Education:* 6% grade 7 to 9, 6% grade 10 to 11, 19% high school, 56% 1 to 4 years college, 13% college graduate, 0% professional/graduate school *Socioeconomic status:* 25% < $10,000 per year, 19% $10,000 to 19,999 per year, 31% $20,000 to 29,999 per year, 6% $30,000 to 39,999 per year, 13% $40,000 to 49,999 per year, 6% > $50,000 per year *Social capital:* 38% married, 13% single, 25% divorced, 25% widowed *Age (mean (SD)), years:* 69.9 (5.9) *Disability:* 19% had joint replacement surgery, mean (SD) number of comorbid conditions 2.3 (1.2)
Interventions	***Intervention: telephone health education strategy*** *Description:* The intervention was a telephone health education strategy, which involved six weekly mailings of OA self-management modules (adapted from *The Arthritis Helpbook: A Tested Self-Management Program for Coping With Arthritis and Fibromyalgia*). The modules covered (1) pathology, (2) OA medications, (3) interrelationship between emotional and physical components of pain and the importance of relaxation techniques, (4) depression, (5) importance of regular exercise and (6) weight management. Participants also received an audiotape that explained how to perform various relaxation techniques. The content of each module was reinforced by six weekly 45-minute telephone calls conducted by an advanced practice nurse. The nurse reviewed the materials, answered questions and helped participants set self-management goals and learn new skills *Intended audience:* people with OA *Mode:* individual *Personnel:* advanced practice nurse *Delivery method:* phone, audiotape, mailed modules *Language:* English *Format:* tailored to individual needs *Location:* home *Duration:* six weekly sessions of 45 minutes, total duration of six weeks *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* no *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* no *Skill and technique acquisition:* yes *Social integration and support:* no *Health service navigation:* yes ***Comparator: usual care*** *Type:* usual care

Blixen 2004 *(Continued)*

	Description: usual care for their OA with their respective rheumatologist *Additional treatment during trial:* unclear
Outcomes	
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "(…) and was randomly assigned to either to control or the intervention group" Comment: The randomisation method is not described in the article
Allocation concealment (selection bias)	Unclear risk	Comment: Concealment of allocation is not described in the article
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: The types of control and intervention groups make blinding of participants and personnel impossible, which may have introduced a risk of performance bias. Rheumatologists who cared for the control group, however, are unlikely to be influenced by randomisation of a participant to a control group
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "To minimize bias (…) did not participate in the patient self-management programme and who were unaware of treatment assignment" Comment: Outcome assessors were blinded from participants' allocated groups; however, most outcomes are subjective and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "All 32 subjects completed the baseline and three-month telephone interviews. Two of the subjects (one from each group) had protracted periods in hospital at the six-month follow-up and we were unable to interview them" Comment: Although no intention-to-treat-analysis was carried out, risk of bias is low because of the low percentage of dropouts, which was equal in the two groups and was explained by the reasons provided
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in the Methods section are described in the Results section
Other bias	High risk	Comment: Because of the small sample size, a type II error is possible

Buszewicz 2006

Methods	**Study design:**RCT, multi-centre, two arms, non-blinded **Country in which trial was carried out:** United Kingdom (UK) **Method of recruitment of participants:**GP identified potential participants from practice attendees over a six-week period, and computerised records of the practice were searched for relevant OA clinical terms and prescriptions for regular NSAIDs or analgesics **Setting:** primary care **Was the sample size justified with a priori calculation of effect size/power?** yes

Buszewicz 2006 *(Continued)*

Length of follow-up:12 months

Dropouts: 112 (27%) dropped out from the intervention group (five died, 35 withdrew, 72 did not respond), and 81 (20%) dropped out from the control group (two died, 23 withdrew, 56 did not respond)

Participants	**Inclusion criteria** • Clinical diagnosis of knee and/or hip OA by GP • Aged 50 years or older • Diagnosis of OA for at least a year • Associated pain and/or functional disability during past month **Exclusion criteria** • Recommendation for surgery for arthritis • Poor mobility • Poor understanding of English • Associated neurological signs • Cognitive impairment **Baseline characteristics** Baseline characteristics were similar in all treatment groups ***Intervention group: self-management programme (N = 406 randomly assigned, 406 included in analysis)*** PROGRESS-Plus *Place of residence:* UK *Race, ethnicity and culture:* 100% white, 0% Black African, < 1% Black Caribbean *Sex:* 37% male, 63% female *Education:* 72% no higher education, 28% higher education *Social capital:* 83% house owner/occupier, < 1% staying with family or friends, 17% rented accommodation *Age (mean (SD)), years:* 68.4 (8.2) ***Control group: education (N = 406 randomly assigned, 406 included in analysis)*** PROGRESS-Plus *Place of residence:* UK *Race, ethnicity and culture:* 99% white, < 1% Black African, 1% Black Caribbean *Sex:* 37% male, 63% female *Education:* 73% no higher education, 27% higher education *Social capital:* 79% house owner/occupier, 1% staying with family or friends, 20% rented accommodation *Age (mean (SD)), years:* 68.7 (8.6)
Interventions	***Intervention: self-management programme*** *Description:* The intervention consisted of group sessions provided by the Arthritis Care organisation. Components of the group sessions included information about arthritis, information about principles

Buszewicz 2006 *(Continued)*

	of self-help, exercises (breathing, physical exercise, muscle relaxation, distraction techniques, communication, guided imagery, dealing with fatigue), healthy eating, strategies for accessing resources, etc *Intended audience:* people with arthritis *Mode:* group sessions (12 to 18 participants) *Personnel:* a trained volunteer from the voluntary organisation Arthritis Care *Delivery method:* face-to-face *Language:* English *Format:* standard format *Location:* - *Duration:* six weekly sessions of 2½ hours each *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* yes *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* yes *Skill and technique acquisition:* yes *Social integration and support:* unclear *Health service navigation:* yes ***Comparator: education*** *Type:* information only *Description:* - *Frequency:* one single education booklet was provided at the start of the trial *Additional treatment during trial:* unclear
Outcomes	Outcomes assessed at: baseline and four and 12 months of follow-up **Primary outcomes of study** • Quality of life (SF-36 subscale mental and physical health, 0 to 100, higher score is better) **Secondary outcomes of study** • Pain (WOMAC, 0 to 20, lower score is better) • Stiffness (WOMAC, 0 to 8, lower score is better) • Self-reported function (WOMAC, 0 to 68, lower score is better) • Emotional distress (Hospital Anxiety and Depression Scale (HADS), 0 to 21, lower score is better) • Self-management (ASES pain and other symptoms subscales, 5 to 35 and 6 to 42, higher score is better)
Notes	We extracted the following outcomes at 12 months (intermediate term) for the analyses in this review: self-management (ASES), pain (WOMAC), function self-reported (WOMAC), emotional distress (HADS

Buszewicz 2006 *(Continued)*

subscale depression), quality of life (EQ-5D utility score) and dropouts (proportion of missing participants)

Funding was provided by the Medical Research Council (reference G9900306)

Author (M Buszewicz) sent end of treatment mean and SD scores for the intervention and control groups at all follow-up times on request

Data analysis: For the outcome self-management in OA, we could choose between ASES subscale pain and ASES subscale other and chose ASES subscale pain, as we judged that pain was a more measurable aspect of self-management. For the outcome quality of life, we could choose between SF-36 mental health subscale and EQ-5D utility score and chose EQ-5D utility score, as we judged this to be a better measure for quality of life. For the outcome global OA scores, we added the subscales of WOMAC (pain, stiffness and physical function) provided by the author to get the WOMAC total score; we estimated the SD with the formula provided in Table 7.7a in the *Cochrane Handbook for Systematic Reviews of Interventions* and the SDs from the subscales (this method was chosen in close consultation with a biostatistician)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "We used an independent centralised computerised system to randomise the participants. Practices were stratified by area and we used minimisation to balance for differences in age and sex" Comment: adequate sequence generation resulting in low risk of selection bias
Allocation concealment (selection bias)	Low risk	Quote: "Research nurses faxed details of consenting participants to the trial manager, who passed the information on to the randomisation centre that allocated participants to their experimental group" Comment: Because the randomisation centre allocated participants to their groups, allocation seemed to have remained concealed and the risk of selection bias low
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Study participants were not blinded during the trial; therefore outcomes in this trial may have been biased Comment: Although the volunteers who give the intervention (SMP) are not blinded to the allocation, this probably did not result in a higher risk of performance bias because they did not have contact with the control arm
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Outcomes were assessed through questionnaires completed by the participants (…)." Comment: Outcomes were self-reported by participants who were aware of their allocated treatment. Therefore detection bias was not sufficiently accounted for in this trial
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "Questionnaire response rates were 95% at baseline, 80% at four months, and 76% at 12 months (…)" Quote: "Analysis was based on intention to treat" Comment: Both ITT and per-protocol analyses were carried out. Although correct imputing techniques were used, the response rates per time point were low, which means that a lot of data had to be imputed (24%), thus introducing a risk of attrition bias. Also, the numbers of withdrawals and non-responders

Buszewicz 2006 *(Continued)*

		are extensive in both treatment groups and are greater in the intervention arm (27% vs 20%)
Selective reporting (reporting bias)	Low risk	Comment: All prespecified outcomes were reported in the article
Other bias	High risk	Quote: "In the intervention group more than half (56%; 219/392) attended four or more sessions, 9% (37/392) attended only one or two sessions, and 29% (115/392) attended none" Comment: Overall adherence in this trial was low, which results in high risk for biased results

Calfas 1992

Methods	**Study design:**RCT, single centre, two arms, outcome assessment blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**Individuals were recruited by small newspaper ads, physician referral and public service announcements **Setting:**outpatient **Was the sample size justified with a priori calculation of effect size/power?**no justification of sample size provided **Length of follow-up:** 12 months **Dropouts:**Five (25%) dropped out from the intervention group (four refused or dropped out, one had transportation problems or died or fell ill); five (25%) dropped out from the control group (four refused or dropped out, one had transportation problems or died or fell ill)
Participants	**Inclusion criteria** • Diagnosis of OA verified by each participant's physician according to the ARA criteria for OA • Individuals had to report at least some functional impairment (i.e. answered at least one of the functional questions as "modest" or above) **Exclusion criteria:** none **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: cognitive-behaviour modification (N = 20 randomly assigned, 17 analysed post-test and after two months of FU, 15 analysed after six and 12 months of FU)*** *Duration of OA:* 85% longer than five years, 5% one to five years, 10% less than one year PROGRESS-Plus *Place of residence:* San Diego, urban *Race, ethnicity and culture:* 100% Caucasian *Occupation:* 85% retired, 5% full-time employment, 10% part-time employment *Sex:* 25% male, 75% female *Social capital:* 70% married, 20% divorced, 10% widowed

Calfas 1992 *(Continued)*

Age (mean (SD)), years: 66.7 (8.1)

Control group: traditional education intervention (N = 20 randomly assigned, 18 analysed post-test, 16 analysed after two months of FU, 15 analysed after six and 12 months of FU)

Duration of OA: 100% longer than five years

PROGRESS-Plus

Place of residence: San Diego, urban

Race, ethnicity and culture: 95% Caucasian, 5% non-Caucasian

Occupation: 80% retired, 5% full-time employment, 15% part-time employment

Sex: 30% male, 70% female

Social capital: 70% married, 15% divorced, 10% widowed, 5% other

Age (mean (SD)), years: 67.3 (6.5)

Interventions

Intervention: cognitive-behaviour modification

Description: Participants were introduced to the rationale for a cognitive approach to pain management. Goals of the cognitive intervention were to (a) teach participants to reconceptualise their pain; (b) teach them to monitor their thoughts or beliefs, feelings and actions, and to recognise the relationships among them; (c) strengthen participants' belief that they can cope effectively; and (d) help participants make behavioural changes to improve their functioning. Some of the techniques used included cognitive reappraisal, imagery, relaxation training, mental distraction and goal setting. Each session was videotaped. Each session began with group members reporting on goals and homework from the previous session. Most of the sessions were spent on one topic. The facilitator used examples from the participants to illustrate points, emphasising group participation. At the end of each session, group members were given a homework assignment and set goals that were shared with the group

Intended audience: people with OA

Mode: group sessions

Personnel: trained facilitators

Delivery method: face-to-face

Language: English

Format: tailored to individual's needs

Location: Arthritis Centre at San Diego State University

Duration: one session per week, for 10 weeks

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: yes

Emotional well-being: yes

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Calfas 1992 *(Continued)*

	Social integration and support: no *Health service navigation:* no ***Comparator: traditional education intervention*** *Type:* attention control *Description:* The education control group consisted of a series of didactic lectures given by healthcare professionals. Topics for the series included rheumatology, pharmacology, nutrition, joint replacement surgery, fibromyalgia, occupational therapy, physical therapy and recreational therapy, and a presentation given by the Arthritis Foundation. Specific behavioural instructions were not given during any of the lectures *Additional treatment during trial:* unclear
Outcomes	*Outcome assessed at:* baseline, post-test and two, six and 12 months of follow-up **Primary outcomes of study** • Quality of life (Quality of Well-Being scale (QWB)) • Physical health, social and psychological well-being on the AIMS • Emotional distress (Beck Depression Inventory (BDI)) **Secondary outcomes of study** • Social support (Social Support Questionnaire (SSQ))
Notes	Only outcomes on dropouts could be extracted because presentation of data was insufficient Funding was provided by a grant from the Arthritis Foundation and by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (AR 400770-01A1)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “After screening and recruitment, each subject was randomly assigned to one of two groups (…)” Comment: Authors do not describe the method used for randomisation
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants and personnel were not blinded during the study; therefore results are possibly biased
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: “Subjects were assessed by interviewers who were blind to treatment condition” Comment: Outcome assessors were blinded at each outcome assessment; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: “At the 6- and 12-month follow-up, 75% of the subjects in each group completed testing”

Calfas 1992 *(Continued)*

		Quote: "Seventy-five percent of the pretest scores of those who dropped out or had missing data were significantly different from the scores of those who stayed in the study" Quote: "Tests are based on unequal n's because some cases were lost to follow-up" Comment: Dropout rates in both groups are equal but substantial. A difference in pretest scores between dropouts and participants who completed follow-up introduces a high risk of bias. No intention-to-treat analysis was performed
Selective reporting (reporting bias)	Unclear risk	Only outcomes on dropouts could be extracted because data presentation was insufficient. Comment: In Methods, the authors list social support using the Social Support Questionnaire (SSQ), but this is not reported in the Results section. No separate group data are presented in Table 3
Other bias	Low risk	Comment: No other potential sources of bias were identified

Cronan 1997

Methods	**Study design:**RCT, single-centre, four arms, non-blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**Letters explaining the study and inviting people to participate were sent to 3,000 randomly selected persons from a large HMO membership list of 50,450 people who were 60 years of age or older **Setting:**general population **Was the sample size justified with a priori calculation of effect size/power?**No justification was provided for the sample size **Length of follow-up:**three years **Dropouts** • After one year: 24 (27.6%) dropped out from the social support group, 11 (11.3%) dropped out from the education group, 16 (18.0%) dropped out from the combination group, 13 (14.4%) dropped out from the control group • After two years: 32 (36.8%) dropped out from the social support group, 17 (17.5%) dropped out from the education group, 26 (29.2%) dropped out from the combination group, 19 (21.1%) dropped out from the control group • After three years: 35 (40%) dropped out from the social support group, 27 (28%) dropped out from the education group, 33 (37%) dropped out from the combination group, 24 (27%) dropped out from the control group No reasons were provided for dropout
Participants	*Criteria for defining the condition being treated (OA)* • Self-reported symptoms of OA (in 90.3%, the diagnosis was confirmed in the medical records) **Inclusion criteria** • Self-reported symptoms of OA (confirmed in medical records in 90.3%) • Aged 60 or older • Willingness to attend 10 weekly and 10 monthly sessions

Cronan 1997 *(Continued)*

Exclusion criteria: none

Baseline characteristics

Baseline characteristics were similar among all treatment groups

Intervention group: education group (N = 97 randomly assigned, 86 analysed after one year of FU, 80 analysed after two years of FU, 70 analysed after three years of FU)

Duration of OA (mean (SD)): 7.48 (5.93) years since diagnosis (N = 83)

PROGRESS-Plus

Place of residence: San Diego County

Race, ethnicity and culture: 94.8% Caucasian, 3.1% African American, 1.0% Hispanic, 1.0% Asian, 0% other

Sex: 29.9% male, 70.1% female

Education: 29.9% high school, 43.3% some college, 26.8% college degree

Socioeconomic status: income 26.2% under $20,000, 41.7% $20,000 to $40,000, 21.4% $40,000 to $60,000, 10.8% over $60,000

Age (mean (SD)), years: 69.68 (5.79)

Disability: 78.4% had other medical conditions, 48.5% had another serious comorbidity

Intervention group: combination group (N = 89 randomly assigned, 73 analysed after one year of FU, 63 analysed after two years of FU, 56 analysed after three years of FU)

Duration of OA (mean (SD)): 6.56 (5.09) years since diagnosis (N = 75)

PROGRESS-Plus

Place of residence: San Diego County

Race, ethnicity and culture: 92.1% Caucasian, 2.2% African American, 2.2% Hispanic, 2.2% Asian, 1.1% other

Sex: 34.8% male, 65.2% female

Education: 34.8% high school, 40.5% some college, 24.7% college degree

Socioeconomic status: income 31.8% under $20,000, 43.5% $20,000 to $40,000, 17.6% $40,000 to $60,000, 7.1% over $60,000

Age (mean (SD)), years: 68.77 (5.70)

Disability: 66.3% had other medical conditions, 47.2% had another serious comorbidity

Control group: social support group (N = 87 randomly assigned, 63 analysed after one year of FU, 55 analysed after two years of FU, 52 analysed after three years of FU)

Duration of OA (mean (SD)): 6.47 (5.55) years since diagnosis (N = 75)

PROGRESS-Plus

Place of residence: San Diego County

Race, ethnicity and culture: 92.0% Caucasian, 3.5% African American, 1.1% Hispanic, 3.4% Asian, 0% other

Sex: 43.7% male, 56.3% female

Education: 31.0% high school, 36.8% some college, 32.2% college degree

Cronan 1997 *(Continued)*

Socioeconomic status: income 20.0% under $20,000, 44.0% $20,000 to $40,000, 25.4% $40,000 to $60,000, 10.7% over $60,000

Age (mean (SD)), years: 69.14 (5.54)

Disability: 73.6% had other medical conditions, 50.6% had another serious comorbidity

Control group: no treatment (N = 90 randomly assigned, 77 analysed after one year of FU, 71 analysed after two years of FU, 66 analysed after three years of FU)

Duration of OA (mean (SD)): 7.34 (5.19) years since diagnosis (N = 79)

PROGRESS-Plus

Place of residence: San Diego County

Race, ethnicity and culture: 90.1% Caucasian, 1.1% African American, 4.4% Hispanic, 0% Asian, 4.4% other

Sex: 35.6% male, 64.4% female

Education: 28.9% high school, 30.0% some college, 31.1% college degree

Socioeconomic status: income 30.8% under $20,000, 43.6% $20,000 to $40,000, 14.1% $40,000 to $60,000, 11.5% over $60,000

Age (mean (SD)), years: 69.21 (5.51)

Disability: 61.1% had other medical conditions, 43.3% had another serious comorbidity

Interventions	***Intervention: education group*** *Description:* The education intervention was provided in a presentation format by professional health educators. Many presentations included active involvement of the participants. Presentations contained information about preventive health behaviours and self-management strategies, in addition to information about when to see a healthcare provider for ailments related to OA. The presentations emphasised appropriate healthcare usage, which is not always less healthcare usage. Participants were taught to recognise signs that indicate the need for quick medical attention to avoid future problems, in addition to learning to eliminate unnecessary healthcare utilisation by developing self-management skills. All materials were written using a self-help, consumer-centred approach. The materials encouraged people to play an important role in the decision-making process involving their health. In addition, materials on psychology, coping strategies, diet and exercise were presented. All materials were based on programmes developed by Lorig et al. *Intended audience:* people with OA *Mode:* group *Personnel:* professional health educators *Delivery method:* face-to-face *Language:* English *Format:* standard format *Location:* central site (median 10 miles from participant's home (range 2 to 75 miles)) *Duration:* 10 weekly sessions followed by 10 monthly sessions; each session took two hours *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes

Cronan 1997 *(Continued)*

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: yes

Intervention: combination group

Description: The combined intervention included both educational classes and social support, with the first hour dedicated to education and the second to social support. During the second hour, no staff members were present (see social support group)

Intended audience: people with OA

Mode: group

Personnel: professional health educators (first hour only)

Delivery method: face-to-face

Language: English

Format: standard format and tailored to individual needs

Location: central site (median 10 miles from participant's home (range 2 to 75 miles))

Duration: 10 weekly sessions followed by 10 monthly sessions; each session took two hours

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: yes

Comparator: social support group

Type: alternate intervention

Description: The social support intervention involved group discussions prompted by weekly task assignments aimed at promoting empathy and sharing coping techniques between group members. Group members were told that support groups can be effective in helping people deal with their arthritis. Tasks ranged from "elect a chair and secretary of the group and discuss what these people should do" to "divide the group into sets of 2 – 4 people, each of whom will agree to contact the others on a daily basis; work out a way a time that each member of a set can contact other members as near to

Cronan 1997 *(Continued)*

	every day as possible." Other topics include problem solving, the role of emotion, accessing community resources and developing a group project. No staff members were present during these meetings *Frequency:* 10 weekly sessions followed by 10 monthly sessions; each session took two hours *Additional treatment during trial:* unclear ***Comparator: control group*** *Type:* no treatment *Description:* Participants assigned to the control group were told that they would not need to attend regular meetings, that they would be contacted for periodic assessments and that they would receive quarterly newsletters. The quarterly newsletters focused on events in the area, offered helpful hints for issues not related to health and provided updates on the status of the study. Control group participants were also told that they would be invited to a presentation of the results at the end of the study and that they would be invited to participate in any continuing social support groups. In addition, they were told that their participation was vital in determining whether these programmes would be effective in helping people with osteoarthritis *Additional treatment during trial:* unclear
Outcomes	*Outcome assessed at:* baseline and one, two and three years of follow-up **Primary outcomes of study** • Cohesiveness (sociometric questionnaire) • Helplessness (arthritis helplessness index (AHI)) • Knowledge of OA (true/false questionnaire) • Self-management (ASES, 10 to 100, higher score is better) • Healthcare utilisation and costs (data from HMO database) • Quality of life (quality of well-being scale (QWB), 0 to 1.0, higher score is better) • Intervention evaluation (four open-ended questions) • Utilisation rates (participants' medical records) • Comorbidity (10 yes/no questions from AIMS)
Notes	We extracted the following outcomes at 12 months (intermediate term) and at three years (long term) for the analyses in this review: quality of life (QWB) and dropouts (proportion of missing participants) Funding was provided by several grants (NIH Grant AH-40423, NIH Grant 5P60 AR-40770, NIAMS Grant AR-33489) Data analysis: Intervention groups were judged to be similar enough to combine by pooling data

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "After the interview, participants were randomly assigned to one of three health intervention groups or to a control group" Comment: No information was provided on the method of randomisation
Allocation concealment (selection bias)	Unclear risk	Comment: No information was provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants could not be blinded to their treatment allocation because of the nature of the intervention, thus introducing a risk of bias. Although personnel could not be blinded to treatment allocation either, the risk

Cronan 1997 *(Continued)*

		of bias seems to remain low because personnel were involved in only one intervention group (health education intervention)
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: No information was provided on the blinding of outcome assessors; however, most outcomes are subjective, and participants are not blinded to group allocation, thus introducing a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "Attrition from the study after the intervention period was more likely to occur in the social support group than in the education group at the 1-year and 2-year assessment periods" Quote: "Participants were allowed to skip items that they felt uncomfortable answering, resulting in variability in sample sizes between measures" Comment: After three years, 35 participants (40%) dropped out from the social support group, 27 (28%) dropped out from the education group, 33 (37%) dropped out from the combination group and 24 (27%) dropped out from the control group. Dropout rates are high and differ among treatment groups. This, combined with the fact that participants were allowed to skip items they did not want to answer, results in a very high risk of attrition bias
Selective reporting (reporting bias)	High risk	Comment: Only quality of life and dropouts could be extracted from this trial
Other bias	High risk	Quote: "Participants were present at an average of 6.59 of the 10 weekly meetings and 4.32 of the 10 monthly meetings, for a total of 10.9 of the 20 intervention sessions" Comment: Attendance at meetings is on average very low

Crotty 2009

Methods	**Study design:** RCT, single-centre, two arms, non-blinded **Country in which trial was carried out:** Australia **Method of recruitment of participants:**Participants who had an initial consultation with an orthopaedic surgeon concerning a potential hip or knee replacement and who were then added to the waiting list for joint replacement surgery were invited to take part in the study **Setting:**outpatients added to waiting list for joint replacement surgery **Was the sample size justified with a priori calculation of effect size/power?**yes **Length of follow-up:** six months **Dropouts:**no information provided
Participants	**Inclusion criteria** • Mini-Mental Score ≥ 24 • Able to read and speak English • Living in the southern region of Adelaide • Not classified as requiring urgent surgery • No significant frailty or illness that precluded completion of the protocol **Exclusion criteria:**none **Baseline characteristics**

Crotty 2009 *(Continued)*

	Baseline characteristics were similar among all treatment groups ***Intervention group: patient education self-management (N = 75 randomly assigned, 75 analysed)*** *Location of OA:* 33.3% hip *Height (mean (SD)):* 168.0 (10.3) cm *Weight (mean (SD)):* 89.0 (25.4) kg PROGRESS-Plus *Place of residence:* urban (Adelaide) *Occupation:* 12.4% working full-time or part-time, 4.0% home duties, 82.7% retired *Sex:* 40% male, 60% female *Education:* 5.5% none or some primary school, 13.7% primary school, 30.1% high school to year eight, 23.3% high school to year 12, 21.3% TAFE/trade, 5.5% university or above *Social capital:* 42.7% lives alone *Age (mean (SD)), years:* 68.1 (10.6) *Disability:* two (one to three) medical conditions (mean (IQR)); 33.3% on waiting list for hip replacement ***Control group: usual care (N = 77 randomly assigned, 77 analysed)*** *Location of OA:* 32.5% hip *Height (mean (SD)):* 166.4 (9.9) cm *Weight (mean (SD)):* 82.6 (16.4) kg PROGRESS-Plus *Place of residence:* urban (Adelaide) *Occupation:* 10.7% working full-time or part-time, 6.7% home duties, 82.7% retired *Sex:* 39% male, 61% female *Education:* 1.4% none or some primary school, 16.2% primary school, 31.1% high school to year eight, 31.1% high school to year 12, 13.5% TAFE/trade, 6.8% university or above *Social capital:* 53.9% lives alone *Age (mean (SD)), years:* 67 (11.0) *Disability:* two (one to three) medical conditions (mean (IQR)); 32.5% on waiting list for hip replacement
Interventions	***Intervention: patient education self-management*** *Description:* Flinders University Chronic Disease Self-Management Model, which incorporates assessment of self-management knowledge, behaviours, attitudes, strengths and barriers, was used as a basis for the following interventions: • Face-to-face interview using the "Partners in Health"-scale, which identifies strengths of and barriers to self-management • Problems and goals interview that elicits participants' main life problem and medium-term goals; these are rated • Completion of a self-management action plan Monthly telephone calls were made to check health status, self-management strategies reinforcement and progress with goal attainment monitoring. Encouraging to participate in self-management pro-

Crotty 2009 *(Continued)*

gramme "Moving Towards Wellness" and in two joint replacement–specific education modules run by Arthritis SA

Intended audience: people with knee or hip OA on the waiting list for arthroplasty

Mode: individual

Personnel: peer support volunteers

Delivery method: face-to-face, by phone

Language: English

Format: tailored to individual needs

Location: -

Duration: telephone calls conducted monthly, 2.5 hours weekly for six weeks ("Moving Towards Wellness")/2.5 hours weekly for two weeks (joint replacement education programme)

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: no

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: no

Social integration and support: no

Health service navigation: no

Comparator: usual care

Type: usual care

Description: management by primary care physician; participant self-initiates appointments. Participant has access to a generic chronic disease self-management course, “Moving Towards Wellness” (2.5 hours weekly for six weeks), as advertised to the community

Additional treatment during trial: unclear

Outcomes

Outcome assessed at: baseline and six months of follow-up

Primary outcomes of study

- Health-directed activities (heiQ health-directed activities subscale, 1 to 6, higher score is better)

Secondary outcomes of study

- Quality of life (AQoL, -0.04 to 1.00, higher score is better)
- Beliefs about medication (beliefs about medication questionnaire (BMQ))
- Pain (WOMAC, 0 to 20, lower score is better)
- Stiffness (WOMAC, 0 to 8, lower score is better)
- Function self-reported (WOMAC, 0 to 68, lower score is better)
- Emotional distress (CES-D, 0 to 60, lower score is better)

Crotty 2009 *(Continued)*

- Positive and active engagement in life (heiQ positive and active engagement in life subscale, 1 to 6, higher score is better)
- Skill and technique acquisition (heiQ skill and technique acquisition subscale, 1 to 6, higher score is better)
- Constructive attitudes and approaches (heiQ constructive attitudes and approaches subscale, 1 to 6, higher score is better)
- Self-management (heiQ self-monitoring and insight subscale, 1 to 6, higher score is better)
- Health service navigation (heiQ health service navigation subscale, 1 to 6, higher score is better)
- Social integration and support (heiQ social integration and support subscale, 1 to 6, higher score is better)
- Emotional distress (heiQ emotional distress subscale, 1 to 6, higher score is better)

Notes	We extracted the following outcomes at six months (intermediate term) for the analyses in this review: health-directed activities (heiQ health-directed activities subscale), quality of life (AQoL), pain (WOMAC), function self-reported (WOMAC), emotional distress (CES-D), positive and active engagement in life (heiQ positive and active engagement in life subscale), skill and technique acquisition (heiQ skill and technique acquisition subscale), constructive attitudes and approaches (heiQ constructive attitudes and approaches subscale), self-management (heiQ self-monitoring and insight subscale), health service navigation (heiQ health service navigation subscale), social integration and support (heiQ social integration and support subscale) and dropouts (proportion of missing participants) The project was funded by the Commonwealth Department of Health and Ageing as part of its Better Arthritis Care budget initiative. Funding was administered through the Royal Australian College of Physicians, which was contracted to manage the Arthritis and Musculoskeletal Conditions Quality Improvement Program (AMQuIP) Data analysis: For all outcomes, 95% confidence intervals were converted into SD

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “A statistician external to the study generated the randomisation sequence using the random number generator in Microsoft Excel…” Comment: Appropriate method of randomisation with low risk of bias
Allocation concealment (selection bias)	Low risk	Quote: “…sequentially numbered, opaque, sealed envelopes containing group allocation for participants” Comment: Allocation was concealed during recruitment and randomisation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: In this type of intervention and control, it is impossible to blind participants to their treatment group, thus introducing a risk of bias. The nurses conducting the intervention treatment, although also not blinded, have no influence on the treatment given to the control group (usual care by their own physician), leading to low risk of introducing bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Six months after randomisation all participants were mailed a set of the same questionnaires that they completed at baseline and asked to complete and return them in the included reply paid envelope" Comment: Outcome assessment was self-reported, and as participants were not blinded to treatment allocation, a possible risk of bias is present
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: “Data were analysed by intention to treat according to the random allocation”

Crotty 2009 *(Continued)*

		Comment: No information provided on loss to follow-up or dropouts. Appropriate statistical analyses were performed though, with the number of participants analysed at baseline equal to the number analysed at follow-up
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in Methods-section are reported in the Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Hansson 2010

Methods	
Participants	**Inclusion criteria** • Clinical diagnosis of OA in the knee, hip or hand • Pain, stiffness and limitation of movement in the affected joint **Exclusion criteria** • Inability to speak and understand Swedish **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: PEPOA (N = 61 randomly assigned, 61 analysed)*** *Location of OA:* 5% hip, 34% knee, 34% hand, 26% more locations *BMI:* 36% BMI 20 to 25, 39% BMI 25 to 30, 26% BMI > 30 PROGRESS-Plus *Place of residence:* Sweden *Age (mean (SD)), years:* 62 (9.43) ***Control group: usual care (N = 53 randomly assigned, 53 analysed)*** *Location of OA:* 4% hip, 34% knee, 30% hand, 32% more locations *BMI:* 34% BMI 20 to 25, 38% BMI 25 to 30, 28% BMI > 30 PROGRESS-Plus *Place of residence:* Sweden *Age (mean (SD)), years:* 63 (9.51)
Interventions	***Intervention: PEPOA*** *Description:* PEPOA consists of five group sessions. The first session included information about anatomy, physiology of pain, coping with pain and brainstorming about what participants find hard to do. Next sessions were about exercise and physical activity (plus practical demonstration of home-training exercises and orthopaedic aids), current research in OA, medication, appropriate diet, ergonomics and practical instructions about equipment and technical aids and surgery. In the fourth session, feedback on the brainstorming of the first session was provided *Intended audience:* people with OA *Mode:* group sessions (eight to 10 participants)

Hansson 2010 *(Continued)*

Personnel: healthcare professionals (different each session)

Delivery method: face-to-face

Language: Swedish

Format: standard format

Location: -

Duration: five group sessions, three hours for each session, once a week, for five weeks

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: no

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator: usual care

Type: usual care

Description: Control group continued living as usual

Additional treatment during trial: unclear

Outcomes	
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Then an independent person randomised the patient to either the intervention group or the control group, using a random number list and sealed envelopes" Comment: A random number list is an appropriate randomisation method
Allocation concealment (selection bias)	Unclear risk	Quote: "Then an independent person randomised the patient to either the intervention group or the control group, using a random number list and sealed envelopes" Comment: It is unclear whether the sealed envelopes were opaque and thus appropriately concealed
Blinding of participants and personnel (performance bias)	High risk	Quote: "As in all randomised controlled trials concerning rehabilitation, it is not possible to use a double-blind design, since the patients always know whether they are in the intervention group or the control group"

Hansson 2010 *(Continued)*

All outcomes		Comment: Participants were not blinded during the trial for their allocation, possibly introducing a risk of bias. Although personnel also were not blinded to treatment allocation, it is unlikely that this has influenced the results
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The present study has a single-blind design, with measurements at baseline and after six months performed by either a physiotherapist or an occupational therapist, who did not know whether the patient had been in the intervention group or the control group" Comment: Outcome assessors were blinded at the six-month follow-up measurement; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Data were analysed on an intention-to-treat basis, with the dropouts included and the last observation carried forward" Comment: Although the numbers of dropouts between intervention and control groups are not equal (respectively, 16% and 8%), six of the 10 dropouts in the intervention group dropped out before receiving the intervention. Furthermore, reasons for dropping out were similar between groups. Appropriate methods were used to correct for incomplete outcome data
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in Methods section were reported in Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Heuts 2005

Methods	
Participants	**Inclusion criteria** • OA conforms to ICHPP-2 criteria • Aged 40 to 60 years **Exclusion criteria** • A diagnosis of rheumatoid arthritis, ankylosing spondylitis or gout **Baseline characteristics** Baseline characteristics were similar in all treatment groups ***Intervention group: self-management programme (N = 149 randomly assigned, 132 analysed)*** *BMI (mean (SD)):* 28.0 (4.8) kg/m^2 PROGRESS-Plus *Place of residence:* The Netherlands *Occupation:* 42% paying job, 33% no paying job (data from N = 99) *Sex:* 41% male, 59% female *Education:* 24% low, 33% middle, 18% high (data from N = 99) *Age (mean (SD)), years:* 51.0 (5.0) ***Control group: care as usual (N = 148 randomly assigned, 141 analysed)***

Heuts 2005 *(Continued)*

	BMI: 28.3 (5.2) kg/m^2 PROGRESS-Plus *Place of residence:* The Netherlands *Occupation:* 39% paying job, 38% no paying job (data from N = 108) *Sex:* 40% male, 60% female *Education:* 27% low, 28% middle, 22% high (data from N = 108) *Age (mean (SD)), years:* 52.2 (5.1) **Note:** After randomisation, before the start of the intervention, 24 participants withdrew for practical reasons (17 from the intervention group, seven from the control group); some were not able to participate in the intervention schedule, and some were disappointed about the result of the randomisation
Interventions	***Intervention: self-management programme*** *Description:* People were taught how to take initiative in their personal health and functioning. Participants learned to use adequate goal setting in combination with self-incentives as motivators to optimise their level of activity. Rational use of prescribed medication and other treatments was discussed. Self-relaxation training was given for pain control, as well as for improvement of overall well-being. Problem solving was part of the programme for empowering the participant in handling daily hassles. Self-diagnostic skills were taught for monitoring and interpreting changes in one's health status. Participants also received information about community resources and were trained to optimise use of healthcare services *Intended audience:* persons with chronic complaints ascribed to OA *Mode:* group sessions *Personnel:* trained physiotherapists *Delivery method:* face-to-face, written and audiovisual material *Language:* Dutch *Format:* standard format *Location:* physiotherapists' clinic *Duration:* six sessions of two hours each *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* no *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* no *Skill and technique acquisition:* yes *Social integration and support:* unclear *Health service navigation:* yes ***Comparator: care-as-usual***

Heuts 2005 *(Continued)*

	Type: usual care *Description:* What was prescribed by a family physician or a consulted specialist remained unchanged *Additional treatment during trial:* unclear
Outcomes	
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "A computer generated randomization scheme was prepared (…)" Comment: Appropriate method for generating a random sequence was used
Allocation concealment (selection bias)	Low risk	Quote: "(…) and managed by a secretary, who was not involved in patient selection, treatment and data analysis" Comment: Allocation was concealed from study personnel during selection
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants could not be blinded from their treatment, thus participant-reported outcomes are at high risk of performance bias Comment: No blinding of personnel (physiotherapists) does not seem to introduce a risk of performance bias because the comparator is usual care
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "(…) assessments were performed by an independent research assistant who was blinded to treatment assignment and was not involved in the treatment (…)" Comment: Outcome was assessed by blinded personnel who were not involved in treatment; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "(…) all participants were analysed in the intervention group they entered by randomization (…) no imputation method was administered for missing questionnaires" Quote: "Five participants withdrew during the intervention from the self-management program: 3 were not satisfied by the program, one because of knee pain, and one because of the situation at home" Comment: Participants were analysed in the group to which they were randomly assigned. Although missing data are not treated according to the intention-to-treat principle, the risk of attrition bias remains low, as the dropout rate is low with reasons provided
Selective reporting (reporting bias)	Low risk	Comment: All outcomes mentioned in the Methods section were reported in the Results. No prepublished protocol is available
Other bias	Low risk	Comment: No other potential sources of bias were identified

Hopman-Rock 2000

Methods	**Study design:**RCT, single-centre, two arms, outcome assessment blinded

Hopman-Rock 2000 *(Continued)*

Country in which trial was carried out: The Netherlands

Method of recruitment of participants: Participants were recruited by announcements in newspapers and on television in the area around the research centre

Setting: general population

Was the sample size justified with a priori calculation of effect size/power?yes

Length of follow-up:six months

Dropouts: One (2%) dropped out from the control group (personal reasons), no one dropped out from the intervention group

Note: "During the statistical analyses after the follow-up assessment, subjects without a confirmed diagnosis of OA were excluded (4 in the experimental group, 10 in the control group). (...) The results for the total group (...) led to the same conclusions"

Participants

Inclusion criteria:

- Self-reported OA of hip or knee, confirmed by clinical (only for knee) or radiographic ACR criteria (hip and knee, Kellgren-Lawrence score of 2 or higher)
- Aged 55 to 75 years

Exclusion criteria

- People on a waiting list for knee or hip replacement

Baseline characteristics

Baseline characteristics were similar among all treatment groups

Intervention group: health educational and exercise programme (N = 60 randomly assigned, 56 analysed)

BMI (mean (SD)): 28.4 (4.8) kg/m^2

Duration of OA: 4% less than one year, 25% one to three years, 30% three to 10 years, 18% 10 to 20 years, 18% longer than 20 years

PROGRESS-Plus

Place of residence: around Leiden, The Netherlands

Sex: 22% male, 78% female

Education: 17% primary, 54% secondary, 27% college/university

Social capital: 68% living together/married, 29% living alone

Age (mean (SD)), years: 65.4 (5.3)

Disability: 2.2 (1.5) other chronic conditions (SD)

Control group (N = 60 randomly assigned, 49 analysed)

BMI (mean (SD)): 26.8 (3.5) kg/m^2

Duration of OA: 2% less than one year, 16% one to three years, 37% three to 10 years, 18% 10 to 20 years, 16% longer than 20 years

PROGRESS-Plus

Place of residence: around Leiden, The Netherlands

Sex: 12% male, 88% female

Hopman-Rock 2000 *(Continued)*

Education: 26% primary, 45% secondary, 20% college/university

Social capital: 67% living together/married, 25% living alone

Age (mean (SD)), years: 65.2 (5.7)

Disability: 2.8 (1.7) of other chronic conditions (SD)

Interventions	***Intervention: health educational and exercise programme***

Description: The first hour of each session was guided by a peer educator, and the following topics were discussed: pathophysiology of OA, lifestyle and physical activity, pain management, the importance of weight reduction and diet, ergonomic aspects and medical aspects of OA (treatments, radiographs). Additionally, questions were answered by an invited occupational therapist and a GP. The course included the use of a pain diary and personal goal planning, as well as interactive methods in the group. The second hour, participants were taught an exercise programme by a physical therapist. Fifteen minutes of each session was spent on education about the balance between rest and activity, preferable types of activity and how to incorporate them in a daily lifestyle, along with practical advice on physical activity, such as the benefits of walking. Participants were encouraged to do the exercises at home at least three times a week

Intended audience: people with hip or knee OA

Mode: group (max 15 participants)

Personnel: first hour, a peer educator and an invited occupational therapist and a GP; second hour, a physical therapist

Delivery method: face-to-face

Language: Dutch

Format: tailored to individual needs and standard format

Location: -

Duration: six weekly sessions of two hours

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator

Type: no treatment

Description: control group without intervention. The programme book (together with a gift voucher for $25) was offered to the control group after the follow-up test was finished

Additional treatment during trial: unclear

Hopman-Rock 2000 *(Continued)*

Outcomes	*Outcome assessed at:* baseline, post-treatment, and six months of follow-up **Primary outcomes of study** • Pain (pain subscale of IRGL (Dutch AIMS), 6 to 25, lower score is better) • Pain (pain intolerance on 10-cm VAS, 0 to 10, lower score is better) • Quality of life (VAS, 0 to 100, higher score is better) • Quality of life (sum score seven questions, 7 to 39, higher score is better) • Functional performance (range of motion with goniometer) • Functional performance (isometric muscle strength with dynamometer) • Functional performance (20-m walking test, timed up-and-go test, stair walking, reaching for toes in sitting position, lower scores are better) • Knowledge about OA (20 statements) • Self-management (self-efficacy on five 10-cm VAS, 0 to 100, higher score is better) • Function self-reported (mobility subscale of IRGL (Dutch AIMS), 7 to 28) • Healthcare utilisation (use of medication, number of visits to GP, number of visits to physical therapist)
Notes	We extracted the following outcomes at post-treatment (short term) and at six months (intermediate term) for the analyses in this review: self-management (five 10-cm VAS), pain (IRGL subscale pain), function self-reported (IRGL subscale mobility), functional performance (20-m walking test), quality of life (sum score seven questions) and dropouts (proportion of missing participants) Funding provided by a grant from The Netherlands Health Research and Development Council Data analysis: For the outcome function—performance—we could choose between the 20-m walking test and the timed up-and-go test and chose for the 20-m walking test, as we judged that more studies reported performance measures similar to the 20-m walking test

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “Finally, 120 subjects were randomly assigned to one of the 2 conditions (…)” Comment: no information provided on method of random sequence generation
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: With this kind of intervention and control, it is not possible to blind participants from their allocated treatment, possibly introducing a risk of bias. Although personnel were not blinded during the trial, it is unlikely that this has biased the results
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: “The personal interview (by a trained interviewer who was blinded) contained (…). The examination was carried out by 3 trained physical therapists who were blinded for the condition” Comment: Personnel who assessed several outcomes were blinded to treatment allocation; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias)	Low risk	Comment: Although it is unclear whether an intention-to-treat analysis was performed, risk of bias is low because dropout is minimal (N = 1)

Hopman-Rock 2000 *(Continued)*
All outcomes

Selective reporting (reporting bias)	Low risk	Comment: All outcomes reported in the Methods section are reported in the article
Other bias	Low risk	Comment: No other potential sources of bias were identified

Hughes 2004

Methods	**Study design:** RCT (block-randomisation (30), stratified for ACR functional class), single-centre, two arms, non-blinded **Country in which trial was carried out:**US **Method of recruitment of participants:** by newsletter, announcements in the local media and presentations to local senior groups **Setting:**general population **Was the sample size justified with a priori calculation of effect size/power?**no justification of sample size given **Length of follow-up:** 12 months **Dropouts:**After 12 months, 57 (50%) dropped out from the intervention group (38 unable to contact, eight refused interview, five illness, three moved out of state, two caregiving responsibilities, one disconnected phone), 68 (68%) dropped out from the control group (43 unable to contact, 11 refused interview, six illness, three moved out of state, three caregiving responsibilities, two disconnected phone)
Participants	**Inclusion criteria** • Clinical criteria of the ACR for knee OA or hip OA **Exclusion criteria** • Younger than age 60 • Currently participating in an aerobic exercise programme • Uncomplicated hip or knee surgery in the past six months • Complicated surgery within the past year • Steroid injections in either knee or hip within the previous three months • Diagnosed with RA • Acutely inflamed or significantly swollen joint • Severe limiting cardiovascular disease • Active thrombophlebitis • Recent pulmonary embolus • An acute systemic illness • Poorly controlled diabetes • Other health conditions that might preclude exercise training **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: fit and strong! (N = 115 randomly assigned, 58 analysed)*** PROGRESS-Plus *Place of residence:* US

Hughes 2004 *(Continued)*

Race, ethnicity and culture: 69.4% white-Caucasian, 27.8% African American, 1.9% Hispanic, 0.9% Asian-Pacific Islander, 0% other

Sex: 19.4% male, 80.6% female

Education: 12.1% less than high school, 21.5% high school, 66.4% more than high school

Socioeconomic status: 32.4% income less than $20,000

Age, years: 73.3

Disability: 22.6% ARA class I, 64.5% ARA class II, 12.9% ARA class III; 51.4% hypertension, 46.7% cardiovascular disease, 6.1% asthma, 3.5% emphysema, 14.6% diabetes, 6.1% cancer

Control group: information only (placed on waiting list for Fit and Strong! programme) (N = 100 randomly assigned, 32 analysed)

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 75.0% white-Caucasian, 16.3% African American, 3.3% Hispanic, 3.3% Asian-Pacific Islander, 2.2% other

Sex: 14.1% male, 85.9% female

Education: 8.8% less than high school, 18.5% high school, 72.7% more than high school

Socioeconomic status: 33.7% income less than $20,000

Age, years: 73.4

Disability: 22.2% ARA class I, 64.2% ARA class II, 13.6% ARA class III; 58.3% hypertension, 42.3% cardiovascular disease, 7.0% asthma, 5.0% emphysema, 12.8% diabetes, 2.0% cancer

Interventions

Intervention: Fit and Strong!

Description: The first 60 minutes of the intervention included both resistance training and fitness walking. In resistance training, resistance was progressively increased throughout the programme by adding weight. Fitness walking progressed from maximum duration at baseline to 30 minutes over time. The last 30 minutes included an adapted version of the group discussion-educational component by Kovar and colleagues (1992) to enhance adherence efficacy. Self-efficacy for exercise (confidence in the ability to conduct the exercises in a safe and effective manner) and self-efficacy for exercise adherence (confidence in the ability to adhere to exercise participation over time and in the presence of barriers) were addressed, as was self-efficacy to manage pain and other arthritis-related symptoms. Therapists provided systematic feedback to participants on progress made. The emphasis was on building skills and identifying strategies that will assist the participant in maintaining adherence. Staff used group and individual sessions to inform participants about opportunities for maintaining exercise within the community or in the individual's home. All graduates are also given a copy of *The Arthritis Helpbook*, a graduation certificate and tapes of music used during the class at a graduation ceremony at 8 weeks.

Intended audience: people with OA of hip or knee

Mode: group sessions (15 participants)

Personnel: physical therapist

Delivery method: face-to-face

Language: English

Format: tailored to individual needs

Location: senior centres and senior housing residences

Hughes 2004 *(Continued)*

Duration: 90-minute sessions, held three times a week, for eight weeks

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: no

Constructive attitudes and approaches: no

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator

Type: Information only

Description: Control group participants are given a copy of *The Arthritis Helpbook* and a list of exercise programmes in the community that they can access. They are also given a variety of self-care materials and handouts at each post-test.

Additional treatment during trial: unclear

Outcomes	*Outcome assessed at:* baseline and two, six and 12 months of follow-up **Primary outcomes of study** • Self-management (self-efficacy (ASES) subscales exercise (1 to 10), pain management (10 to 100), other symptoms (10 to 100)) • Self-management (barriers adherence efficacy scale, time adherence self-efficacy) • Adherence • Functional performance (lower extremity muscle strength, timed sit-to-stand in 60 seconds, six-minute distance walk) • Pain (WOMAC-pain, 0 to 20, lower score is better) • Function self-reported (WOMAC-stiffness (0 to 8), physical function (0 to 68), lower scores are better) • Maintenance of physical activity • Pain (geri-AIMS pain, 0 to 10)
Notes	We extracted the following outcomes at 12 months (intermediate term) for the analyses in this review: self-management (ASES subscale pain), pain (WOMAC subscale pain), global OA scores (WOMAC total), function self-reported (WOMAC subscale physical function), functional performance (six-minute distance walk) and dropouts (proportion of missing participants) The development of Fit and Strong! was made possible by a grant from the Chicago Chapter of the Arthritis Foundation. The research was also supported by funding from the National Institute on Arthritis and Musculoskeletal Disease (Grant AR60692) and by the National Institute on Aging and the Royal Center for Research on Applied Gerontology (Grant AG 15890) Data analysis: For all outcomes, 95% confidence intervals were converted into SD. For the outcome self-management in OA, we could choose between ASES subscale pain and ASES subscale other and chose for ASES subscale pain, as we judged that pain was a more measurable aspect of self-management. For the outcome global OA scores, we added the subscales of WOMAC (pain, stiffness and physical function) to get the WOMAC total score; we estimated the SD with the formula provided in Table

Hughes 2004 *(Continued)*

7.7a in the *Cochrane Handbook for Systematic Reviews of Interventions* and the SDs from the subscales (this method was chosen in close consultation with a biostatistician)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "At that point, we randomised the participant to either the treatment or the control group" Comment: no information provided on the method of random sequence generation
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Third, participants were not blinded regarding their treatment status. It is not possible in an exercise study to blind the instructor to the participants, nor is it possible to blind the exercise participants to the fact that they are receiving a treatment" Comment: Participants were not blinded, which may have introduced a risk of bias. Although personnel were not blinded either, it is unlikely that this introduced a risk of bias because no intervention staff came in contact with the control group
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "In this comparatively small study, it also was not possible to blind the research staff regarding group assignment because many of the staff also helped to set up the class, assisted the physician to conduct the physical examinations, and other activities" Comment: Outcome assessors were not blinded, which resulted in a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "These analyses demonstrated no significant differences between responders and non-responders on any of the outcome measures, (…) in terms of demographic characteristics or level of arthritis severity. In both groups, participants who attrited from post-test measurement had slightly worse scores on the study outcome measures at baseline" Comment: The dropout rate in this study is very high and is not equally divided between treatment arms (50% in the intervention group vs 68% in the control group). Although not statistically significant, an apparent trend suggested a slightly worse score on study outcome measures at baseline among participants who dropped out. No intention-to-treat analysis was performed, but a completers-only analysis was done
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in the Methods were reported in the Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Hurley 2007

Methods	**Study design:** cluster-RCT, multi-centre, three arms, outcome assessment blinded **Country in which trial was carried out:**UK

Hurley 2007 *(Continued)*

Method of recruitment of participants: Potential participants were identified from databases of primary care practices

Setting:primary care

Was the sample size justified with a priori calculation of effect size/power?yes

Length of follow-up:six months

Dropouts:27 (19%) in the usual primary care group dropped out, 25 (17%) in the individual rehabilitation group dropped out, 24 (18%) in the group rehabilitation group dropped out (five exercise-related events, 36 stopped attending without reason, 15 lost interest, 10 were unable to fulfil time commitment, five had unrelated medical problems, five moved away)

Participants

Inclusion criteria

- Knee pain due to OA based on clinical presentation and history (no attempts were made to identify the cause of the pain using investigations not routinely available to primary care physicians (e.g. radiographs))
- Age 50 years or older
- Consulted primary care physician for mild, moderate or severe knee pain of > six months' duration

Exclusion criteria

- Lower limb arthroplasty
- Physiotherapy for knee pain in the preceding 12 months
- Intra-articular injections in the preceding six months
- Unstable medical conditions
- Inability/unwillingness to exercise
- Wheelchair dependence
- Inability to understand English

Baseline characteristics

Baseline characteristics were similar among all treatment groups

Intervention group: individual rehabilitation (N = 186 randomly assigned, 121 analysed)

Location of OA: 100% knee

BMI (mean (range)): 30.0 (18 to 45) kg/m^2

Duration of OA (mean (IQR)): 7 (3 to 15) years

PROGRESS-Plus

Place of residence: South-East London

Sex: 28.8% male, 71.2% female

Age (mean (range)), years: 66 (50 to 91)

Intervention group: group rehabilitation (N = 132 randomly assigned, 108 analysed)

Location of OA: 100% knee

BMI (mean (range)): 30.18 (20 to 50) kg/m^2

Duration of OA (mean (IQR)): 5 (2.5 to 11) years

PROGRESS-Plus

Place of residence: South-East London

Hurley 2007 *(Continued)*

	Sex: 28.8% male, 71.2% female *Age (mean (range)), years:* 68 (51 to 84) ***Control group (N = 140 randomly assigned, 113 analysed)*** *Location of OA:* 100% knee *BMI (mean (range)):* 30.3 (20 to 51) kg/m^2 *Duration of OA (mean (IQR)):* 6 (3 to 15) years PROGRESS-Plus *Place of residence:* South-East London *Sex:* 31.4% male, 68.6% female *Age (mean (range)), years:* 67 (51 to 89)
Interventions	***Intervention (individual/group rehabilitation)*** *Description:* The intervention comprised integrated patient education, with simple self-management and pain coping strategies, delivered in the first 15 to 20 minutes of each rehabilitation session. Sessions were designed to be interactive, including active problem solving when appropriate. This was followed by a 35 to 45-minute individualised progressive exercise programme. The order in which the exercises were performed varied because of the circuitous regimen, and exercise specificity varied between participants and within participants over time, depending on their ability, rate of progression and identified areas of disability. Exercise complexity and intensity were increased through mutual agreement between physiotherapist and participant. The content was similar in individual or group rehabilitation *Intended audience:* people with chronic knee pain *Mode:* individual/group sessions (~eight participants) *Personnel:* experienced physiotherapist *Delivery method:* face-to-face *Language:* English *Format:* tailored to individual needs *Location:* physiotherapy outpatient department *Duration:* 12 sessions (twice weekly for six weeks), one hour per session (15 to 20 minutes + 35 to 45 minutes) *Additional treatment during trial:* Management of all participants' knee and coexistent medical problems continued at the primary care physician's discretion Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* yes *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* yes *Skill and technique acquisition:* yes

Hurley 2007 *(Continued)*

	Social integration and support: no *Health service navigation:* no ***Comparator*** *Type:* usual care *Description:* Primary care physicians were free to prescribe or refer participants (to any services or interventions they considered appropriate), whatever intervention they considered appropriate, and followed up as necessary. Management was typical of other primary care reports *Additional treatment during trial:* unclear
Outcomes	*Outcome assessed at:* baseline, post-treatment (or six weeks after recruitment to usual care arm) and six months after completion intervention (or 7.5 months after recruitment to usual care arm) **Primary outcomes of study** • Function self-reported (WOMAC subscale function, 0 to 68, lower score is better) • Costs of intervention **Secondary outcomes of study** • Pain (WOMAC subscale pain, 0 to 20, lower score is better) • Global OA scores (WOMAC-total score, 0 to 96, lower score is better) • Functional performance (aggregated functional performance time (AFPT) in seconds, lower score is better) • Exercise health beliefs (ExBeliefs questionnaire) • Emotional distress (HADS-anxiety (0 to 21), -depression subscales (0 to 21), lower score is better) • Quality of life (EQ-5D, higher score is better) • Quality of life (MACTAR, higher score is better) • Functional performance (quadriceps strength and voluntary activation, higher score is better)
Notes	We extracted the following outcomes at six months (intermediate term) for the analyses in this review: pain (WOMAC subscale pain), global OA scores (WOMAC-total), function self-reported (WOMAC subscale function), quality of life (EQ-5D), emotional distress (HADS subscale depression) and dropouts (proportion of missing participants) Research funded by Arthritis Research Campaign Research Fellowship Study author (M Hurley) provided additional information on request Data analysis: Intervention groups were judged to be similar enough to combine by pooling data

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “The randomization list was generated at a central location away from the research centre by an author (BR), who was not involved in the execution of the trial” Comment: This method of random sequence generation has a low risk of bias
Allocation concealment (selection bias)	Unclear risk	Comment: no information on allocation concealment given
Blinding of participants and personnel (performance bias)	High risk	Comment: Blinding of participants and personnel to treatment allocation is not possible because of the types of intervention and control in this trial

Hurley 2007 *(Continued)*

All outcomes		
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Outcome assessors were blinded to a participant's allocation. Success of blinding was evaluated (…)" Comment: In 'discussion', authors explain that subgroup analysis suggested that unblinding was associated with slightly better outcome, but that this difference was not statistically significant. Most outcomes are self-reported, and participants were not blinded to treatment allocation, so high risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "There was no evidence of differential attrition" Quote: "Statistical analysis followed a prespecified protocol, based on intent-to-treat with no interim or post hoc analyses" Quote: "By 6 months, 76 (18%) participants had withdrawn" Comment: Loss to follow-up was low and equal among all treatment groups (19%, 17% and 18%). Reasons for loss to follow-up are stated; only 1% withdrew because of exercise-related adverse events
Selective reporting (reporting bias)	Low risk	All planned outcomes were reported
Other bias	High risk	Quote: "Of the participants who attended 6-month follow-up, 105 (85%) of 120 Indiv-rehab participants and 59 (55%) of 107 Grp-rehab participants attended >= 10 of the 12 sessions" Comment: Adherence in the group-rehabilitation arm was lower than in the individual-rehabilitation arm Quote: "Multilevel modeling was used to adjust for the intracluster correlations (...)" Comment: The cluster design was accounted for in the statistical analysis

Jessep 2009

Methods	**Study design:** RCT, single-centre, two arms, outcome assessment was blinded **Country in which trial was carried out:** UK **Method of recruitment of participants:** Potential participants were identified from two general practice databases and were sent a letter outlining the study and inviting to participate **Setting:** outpatient physiotherapy department **Was the sample size justified with a priori calculation of effect size/power?**no (no data available on which to base a power calculation; sample size of 60 was considered convenient and adequate) **Length of follow-up:** 12 months **Dropouts:**Eight (23%) dropped out from the physiotherapy group (one withdrew, one was diagnosed with spinal stenosis, two developed unrelated health problems, one had heart surgery, one moved away, two stopped attending), eight (28%) dropped out from the self-management group (one withdrew, one developed heart problems, one was diagnosed with polymyalgia rheumatica, one developed hip complications, one had related knee surgery, one moved away, one stopped attending)
Participants	**Inclusion criteria**

Jessep 2009 *(Continued)*

- Clinical OA based on clinical presentation and history (mild, moderate or severe non-specific knee pain, lasting > six months with no identifiable recent cause)
- Aged over 50
- Presented to GP with chronic knee pain

Exclusion criteria

- Reporting that knee pain emanated from knee trauma within the past year
- Lower limb arthroplasty
- Unstable, coexisting medical or psychological conditions
- Physiotherapy for knee pain in the previous 12 months
- Intra-articular injections in the previous six months
- Unable or unwilling to exercise
- Unable to walk 100 metres
- Insufficient command of English to complete the assessment and undertake the intervention
- Other joint pain that would prevent them from participating in an exercise programme
- Taking steroids
- Wheelchair bound

Baseline characteristics

Baseline characteristics were similar in all treatment groups

Intervention group: ESCAPE-Knee pain (N = 29 randomly assigned, 26 analysed post-treatment, 21 analysed after 12 months of FU)

Location of OA: 100% knee

BMI (mean (range)): 30 (20 to 42) kg/m^2

Duration of OA (mean (range)): 13 (1 to 30) years

PROGRESS-Plus

Place of residence: UK

Sex: 24% male, 76% female

Age (mean (range)), years: 66 (53 to 81)

Control group: physiotherapy (N = 35 randomly assigned, 31 analysed post-treatment, 27 analysed after 12 months of FU)

Location of OA: 100% knee

BMI (mean (range)): 29 (20 to 47) kg/m^2

Duration of OA (mean (range)): 12 (0.5 to 55) years

PROGRESS-Plus

Place of residence: UK

Sex: 37% male, 63% female

Age (mean (range)), years: 67 (51 to 76)

Interventions	***Intervention: ESCAPE-Knee pain*** *Description:* The programme aims to change people's behaviour by challenging inappropriate beliefs regarding their condition and physical activity, encouraging regular exercise and enabling self-management. It exists of informal group discussions that promote shared learning, information and advice about simple coping strategies, problem-solving and planning skills and active participation in a pro-

Jessep 2009 *(Continued)*

	gressive exercise regimen. By the end of the programme, participants have learnt how to utilise physical activity to self-manage their symptoms *Intended audience:* people with chronic knee pain (ascribed to OA) *Mode:* group sessions (six participants) *Personnel:* physiotherapist *Delivery method:* face-to-face, written information summary *Language:* English *Format:* standard format and tailored to individual needs (exercises) *Location:* local authority adult education centre *Duration:* 10 sessions of one hour, held twice a week (total duration of five weeks) *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* yes *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* yes *Skill and technique acquisition:* yes *Social integration and support:* no *Health service navigation:* no ***Comparator: physiotherapy*** *Type:* alternate intervention *Description:* usual care physiotherapy, with the treatment modalities that the physiotherapist believed were necessary *Frequency:* 30 to 45-minute assessment, and after that, up to a maximum of 10 sessions *Additional treatment during trial:* unclear
Outcomes	*Outcome assessed at:* baseline, post-treatment and 12 months of follow-up **Primary outcomes of study** • Function self-reported (WOMAC subscale physical function, 0 to 68, lower score is better) **Secondary outcomes of study** • Pain (WOMAC subscale pain, 0 to 20, lower score is better) • Functional performance (Aggregate Functional Performance Test (AFPT), 0 to infinite, lower score is better) • Quality of life (EQ-5D, 0 to 1, higher score is better) • Exercise-related health beliefs (0 to 68, lower score is better) • Emotional distress (Hospital Anxiety and Depression scale (HADS), 0 to 21, lower score is better) • Healthcare utilisation (Client Services Receipt Inventory (CSRI))

Jessep 2009 *(Continued)*

Notes	We extracted the following outcomes at post-treatment (short term) and 12 months (intermediate term) for the analyses in this review: pain (WOMAC subscale pain), function self-reported (WOMAC subscale function), quality of life (EQ-5D), emotional distress (HADS subscale depression) and dropouts (proportion of missing participants) Funding provided by the Physiotherapy Research Foundation Project (number PRF/03/3)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "(…) participants were allocated to receive outpatient physiotherapy or ESCAPE-knee pain using a randomisation list generated (…) at a centre away from Sevenoaks Hospital (…)" Comment: It remains unclear how the random sequence was generated
Allocation concealment (selection bias)	Low risk	Quote: "(…) and held at a centre away from Sevenoaks Hospital to ensure concealed allocation" Comment: Allocated treatment was concealed from personnel during inclusion of participants
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants were not blinded from their allocated treatment, and therefore outcomes might be biased. Physiotherapists treating the participants (both intervention and control groups) were not blinded from the allocated treatment group
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The assessor was unaware of each participant's treatment allocation" Comment: Outcome assessors were blinded to treatment allocation; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "All clinical and cost data analyses were by intention-to-treat (i.e. participant data were analysed in the groups to which they were randomised) (…)" Comment: Although dropout rates are equal between groups and the reasons given indicate no withdrawals due to treatment allocation, the rate of dropout is high in both groups. Authors report using an intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Comment: All prespecified outcomes in the protocol were reported in the Outcomes section
Other bias	Low risk	Comment: No other potential sources of bias were identified

Keefe 1990

Methods	**Study design:**RCT, single-centre, three arms, non-blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**unclear **Setting:**outpatients **Was the sample size justified with a priori calculation of effect size/power?**no justification provided

Keefe 1990 *(Continued)*

Length of follow-up: six months

Dropouts: After six months, three (10%) dropped out from usual care, one (3%) dropped out from arthritis education and two (6%) dropped out from pain coping skills training. No reasons provided

Participants	**Inclusion criteria** • Diagnosed as having OA of the knee on the basis of medical evaluation and radiographic examination **Exclusion criteria** • Having arthritic disorders other than OA • Having other known organic disease that would significantly affect function (e.g. chronic obstructive pulmonary disease) • Patients receiving disability support payments for OA **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: pain coping skills (N = 32 randomly assigned, 31 analysed)*** *Location of OA:* 100% knee *Mean percentage above ideal weight (mean (SD)):* 21.0 (17.6) PROGRESS-Plus *Place of residence:* US *Sex:* 22% male, 78% female *Age (mean (SD)), years:* 62.4 (12.0) ***Control group: arthritis education (N = 36 randomly assigned, 35 analysed)*** *Location of OA:* 100% knee *Mean percentage above ideal weight (mean (SD)):* 27.2 (29.0) PROGRESS-Plus *Place of residence:* US *Sex:* 31% male, 69% female *Age (mean (SD)), years:* 66.0 (9.5) ***Control group: standard care control (N = 31 randomly assigned, 28 analysed)*** *Location of OA:* 100% knee *Mean percentage above ideal weight (mean (SD)):* 24.1 (22.2) PROGRESS-Plus *Place of residence:* US *Sex:* 32% male, 68% female *Age (mean (SD)), years:* 63.0 (13.0)
Interventions	***Intervention: pain coping skills*** *Description:* Participants in this condition received a cognitive-behavioural intervention. To help participants understand the rationale for training in pain coping skills, a simplified version of Melzack and

Keefe 1990 *(Continued)*

Wall's gate control model was used to show that pain is a complex experience affected by thoughts, feelings and behaviours. Relaxation, imagery and distraction techniques were introduced as methods for controlling pain through attention diversion. Cognitive restructuring was utilised to help participants recognise and modify irrational cognitions related to pain

During the six months of follow-up, participants were called three times by their group therapists (one month, two months and four months after completion of treatment). The calls focused on a review of progress since completion of the group and a discussion of participants' use of pain coping skills

Intended audience: people with knee OA

Mode: group sessions(six to nine participants)

Personnel: psychologist and nurse

Delivery method: face-to-face

Language: English

Format: standard format

Location: -

Duration: 10 weekly sessions, each of 90-minute duration

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: no

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator: arthritis education

Type: attention control

Description: The arthritis education intervention used a lecture-discussion format developed by Lorig. Four topics were discussed: the nature of OA, treatment methods, exercise and mobility and function

During the six months of follow-up, participants were called three times by their group therapists (one month, two months and four months after completion of treatment). The calls focused on a review of progress since completion of the group and answering any questions that participants had about their arthritis and its treatment

Frequency: 10 weekly sessions of 90-minute duration

Additional treatment during trial: unclear

Comparator: standard care control

Type: usual care

Description: Participants in this condition continued with their routine care for OA

Keefe 1990 *(Continued)*

	Additional treatment during trial: Participants did not attend any group sessions for arthritis education or pain coping skills training
Outcomes	*Outcome assessed at: b*aseline, post-treatment and six months of follow-up **Primary outcomes of study** • Constructive attitudes and approaches (Coping Strategies Questionnaire (CSQ) subscales) • Pain (AIMS subscale pain, 0 to 10, lower score is better) • Function self-reported (AIMS subscale physical disability, 0 to 10, lower score is better) • Emotional distress (AIMS subscale psychological disability, 0 to 10, lower score is better) • Medication use • Direct observation of motor pain behaviour (videotapes)
Notes	We extracted the following outcomes at post-treatment (short term) and at six months (intermediate term) for the analyses in this review: pain (AIMS subscale pain), function self-reported (AIMS subscale physical disability), emotional distress (AIMS subscale psychological disability) and dropouts (proportion of missing participants) The trial was supported by NIAMS Grant No. RO1 AM NS 35270

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "(…) and were randomly assigned (using a random number table) (…)" Comment: Method of randomisation with low risk of bias is used
Allocation concealment (selection bias)	Unclear risk	Comment: No information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: These types of interventions and controls make it impossible for authors to blind participants and personnel, which introduces a risk of performance bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: Outcomes were assessed through self-report and might therefore have been biased, as participants were not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: The authors provide no information on the manner in which the analyses were performed and how missing data were handled. However, the rate of dropout was low and equal among treatment groups
Selective reporting (reporting bias)	Unclear risk	Comment: Constructive attitudes and approaches and medication use not reported
Other bias	Low risk	Comment: No other potential sources of bias were identified

Keefe 1996

Methods	**Study design:**RCT, three arms, non-blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**unclear

Keefe 1996 *(Continued)*

	Setting:unclear
	Was the sample size justified with a priori calculation of effect size/power?no justification for sample size provided
	Length of follow-up:12 months
	Dropouts: Nine (30%) dropped out from SA-CST, five (17%) dropped out from CST, four (14%) dropped out from AE-SS. No reasons provided
Participants	**Inclusion criteria** • Diagnosed as having OA of the knee • Being married • No arthritic disorder other than OA • No other known organic disease that would significantly affect function (e.g. chronic obstructive pulmonary disease) • Were not receiving disability support payments for OA **Exclusion criteria:**none **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: coping skills training (CST) (N = 29 randomly assigned, 26 analysed post-treatment, 24 analysed after six and 12 months of FU)*** *Location of OA:* 100% knee PROGRESS-Plus *Place of residence:* US *Sex:* 48% males, 52% females *Age (mean), years:* 61.4 ***Intervention group: spouse-assisted CST (N = 30 randomly assigned, 27 analysed post-treatment, 25 analysed after six months, 21 analysed after 12 months of FU)*** *Location of OA:* 100% knee PROGRESS-Plus *Place of residence:* US *Sex:* 40% males, 60% females *Age (mean), years:* 63.5 ***Control group: arthritis education –spousal support (N = 29 randomly assigned, 28 analysed post-treatment, 26 analysed after six months, 25 analysed after 12 months of FU)*** *Location of OA :* 100% knee PROGRESS-Plus *Place of residence:* US *Sex:* 32% male, 68% female *Age (mean), years:* 62.8
Interventions	***Intervention: coping skills training (CST)***

Keefe 1996 *(Continued)*

Description: "The rationale for the CST intervention was based on Melzack and Wall's gate control model of pain, which views pain as a complex experience, affected by thoughts, feelings, and behaviours. Patients were told that a major focus of the group was developing a menu of pain-coping skills. Three sets of coping skills were included: attention diversion skills, activity-based skills, and cognitive coping strategies"

Intended audience: married participants with knee OA

Mode: group sessions (four to six participants, without their spouses)

Personnel: nurse and a psychologist, both trained in the CST programme

Delivery method: ace-to-face

Language: English

Format: standard format

Location: -

Duration: 10 weekly sessions of two hours

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: no

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Intervention: spouse assisted –CST

Description: Each session included training in pain coping skills and couples skills. Participants and their spouses were provided with a rationale that emphasised (1) OA is a couples issue that can affect each partner and his or her relationships, (2) developing ways to cope with OA is an important task for each couple and (3) the reactions of the spouse can influence the participant's ability to cope with pain. Pain coping strategies were the same as in the CST group. Couples skills were designed to supplement and reinforce the participant's pain coping skills; they included communication skills, behavioural rehearsal, mutual goal setting, joint home and in vivo practice and maintenance training

Intended audience: married participants with knee OA

Mode: group sessions (four to six participants and their spouses)

Personnel: nurse and psychologist, both trained in the CST programme

Delivery method: face-to-face

Language: English

Format: standard format

Location: -

Duration: ten weekly sessions of two hours

Keefe 1996 *(Continued)*

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: no

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: yes

Health service navigation: no

Comparator: arthritis education –spousal support

Type: alternate intervention

Description: The group was provided a detailed overview of the nature of OA, methods of diagnosis, medical and surgical treatments, home remedies and methods for maintaining mobility and flexibility. Spouses attended each session along with the participant. Spouses were encouraged to participate fully in all discussions of educational material

Frequency: 10 weekly group sessions (four to six participants and their spouses) of two hours

Additional treatment during trial: unclear

Outcomes

Outcome assessed at: baseline, post-treatment and six and 12 months of follow-up

Primary outcomes of study

- Marital adjustment (Dyadic Adjustment Scale (DAS))
- Self-management (Arthritis Self-Efficacy Scale (ASES) subscales pain (10 to 100), function (10 to 100) and other symptoms (10 to 100), higher scores are better)
- Pain (AIMS subscale pain, 0 to 10, lower score is better)
- Function self-reported (AIMS subscale physical disability, 0 to 10, lower score is better)
- Emotional distress (AIMS subscale psychological disability, 0 to 10, lower score is better)
- Constructive attitudes and approaches (Coping Strategies Questionnaire (CSQ) subscales coping attempts, pain control and rational thinking)
- Pain behaviour (videotape)

Notes

We extracted the following outcomes at post-treatment (short term) and at 12 months (intermediate term) for the analyses in this review:

self-management (ASES mean), pain (AIMS subscale pain), function self-reported (AIMS subscale physical disability), emotional distress (AIMS subscale psychological disability), constructive attitudes and approaches (CSQ subscale coping attempts) and dropouts (proportion of missing participants)

The trial was supported by NIAMS Grant AR-35270

Data analysis: Intervention groups were judged to be similar enough to combine by pooling data

Risk of bias

Bias	Authors' judgement	Support for judgement

Keefe 1996 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Quote: "All patients were randomly assigned to 1 of 3 treatment conditions (…)" Comment: no information provided on randomisation method
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	These types of controls and interventions make blinding of participants and personnel not possible, which may have introduced bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: no information provided on blinding of outcome assessors; however, most outcomes are subjective, and participants are not blinded to group allocation, introducing a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: Dropout rate immediately post-treatment showed no significant differences. After 12 months, dropout seems unequal and is moderate to high (30%, 17% and 14%). Reasons are not provided. Furthermore, only completers are included in the analyses
Selective reporting (reporting bias)	Low risk	Comment: All outcome measures listed in the Methods section are reported in Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Keefe 2004

Methods	**Study design:**RCT, multi-centre, four arms, non-blinded **Country in which trial was carried out:** US **Method of recruitment of participants:**Participants were recruited from rheumatology clinics and advertisements placed in newspapers **Setting:**outpatients and general population **Was the sample size justified with a priori calculation of effect size/power?** no, sample size not justified **Length of follow-up:** 12 weeks (as long as intervention takes) **Dropouts:**Two (11%) dropped out from SA-CST, one (5%) dropped out from SA-CST + Ex, two (11%) dropped out from usual care, no one dropped out from the exercise group. No reasons provided
Participants	**Inclusion criteria** • Diagnosed with knee OA by a board-certified rheumatologist • Being married **Exclusion criteria** • Comorbid medical conditions that could affect their health status over the course of the trial (e.g. a recent myocardial infarction) • An abnormal cardiac response to exercise (e.g. exercise-induced ventricular tachycardia, abnormal blood pressure response)

Keefe 2004 *(Continued)*

- Other known organic disease that would contraindicate safe participation in the study (e.g. chronic obstructive pulmonary disease, congestive heart failure, cancer)

Baseline characteristics

Participants in the spouse-assisted CST alone and spouse-assisted CST plus exercise conditions had higher pretreatment levels of pain than those in the two other conditions

Intervention group: spouse-assisted CST (N = 18 randomly assigned, 18 analysed)

Location of OA: 100% knee

PROGRESS-Plus

Place of residence: US

Sex: 50% male, 50% female

Age (mean (SD)), years: 60.0 (12.2)

Intervention group: spouse-assisted CST + exercise (N = 20 randomly assigned, 20 analysed)

Location of OA: 100% knee

PROGRESS-Plus

Place of residence: US

Sex: 35% male, 65% female

Age (mean (SD)), years: 60.2 (9.1)

Control group: exercise (N = 16 randomly assigned, 16 analysed)

Location of OA: 100% knee

PROGRESS-Plus

Place of residence: US

Sex: 62.5% male, 37.5% female

Age (mean (SD)), years: 60.3 (8.7)

Control group: standard care (N = 18 randomly assigned, 18 analysed)

Location of OA: 100% knee

PROGRESS-Plus

Place of residence: US

Sex: 38.9% male, 61.1% female

Age (mean (SD)), years: 57.6 (14.3)

Interventions	***Intervention: SA-CST*** *Description:* Couples were provided with a rationale that emphasised that (1) pain is a complex experience, which, as the Gate Control Theory suggests, can be influenced by thoughts, feelings and behaviours; (2) participants and their spouses can acquire and maintain skills for managing pain through frequent practice; and (3) because OA is a couples issue that affects each partner and his or her relationships, involving the spouse in training can be quite helpful. Training sessions emphasised active learning. Group leaders provided feedback and suggestions to enhance the efficacy of skills practice. Each session involved two major components: (1) training in pain coping skills, and (2) training in couples skills designed to supplement and reinforce the participant's pain coping skills

Keefe 2004 *(Continued)*

Intended audience: participants with knee OA and their spouses

Mode: group sessions (three to five couples)

Personnel: PhD-level psychologists who had been trained in SA-CST

Delivery method: face-to-face

Language: English

Format: standard format

Location: -

Duration: 12 once-weekly, two-hour sessions (12 weeks) (total: 24 hours)

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: no

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: yes

Health service navigation: no

Comparator: exercise

*Type: a*lternate intervention

Description: The exercise programme included (1) cardiopulmonary endurance training, (2) strength training and (3) flexibility/range of motion training

Frequency: 60-minute sessions, three times a week, for 12 weeks (total: 26 hours)

Additional treatment during trial: unclear

Comparator: standard care

Type: usual care

Description: Participants continued to receive their routine care; neither they nor their spouses attended coping skills training sessions or exercise sessions

Additional treatment during trial: unclear

Note: The SA-CST + exercise programme is a combination of two programmes described above, with a total duration of 50 hours

Outcomes	*Outcome assessed at:* baseline and post-treatment

Primary outcomes of study

- Functional performance (aerobic fitness (maximal effort bicycle ergometry exercise test))
- Functional performance (muscle strength (maximal effort leg extensions, leg flexions and bicep curls)
- Constructive attitudes and approaches (Coping Strategies Questionnaire (CSQ) subscales)

Keefe 2004 *(Continued)*

- Self-management (Arthritis Self-Efficacy Scale (ASES), 30 to 300, higher score is better)
- Marital adjustment (Dyadic Adjustment Scale (DAS))
- Pain (AIMS subscale pain, 0 to 10, lower score is better)
- Emotional distress (AIMS subscale psychological disability, 0 to 10, lower score is better)

Notes	We extracted the following outcomes at post-treatment (short term) for the analyses in this review: self-management (ASES), pain (AIMS subscale pain), emotional distress (AIMS subscale psychological disability), constructive attitudes and approaches (CSQ subscale coping attempts) and dropouts (proportion of missing participants) This research was supported by National Institute of Arthritis and Musculoskeletal Diseases Grant No. AR-35270 Data analysis: Intervention groups were judged to be similar enough to combine by pooling data

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “Patients completed a baseline evaluation and were then randomly assigned to one of four conditions” Comment: no information on method of randomisation provided
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: This type of control and intervention makes blinding of participants impossible, and this may have influenced the results Comment: This type of control and intervention makes blinding of personnel impossible; usual care probably is not influenced by performance bias, but not blinding the personnel in the other intervention groups may have introduced bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: No information was provided on blinding of outcome assessors; however, most outcomes are subjective, and participants were not blinded to group allocation, introducing a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: “Over the course of the study, five subjects dropped out of the study, two from the spouse-assisted coping skills training condition, one from the spouse-assisted coping skills training plus exercise training condition, and two from the standard care control condition” Comment: Low dropout rate, equally distributed among treatment groups. No information was provided on the analysis techniques used
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in the Methods section are reported in the Results
Other bias	High risk	Quote: “No group differences were found on any of the pre-treatment measures except for the AIMS Pain measure. (…) Patients in the spouse-assisted CST alone and spouse-assisted CST plus exercise conditions had higher pre-treatment levels of pain than those in the two other conditions” Comment: A higher baseline pain level in both intervention treatment arms could possibly enlarge the observed treatment effect (type I error)

Lorig 2008

Methods	**Study design:**RCT, single-centre, two arms, non-blinded **Country in which trial was carried out:** US **Method of recruitment of participants:**Participants were recruited by links to the study website placed on established websites, online newsletters and discussion groups. Calendar announcements and articles in newspapers also directed subjects to the study website. Potential participants could leave their e-mail address **Setting:**general population **Was the sample size justified with a priori calculation of effect size/power?**No justification of the sample size was provided **Length of follow-up:**one year **Dropouts:**126 (29%) dropped out of the intervention group; 78 (18%) dropped out of the control group. No reasons provided
Participants	**Inclusion criteria** • A self-reported diagnosis of OA, RA or fibromyalgia (confirmation requested from participants' physicians, of whom 68% replied and confirmed diagnosis in all but six cases) • Aged 18 years or older • Access to a computer with Internet and e-mail capabilities • Agreed to one to two hours per week of log-on time spread over at least three sessions per week for six weeks • Able to complete the online questionnaire **Exclusion criteria** • Being in active treatment for cancer for one year or less • Participating in the small-group ASMP or the Chronic Disease Self-Management Program **Baseline characteristics (including RA and FM participants)** Usual care control participants had an average of ˜ one chiropractic visit more in the past six months ***Intervention group: Internet-based ASMP (OA subgroup: N = 275 randomly assigned, 134 analysed)*** PROGRESS-Plus *Place of residence:* US *Race, ethnicity and culture:* non-Hispanic white 90.9% *Sex:* 10.2% male, 89.8% female *Education (mean (SD)):* 15.6 (3.09) years of education *Social capital:* 65.5% married *Age (mean (SD)), years:* 52.2 (10.9) ***Control group: usual care (OA subgroup: N = 276 randomly assigned, 158 analysed)*** PROGRESS-Plus *Place of residence:* US *Race, ethnicity and culture:* non-Hispanic white 93.7% *Sex:* 9.5% male, 90.5% female

Lorig 2008 *(Continued)*

	Education (mean (SD)): 15.7 (3.11) years of education *Social capital:* 71.1% married *Age (mean (SD)), years:* 52.5 (12.2) **Note:**The six participants for whom diagnosis was not confirmed by their physician were dropped from analysis.
Interventions	***Intervention: Internet-based ASMP*** *Description:* The Internet ASMP consists of password-protected, interactive, Web-based instruction; Web-based bulletin board discussion; tools that the participants can use individually (e.g. exercise logs, medication diaries, tailored exercise programmes); and the *Arthritis Helpbook*. The programme focused on reduction of pain and improvement of function. Topics covered were self-management principles, goal setting/action plans, pain management, relaxation/cognitive pain management, problem-solving steps, fitness/exercise, feedback/problem solving, difficult emotions, healthy eating, osteoporosis, fatigue and energy conservation, medication, depression, working with your healthcare professional, evaluating treatment plans and sleep *Intended audience:* people with OA, RA and/or fibromyalgia *Mode:* individual *Personnel:* a pair of peer moderators *Delivery method:* Internet *Language:* English *Format:* tailored to individual needs and standard format *Location:* anywhere with access to computer with Internet *Duration:* six weeks, at least three times per week for one to two hours *Additional treatment during trial:* Participants were allowed to continue usual care Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* yes *Emotional well-being:* yes *Self-monitoring and insight:* yes *Constructive attitudes and approaches:* yes *Skill and technique acquisition:* yes *Social integration and support:* no *Health service navigation:* yes ***Comparator: usual care*** *Type:* usual care *Description:* not specified *Additional treatment during trial:* unclear
Outcomes	*Outcome assessed at:* baseline and six and 12 months of follow-up

Lorig 2008 *(Continued)*

Primary outcomes of study

- Pain (10-cm VAS, 0 to 10, lower score is better)
- Function self-reported (fatigue on 10-cm VAS, 0 to 10, lower score is better)
- Function self-reported (Activities Limitation Scale, 0 to 4, lower score is better)
- Emotional distress (Health Distress Scale, 0 to 5, lower score is better)
- Function self-reported (Health Assessment Questionnaire (HAQ), 0 to 3, lower score is better)
- Global OA scores (Self-Rated Global Health Scale, 0 to 5, lower score is better)
- Health-directed activity (aerobic exercise (minutes/wk), higher is better)
- Health-directed activity (stretching and strengthening exercise (minutes/wk), higher is better)
- Health-directed activity (practice of stress management (times/wk), higher is better)
- Health service navigation (communication with physician, 0 to 5, higher score is better)
- Health service utilisation (visits to physician, emergency room, chiropractor, physical therapist and nights in hospital)
- Self-management (Arthritis Self-Efficacy Scale (ASES), 1 to 10, higher score is better)

Notes	We extracted the following outcomes at 12 months (intermediate term) for the analyses in this review: self-management (ASES), pain (VAS), global OA scores (Self-Rated Global Health Scale), function self-reported (HAQ), emotional distress (health distress scale), health-directed activity (practice of stress management), health service navigation (communication with physician) and dropouts (proportion of missing participants) This study was supported by NIH Center for Disease Control's Arthritis Grant (AR-43538) Control participants were sent a $10 Amazon.com certificate for each questionnaire completed Data analysis: Change scores were combined with end point scores using generic inverse variance

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Following completion of the online questionnaire, participants were randomized to either the intervention group or to a control group" Comment: no information on the method of randomisation provided
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on the allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants were not blinded during the trial, which introduces a risk of performance bias. Although personnel also were not blinded, it is unlikely that this has biased the results because the physicians providing usual care were not involved in the trial
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: Participants self-reported the outcomes and were not blinded from their allocated treatment
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "All analyses were also run substituting last-reported value for any missing cases (intent-to-treat [ITT] analyses)" Comment: A large number of participants dropped out with no reasons provided. Numbers of dropouts differed between groups (29% in the intervention group vs 18% in the control group). Also, non-completers at six and 12 months differed significantly from completers (non-completers e.g. were younger and had significantly higher levels of health distress and activity limitation at baseline)

Lorig 2008 *(Continued)*

Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in the Methods section were reported in the Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Maisiak 1996

Methods	**Study design:** RCT, single-centre, three arms, outcome assessment blinded **Country in which trial was carried out:** US **Method of recruitment of participants:** Recruitment over a two-year period from lists of previous callers to the Arthritis Information Service (AIS) of Alabama and subscribers to an arthritis newsletter, and from newspaper advertisements **Setting:** primary care and outpatients **Sample size justified with a priori calculation of effect size/power?** yes **Length of follow-up:** nine months **Dropouts:** After nine months of follow-up, eight (6%) dropped out from usual care, 11 (8%) dropped out from symptom monitoring and seven (7%) dropped out from treatment counselling (11 of 26 were in the OA subgroup)
Participants	**Inclusion criteria** • A diagnosis of primary OA of the hip or knee or a diagnosis of primary RA made by the patient's physician (letter documenting diagnosis and physician's confidence in that diagnosis (no ACR criteria had to be provided) • Some current pain or some current disability due to arthritis • ≥ 21 years of age • Able to communicate by telephone over a nine-month period • Residing in Alabama **Exclusion criteria** • Serious (non-rheumatologic) comorbidities that might affect study participation **Baseline characteristics (include RA patients ~ 55%)** Baseline characteristics were similar in all treatment groups ***Intervention group: treatment counselling (N = 135 randomly assigned, 128 included in analysis, including RA patients; 62 in OA subgroup)*** *Mean disease duration:* 16.5 years (including RA patients) PROGRESS-Plus *Place of residence:* Alabama *Race, ethnicity and culture:* 87% white *Sex:* 4% male, 96% female *Education:* 12.1 mean years of schooling (including RA patients) *Age (mean), years:* 60.1 (SD not specified) *Disability:* 43% see a specialist (including RA patients)

Maisiak 1996 *(Continued)*

Control group: symptom monitoring (N = 135 randomly assigned, 124 included in analysis, including RA patients; 54 in OA subgroup)

Mean disease duration: 15.9 years (including RA patients)

PROGRESS-Plus

Place of residence: Alabama

Race, ethnicity and culture: 85% white

Sex: 6% male, 94% female

Education: 12.0 mean years of schooling (including RA patients)

Age (mean), years: 60.7

Disability: 51% see specialist (including RA patients)

Control group: usual care (N = 135 randomly assigned, 127 included in analysis, including RA patients; 54 in OA subgroup)

Mean disease duration: 15.7 years (including RA patients)

PROGRESS-Plus

Place of residence: Alabama

Race, ethnicity and culture: 82% white

Sex: 13% male, 87% female

Education: 12.2 mean years of schooling (including RA patients)

Age (mean), years: 60.5

Disability: 50% see specialist (including RA patients)

Interventions	**Intervention group: treatment counselling** *Description:* Six categories of patient behaviour were targeted for potential change, including patient-physician communication, medication compliance, removing barriers to medical care, symptom reviews, self-care activities and stress control using a 13-page written, structured protocol especially designed for counselling patients *Intended audience:* people with OA or RA *Mode:* individual *Personnel:* trained counsellors *Delivery method:* telephone and mail *Language:* English *Format:* structured protocol with options for tailoring to individual needs *Location:* home *Duration:* 20 minutes per session, five times at two-week intervals during first three months, and six more times at four-week intervals during the last six months *Additional treatment during trial:* unclear Were the following heiQ components addressed? *Health-directed behaviour:* yes

Maisiak 1996 *(Continued)*

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: no

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: yes

Control group: symptom monitoring

Type: attention control

Description: The purpose was to provide a detailed review of the participant's symptoms and to provide attention to the participant in an amount equal to that provided to participants in the treatment counselling group. During each session, the specialist would administer questions regarding symptom assessment from the AIMS2. The symptom monitoring specialist was not allowed to ask any other questions concerning arthritis and did not provide any advice to the participant

Frequency: 20 minutes per session, five times at two-week intervals during first three months, and six more times at four-week intervals during the last six months

Additional treatment during trial: unclear

Other relevant information: The symptom monitoring specialists were temporary, part-time staff who tended to be college students with little or no training in arthritis education or counselling. Each received two hours of training on administering the AIMS2 questions by phone

Control group: usual care

Type: usual care

Description: Participants in the usual care control group were not contacted by the study staff outside of their three assessments (baseline, six and nine months)

Additional treatment during trial: no restriction in using any other outside sources of assistance, including the Arthritis Information Service in Alabama

Outcomes	*Outcomes assessed at:* baseline and three, six and nine months of follow-up **Primary outcome** • Global OA scores (AIMS-2 total health status, 0 to 10, lower score is better) **Secondary outcome** • Number of visits to physician
Notes	We extracted the following outcome at nine months (intermediate term) for the analyses in this review: global OA scores (AIMS-2 total health status) and dropouts (proportion of missing participants) Supported by a Multipurpose Arthritis and Musculoskeletal Diseases Center grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases Author (R Maisiak) sent additional information about the trial on request

Risk of bias

Bias	Authors' judgement	Support for judgement

Maisiak 1996 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote: "The random sequence was generated by a standard random number software program on a computer" [additional information provided by author] Comment: This method of randomisation has a low risk of introducing bias
Allocation concealment (selection bias)	Unclear risk	Quote: "The sequence of assignments was printed on paper and stored in a locked cabinet" [additional information provided by author] Comment: Allocation seems appropriately concealed; however, it is not clear whether allocation was concealed for the telephone interviewers who assigned participants to treatment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants were not blinded to treatment allocation, which might have introduced bias Comment: Personnel were not blinded to treatment allocation, which might have introduced bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The assessment interviewers were mostly college students who were not told of the overall purpose of the study, and were blinded to the group assignment of the patients they were interviewing" Comment: Outcomes were assessed by blinded personnel; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: Although no intent-to-treat analysis seems to be performed, the quantity of missing data is very low (6%) for the OA group, and reasons for dropout are unlikely to have biased the results significantly. It is unclear how missing data were divided between treatment groups
Selective reporting (reporting bias)	Unclear risk	Quote: "Other measures were also included in the assessment and were unrelated to this report" Comment: Several other measures were assessed but not reported. It is unclear what these measures were
Other bias	Low risk	Comment: No other potential sources of bias were identified

Martire 2007

Methods	**Study design:**RCT, single-centre, three arms, non-blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**not specified **Setting:**outpatients **Was the sample size justified with a priori calculation of effect size/power?**no justification of sample size provided **Length of follow-up:**six months **Dropouts:**21 (39%) dropped out of usual care, 12 (13%) dropped out of PES and 17 (17%) dropped out of CES. Reasons for dropout in the usual care group were dissatisfaction with assignment to this group (six), health problems in partner or family members (four) and unknown (11).
Participants	**Inclusion criteria**

Martire 2007 *(Continued)*

- Diagnosed with hip or knee OA
- 50 years of age or older
- Married
- Pain of at least moderate intensity on most days over the past month
- Difficulty with at least one instrumental activity of daily living (e.g. household tasks, driving)
- Received assistance from the spouse with at least one instrumental activity of daily living

Exclusion criteria

- Attended the Arthritis Self-Management Program in the past five years
- Had a comorbid diagnosis of fibromyalgia or rheumatoid arthritis

Baseline characteristics

Participants in the CES group had more depressive symptoms than participants in the PES group

Intervention group: PES (N = 89 randomly assigned, 89 analysed)

Location of OA: hip and knee

Duration of OA (mean (SD)): 15.3 (11.8) years

PROGRESS-Plus

Place of residence: US

Sex: 28% male, 72% female

Education (mean (SD)): 14.6 (1.7) years of education

Age (mean (SD)), years: 68.0 (8.0)

Intervention group: CES (N = 99 randomly assigned, 99 analysed)

Location of OA: hip and knee

Duration of OA (mean (SD)): 14.3 (9.4) years

PROGRESS-Plus

Place of residence: US

Sex: 26% male, 74% female

Age (mean (SD)), years: 69.2 (7.2)

Control group (N = 54 randomly assigned, 54 analysed)

Location of OA: hip and knee

Duration of OA (mean (SD)): 16.1 (12.0) years

PROGRESS-Plus

Place of residence: US

Sex: 28% male, 72% female

Age (mean (SD)), years: 68.4 (7.5)

Interventions	***Intervention: PES*** *Description:* The protocol was based on the Arthritis Self-Management Program (Lorig). No spouses, family members, or friends participated in PES sessions. Major components were information about arthritis, self-management strategies, benefits of exercise, communication skills and effectively coping

Martire 2007 *(Continued)*

with negative emotions. At the end of the session, each participant set a health-related goal; the next session, each participant reported on his or her success in meeting this goal and received feedback from the group. Up to five monthly booster sessions were conducted via telephone in the six months of follow-up after treatment

Intended audience: people with hip or knee OA

Mode: group sessions (four to six participants)

Personnel: trained facilitator

Delivery method: face-to-face

Language: English

Format: tailored to individual needs and standard format

Location: -

Duration: weekly sessions of two hours for six weeks

Additional treatment during trial: Participants also received usual care

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: yes

Emotional well-being: yes

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Intervention: CES

Description: same as PES; topics were framed as couple's issues whenever possible. Up to five monthly booster sessions were conducted via telephone in the six months of follow-up after treatment

Intended audience: people with hip and knee OA

Mode: group sessions (four to six participants and their spouses)

Personnel: trained facilitator

Delivery method: face-to-face

Language: English

Format: tailored to individual needs and standard format

Location: -

Duration: once a week, session of two hours, six weeks

Additional treatment during trial: Participants also received usual care

Were the following heiQ components addressed?

Health-directed behaviour: yes

Martire 2007 *(Continued)*

Positive and active engagement in life: yes

Emotional well-being: yes

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: yes

Health service navigation: no

Comparator

Type: usual care

Description: Individuals in this group maintained their regular medical regimens and saw their rheumatologist as needed, and no attempt was made to alter the usual care that these individuals received from their physicians

Additional treatment during trial: Couples in this group did not participate in any arthritis self-management programme or receive any surgical interventions during their study participation

Outcomes

Outcome assessed at: baseline, post-treatment and six months of follow-up

Primary outcomes of study

- Global OA scores (WOMAC total score 0 to 96, lower score is better)
- Pain (WOMAC subscale pain, 0 to 20, lower score is better)
- Function self-reported (WOMAC subscale physical function, 0 to 68, lower score is better)
- Emotional distress (CES-D, 0 to 60, lower score is better)
- Self-management (Arthritis Self-Efficacy Scale (ASES), 30 to 300, higher score is better)
- Marital satisfaction (marital adjustment test)

Secondary outcomes of study

- Measures for spouses: perceived stress, depressive symptoms, caregiver mastery, critical attitudes and marital satisfaction

Notes

We extracted the following outcomes at post-treatment (short term) and at six months (intermediate term) for the analyses in this review: self-management (ASES), pain (WOMAC subscale pain), global OA scores (WOMAC), function self-reported (WOMAC subscale function) and dropouts (proportion of missing participants)

The trial was supported by several grants (Grant P50 HL65111-65112; (Pittsburgh Mind-Body Center) from the National Heart, Lung, and Blood institute; Grants K01 MH065547 and P30 MH52247 from the National institute of Mental Health; Grant K07 AG000923 from the National Institute on Aging; Grant R01 NR008272 from the National Institute of Nursing Research; Grant P01 AR50245 from the National Institute of Arthritis and Musculoskeletal Diseases; and Grant to F.J. Keefe from the Arthritis Foundation)

The author (L Martire) provided additional information about the trial on request

Data analysis: Intervention groups were judged to be similar enough to combine by pooling data

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "The first group of accumulated couples was randomly assigned to conditions 1, 2, or 3 using a lottery method. The next group of couples at that

Martire 2007 *(Continued)*

		site was randomly assigned to the two remaining two study conditions, and so on" [information provided by the author] Comment: Method of random sequence generation has high risk of bias
Allocation concealment (selection bias)	Low risk	Quote: "This information was kept in a locked drawer of my office" [information provided by the author] Comment: Allocation was concealed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: This type of intervention makes blinding of participants and personnel to treatment allocation not possible, introducing a risk of bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "All assessments were carried out by interviewers who were not involved in intervention implementation" Comment: Blinding of outcome assessors remains unclear; however, most outcomes are subjective, and participants are not blinded to group allocation, introducing a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "There was a different rate of attrition across the three groups (…). The reasons for attrition from the usual care group were dissatisfaction with assignment to this group (n=6) (…)." Comment: A significantly different rate of attrition among treatment groups is noted. Dropout rate is acceptable in both intervention groups (13% in PES, 17% in CES) but very high in the usual care group (39%). An intention-to-treat approach is used to analyse data
Selective reporting (reporting bias)	Unclear risk	Comment: Emotional distress and marital satisfaction were not reported
Other bias	High risk	Quote: "Post hoc analyses (…) indicated that individuals with OA in the CES group had more pre-intervention depressive symptoms than those in the PES group. Thus, the effects of intervention were not examined for this outcome (…)" Comment: Difference in baseline depressive symptoms may have introduced bias Comment: Compliance was equal in intervention groups but not very high (24% in PES and 28% in CES did not receive intervention; 33% in both PES and CES attended all sessions)

Maurer 1999

Methods	**Study design:** RCT (stratified by disease severity), multi-centre, two arms, outcome assessment blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**Several methods were used to recruit participants; most patients were involved from direct physician referrals from with study-affiliated clinics **Setting:**outpatients **Was the sample size justified with a priori calculation of effect size/power?** yes

Maurer 1999 *(Continued)*

Length of follow-up:four weeks

Dropouts: Seven (13%) dropped out from the education group; eight (14%) dropped out from the exercise group. No reasons were provided

Participants	**Inclusion criteria** • Knee OA according to the current ACR criteria (1999) • A score of 1, 2 or 3 on the Kellgren radiographic scale • At least one osteophyte other than at the tibial spine (with severity below Kellgren grade 4) • Between 50 and 80 years of age • No drugs for their arthritis other than stable doses of analgesics or NSAIDs • Mild to moderate knee pain for at least the previous three months **Exclusion criteria** • Concurrently receiving physical therapy • Actively involved in any other pharmaceutical or exercise study • Had undergone isokinetic strength training within the previous three years • Had significant cardiovascular disease • More than mild knee swelling • Large popliteal cysts • Knee instability • Major knee or hip surgery on the side to be treated • Systemic disease other than OA that might affect muscle function • Severe osteopenia • A history of fracture in the area of the joint to be treated • Paresis of the lower extremity **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: education group (N = 56 randomly assigned, 49 analysed)*** *Location of OA:* 100% knee *Body weight (mean (SD)):* 190.4 (35.0) lb *Duration of OA (mean (SD)):* 13.1 (11.7) years PROGRESS-Plus *Place of residence:* US *Sex:* 64% male, 36% female *Age (mean (SD)), years:* 64.5 (8.4) ***Control group: exercise group (N = 57 randomly assigned, 49 analysed)*** *Location of OA:* 100% knee *Body weight (mean (SD)):*183.8 (32.8) lb *Duration of OA (mean (SD)):* 9.7 (9.0) PROGRESS-Plus *Place of residence:* US *Sex:* 53% male, 47% female

Maurer 1999 *(Continued)*

	Age (mean (SD)), years: 66.3 (8.8)
Interventions	***Intervention: education group***
	Description: All participants were provided with educational information about OA. Several Arthritis Foundation pamphlets were distributed, and four lectures and discussions were conducted. The programme consisted of (1) a lecture by a rheumatologist on the disease process of OA and its clinical characteristics, (2) a video discussing joint protection and OA self-management techniques, (3) a session on nutrition guidelines and a guide to relevant community services (provided by a dietician and a social worker, respectively) and (4) a discussion on coping with pain and disability led by a psychologist
	Intended audience: people with knee OA
	Mode: group sessions
	Personnel: healthcare professionals
	Delivery method: face-to-face
	Language: English
	Format: standard format
	Location: -
	Duration: four sessions in eight weeks
	Additional treatment during trial: unclear
	Were the following heiQ components addressed?
	Health-directed behaviour: yes
	Positive and active engagement in life: no
	Emotional well-being: no
	Self-monitoring and insight: yes
	Constructive attitudes and approaches: no
	Skill and technique acquisition: yes
	Social integration and support: no
	Health service navigation: no
	Comparator: exercise group
	Type: alternate intervention
	Description: Participants underwent strength training of the knee extensor muscle groups unilaterally with the dynamometer three times a week for eight weeks. During the exercise sessions, a total of 27 repetitions were performed as three sets of three extensions at each of the following angular velocities: 90 gr/s, 120 gr/s and 150 gr/s. Between each set of three repetitions, the velocity was not adjusted while the participant rested for one minute
	Frequency: three times a week for eight weeks
	Additional treatment during trial: unclear
Outcomes	*Outcome assessed at:* baseline, post-treatment and four weeks of follow-up
	Primary outcomes of study

Maurer 1999 *(Continued)*

- AIMS-2 subscales (including mobility, walk and bend, extremity function, household and self-care, bodily pain, social and psychological factors)
- SF-36 MOS total score and subscales (including pain, physical functioning, role-physical, role-emotional, general health, vitality, social functioning, mental health)
- Global OA score (WOMAC VAS, 0 to 240, lower score is better)
- Pain (WOMAC VAS, 0 to 50, lower score is better)
- Function self-reported (WOMAC VAS, 0 to 170, lower score is better)
- Pain during 50 feet walking with moderate pace (10-point scale, 0 to 10, lower score is better)
- Pain during 50 feet walking with maximum pace (10-point scale, 0 to 10, lower score is better)
- Pain during ascending and descending half a flight of stairs (10-point scale, 0 to 10, lower score is better)
- How much does your knee limit your ability to perform your daily activities (open-ended question)
- Treatments before and at time of investigation
- Clinical findings: disease duration, pattern of involved joints, duration of morning stiffness (as assessed by physician)
- Knee range of motion (goniometer)
- Heel-to-buttocks distance at maximal knee flexion
- Joint effusion (palpation, yes or no)
- Isokinetic peak torque at 90 degrees/s and 120 degrees/s (dynamometer)
- Isometric strength (dynamometer)
- Participant evaluation based on the extent to which the participant's pain limited his or her activity level
- Physician evaluation based on the extent to which the participant's pain limited his or her activity level (according to the physician)

Notes	We extracted the following outcomes at post-treatment (short term) and at six months (intermediate term) for the analyses in this review: global OA scores (WOMAC), pain (WOMAC subscale pain), function self-reported (WOMAC subscale function) and dropouts (proportion of missing participants) This trial was supported by a Health Services Research and Development grant from the Department of Veterans Affairs The author (R Schumacher) provided additional information about the trial on request Data analysis: As this study provided no information on variance, we used the SD from WOMAC total score, WOMAC subscale pain and WOMAC subscale function from an observational trial that was judged to have a population similar enough to the population of this study (Wolfe 1999)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Patients were assigned by a random number generator to either the exercise group (…) or to the education group (…)” Comment: This method of randomisation has a low risk of selection bias
Allocation concealment (selection bias)	Unclear risk	Comment: No information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: In this type of intervention and control, participants and personnel cannot be blinded from treatment allocation, possibly introducing a risk of bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: Self-reported outcomes were assessed with the use of validated questionnaires. Other outcomes were assessed by blinded personnel; howev-

Maurer 1999 *(Continued)*

		er, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Of the 113 who enrolled, 15 either dropped out or were excluded (…) In none of these cases termination believed to be caused by a study-related event; (…)" Comment: The numbers of dropouts are equal in the two groups, and the rate of dropout is not too high (13% and 14%). Reasons for dropout remain unclear. No information is provided on the type of analysis performed
Selective reporting (reporting bias)	Low risk	All important outcomes for this review were reported
Other bias	Low risk	Comment: No other potential sources of bias were identified

Mazzuca 1997

Methods	**Study design:**RCT, single-centre, two arms, non-blinded **Country in which trial was carried out:**US **Method of recruitment of participants:**Participants were identified by a computerised medical record system containing all outpatient information for the host facility **Setting:**primary care **Was the sample size justified with a priori calculation of effect size/power?**no justification provided **Length of follow-up:** 12 months **Dropouts:**19 (18%) dropped out of the intervention group, and 21 (20%) dropped out of the control group. Reasons for dropout were moving without forwarding address (18), inconvenience (15), unrelated illness (five) and death (two)
Participants	**Inclusion criteria** • Radiographically confirmed OA of the knee (knee radiograph reported as showing "OA" or "osteophyte") • With a clinic visit in the past year • A pharmacy record indicating a current prescription of a salicylate or other NSAID or a pure analgesic (acetaminophen or opioid analgesic) • Participants needed to be accessible by telephone • Participants needed to score seven (of a possible 10) on the Mini-Mental State Examination **Exclusion criteria** • Concurrent systemic inflammatory rheumatic disease • Medical comorbidity that would render the participant unable to participate fully in study procedures (e.g. terminal conditions such as chronic obstructive pulmonary disease, end-stage renal disease, heart failure, or malignancy with anticipated life expectancy less than one year) • Alcohol abuse or a psychiatric disorder • Previous or planned knee arthroplasty • Treatment by a rheumatologist in the past year **Baseline characteristics** Baseline characteristics were similar among all treatment groups

Mazzuca 1997 *(Continued)*

Intervention group: self-care education (N = 105, 82 analysed)

Location of OA: 100% knee

Duration of OA (mean (SD)): 14.6 (17.9) years

PROGRESS-Plus

Place of residence: inner city, US

Race, ethnicity and culture: 71% African American

Occupation: 50% unemployed, 19% employed, 30% retired

Sex: 15% male, 85% female

Education (mean (SD)): 9.7 (2.6) years of education

Socioeconomic status: 97% annual income ≤ $20,000

Social capital: 75% living alone

Age (mean (SD)), years: 62.8 (12.2)

Disability (mean (SD)): 1.5 (1.0) comorbid conditions

Control group: attention control (N = 106 randomly assigned, 83 analysed)

Location of OA: 100% knee

Duration of OA (mean (SD)): 13.5 (13.7) years

PROGRESS-Plus

Place of residence: inner city, US

Race, ethnicity and culture: 67% African American

Occupation: 56% unemployed, 11% employed, 33% retired

Sex: 15% male, 85% female

Education (mean (SD)): 9.7 (3.5) years

Socioeconomic status: 96% annual income ≤ $20.000

Social capital: 72% living alone

Age (mean (SD)), years: 62.0 (11.0)

Disability (mean (SD)): 1.7 (1.1) comorbid conditions

Interventions	***Intervention: self-care education*** *Description:* Participants received a course of individualised arthritis self-care instructions based on needs demonstrated in the diagnostic assessment and through preliminary communication with the primary care physician. Core content areas included quadriceps strengthening exercises, control of joint pain with thermal modalities and joint protection. Participants for whom NSAIDs or salicylates were prescribed were instructed on the importance of taking their medication with food or milk. Toward the end of the teaching session, all participants were encouraged to identify the vocational and avocational roles that were most threatened by their knee OA. A brief problem-solving exercise then occurred. At the end of the session, each participant was given a copy of the Arthritis Foundation Information series pamphlet, "Osteoarthritis", and a printed set of instructions for isometric quadriceps exercises. Participants were contacted by phone at one week and one month after initial instruction. The phone contacts were unscripted but were structured to ensure that (1) compliance with self-care rec-

Mazzuca 1997 *(Continued)*

ommendations was assessed and reinforced, as appropriate, (2) misconceptions were clarified and (3) continued participation in the study was encouraged

Intended audience: people with knee OA

Mode: individual

Personnel: experienced arthritis nurse educator under the supervision of a rheumatologist

Delivery method: face-to-face and by telephone

Language: English

Format: tailored to individual needs

Location: general practitioner clinic

Duration: one session of 30 to 60 minutes and thereafter a telephone call (five to 10 minutes) at one week and at one month after initial instruction

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: no

Constructive attitudes and approaches: no

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator: attention control

Type: attention control

Description: All participants viewed the synchronised slide-tape presentation “Arthritis, it’s not just growing old”. This 20-minute audiovisual presentation was designed for the general public, with the primary purposes of defining common types of arthritis in adults and encouraging those who think they may have arthritis to seek medical care. Participants also were given a current issue of the "IU-MAMDC Newsletter". A brief follow-up telephone contact was made one week and one month after attention-placebo treatment but only for the purpose of reinforcing continued participation in the study

Frequency: once at 20 minutes, followed by two phone calls, each of five to 10 minutes

Additional treatment during trial: unclear

Outcomes

Outcome assessed at: baseline and four, eight and 12 months of follow-up

Primary outcomes of study

- Function self-reported (HAQ disability, 0 to 3, lower score is better)
- Pain (Health Assessment Questionnaire (HAQ) subscale pain on VAS, 0 to 10, lower score is better)
- Pain walking and at rest (VAS, 0 to 10, lower score is better)
- Quality of life (Quality of Well-Being Scale (QWB), 0 to 1, lower score is better)

Secondary outcomes of study

Mazzuca 1997 *(Continued)*

	• Healthcare utilisation and costs
Notes	We extracted the following outcomes at 12 months (intermediate term) for the analyses in this review: pain (HAQ subscale pain), function self-reported (HAQ subscale disability), quality of life (QWB) and dropouts (proportion of missing participants) The trial was supported in part by a grant from NIAMS (AR-20582)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "If contact occurred during a time when the nurse educator was available, the subject was enrolled in the E-group and received self-care education. Subjects recruited at other times were assigned to the AC condition" Comment: This method of random sequence generation has a high risk of bias
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants were not blinded, which may have introduced bias. Although personnel also were not blinded, risk of bias remains low because participants in the control group seem to have no contact with the study personnel
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: Outcomes were assessed through self-report; therefore outcome assessment was not blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "(…) omission from the analyses of any subjects who did not complete all assessments (i.e., those who were lost to followup or who missed interim assessments)." Quote: "Specifically, the 23 subjects in group AC with incomplete followup data tended to have higher baseline HAQ disability scores (i.e., they had poorer function) than the 23 dropouts from group E" Comment: No intention-to-treat analysis was performed. Although numbers of dropouts were equal in the two groups (18% vs 20%), differences in characteristics of dropouts were evident between these groups, which may have biased results
Selective reporting (reporting bias)	Low risk	Comment: Outcomes listed in the Methods were reported in the Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Mazzuca 2004

Methods	**Study design:**cluster-RCT, multi-centre, two arms, outcome assessment blinded **Country in which trial was carried out:** US **Method of recruitment of participants:**screening of the computer records of a large HMO **Setting:**primary care **Was the sample size justified with a priori calculation of effect size/power?**no justification provided

Mazzuca 2004 *(Continued)*

Length of follow-up:12 months

Dropouts:31 (28%) dropped out from the intervention group, and 21 (28%) dropped out from the control group. No reasons provided

Participants	**Inclusion criteria** • A clinical diagnosis of knee OA in the participant's medical record that satisfied the ACR clinical criteria for knee OA • Under treatment for chronic knee pain by primary care physician **Exclusion criteria:**none **Baseline characteristics** Women constituted a marginally larger proportion of participants in group E than in group C (78% vs 65%) ***Intervention group: nurse-directed intervention (N = 111 randomly assigned, 80 analysed)*** *Location of OA:* 100% knee PROGRESS-Plus *Place of residence:* US *Race, ethnicity and culture:* 72% white *Sex:* 22% male (n = 24), 78% female (n = 87) *Education:* 86% 12 or more years *Social capital:* 64% married *Age (mean (SD)), years:* 61.8 (12.5) ***Control group: waiting list (N = 75 randomly assigned, 54 analysed)*** *Location of OA:* 100% knee PROGRESS-Plus *Place of residence:* US *Race, ethnicity and culture:* 69% white *Sex:* 35% male (n = 26), 65% female (n = 49) *Education:* 89% 12 or more years *Social capital:* 60% married *Age (mean (SD)), years:* 61.8 (11.9)
Interventions	***Intervention: nurse-directed intervention*** *Description:* Care for knee OA in group E was guided by an algorithm designed to achieve the following two goals: (1) first-line management with non-pharmacological treatment modalities and acetaminophen; and (2) pharmacological management of knee pain according to a stepped protocol designed to minimise the risk of NSAID therapy. The algorithm prescribed a set of non-pharmacological measures, including isometric quadriceps exercises, application of heat or cold, instruction in behavioural principles of joint protection (including weight loss), use of shoes with well-cushioned soles and a cane or other assistive device for ambulation (if necessary). Participants were advised to supplement these measures with acetaminophen (1 g up to three to four times daily) for relief of knee pain

Mazzuca 2004 *(Continued)*

After 30 to 60 minutes of self-care instructions, participants were given ≥ two weeks to adopt these measures and evaluate their effects, at which point they were contacted by telephone by the nurse to assess their progress. Participants who reported a satisfactory response had their self-care efforts reinforced; continued progress was followed biweekly for the remainder of the study period

The algorithm for participants with newly diagnosed knee OA (18 weeks) differed from the algorithm for participants who already took an NSAID (10 weeks)

Intended audience: people with OA of the knee

Mode: individual

Personnel: arthritis nurse educator, with concurrence of the primary care physician

Delivery method: face-to-face and telephone

Language: English

Format: standard format

Location: GP clinic and at home

Duration: one session of 30 to 60 minutes; after two weeks biweekly telephone calls (five to 10 minutes) for 10 or 18 weeks

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: no

Constructive attitudes and approaches: no

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator

Type: waiting list (usual care)

Description: Participants in the control group (delayed intervention) received routine OA care. In addition, their primary care physicians received the care algorithm as part of the HMO's in-service education programme. No resource or consultation routinely available to participants with OA within the HMO and deemed appropriate by the primary care physician was denied to participants in the control group

Additional treatment during trial: unclear

Outcomes	*Outcome assessed at:* baseline and three, six and 12 months of follow-up **Primary outcomes of study** • Pain (WOMAC subscale pain, 5 to 25, lower score is better) • Function self-reported (WOMAC subscale function, 17 to 85, lower score is better)

Mazzuca 2004 *(Continued)*

Notes	We extracted the following outcomes at 12 months (intermediate term) for the analyses in this review: pain (WOMAC subscale pain), function self-reported (WOMAC subscale function) and dropouts (proportion of missing participants) This study was supported in part by a grant from the National Institutes of Health (P60 AR20582) Data analysis: For all outcomes, standard errors were converted into SD

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: “One hundred eleven patients were seen at HMO sites randomized to group E; 75 received care at sites randomize to group C” Comment: no information provided on the method of random sequence generation
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on the concealment of allocation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants could not be blinded, with the possible risk of performance bias. The primary physicians also were not blinded to treatment allocation, thereby introducing a risk of performance bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: “Study personnel who were blind to group assignment conducted the interviews” Comment: Outcome assessors were blinded to group assignment; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: “Subjects who discontinued the interview schedule prematurely did not differ from those completing the study with respect to sex, age, ethnicity, education level, marital status or WOMAC scores” Comment: Although dropouts did not differ from completers, the number of participants dropping out is high (28% in both groups), and it is unclear whether an intention-to-treat analysis is performed
Selective reporting (reporting bias)	Low risk	Comment: Outcomes listed in the Methods section were reported in the Results
Other bias	High risk	Quote: “In addition, their PCP received the care algorithm as part of the HMO’s in-service education program” Comment: The primary care physicians who provided the usual care to the control group participants also received the intervention algorithm, possibly introducing contamination Quote: "Because the HMO site was the unit of randomization and the 2 nurses worked at different care sites, patient and site were included in the models as random effects" Comment: The cluster design was accounted for in the statistical analysis

McKnight 2010

Methods	**Study design:**RCT (stratified by sex), single-centre, three arms, non-blinded. **Country in which trial was carried out:** US **Method of recruitment of participants:** recruitment from the local community by mass mailings, television/newspaper advertisements and flyers **Setting:** general population **Was the sample size justified with a priori calculation of effect size/power?** yes **Length of follow-up:** 24 months **Dropouts:**27 (30%) dropped out from the strength training group (21 discontinued because of no interest, personal reasons, knee replacement, time commitment or other reasons; six were lost-to-follow-up), 20 (23%) dropped out from the self-management group (10 discontinued because of no compliance, time commitment, inflammatory arthritis; 10 were lost-to-follow-up), 25 (26%) dropped out from the combination group (20 discontinued because of health problems, no interest, time commitment, no compliance and other reasons; five were lost-to-follow-up)
Participants	**Inclusion criteria** • Kellgren/Lawrence classification grade 2 radiographic evidence of knee OA in one or both knees • Age 35 to 64 years • History of pain on most days (i.e. four or more days in a week) in one or both knees for at least four months during the year before study entry • Duration of symptoms (defined as pain on most days for at least four months in one year) of less than five years • Some level of disability due to knee pain for at least three of the following items: descending or ascending stairs, walking, kneeling or performing daily activities **Exclusion criteria** • Any uncontrolled medical condition that could prevent safe participation in the study (e.g. uncontrolled heart disease, blood pressure or respiratory condition) • Any neurological condition that could affect coordination • Inflammatory arthritis (e.g. rheumatoid or psoriatic arthritis) • Participates in vigorous (e.g. exercise, walking, household chores) physical activity for longer than 120 minutes per week • Participates in any form of resistance training • History of knee surgery • Kellgren and Lawrence grade 3 or 4 radiographic evidence of OA in one or both knees • Body mass index greater than 37.5 kg/m^2 • History of a knee corticosteroid injection in the three months before study entry • Plans to move from the local area • Plans to become pregnant during the study **Baseline characteristics** Baseline characteristics were similar in all treatment groups ***Intervention group: self-management group (N = 87 randomly assigned, 87 analysed)*** *Location of OA:* 100% knee *BMI (mean (SD)):* 27.9 (4.1) kg/m^2 PROGRESS-Plus *Place of residence:* US

McKnight 2010 *(Continued)*

Race, ethnicity and culture: 96.3% white

Sex: 26.3% male, 74.7% female

Education: 55.9% college educated

Age (mean (SD)), years: 52.6 (6.5)

Intervention group: combination group (N = 95 randomly assigned, 95 analysed)

Location of OA: 100% knee

BMI (mean (SD)): 27.4 (4.1) kg/m^2

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 86.3% white

Sex: 24.0% male, 76.0% female

Education: 59.1% college educated

Age (mean (SD)), years: 51.9 (7.7)

Control group: strength training (N = 91 randomly assigned, 91 analysed)

Location of OA: 100% knee

BMI (mean (SD)): 27.9 (4.5) kg/m^2

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 92.6% white

Sex: 19.8% male, 80.2% female

Education: 74.1% college educated

Age (mean (SD)), years: 53.3 (7.2)

Interventions	***Intervention: self-management*** *Description:* The two-phase self-management intervention targeted coping and self-efficacy skills. Phase 1 (nine months) consisted of classroom sessions. These were followed by telephone calls designed to boost knowledge and behaviours from classroom sessions, as well as to provide practical, one-on-one problem-solving discussions to tailor the treatment to each participant's needs. The telephone sessions continued into phase 2, when they were staggered. Coping skills focused on promoting increased adaptive strategies and reducing avoidant or passive strategies. Self-efficacy skills focused on increasing perceptions of control for physical functioning, pain management and other ancillary arthritis symptoms. Staff taught self-management skills using educational and behavioural methods, including homework assignments and active involvement/practice during treatment sessions. *Intended audience:* People with knee OA *Mode:* both group sessions and individual phone calls *Personnel:* programme manager and local health professionals *Delivery method:* face-to-face and telephone *Language:* English

McKnight 2010 *(Continued)*

Format: tailored to individual needs

Location: 'classroom'

Duration: Phase 1 consisted of nine months of 12 weekly sessions, each of 90 minutes (60% didactic, 40% interactive) and (once-)weekly telephone calls; phase 2 consisted of telephone calls staggering to biweekly, monthly and then bimonthly

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Intervention: combination

Description: The combined group concurrently participated in both strength training and self-management courses, with slight alterations to ensure equivalence of contact time across the treatment groups. Specifically, the staff contacted participants in the combined treatment group less often than participants in the strength training and self-management programmes during phase 2. Otherwise, the combined group participated in the full, independent treatment protocols for both strength training and self-management programmes

Intended audience: people with knee OA

Mode: both group sessions and individual phone calls

Personnel: programme manager, local health professionals and physiotherapists

Delivery method: face-to-face and telephone

Language: English

Format: tailored to individual needs and standard format

Location: 'classroom'

Duration: Phase 1 consisted of nine months of 12 weekly sessions, each of 90 minutes (60% didactic, 40% interactive) and (once-)weekly telephone calls for the SMP-programme and three sessions per week, each of 60 minutes, for strength training; phase 2 consisted of telephone calls staggering to bi-weekly, monthly and then bimonthly for the SMP-programme and strength training sessions every two weeks in the first six weeks and then monthly for a total of 15 months

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

McKnight 2010 *(Continued)*

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator: strength training

Type: alternate intervention

Description: Strength training consisted of two phases. Phase 1 (nine months) consisted of sessions supervised by expert physical trainers, targeted to improve stretching and balance, range of motion and flexibility, as well as isotonic muscle strengthening. Phase 2 (15 months) focused on developing self-directed long-term exercising habits

Frequency: three sessions per week for nine months (phase 1), then every two weeks in first six weeks and then monthly for a total of 15 months (phase 2). Each session took 60 minutes

Additional treatment during trial: unclear

Outcomes	Outcome assessed at: baseline and three, nine, 18 and 24 months of follow-up **Primary outcomes of study** • Functional performance (maximum voluntary isometric lower body strength, range of motion (FOCUS), work capacity (ERGOS work simulator), timed get-up-and-go test, stair climbing) • Pain (10-cm VAS, 0 to 100, lower score is better) • Pain (SF-36 subscale pain and WOMAC subscale pain) • Function self-reported (WOMAC subscales stiffness and disability) • Function self-reported (SF-36 subscale function) • Function self-reported (VAS for disability, 0 to 100, lower score is better) **Secondary outcomes of study** • Coping efficacy, self-management and health-related quality of life (Client Satisfaction Questionnaire (CSQ)) • Self-management (Arthritis Self-Efficacy Scale (ASES)) • Emotional distress (Positive and Negative Affect Schedule (PANAS)) • Quality of life (EuroQol) • Social integration and support (Medical Outcomes Social Support Survey)
Notes	Only outcomes on dropouts could be extracted because presentation of data was insufficient Funding provided by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (Grant R01-AR-047595)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Two-hundred seventy-three participants were stratified by sex and randomly assigned by the project coordinator via a random number table to 1 of the 3 treatment groups" Comment: This method of randomisation has low risk of introducing bias

McKnight 2010 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants and personnel were not blinded to treatment allocation, which may have introduced bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "The Knee Study was a 24-month unblinded RCT" Comment: The authors specifically report that this was an unblinded study, which means that outcome assessors were not blinded, and risk of detection bias is present
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Missing data were handled with a multiple-imputation procedure, imputing 5 complete data sets to provide complete data for our intent-to-treat analysis" Comment: An ITT analysis was performed with multiple imputation of missing data. Loss to follow-up was similar among treatment groups, with reasons provided and similar, although the dropout rate was fairly high (30%, 23% and 26%)
Selective reporting (reporting bias)	High risk	Comment: Only outcomes on dropouts could be extracted because data presentation was insufficient
Other bias	High risk	Quote: "Overall compliance was higher during phase 1 (67.5%) compared with phase 2 (50.3%), with negligible differences between the groups" Comment: Overall and group-specific adherence to the allocated treatment was low

Murphy 2008

Methods	**Study design:** RCT, multi-centre, two arms, outcome assessment blinded **Country in which trial was carried out:** US **Method of recruitment of participants:** Participants were recruited through fliers and on-site presentations **Setting:** senior housing facilities/senior centres **Was the sample size justified with a priori calculation of effect size/power?**no justification for sample size provided **Length of follow-up:**six months **Dropouts:**Four (14%) dropped out of intervention group (one because of medical problems, three before intervention with no baseline data), and one (4%) dropped out of control group (medical problems)
Participants	**Inclusion criteria** • Symptomatic hip or knee OA as determined by ACR clinical criteria through an on-site examination by one of the study rheumatologists • Reporting of OA symptoms (i.e. pain, stiffness, fatigue) that caused difficulty or the need for personal assistance in at least one of four activities of daily living (i.e. bathing, transferring, toileting, walking) • Age 62 or older

Murphy 2008 *(Continued)*

- Ambulating with or without walking aid
- English-speaking
- No significant cognitive impairment (score of 5 or greater on the six-item screener)

Exclusion criteria

- Hip or knee surgery within the previous nine months
- A condition in which exercise would be contraindicated (e.g. uncontrolled hypertension, recent surgery, severe pain during exercise)
- Current participation in physical/occupational therapy
- Dementia
- Unable to give consent

Baseline characteristics

Baseline characteristics were similar among all treatment groups

Intervention group: exercise + activity strategy training (N = 28 randomly assigned, 25 analysed)

Location of OA: 67% knee, 22% hip and knee, 11% hip [total study population]

BMI (mean (SD)): 30.1 (6.5) kg/m^2

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 93% white

Sex: 7% male, 93% female

Education: 75% had some college to advanced degree

Social capital: 25% married

Age (mean (SD)), years: 75.8 (7.1)

Disability (mean (SD)): 1.5 (1.4) total chronic conditions, 4.4 (2.1) painful or stiff joints

Control group: exercise + education (N = 26 randomly assigned, 26 analysed)

Location of OA: 67% knee, 22% hip and knee, 11% hip [total study population]

BMI (mean (SD)): 30.0 (4.8) kg/m^2

PROGRESS-Plus

Place of residence: US

Race, ethnicity and culture: 89% white

Sex: 15% male, 85% female

Education: 58% had some college to advanced degree

Social capital: 19% married

Age (mean (SD)), years: 74.8 (7.3)

Disability (mean (SD)): 1.0 (1.2) total chronic conditions, 4.6 (2.1) painful or stiff joints

Interventions	***Intervention: exercise + activity strategy training*** *Description:* The AST sessions involved education, group discussion and demonstration and practice of techniques to facilitate activity performance. Participants practiced strategies for symptom man-

Murphy 2008 *(Continued)*

agement. Physical activity enhancement was encouraged by addressing individual barriers and group problem solving to build in additional physical and other valued activities into daily routines. Once an occupational therapist went to the resident of the participant to guide individualised instruction on in-home strategies

Exercise was equal in the two groups (control and intervention) and consisted of progressive resistance exercises using ankle cuff weights by which extra weight could be added. The programme was tailored to individual participants as needed. The programme took 45 minutes to perform, including warm-up and cool-down

During the next six months, participants met for two additional group sessions (spaced two months apart) to review the exercise programme and highlight main points from the AST components

Intended audience: people with knee and hip OA

Mode: group sessions (and one individual session)

Personnel: occupational therapists

Delivery method: face-to-face

Language: English

Format: tailored to individual needs

Location: at the housing site of the participants (senior homes)

Duration: two sessions per week, 1.5-hour sessions, for four weeks

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: no

Self-monitoring and insight: no

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: no

Comparator

Type: alternate intervention

Description: Exercise plus education group; exercise was the same as in the intervention group, the health education programme was based on educational materials from the Arthritis Foundation on topics such as managing pain, importance of exercise, diet, arthritis and medication options

During the next six months, participants met for two additional group sessions (spaced two months apart) to review the exercise programme and highlight main points from the education components

Frequency: two sessions of 1.5 hours per week for four weeks

Additional treatment during trial: unclear

Outcomes	*Outcome assessed at:* baseline, post-treatment and after six months of follow-up

Murphy 2008 *(Continued)*

Primary outcomes of study

- Pain (WOMAC subscale pain, 0 to 20, lower score is better)
- Functional performance (physical activity (CHAMPS, Actiwatch-S))

Secondary outcomes of study

- Self-management (Arthritis Self-Efficacy Scale (ASES) subscales pain and other symptoms)
- Functional performance (six-minute walk test, timed up-and-go test)

Notes	We extracted the following outcomes at post-treatment (short term) for the analyses in this review: self-management (ASES subscales pain and other symptoms), pain (WOMAC subscale pain), functional performance (six-minute walking distance) and dropouts (proportion of missing participants) The trial was supported by the National Centre for Medical Rehabilitation Research (Grant K01-HD-045293) and the Office of the Vice President for Research at the University of Michigan The author (S Murphy) provided additional information on the trial on request

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “Participants were randomly assigned after baseline assessment at each site into 1 of the 2 interventions in blocks of 2 using a random digit table” Comment: This method of randomisation has a low risk of introducing bias
Allocation concealment (selection bias)	Unclear risk	Comment: no information given on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Blinding of participants and personnel to treatment allocation is not possible in this type of control and intervention, thereby introducing a risk of performance bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: “All testing was done at each site by trained assessors blinded to the group assignment of participants” Comment: Outcome assessors were blinded; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: “An intent-to-treat analysis was performed using the last observation carried forward (in this case, baseline values when post-test values were missing)” Comment: Intention-to-treat approach was used in analyses. Dropout rate was low both groups (4/28 in intervention group, 1/26 in control group) for similar reasons
Selective reporting (reporting bias)	Unclear risk	Quote: "Participants then returned for follow up testing 6 months after baseline. Only baseline to posttest outcomes are presented here" Comment: Six-month outcomes were not reported
Other bias	Low risk	Comment: No other potential sources of bias were identified

Nunez 2006

Methods	**Study design:**RCT, single-centre, two arms, outcome assessment blinded **Country in which trial was carried out:**Spain **Method of recruitment of participants:**Participants were referred by the orthopaedic surgery department (100 consecutive patients meeting the inclusion criteria were enrolled) **Setting:** outpatients **Was the sample size justified with a priori calculation of effect size/power?** yes **Length of follow-up:**six months **Dropouts:** Eight (16%) dropped out of the intervention group (two died, two had severe pathology, one lost contact, three dropped out), 12 (24%) dropped out of the control group (two died, one had severe pathology, two lost contact, two transferred to other communities, five dropped out)
Participants	**Inclusion criteria** • Knee OA according to the Kellgren and Lawrence criteria • On waiting list for total knee replacement less than six months **Exclusion criteria** • Functional illiteracy • Inflammatory musculoskeletal disease • Metabolic or neoplastic disease • Severe psychopathology or comorbidity, defined as a diagnosis in the medical record severe enough that the participant could not complete the TEFR programme **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: TEFR (N = 51 randomly assigned, 43 analysed)*** *Location of OA:* 100% knee *Duration of OA (mean (SD)):* 12.06 (9.65) months PROGRESS-Plus *Place of residence:* Spain, urban *Occupation:* 88% retired or housewives, 10% permanently disabled, 2% active (sick leave) *Sex:* 24% male, 76% female *Social capital:* 63% had a family, 35% was alone, 2% lived with a carer *Age (mean (SD)), years:* 72.59 (6.20) *Disability:* 86% had comorbidities, 37% had prior prostheses ***Control group (N = 49 randomly assigned, 37 analysed)*** *Location of OA:* 100% knee *Duration of OA (mean (SD)):* 11.61 (11.45) months PROGRESS-Plus *Place of residence:* Spain, urban *Occupation:* 80% retired or housewives, 16% permanently disabled, 4% active (sick leave)

Nunez 2006 *(Continued)*

	Sex: 25% male, 65% female *Social capital:* 72% had a family, 24% were alone, 4% lived with a carer *Age (mean (SD)), years:* 69.45 (6.79) *Disability:* 82% had a comorbidity, 29% had a prior prosthesis
Interventions	***Intervention*** *Description:* The TEFR programme was based on theories of social learning and self-management and was carried out using active learning strategies. The programme consisted of two individual visits at first week and at three months, and group group sessions in weeks three and four. In group sessions, when possible, a relative or a significant other accompanied the participants. All participants were provided with written information on the contents of the sessions. The contents were centred on consequences of the disease in daily life and included principles of economy/energy conservation and joint protection; evaluation and control of pain; treatment recommended for the management of knee OA; demonstration and use of assistive devices and tables of physical exercises with no burden on the lower limbs; and general exercises to mobilise the joints and strengthen the musculature of the rest of the body. Participants were instructed to practise the exercises at home during the week before the second group session *Intended audience:* people with knee OA *Mode:* group sessions (10 to 12 participants) and individual sessions (both two sessions) *Personnel:* trained health educator *Delivery method:* face-to-face *Language:* Spanish *Format:* both tailored to individual needs and standard format *Location:* - *Duration:* Total duration of the programme was three months; at week one and month three, individual visits lasted 30 minutes, at weeks three and four, group visits lasted 90 minutes *Additional treatment during trial:* Both groups of participants received 3 to 4 g/d of paracetamol alone or ≤ 2 g/d of paracetamol combined with 2,400 mg/d of ibuprofen or NSAIDs. The dosage of NSAIDs used was varied according to individual participant needs Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* no *Emotional well-being:* no *Self-monitoring and insight:* no *Constructive attitudes and approaches:* no *Skill and technique acquisition:* yes *Social integration and support:* no *Health service navigation:* no ***Comparator*** *Type:* usual care

Nunez 2006 *(Continued)*

	Description: Participants received conventional (pharmacological) treatment only. Participants received 3 to 4 g/d of paracetamol alone or ≤ 2 g/d of paracetamol combined with 2,400 mg/d of ibuprofen or NSAIDs. The dosage of NSAIDs used was varied according to individual participant needs *Additional treatment during trial:* none
Outcomes	*Outcome assessed at:* baseline and after six months of follow-up **Primary outcomes of study** • Pain (WOMAC subscale pain, 0 to 20, lower score is better) • Function self-reported (WOMAC subscale function, 0 to 68, lower score is better) • Stiffness (WOMAC subscale stiffness, 0 to 8, lower score is better) **Secondary outcomes of study** • Quality of life (SF-36 subscales, 0 to 100, higher score is better) • Medical treatment (dose of analgesics and NSAIDs per week, number of visits to general physicians, cost of visits to general physicians)
Notes	We extracted the following outcomes at six months (intermediate term) for the analyses in this review: positive and active engagement in life (SF-36 subscale role emotional), pain (WOMAC subscale pain), global OA scores (WOMAC), function self-reported (WOMAC subscale function), quality of life (SF-36 subscale general health perception), emotional distress (SF-36 subscale mental health), social integration and support (SF-36 subscale social function) and dropouts (proportion of missing participants) Data analysis: For the outcome global OA scores, we added the subscales of WOMAC (pain, stiffness and physical function) provided by the author to get the WOMAC total score; we estimated the SD with the formula provided in Table 7.7a in the *Cochrane Handbook for Systematic Reviews of Interventions* and the SDs from the subscales (this method was chosen in close consultation with a biostatistician)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "A randomization table generated by an ad hoc program based on the pseudorandomized routine of the STATA 5.0 statistical package was used" Comment: appropriate method of randomisation, with low risk of selection bias
Allocation concealment (selection bias)	Unclear risk	Comment: no information provided on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: The types of interventions and controls make it impossible for authors to blind participants to their allocated treatment, which might have introduced performance bias Personnel also were not blinded during the trial, but it is unlikely that physicians providing usual care would have been influenced by the fact that participants were randomly assigned to a control group
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "All assessments were performed by an independent, blinded investigator" Comment: Outcome assessors were blinded; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "The crude logistics regression analysis showed that there was no evidence that the loss of patients was related to any of the baseline characteristics or to either of the groups (…)"

Nunez 2006 *(Continued)*

		Comment: The rate of dropout is moderately high but unequal between groups (16% in the intervention group, 24% in the control group). However, analysis showed that dropout rates were not dependent on baseline characteristics or group allocation. No intention-to-treat analysis was performed
Selective reporting (reporting bias)	Low risk	Comment: All outcomes listed in the Methods were reported in the Results
Other bias	Low risk	Comment: No other potential sources of bias were identified

Victor 2005

Methods	
Participants	**Inclusion criteria** • A diagnosis of OA made clinically and confirmed with X-ray • Aged 45 years or older • Knee pain due to OA **Exclusion criteria** • No informed consent given • Unable to speak or understand English • Psychiatric illness • Other diseases (e.g. cardiac) **Baseline characteristics** In the control group, more participants lived alone, more participants came from a non-white race and the levels of disability in this group were significantly lower than in the intervention group ***Intervention group: self-management group (N = 120, 87 analysed after one month of FU, 72 analysed after 12 months of FU)*** *Location of OA:* 100% knee *Duration of OA:* 55% had OA for longer than three years PROGRESS-Plus *Place of residence:* UK *Race, ethnicity and culture:* 48% non-white ethnic group *Occupation:* 32% employed, 34% professional or managerial job *Sex:* 25% male, 75% female *Education:* 40% higher education *Socioeconomic status:* 25% living alone, 70% home owners *Social capital:* 57% married *Age (mean (SD)), years:* 62 (11) *Disability:* 57% OA in both knees, 61% OA in other joints, 24% limiting long-term illness ***Control group: waiting list (N = 73 randomly assigned, 56 analysed after one month of FU, 53 analysed after 12 months of FU)***

Victor 2005 *(Continued)*

	Location of OA: 100% knee *Duration of OA:* 55% OA for longer than three years PROGRESS-Plus *Place of residence:* UK *Race, ethnicity and culture:* 80% non-white ethnic group *Occupation:* 28% employed, 33% professional or managerial job *Sex:* 31% male, 69% female *Education:* 34% higher education *Socioeconomic status:* 46% living alone, 62% home owners *Social capital:* 45% married *Age (mean (SD)), years:* 65 (11) *Disability:* 63% OA in both knees, 67% OA in other joints, 25% limiting long-term illness
Interventions	***Intervention: self-management group*** *Description:* Goals were to inform participants about OA, its causes and effects; to increase self-efficacy by developing strategies and skills in coping with pain, joint protection and exercise; and to improve self-esteem and quality of life by sharing experiences and group support. The structured programme covered clinical information, participatory activities to promote increased function and skills development in coping *Intended audience:* people with OA of the knee *Mode:* group sessions (six to eight participants) *Personnel:* research nurses *Delivery method:* face-to-face and written (booklet) *Language: English* *Format:* structured format *Location:* space made available by the GP (e.g. waiting room, treatment or consulting room) *Duration:* four sessions, each of one hour *Additional treatment during trial:* usual care Were the following heiQ components addressed? *Health-directed behaviour:* yes *Positive and active engagement in life:* no *Emotional well-being:* no *Self-monitoring and insight:* no *Constructive attitudes and approaches:* yes *Skill and technique acquisition:* yes *Social integration and support:* yes *Health service navigation:* no

Victor 2005 *(Continued)*

Comparator: waiting list

Type: waiting list

Description: Participants from the control practices received only the booklet and were put on a waiting list for the PEP. After completion of the trial, they were offered the full intervention

Additional treatment during trial: usual care

Outcomes		
Notes		

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: “The unit of minimisation was the practice rather than individual patients” Comment: Trial used cluster-randomisation (per practice) instead of individual participant randomisation, which introduces a risk of bias
Allocation concealment (selection bias)	Low risk	Quote: “Minimisation status as either waiting list or intervention was concealed from practices until they had achieved their patient recruitment quota” Comment: Allocation was concealed during participant recruitment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants and personnel were not blinded from treatment allocation during the intervention period, introducing a risk of bias
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: “At baseline and 12 months direct interviews were undertaken by interviewers ‘blind’ to the status of participants” Comment: Outcome assessors were blinded to treatment allocation; however, most outcomes were subjective, and participants were not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: “Loss to follow-up was significantly greater amongst those with lower ‘coping’ scores as measured by the AHI, lower socio-domain SF-36 scores and lower physical health status” Comment: Loss to follow-up is unequally distributed (48% in intervention group vs 27% in control group), large in both groups and significantly greater among persons with specific characteristics. No intention-to-treat analysis was performed but a per-protocol analysis was performed (only completers)
Selective reporting (reporting bias)	Unclear risk	Comment: outcomes at three and six months not reported
Other bias	Low risk	Quote: "Thus, we analysed the data at the patient level, but using regression with adjustment for clustering. The Stata statistical software package (Stata Corp) includes a facility for such cluster-adjusted regression, in which the standard error estimates are based on robust estimates of variance" Comment: The cluster design was accounted for in the statistical analysis

Wetzels 2005

Methods	**Study design:** RCT (block-randomisation), multi-centre, two arms, outcome assessment blinded **Country in which trial was carried out:** The Netherlands **Method of recruitment of participants:** Participants were recruited through their GP from their practice medical records **Setting:** primary care **Was the sample size justified with a priori calculation of effect size/power?** yes **Length of follow-up:**six months **Dropouts:** Six dropped out from the intervention (12%), and one dropped out from the control group (2%). Reasons were motivation problems, moved elsewhere, hip/knee surgery, too severe problems of comorbidity and treatment by a geriatric specialist
Participants	**Inclusion criteria** • A clinical diagnosis of hip or knee OA, registered in the participant's practice medical history record as free text or as ICPC-code L89 (OA of the knee) or L90 (OA of the hip) • Aged 65 years or older **Exclusion criteria** • Had undergone hip or knee replacement operation • Had been referred for a replacement operation • GP thought patient was not suitable for participating (because of severe psychosocial circumstances or a terminal disease) **Baseline characteristics** Baseline characteristics were similar in all treatment groups ***Intervention group: self-management programme (N = 51 randomly assigned, 40 analysed)*** *Location of OA:* 52.9% knee, 17.6% hip, 29.4% both hip and knee PROGRESS-Plus *Place of residence:* urban, The Netherlands *Sex:* 23.5% male, 76.5% female *Education:* 54% primary or lower secondary, 46% upper secondary or further *Age (mean (SD)), years:* 75.63 (6.68) ***Control group (N = 53 randomly assigned, 48 analysed)*** *Location of OA:* 54.7% knee, 22.6% hip, 22.6% both hip and knee PROGRESS-Plus *Place of residence:* urban, The Netherlands *Sex:* 24.5% male, 75.5% female *Education:* 50% primary or lower secondary, 50% upper secondary or further *Age (mean (SD)), years:* 73.47 (6.01)
Interventions	***Intervention: self-management programme***

Wetzels 2005 *(Continued)*

Description: The intervention included education and self-management of OA symptoms and consisted of three parts: (1) prepreparation for home visit of the nurse using a booklet and health status charts; (2) 30-minute home visit by a nurse to discuss prepared 'homework' and goal setting (changing behaviour); and (3) follow-up phone call after three months for evaluation and adaptation of goals

Intended audience: people with OA

Mode: individual

Personnel: family practice nurse (certified educated in rheumatology)

Delivery method: face-to-face, follow-up by telephone

Language: Dutch

Format: tailored to individual needs

Location: home

Duration: one 30-minute nurse home visit and one follow-up phone call after three months

Additional treatment during trial: unclear

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: yes

Emotional well-being: no

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: no

Social integration and support: no

Health service navigation: no

Comparator

Type: information only

Description: Participants in the control group received only the educational leaflet about OA

Additional treatment during trial: unclear

Outcomes

Outcome assessed at: baseline and six months of follow-up

Primary outcomes of study

- Function self-reported (Dutch AIMS-SF subscale physical, 7 to 35, lower score is better)
- Global OA scores (Dutch AIMS-SF subscale symptoms, 3 to 15, lower score is better)
- Social integration and support (Dutch AIMS-SF subscale social, 4 to 20, lower score is better)
- Emotional distress (Dutch AIMS-SF subscale affect, 5 to 25, lower score is better)
- Functional performance (timed up-and-go test (TUG) below 12 seconds, percentage of persons (0 to 100), higher score is better)

Secondary outcomes of study

- Participant-reported number of contacts with GP and physiotherapist
- Use of pain medication

Wetzels 2005 *(Continued)*

Notes	We extracted the following outcomes at six months (intermediate term) for the analyses in this review: function self-reported (Dutch AIMS-SF subscale function), emotional distress (Dutch AIMS-SF subscale affect), social integration and support (Dutch AIMS-SF subscale social) and dropouts (proportion of missing participants) Funding was provided by The Netherlands Organisation for Health Research and Development (Zon-MW, number 920-03-252) Data analysis: Change scores were combined with end point scores using generic inverse variance. For the outcome function of performance, we converted dichotomous outcomes to continuous outcomes, so we could combine these data with other available continuous data for this outcome (*Cochrane Handbook for Systematic Reviews of Interventions*, Chapter 9.4.6).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "An independent statistician made randomization lists in advance for each practice. To ensure similar number of patient from different practices in each group, block-randomisation (blocks of two) was used" Comment: Adequate randomisation technique was used
Allocation concealment (selection bias)	Low risk	Quote: "This procedure was performed by a research assistant who was blinded for patients' characteristics" Comment: Participant characteristics were concealed during treatment allocation, reducing the risk of selection bias
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Participants were not blinded from their allocated treatment during trial, introducing a risk of performance bias Comment: Personnel administering the intervention had contact only with the intervention group and not with the control group. Therefore, it is unlikely that this provided risk of bias, even though personnel were not blinded to treatment allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "A research assistant measured in all patients the post-intervention TUG, at this stage he was blinded for intervention-control condition" Comment: The outcome assessor was blinded; however, most outcomes are subjective, and participants are not blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Due to several reasons seven patients withdrew (…) and nine patients did not respond to the final patient self-assessment questionnaire (Figure 1)" Quote: "Data from dropouts and lost to follow-up cases was not available, therefore only cases with data from baseline and after 6 months were included" Comment: Although a strict intention-to-treat analysis was not performed, it is unlikely that this has affected the results significantly, as not many participants dropped out or were lost to follow-up at six months. Although dropout rates were not equally distributed between the two groups (12% in intervention vs 2% in control group), reasons were provided, and so dropout does not seem to introduce a high risk of bias
Selective reporting (reporting bias)	Low risk	Comment: All outcomes specified in the Methods section are reported

Wetzels 2005 *(Continued)*

Other bias	High risk	Quote: "The validity and reliability of the TUG might be compromised by the fact that the test was performed at home, on different chairs, and by different observers. At baseline the assessors of TUG times were not blinded for the assignment of subjects to treatment group" Comment: The risk that results from the timed up-and-go test are not valid or reliable is high

Yip 2007

Methods	**Study design:** RCT, multi-centre, two arms, non-blinded **Country in which trial was carried out:**China **Method of recruitment of participants:** not specified **Setting:**outpatients **Was the sample size justified with a priori calculation of effect size/power?**yes **Length of follow-up:** 12 months **Dropouts** • After four months, 21 (24%) dropped out from the intervention group (10 being busy, three no interest, three walking problems, four cannot contact), 41 (44%) dropped out from the control group (two excluded, 19 being busy, eight no interest, six walking problems, five cannot contact, one passed away) • After one year, 16 (36%) dropped out from the intervention group, and 24 (52%) dropped out from the control group **Note:** In the one-year follow-up paper, only participants who completed at least two of three follow-up assessments were included (N = 45 in intervention group; N = 50 in control group)
Participants	**Inclusion criteria** • Diagnosis of OA of the knee confirmed by medical history and a physical examination (by a registered nurse or a physiotherapist) based on the clinical criteria of the ACR • Capable of completing the questionnaire verbally **Exclusion criteria** • Bed bound, wheelchair bound or experienced loss of balance while standing • Knee replacement • Currently having active physiotherapy such as hydrotherapy or strengthening exercises • Currently receiving acupuncture treatments **Baseline characteristics** Baseline characteristics were similar among all treatment groups ***Intervention group: ASMP + exercise (N = 88 randomly assigned, 88 analysed after four months, 40 analysed after one year of FU)*** *Location of OA:* 100% knee *Duration of OA (mean (SD)):* 8.31 (7.3) years PROGRESS-Plus *Place of residence:* Hong Kong, China

Yip 2007 *(Continued)*

Race, ethnicity and culture: Asian

Occupation: 26.1% housewife, 9.1% professional and administration, 64.8% service provider and non-professional workers

Sex: 15.9% male, 84.1% female

Education: 87.5% Form 3 level or below, 12.5% above Form 3

Social capital: 69.3% married and living together, 30.7% single

Age (mean (SD)), years: 65.6 (9.7)

Disability: number of joints affected: 6.8% one joint, 46.6% two joints, 29.5% three joints, 17.1% four or more joints

Control group (N = 94 randomly assigned, 94 analysed after four months, 37 analysed after one year of FU)

Location of OA: 100% knee

Duration of OA (mean (SD)): 7.85 (6.3) years

PROGRESS-Plus

Place of residence: Hong Kong, China

Race, ethnicity and culture: Asian

Occupation: 26.6% housewife, 8.5% professional and administration, 64.9% service provider and non-professional worker

Sex: 16.0% male, 84.1% female

Education: 87.2% Form 3 level or below, 12.8% above Form 3

Social capital: 70.2% married and living together, 29.8% single

Age (mean (SD)), years: 64.02 (10.3)

Disability: number of joints affected: 14.9% one joint, 51.1% two joints, 18.1% three joints, 15.9% four or more joints

Interventions	***Intervention: ASMP + exercise*** *Description:* The programme was based on the ASMP by Lorig. The programme focused on teaching participants how to cope with and manage common knee OA consequences, such as arthritis pain, fatigue, daily activity limitations and stress. It was designed to give participants skills they could use to optimise their ability to manage their condition. An action plan using three types of exercise was promoted and reinforced weekly during the programme. These included stretching, walking and Tai Chi. A pedometer was given to act as a positive reinforcement in walking (not used as an outcome measure). Participants set their goal on exercise practice and received positive feedback from a nurse every week *Intended audience:* people with knee OA *Mode:* group sessions (10 to 15 participants) *Personnel:* registered nurses, lay person tutor who suffered from knee OA *Delivery method:* face-to-face *Language:* Chinese *Format:* standard format *Location:* -

Yip 2007 *(Continued)*

Duration: once-weekly sessions of two hours' length, six weeks' total duration

Additional treatment during trial: usual care (conventional orthopaedic treatment)

Were the following heiQ components addressed?

Health-directed behaviour: yes

Positive and active engagement in life: no

Emotional well-being: yes

Self-monitoring and insight: yes

Constructive attitudes and approaches: yes

Skill and technique acquisition: yes

Social integration and support: no

Health service navigation: yes

Comparator

Type: usual care

Description: routine orthopaedic treatment (treatment prescribed by orthopaedic doctors or outpatient clinic) with no other treatment

Additional treatment during trial: none

Outcomes	*Outcome assessed at:* baseline, one week, four months and 12 months of follow-up **Primary outcomes of study** • Pain (Visual Analogue Scale (VAS), 0 to 100, lower score is better) • Function self-reported (fatigue intensity on VAS, 0 to 100, lower score is better) • Health-directed activity (frequency, duration of light exercise, hours per week, higher score is better) • Function self-reported (modified Health Assessment Questionnaire (mHAQ), 0 to 100, lower score is better) • Functional performance (range of motion both knees (goniometer)) • Functional performance (muscle strength of hamstrings/quadriceps) • Unplanned arthritis-related medical consultations **Secondary outcomes of study** • Self-management (Arthritis Self-Efficacy Scale (ASES) subscales pain (5 to 50), other symptoms (6 to 60), higher score is better) • Use of self-management techniques (use of cold/hot compresses, share the load among various joints, use large joint to carry heavy load, avoid maintaining weight on the same joint for prolonged periods of time) • Pain at night (VAS, 0 to 100, lower score is better) • Pain during walking (VAS, 0 to 100, lower score is better) • Pain sitting to standing position (VAS, 0 to 100, lower score is better) • Global OA scores (self-rated health, 1 to 5, higher score is better)
Notes	We extracted the following outcomes at one week (short term) and 12 months (intermediate term) for the analyses in this review: self-management (ASES subscale pain), pain (VAS), global OA scores (self-rated health), function self-reported (mHAQ), health-directed activity (hours of light exercise/week) and dropouts (proportion of missing participants)

Yip 2007 *(Continued)*

The trial was funded by partial support of the SN Departmental Research Committee and the Hong Kong Polytechnic University, School of Nursing

The author (J Sit) provided additional information about the trial on request

Data analysis: For the outcome self-management in OA, we could choose between ASES subscale pain and ASES subscale other and chose for ASES subscale pain, as we judged that pain was a more measurable aspect of self-management. The direction of benefit for global OA scores was reversed (i.e. multiplied by -1) in the analysis

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "(…) before being assigned to an intervention or control group by reference to a random number table" Comment: Appropriate method of random sequence generation was used
Allocation concealment (selection bias)	Unclear risk	Comment: no information given on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: Blinding of participants is not possible, and therefore risk of bias is high. Blinding of personnel is not possible either. However, it is unlikely that the treating physician in the control group was influenced by the randomisation of participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Both of them were trained (…) and were not involved in delivering the intervention" Comment: Blinding of outcome assessors remains unclear; however, most outcomes are subjective, and participants are not blinded to group allocation, introducing a risk of detection bias
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "The drop-out group suffered from higher arthritic pain than the participant group. During the previous 16 weeks the drop-out group had visited the doctor for unplanned arthritis-related problems more frequently" Quote: "Analyses of intervention effects with and without an intent-to-treat basis were done and compared. (…) An intent-to-treat basis is presented" Comment: Although the authors state that an ITT analysis is presented, one-year follow-up results are presented only for participants who completed two or more assessments. A high dropout rate was noted in both groups after four months (44% in the control group, 24% in the intervention group). The dropout rates differ between groups, although provided reasons are similar, and differences were found between completers and non-completers. In the one-year analysis, dropout is high as well and differs between groups (52% in the control group, 36% in the intervention group)
Selective reporting (reporting bias)	Low risk	Comment: all outcomes reported
Other bias	High risk	Comment: The cultural difference between the Western population and the Asian population has to be taken into account. This study was performed in China, and assessments were done using face-to-face interviews. It is possible that the outcome improvement was not related to the content of the programme but was simply a result of the participants' behaving according to their expectations of what the researchers were looking for (Hawthorne ef-

Yip 2007 *(Continued)*	fect). Response bias might have occurred as a result of face-to-face interviewing and cultural traditions

List of abbreviations used:
ACR = American College of Rheumatology; AFPT = Aggregated functional performance time; AHI = arthritis helplessness index; AIMS = Arthritis impact measurement scale; AQOL = Assessment of quality of life; ARA = American Rheumatology Association; ASES = Arthritis self-efficacy scale; ASMP = Arthritis self-management education programme; BDI = Beck depression inventory; BMI = body mass index; BMQ = Beliefs about medication questionnaire; CES-D = Center for epidemiologic studies depression scale; CHAMPS = Community healthy activities model program for seniors; CSQ = Coping skills questionnaire; CSRI = Client services receipt inventory; DAS = Dyadic adjustment scale; FM = fibromyalgia; FU = follow-up; GHQ = General health questionnaire; GP = general practitioner; HADS = Hospital anxiety and depression scale; HAQ = Health assessment questionnaire; heiQ = Health education impact questionnaire; ICHPP = international classification of health care problems in primary care; IQR = interquartile range; IRGL = Invloed van reuma op gezondheid en leefwijze; K10 = Kessler psychological distress scale; MACTAR = MacMaster Toronto arthritis patient preference questionnaire; MAPT = Multi-attribute priority tool; MI = myocardial infarction; NSAID = non-steroidal anti-inflammatory drug; OA = osteoarthritis; PANAS = Positive and negative affect schedule; PSFS = participant-specific functional status; QOLS = Quality of life scale; QWB = Quality of well-being scale; RA = rheumatoid arthritis; RCT = randomised controlled trial; SD = standard deviation; SF-36 = Short-form 36; SMP = self-management peducation programme; SOLEO = standing on one leg with eyes open; SOLEC = standing on one leg with eyes closed; SSQ = Social support questionnaire; SSQR = Social support questionnaire revised; TSK = Tampa scale of kinesiophobia; TUG = Timed up-and-go test; UK = United Kingdom; US = United States; VAS = visual analog scale; WOMAC = Western Ontario and McMaster Universities osteoarthritis index.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Barlow 2000	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Bezalel 2010	Intervention did not fulfil our criteria for a self-management programme
Coleman 2010	Study compared two different self-management programmes so did not fulfil our inclusion criteria
Ehrlich-Jones 2001	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Ettinger 1997	Intervention did not fulfil our criteria for a self-management programme
Fernandes 2009	Intervention did not fulfil our criteria for a self-management programme
Fernandes 2010	Intervention did not fulfil our criteria for a self-management programme
Focht 2005	Intervention did not fulfil our criteria for a self-management programme
Goeppinger 1989	Mixed arthritis population without subgroup data for OA
Hoogeboom 2010	Study compared two different self-management programmes so did not fulfil our inclusion criteria
Laforest 2008	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Laforest 2008a	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Lindroth 1989	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Lorig 1985	Mixed arthritis population without subgroup data for OA

Study	Reason for exclusion
Lorig 1998	Study compared two different self-management programmes so did not fulfil our inclusion criteria
Lorig 1999a	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Lorig 1999b	Mixed arthritis population without subgroup data for OA
Lorig 2005	Mixed arthritis population without subgroup data for OA
Martire 2003a	Study compared two different self-management programmes so did not fulfil our inclusion criteria
Martire 2008	Study compared two different self-management programmes so did not fulfil our inclusion criteria
Murphy 2010	Study compared two different self-management programmes so did not fulfil our inclusion criteria
Nour 2006	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)
Solomon 2002	Mixed arthritis population without subgroup data for OA (we attempted to contact the authors but did not receive a response)

Characteristics of studies awaiting assessment *[ordered by study ID]*

Allen 2011

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Coleman 2012

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Hurley 2012

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Schlenk 2011

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Somers 2012

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Stukstette 2011

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Von Korff 2012

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Wu 2011

Methods	
Participants	
Interventions	
Outcomes	
Notes	Awaiting assessment

Characteristics of ongoing studies *[ordered by study ID]*

Allen 2012

Trial name or title	
Methods	
Participants	
Interventions	
Outcomes	
Starting date	
Contact information	
Notes	Awaiting assessment

Bennell 2012

Trial name or title	
Methods	
Participants	

Bennell 2012 *(Continued)*

Interventions	
Outcomes	
Starting date	
Contact information	
Notes	Awaiting assessment

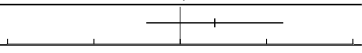
DATA AND ANALYSES

Comparison 1. SMP versus attention control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Self-management of OA	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2 Pain	3		Std. Mean Difference (Random, 95% CI)	Subtotals only
2.1 Short term	1		Std. Mean Difference (Random, 95% CI)	-0.62 [-1.11, -0.13]
2.2 Intermediate term	3		Std. Mean Difference (Random, 95% CI)	-0.26 [-0.44, -0.09]
3 Global OA scores	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
4 Function—self-reported	3		Std. Mean Difference (Random, 95% CI)	Subtotals only
4.1 Short term	1		Std. Mean Difference (Random, 95% CI)	-0.13 [-0.49, 0.23]
4.2 Intermediate term	3		Std. Mean Difference (Random, 95% CI)	-0.19 [-0.50, 0.11]
5 Quality of life	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
6 Withdrawals	5	937	Risk Ratio (M-H, Random, 95% CI)	1.11 [0.78, 1.57]
7 Emotional distress	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
7.1 Short term	1	68	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.85, 0.11]
7.2 Intermediate term	2	409	Std. Mean Difference (IV, Random, 95% CI)	0.02 [-0.18, 0.21]

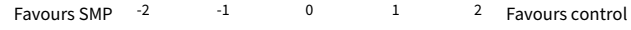
Analysis 1.1. Comparison 1 SMP versus attention control, Outcome 1 Self-management of OA.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Allen 2010	172	6.2 (3.8)	172	5.8 (3.7)		0.4[-0.39,1.19]



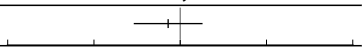
Analysis 1.2. Comparison 1 SMP versus attention control, Outcome 2 Pain.

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
1.2.1 Short term						
Keefe 1990	32	36	-0.6 (0.25)		100%	-0.62[-1.11,-0.13]
Subtotal (95% CI)					**100%**	**-0.62[-1.11,-0.13]**
Heterogeneity: Not applicable						
Test for overall effect: Z=2.48(P=0.01)						
1.2.2 Intermediate term						
Keefe 1990	30	35	-0.3 (0.25)		12.75%	-0.34[-0.83,0.15]
Mazzuca 1997	82	83	-0.3 (0.158)		31.84%	-0.3[-0.61,0.01]
Allen 2010	139	148	-0.2 (0.12)		55.41%	-0.22[-0.45,0.01]
Subtotal (95% CI)					**100%**	**-0.26[-0.44,-0.09]**
Heterogeneity: Tau²=0; Chi²=0.28, df=2(P=0.87); I²=0%						
Test for overall effect: Z=2.92(P=0)						
Test for subgroup differences: Chi²=1.83, df=1 (P=0.18), I²=45.4%						



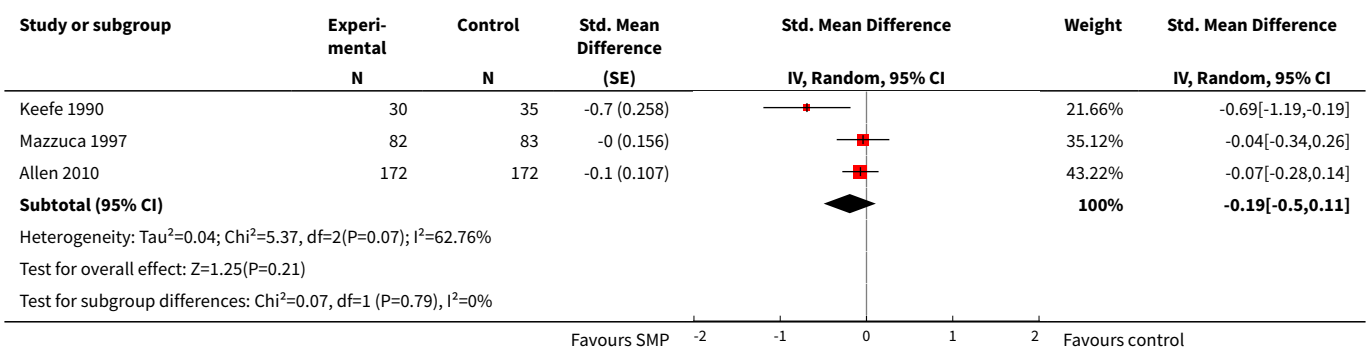
Analysis 1.3. Comparison 1 SMP versus attention control, Outcome 3 Global OA scores.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Maisiak 1996	62	4.1 (1.1)	54	4.2 (1.1)		-0.14[-0.54,0.26]

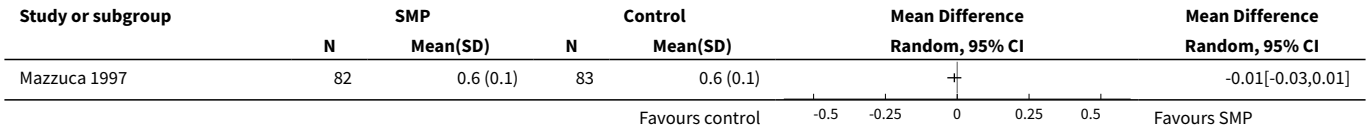


Analysis 1.4. Comparison 1 SMP versus attention control, Outcome 4 Function—self-reported.

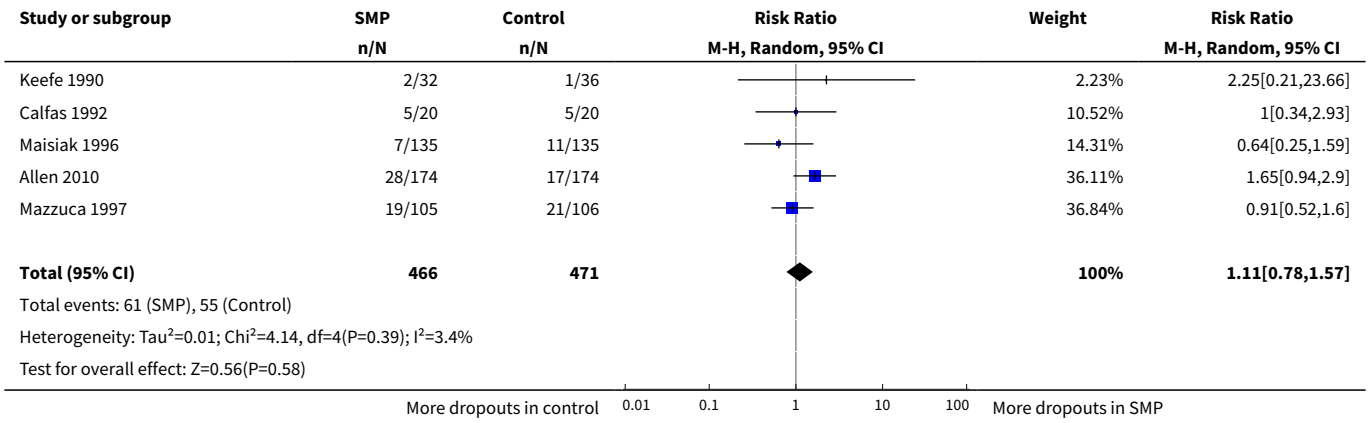
Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
1.4.1 Short term						
Keefe 1990	62	54	-0.1 (0.186)		100%	-0.13[-0.49,0.23]
Subtotal (95% CI)					**100%**	**-0.13[-0.49,0.23]**
Heterogeneity: Not applicable						
Test for overall effect: Z=0.7(P=0.49)						
1.4.2 Intermediate term						



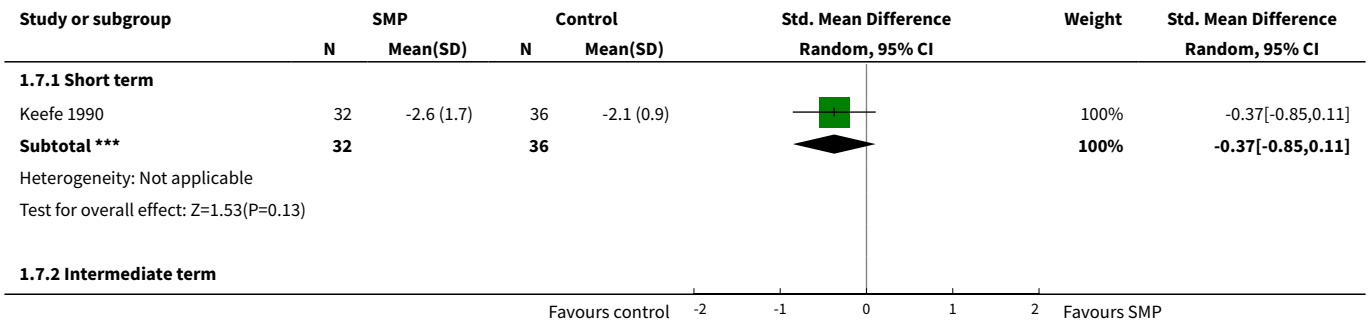
Analysis 1.5. Comparison 1 SMP versus attention control, Outcome 5 Quality of life.

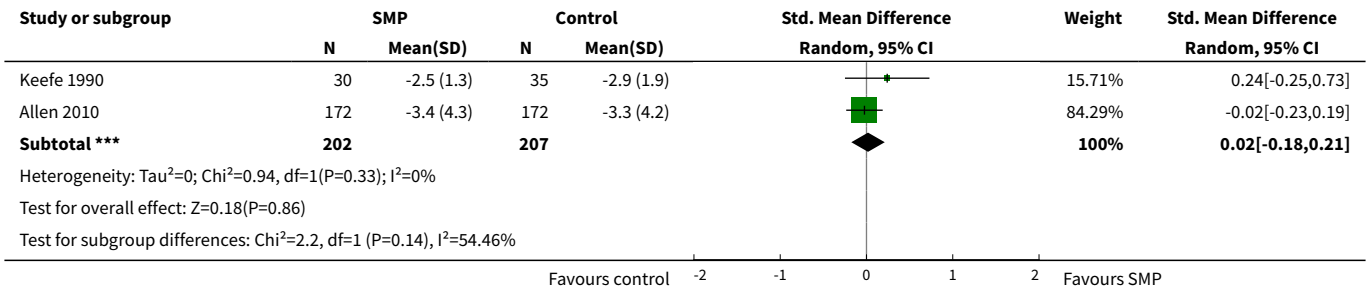


Analysis 1.6. Comparison 1 SMP versus attention control, Outcome 6 Withdrawals.



Analysis 1.7. Comparison 1 SMP versus attention control, Outcome 7 Emotional distress.





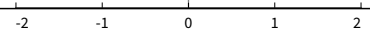
Comparison 2. SMP versus usual care/no treatment/wait list

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Self-management of OA	11		Std. Mean Difference (Random, 95% CI)	Subtotals only
1.1 Short term	5	721	Std. Mean Difference (Random, 95% CI)	0.22 [-0.00, 0.45]
1.2 Intermediate term	10	1647	Std. Mean Difference (Random, 95% CI)	0.14 [0.00, 0.27]
1.3 Long term	1	195	Std. Mean Difference (Random, 95% CI)	0.23 [-0.05, 0.51]
2 Engagement in life	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Short term	1	143	Std. Mean Difference (IV, Random, 95% CI)	-0.23 [-0.57, 0.11]
2.2 Intermediate term	3	357	Std. Mean Difference (IV, Random, 95% CI)	0.01 [-0.20, 0.21]
3 Pain	14		Std. Mean Difference (Random, 95% CI)	Subtotals only
3.1 Short term	6		Std. Mean Difference (Random, 95% CI)	-0.26 [-0.41, -0.10]
3.2 Intermediate term	13		Std. Mean Difference (Random, 95% CI)	-0.17 [-0.26, -0.08]
3.3 Long term	1		Std. Mean Difference (Random, 95% CI)	-0.18 [-0.45, 0.09]
4 Global OA scores	7		Std. Mean Difference (Random, 95% CI)	Subtotals only
4.1 Short term	2		Std. Mean Difference (Random, 95% CI)	-0.34 [-0.59, -0.09]
4.2 Intermediate term	7		Std. Mean Difference (Random, 95% CI)	-0.28 [-0.39, -0.17]
4.3 Long term	1		Std. Mean Difference (Random, 95% CI)	-0.29 [-0.56, -0.02]
5 Function—self-reported	13		Std. Mean Difference (Random, 95% CI)	Subtotals only
5.1 Short term	5		Std. Mean Difference (Random, 95% CI)	-0.01 [-0.19, 0.18]
5.2 Intermediate term	13		Std. Mean Difference (Random, 95% CI)	-0.16 [-0.25, -0.08]
5.3 Long term	1		Std. Mean Difference (Random, 95% CI)	-0.27 [-0.55, 0.01]

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
6 Function—perfor-mance	2		Std. Mean Difference (Random, 95% CI)	Subtotals only
6.1 Short term	1		Std. Mean Difference (Random, 95% CI)	0.33 [-0.07, 0.73]
6.2 Intermediate term	2		Std. Mean Difference (Random, 95% CI)	0.06 [-0.24, 0.36]
7 Quality of life	8		Std. Mean Difference (Random, 95% CI)	Subtotals only
7.1 Short term	2		Std. Mean Difference (Random, 95% CI)	0.14 [-0.47, 0.75]
7.2 Intermediate term	8		Std. Mean Difference (Random, 95% CI)	0.03 [-0.08, 0.14]
7.3 Long term	2		Std. Mean Difference (Random, 95% CI)	0.10 [-0.10, 0.31]
8 Withdrawals	16	3738	Risk Ratio (M-H, Random, 95% CI)	0.99 [0.74, 1.33]
9 Emotional distress	9		Std. Mean Difference (Random, 95% CI)	Subtotals only
9.1 Short term	3		Std. Mean Difference (Random, 95% CI)	0.01 [-0.44, 0.45]
9.2 Intermediate term	8		Std. Mean Difference (Random, 95% CI)	0.11 [-0.06, 0.28]
10 Health-directed ac-tivity	3		Std. Mean Difference (Random, 95% CI)	Subtotals only
10.1 Short term	1		Std. Mean Difference (Random, 95% CI)	0.67 [0.37, 0.97]
10.2 Intermediate term	3		Std. Mean Difference (Random, 95% CI)	0.25 [0.05, 0.46]
11 Skill and technique acquisition	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
12 Constructive atti-tudes and approaches	2		Std. Mean Difference (IV, Random, 95% CI)	Totals not selected
12.1 Short term	1		Std. Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
12.2 Intermediate term	1		Std. Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
13 Social integration and support	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
13.1 Short term	1	143	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.52, 0.15]
13.2 Intermediate term	3	357	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.30, 0.14]
14 Health service navi-gation	2		Std. Mean Difference (Random, 95% CI)	0.15 [-0.03, 0.34]

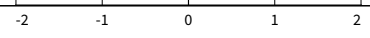
Analysis 2.1. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 1 Self-management of OA.

Study or subgroup	Favours control N	Usual care N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
2.1.1 Short term						
Keefe 2004	38	18	0.3 (0.288)		11.6%	0.29[-0.27,0.85]
Hopman-Rock 2000	54	44	0.4 (0.207)		17.98%	0.4[-0,0.8]
Victor 2005	87	56	-0.2 (0.173)		21.7%	-0.2[-0.54,0.14]
Martire 2007	188	54	0.4 (0.156)		24.01%	0.35[0.05,0.65]
Yip 2007	88	94	0.3 (0.151)		24.71%	0.3[0.01,0.59]
Subtotal (95% CI)					**100%**	**0.22[-0,0.45]**
Heterogeneity: Tau²=0.03; Chi²=7.66, df=4(P=0.1); I²=47.78%						
Test for overall effect: Z=1.92(P=0.06)						
2.1.2 Intermediate term						
Blixen 2004	15	15	0.2 (0.367)		3.1%	0.19[-0.53,0.91]
Yip 2007	40	37	0.4 (0.232)		6.59%	0.39[-0.06,0.84]
Hopman-Rock 2000	45	35	0.2 (0.204)		7.94%	0.19[-0.21,0.59]
Hansson 2010	61	53	0 (0.189)		8.82%	0.03[-0.34,0.4]
Victor 2005	72	53	-0.4 (0.184)		9.15%	-0.4[-0.76,-0.04]
Crotty 2009	75	77	0.2 (0.163)		10.59%	0.15[-0.17,0.47]
Martire 2007	188	54	0.4 (0.156)		11.19%	0.42[0.12,0.72]
Heuts 2005	91	101	0.1 (0.145)		12.07%	0.07[-0.21,0.35]
Lorig 2008	134	158	0.3 (0.12)		14.57%	0.27[0.04,0.5]
Allen 2010	172	171	0.1 (0.107)		15.98%	0.08[-0.13,0.29]
Subtotal (95% CI)					**100%**	**0.14[0,0.27]**
Heterogeneity: Tau²=0.02; Chi²=15.18, df=9(P=0.09); I²=40.72%						
Test for overall effect: Z=1.97(P=0.05)						
2.1.3 Long term						
Heuts 2005	89	106	0.2 (0.145)		100%	0.23[-0.05,0.51]
Subtotal (95% CI)					**100%**	**0.23[-0.05,0.51]**
Heterogeneity: Not applicable						
Test for overall effect: Z=1.58(P=0.11)						
Test for subgroup differences: Chi²=0.6, df=1 (P=0.74), I²=0%						



Analysis 2.2. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 2 Engagement in life.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
2.2.1 Short term							
Victor 2005	87	55 (44)	56	65 (42)		100%	-0.23[-0.57,0.11]
Subtotal ***	**87**		**56**			**100%**	**-0.23[-0.57,0.11]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.34(P=0.18)							
2.2.2 Intermediate term							
Nunez 2006	43	57.7 (47.2)	37	62 (47.3)		22.48%	-0.09[-0.53,0.35]
Victor 2005	72	56 (43)	53	57 (44)		34.55%	-0.02[-0.38,0.33]



Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Crotty 2009	75	4.2 (1)	77	4.1 (1)		42.97%	0.08[-0.24,0.4]
Subtotal ***	**190**		**167**			**100%**	**0.01[-0.2,0.21]**
Heterogeneity: Tau^2=0; Chi^2=0.41, df=2(P=0.81); I^2=0%							
Test for overall effect: Z=0.05(P=0.96)							
Test for subgroup differences: Chi^2=1.36, df=1 (P=0.24), I^2=26.32%							

Favours control -2 -1 0 1 2 Favours SMP

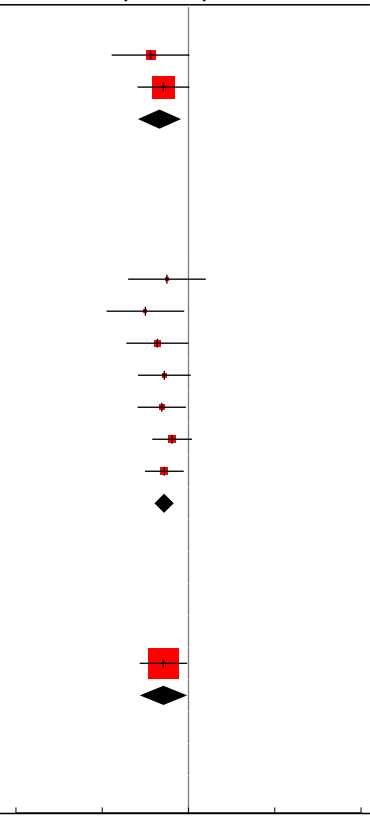
Analysis 2.3. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 3 Pain.

Study or subgroup	Favours SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
2.3.1 Short term						
Keefe 2004	18	18	0.1 (0.286)		7.38%	0.06[-0.5,0.62]
Keefe 1990	32	31	-0.6 (0.26)		8.89%	-0.63[-1.14,-0.12]
Hopman-Rock 2000	45	35	-0.3 (0.227)		11.68%	-0.35[-0.79,0.09]
Victor 2005	87	56	-0.1 (0.171)		20.61%	-0.07[-0.4,0.26]
Martire 2007	188	54	-0.3 (0.156)		24.86%	-0.26[-0.56,0.04]
Yip 2007	88	94	-0.3 (0.151)		26.58%	-0.32[-0.61,-0.03]
Subtotal (95% CI)					**100%**	**-0.26[-0.41,-0.1]**
Heterogeneity: Tau^2=0; Chi^2=4.83, df=5(P=0.44); I^2=0%						
Test for overall effect: Z=3.31(P=0)						
2.3.2 Intermediate term						
Blixen 2004	15	15	0.2 (0.367)		1.51%	0.17[-0.55,0.89]
Keefe 1990	229	113	-0.2 (0.253)		3.2%	-0.21[-0.7,0.28]
Yip 2007	40	37	-0.2 (0.23)		3.87%	-0.23[-0.68,0.22]
Hopman-Rock 2000	45	35	-0 (0.227)		3.96%	-0.02[-0.46,0.42]
Nunez 2006	43	37	-0.2 (0.224)		4.05%	-0.23[-0.67,0.21]
Victor 2005	72	53	-0.1 (0.181)		6.22%	-0.09[-0.44,0.26]
Mazzuca 2004	80	54	0.1 (0.176)		6.58%	0.05[-0.29,0.39]
Crotty 2009	75	77	-0.1 (0.163)		7.65%	-0.1[-0.42,0.22]
Martire 2007	89	54	-0.3 (0.153)		8.71%	-0.27[-0.57,0.03]
Heuts 2005	0	0	0 (0.14)		10.36%	0.03[-0.24,0.3]
Allen 2010	139	136	-0.2 (0.122)		13.6%	-0.22[-0.46,0.02]
Lorig 2008	0	0	-0.3 (0.117)		14.81%	-0.28[-0.51,-0.05]
Hurley 2007	229	113	-0.3 (0.115)		15.48%	-0.29[-0.51,-0.07]
Subtotal (95% CI)					**100%**	**-0.17[-0.26,-0.08]**
Heterogeneity: Tau^2=0; Chi^2=8, df=12(P=0.79); I^2=0%						
Test for overall effect: Z=3.76(P=0)						
2.3.3 Long term						
Heuts 2005	0	0	-0.2 (0.138)		100%	-0.18[-0.45,0.09]
Subtotal (95% CI)					**100%**	**-0.18[-0.45,0.09]**
Heterogeneity: Not applicable						
Test for overall effect: Z=1.31(P=0.19)						
Test for subgroup differences: Chi^2=0.94, df=1 (P=0.62), I^2=0%						

Favours SMP -2 -1 0 1 2 Favours control

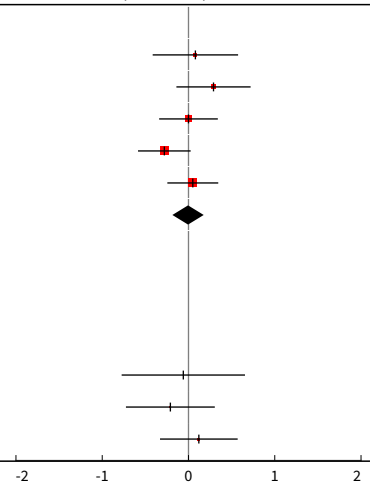
Analysis 2.4. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 4 Global OA scores.

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
2.4.1 Short term						
Yip 2007	40	37	-0.4 (0.23)		30.77%	-0.44[-0.89,0.01]
Martire 2007	99	54	-0.3 (0.153)		69.23%	-0.29[-0.59,0.01]
Subtotal (95% CI)					**100%**	**-0.34[-0.59,-0.09]**
Heterogeneity: Tau^2=0; Chi^2=0.3, df=1(P=0.59); I^2=0%						
Test for overall effect: Z=2.64(P=0.01)						
2.4.2 Intermediate term						
Yip 2007	40	37	-0.2 (0.23)		6.19%	-0.25[-0.7,0.2]
Nunez 2006	43	37	-0.5 (0.23)		6.19%	-0.5[-0.95,-0.05]
Maisiak 1996	62	59	-0.4 (0.184)		9.67%	-0.36[-0.72,-0]
Martire 2007	188	54	-0.3 (0.156)		13.48%	-0.28[-0.58,0.02]
Heuts 2005	0	0	-0.3 (0.143)		15.99%	-0.31[-0.59,-0.03]
Lorig 2008	0	0	-0.2 (0.117)		23.7%	-0.19[-0.42,0.04]
Hurley 2007	229	113	-0.3 (0.115)		24.77%	-0.28[-0.5,-0.06]
Subtotal (95% CI)					**100%**	**-0.28[-0.39,-0.17]**
Heterogeneity: Tau^2=0; Chi^2=1.75, df=6(P=0.94); I^2=0%						
Test for overall effect: Z=4.95(P<0.0001)						
2.4.3 Long term						
Heuts 2005	94	113	-0.3 (0.14)		100%	-0.29[-0.56,-0.02]
Subtotal (95% CI)					**100%**	**-0.29[-0.56,-0.02]**
Heterogeneity: Not applicable						
Test for overall effect: Z=2.07(P=0.04)						
Test for subgroup differences: Chi^2=0.15, df=1 (P=0.93), I^2=0%						

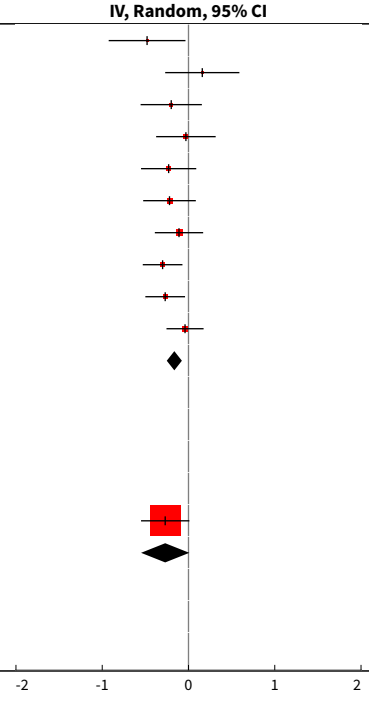


Analysis 2.5. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 5 Function—self-reported.

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
2.5.1 Short term						
Keefe 1990	32	31	0.1 (0.253)		11.69%	0.08[-0.41,0.57]
Hopman-Rock 2000	45	39	0.3 (0.219)		14.86%	0.29[-0.14,0.72]
Victor 2005	87	56	0 (0.173)		21.59%	0[-0.34,0.34]
Martire 2007	89	54	-0.3 (0.156)		25.32%	-0.28[-0.58,0.02]
Yip 2007	88	94	0.1 (0.151)		26.54%	0.05[-0.24,0.34]
Subtotal (95% CI)					**100%**	**-0.01[-0.19,0.18]**
Heterogeneity: Tau^2=0.01; Chi^2=5.17, df=4(P=0.27); I^2=22.6%						
Test for overall effect: Z=0.06(P=0.96)						
2.5.2 Intermediate term						
Blixen 2004	15	15	-0.1 (0.365)		1.5%	-0.06[-0.77,0.65]
Keefe 1990	30	28	-0.2 (0.263)		2.9%	-0.21[-0.72,0.3]
Yip 2007	40	37	0.1 (0.23)		3.79%	0.12[-0.33,0.57]

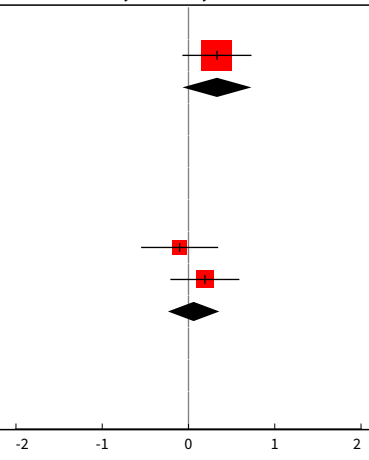


Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
Nunez 2006	43	37	-0.5 (0.227)		3.88%	-0.48[-0.92,-0.04]
Hopman-Rock 2000	45	39	0.2 (0.219)		4.15%	0.16[-0.27,0.59]
Victor 2005	72	53	-0.2 (0.181)		6.1%	-0.2[-0.55,0.15]
Mazzuca 2004	80	54	-0 (0.176)		6.45%	-0.03[-0.37,0.31]
Crotty 2009	75	77	-0.2 (0.163)		7.5%	-0.23[-0.55,0.09]
Martire 2007	99	54	-0.2 (0.156)		8.26%	-0.22[-0.52,0.08]
Heuts 2005	0	0	-0.1 (0.143)		9.8%	-0.11[-0.39,0.17]
Hurley 2007	229	113	-0.3 (0.117)		14.52%	-0.3[-0.53,-0.07]
Lorig 2008	0	0	-0.3 (0.117)		14.52%	-0.27[-0.5,-0.04]
Allen 2010	172	171	-0 (0.11)		16.62%	-0.04[-0.25,0.17]
Subtotal (95% CI)					**100%**	**-0.16[-0.25,-0.08]**
Heterogeneity: Tau²=0; Chi²=10.25, df=12(P=0.59); I²=0%						
Test for overall effect: Z=3.67(P=0)						
2.5.3 Long term						
Heuts 2005	91	108	-0.3 (0.143)		100%	-0.27[-0.55,0.01]
Subtotal (95% CI)					**100%**	**-0.27[-0.55,0.01]**
Heterogeneity: Not applicable						
Test for overall effect: Z=1.89(P=0.06)						
Test for subgroup differences: Chi²=3.22, df=1 (P=0.2), I²=37.94%						



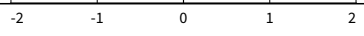
Analysis 2.6. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 6 Function—performance.

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
2.6.1 Short term						
Hopman-Rock 2000	50	46	0.3 (0.204)		100%	0.33[-0.07,0.73]
Subtotal (95% CI)					**100%**	**0.33[-0.07,0.73]**
Heterogeneity: Not applicable						
Test for overall effect: Z=1.62(P=0.11)						
2.6.2 Intermediate term						
Wetzels 2005	0	0	-0.1 (0.228)		44.51%	-0.1[-0.55,0.34]
Hopman-Rock 2000	50	46	0.2 (0.204)		55.49%	0.19[-0.21,0.59]
Subtotal (95% CI)					**100%**	**0.06[-0.24,0.36]**
Heterogeneity: Tau²=0; Chi²=0.92, df=1(P=0.34); I²=0%						
Test for overall effect: Z=0.39(P=0.69)						
Test for subgroup differences: Chi²=1.13, df=1 (P=0.29), I²=11.37%						



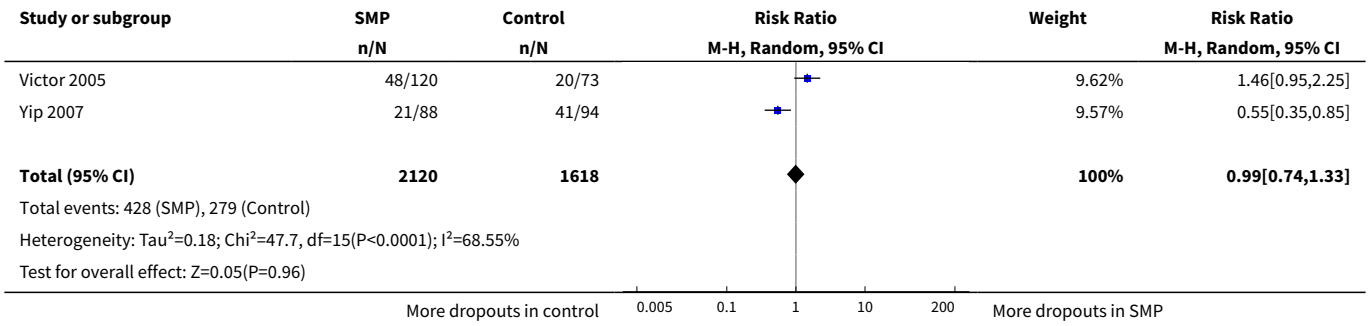
Analysis 2.7. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 7 Quality of life.

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
2.7.1 Short term						
Hopman-Rock 2000	55	46	0.5 (0.204)		48.38%	0.46[0.06,0.86]
Victor 2005	87	56	-0.2 (0.171)		51.62%	-0.16[-0.49,0.17]
Subtotal (95% CI)					**100%**	**0.14[-0.47,0.75]**
Heterogeneity: Tau^2=0.16; Chi^2=5.42, df=1(P=0.02); I^2=81.57%						
Test for overall effect: Z=0.45(P=0.65)						
2.7.2 Intermediate term						
Nunez 2006	43	37	-0.3 (0.227)		6.23%	-0.29[-0.73,0.15]
Hopman-Rock 2000	55	46	0.2 (0.199)		8.11%	0.22[-0.17,0.61]
Hansson 2010	0	0	0 (0.186)		9.26%	0[-0.36,0.36]
Victor 2005	72	53	-0.1 (0.181)		9.78%	-0.1[-0.45,0.25]
Crotty 2009	75	77	0.2 (0.161)		12.43%	0.17[-0.14,0.48]
Cronan 1997	78	69	0.2 (0.148)		14.66%	0.18[-0.11,0.47]
Heuts 2005	0	0	0.1 (0.145)		15.18%	0.11[-0.17,0.39]
Hurley 2007	229	113	-0.1 (0.115)		24.36%	-0.09[-0.31,0.13]
Subtotal (95% CI)					**100%**	**0.03[-0.08,0.14]**
Heterogeneity: Tau^2=0; Chi^2=6.62, df=7(P=0.47); I^2=0%						
Test for overall effect: Z=0.57(P=0.57)						
2.7.3 Long term						
Cronan 1997	56	66	0 (0.153)		47.44%	0.01[-0.29,0.31]
Heuts 2005	87	105	0.2 (0.145)		52.56%	0.19[-0.09,0.47]
Subtotal (95% CI)					**100%**	**0.1[-0.1,0.31]**
Heterogeneity: Tau^2=0; Chi^2=0.73, df=1(P=0.39); I^2=0%						
Test for overall effect: Z=0.99(P=0.32)						
Test for subgroup differences: Chi^2=0.45, df=1 (P=0.8), I^2=0%						

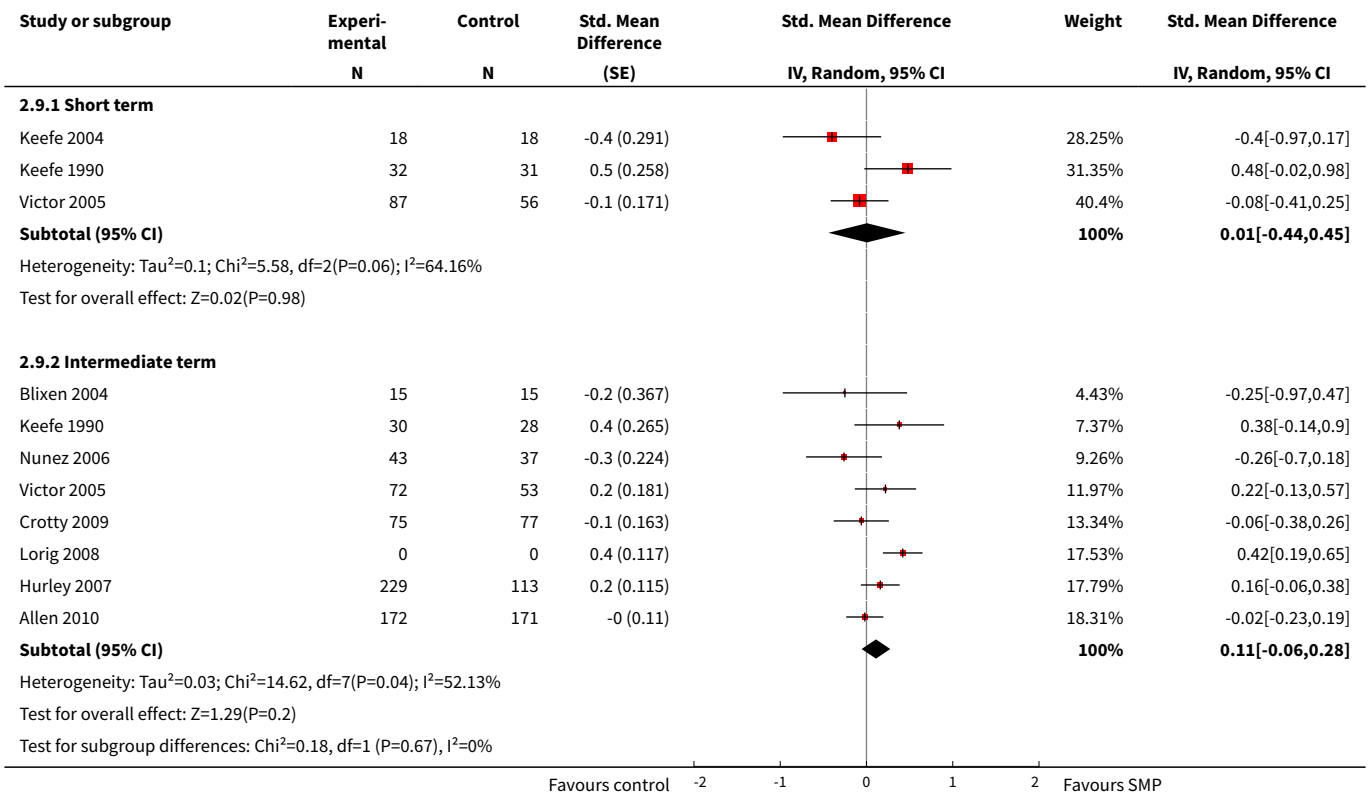


Analysis 2.8. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 8 Withdrawals.

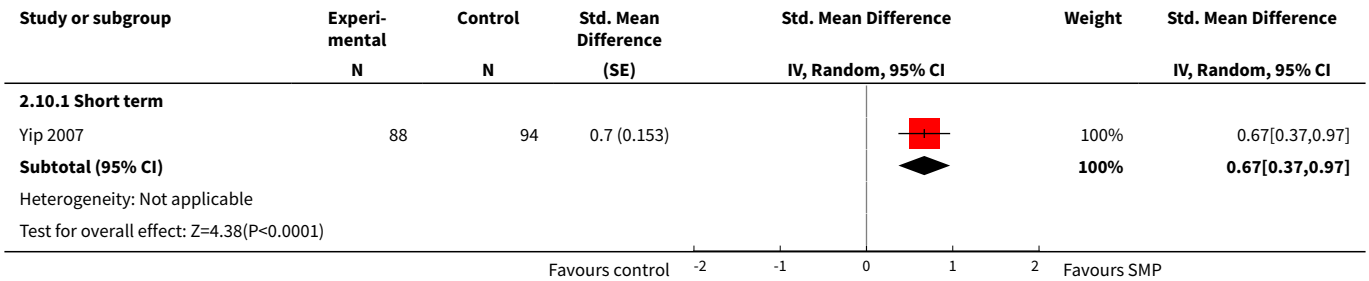
Study or subgroup	SMP n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Allen 2010	28/174	17/175		8.37%	1.66[0.94,2.91]
Blixen 2004	1/16	1/16		1.08%	1[0.07,14.64]
Cronan 1997	60/186	24/90		9.92%	1.21[0.81,1.81]
Hansson 2010	10/61	4/53		4.46%	2.17[0.72,6.52]
Heuts 2005	5/149	0/148		0.94%	10.93[0.61,195.86]
Hopman-Rock 2000	0/60	1/60		0.79%	0.33[0.01,8.02]
Hurley 2007	49/278	27/140		9.71%	0.91[0.6,1.4]
Keefe 1990	2/18	2/16		2.08%	0.89[0.14,5.6]
Keefe 2004	3/52	2/18		2.35%	0.52[0.09,2.86]
Lorig 2008	126/433	78/422		11.21%	1.57[1.23,2.02]
Maisiak 1996	7/135	8/135		5.09%	0.88[0.33,2.35]
Martire 2007	29/188	21/54		9.24%	0.4[0.25,0.64]
Mazzuca 2004	31/111	21/75		9.26%	1[0.62,1.6]
Nunez 2006	8/51	12/49		6.33%	0.64[0.29,1.43]

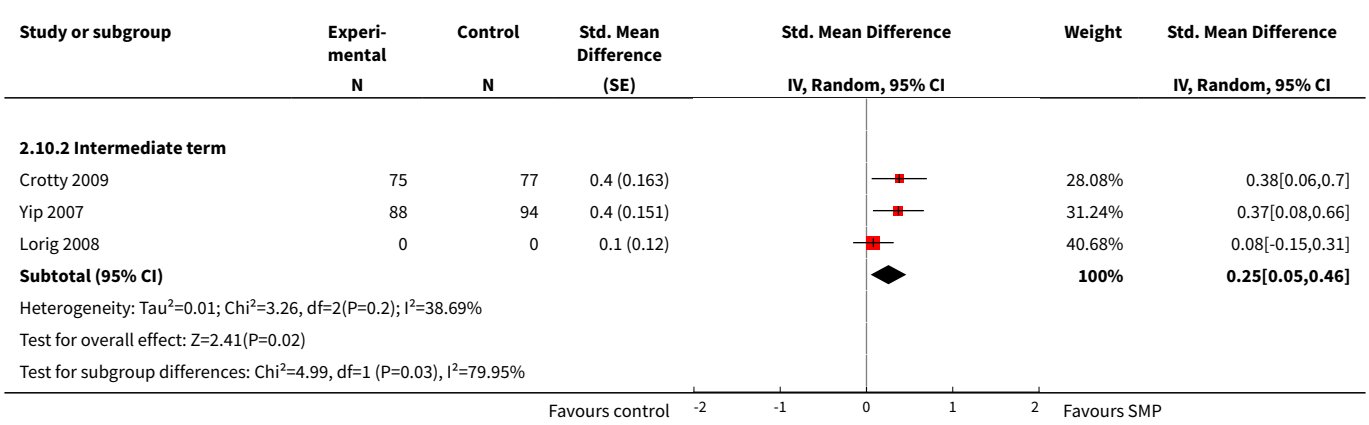


Analysis 2.9. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 9 Emotional distress.

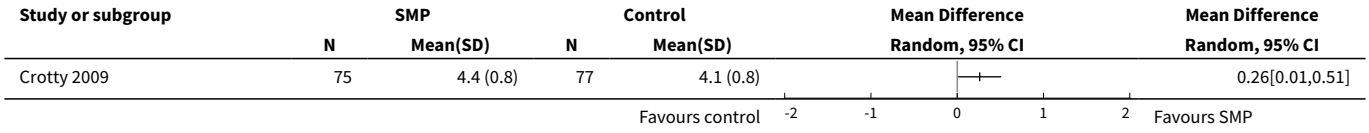


Analysis 2.10. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 10 Health-directed activity.

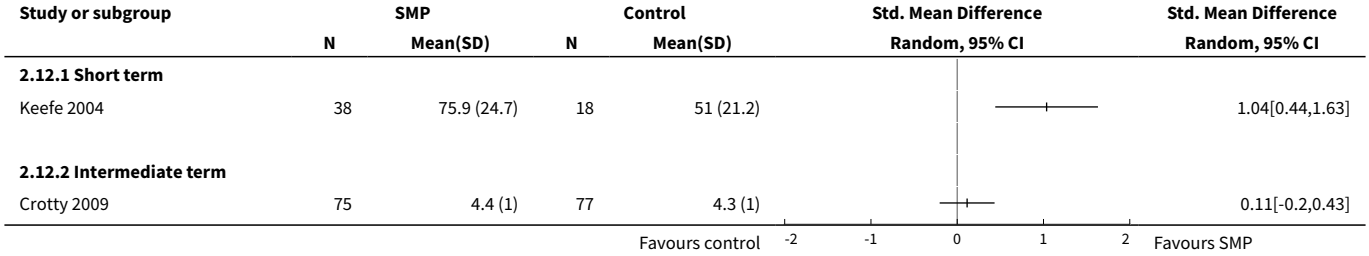




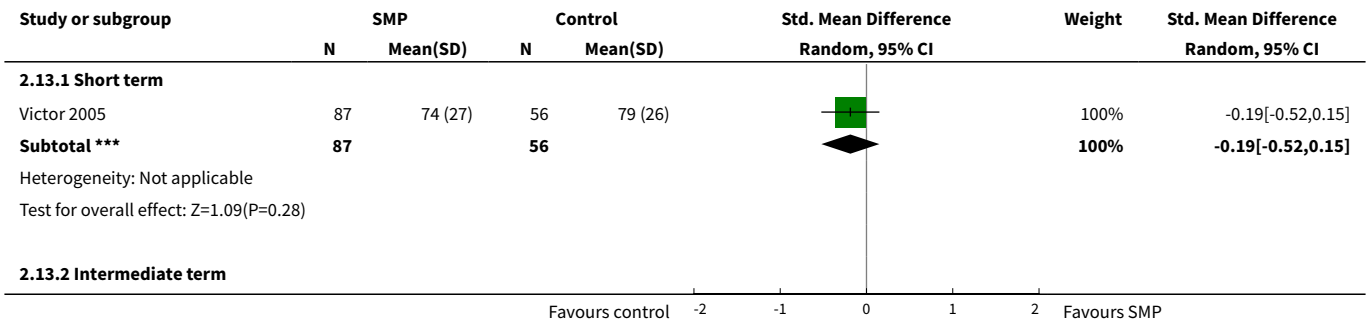
Analysis 2.11. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 11 Skill and technique acquisition.

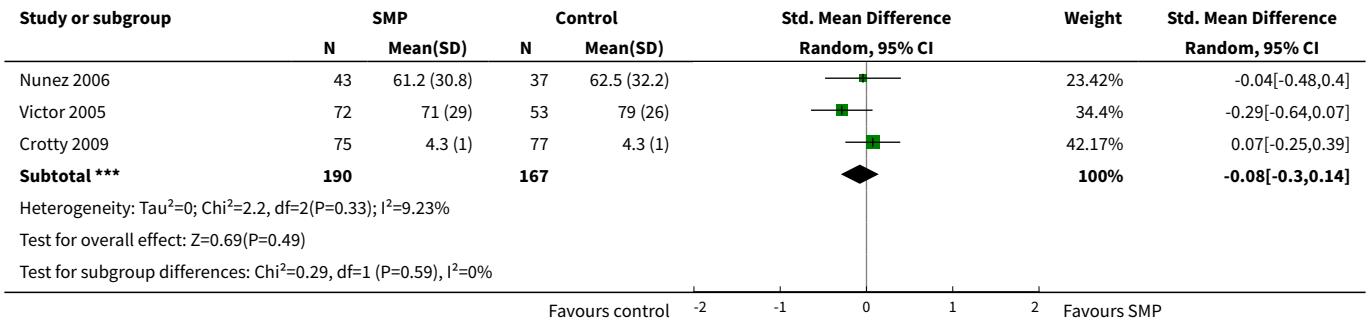


Analysis 2.12. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 12 Constructive attitudes and approaches.

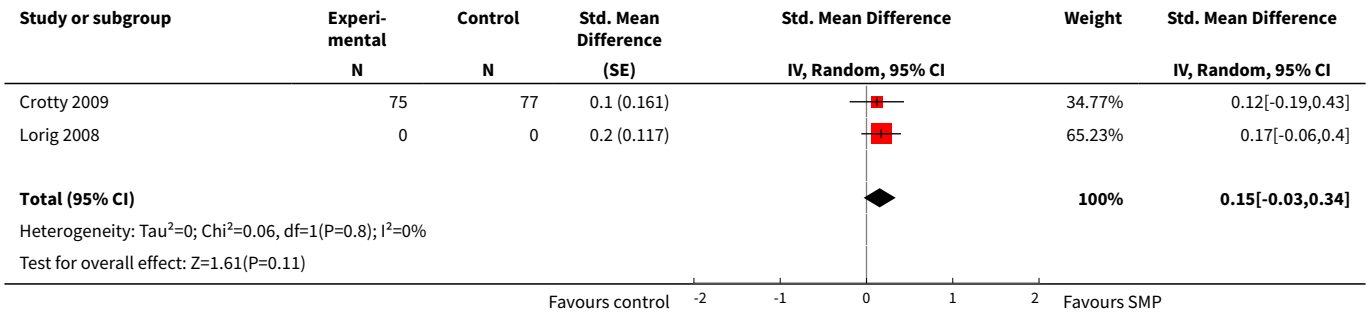


Analysis 2.13. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 13 Social integration and support.





Analysis 2.14. Comparison 2 SMP versus usual care/no treatment/wait list, Outcome 14 Health service navigation.



Comparison 3. SMP versus information only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Self-management of OA	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Short term	1	90	Std. Mean Difference (IV, Random, 95% CI)	0.06 [-0.36, 0.47]
1.2 Intermediate term	3	760	Std. Mean Difference (IV, Random, 95% CI)	0.20 [-0.04, 0.44]
2 Engagement in life	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2.1 Short term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2.2 Intermediate term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
3 Pain	3	751	Std. Mean Difference (IV, Random, 95% CI)	-0.07 [-0.21, 0.08]
4 Global OA scores	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1 Short term	1	89	Std. Mean Difference (IV, Random, 95% CI)	0.09 [-0.33, 0.50]
4.2 Intermediate term	3	751	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.28, 0.16]

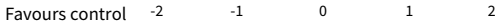
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5 Function—self-reported	4	854	Std. Mean Difference (IV, Random, 95% CI)	-0.09 [-0.22, 0.05]
6 Function—performance	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
7 Quality of life	2	648	Std. Mean Difference (IV, Random, 95% CI)	0.05 [-0.10, 0.21]
8 Withdrawals	4	1251	Risk Ratio (M-H, Random, 95% CI)	1.60 [0.75, 3.40]
9 Emotional distress	3	775	Std. Mean Difference (IV, Random, 95% CI)	-0.00 [-0.30, 0.30]
10 Health-directed activity	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
10.1 Short term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
10.2 Intermediate term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
11 Social integration and support	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
11.1 Short term	1	90	Std. Mean Difference (IV, Random, 95% CI)	-0.02 [-0.44, 0.40]
11.2 Intermediate term	2	181	Std. Mean Difference (IV, Random, 95% CI)	-0.02 [-0.39, 0.35]
12 Health service navigation	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
12.1 Short term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
12.2 Intermediate term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
13 Skill and technique acquisition	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
13.1 Short term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
13.2 Intermediate term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
14 Constructive attitudes and approaches	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
14.1 Short term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
14.2 Intermediate term	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

Analysis 3.1. Comparison 3 SMP versus information only, Outcome 1 Self-management of OA.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
3.1.1 Short term							
Ackerman 2012	39	4.9 (0.4)	51	4.9 (0.5)		100%	0.06[-0.36,0.47]
Subtotal ***	**39**		**51**			**100%**	**0.06[-0.36,0.47]**
Heterogeneity: Tau²=0; Chi²=0, df=0(P<0.0001); I²=100%							
Test for overall effect: Z=0.26(P=0.79)							
3.1.2 Intermediate term							
Hughes 2004	58	74.5 (19.6)	32	64 (20.3)		21.34%	0.53[0.09,0.97]
Ackerman 2012	38	4.9 (0.5)	55	4.9 (0.5)		23.2%	-0.03[-0.44,0.38]
Buszewicz 2006	278	19.9 (6.4)	299	18.8 (6.5)		55.46%	0.17[0.01,0.33]
Subtotal ***	**374**		**386**			**100%**	**0.2[-0.04,0.44]**
Heterogeneity: Tau²=0.02; Chi²=3.41, df=2(P=0.18); I²=41.31%							
Test for overall effect: Z=1.63(P=0.1)							

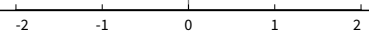
Analysis 3.2. Comparison 3 SMP versus information only, Outcome 2 Engagement in life.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
3.2.1 Short term						
Ackerman 2012	39	4.6 (0.8)	51	4.7 (0.8)		-0.08[-0.41,0.26]
3.2.2 Intermediate term						
Ackerman 2012	38	4.6 (0.9)	55	4.8 (0.9)		-0.2[-0.59,0.18]



Analysis 3.3. Comparison 3 SMP versus information only, Outcome 3 Pain.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Hughes 2004	58	5.4 (3.7)	32	5.3 (4.4)		11.16%	0.02[-0.41,0.45]
Ackerman 2012	38	38.9 (24.6)	55	37.2 (21.9)		12.15%	0.07[-0.34,0.49]
Buszewicz 2006	277	8.1 (3.9)	291	8.5 (3.9)		76.69%	-0.1[-0.27,0.06]
Total ***	**373**		**378**			**100%**	**-0.07[-0.21,0.08]**
Heterogeneity: Tau²=0; Chi²=0.79, df=2(P=0.67); I²=0%							
Test for overall effect: Z=0.91(P=0.36)							



Analysis 3.4. Comparison 3 SMP versus information only, Outcome 4 Global OA scores.

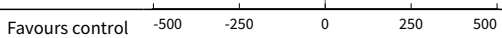
Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
3.4.1 Short term							
Ackerman 2012	39	18.1 (27)	50	16 (22.7)		100%	0.09[-0.33,0.5]
Subtotal ***	**39**		**50**			**100%**	**0.09[-0.33,0.5]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.4(P=0.69)							
3.4.2 Intermediate term							
Hughes 2004	58	25.8 (11.9)	32	28.4 (15.5)		19.87%	-0.2[-0.63,0.24]
Ackerman 2012	37	24.4 (27)	56	18.2 (22.7)		21.03%	0.25[-0.16,0.67]
Buszewicz 2006	277	39.4 (13.6)	291	41.1 (14.1)		59.1%	-0.12[-0.29,0.04]
Subtotal ***	**372**		**379**			**100%**	**-0.06[-0.28,0.16]**
Heterogeneity: Tau^2=0.01; Chi^2=2.99, df=2(P=0.22); I^2=33.08%							
Test for overall effect: Z=0.52(P=0.6)							

Analysis 3.5. Comparison 3 SMP versus information only, Outcome 5 Function—self-reported.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Hughes 2004	58	17.8 (11.2)	32	20.2 (14.7)		9.77%	-0.19[-0.62,0.25]
Wetzels 2005	40	14.6 (4.5)	48	14.4 (4.7)		10.38%	0.03[-0.39,0.45]
Ackerman 2012	38	38.1 (23.8)	54	39.8 (21.5)		10.6%	-0.08[-0.49,0.34]
Buszewicz 2006	279	27.7 (12.9)	305	28.9 (13.4)		69.25%	-0.09[-0.25,0.07]
Total ***	**415**		**439**			**100%**	**-0.09[-0.22,0.05]**
Heterogeneity: Tau^2=0; Chi^2=0.52, df=3(P=0.91); I^2=0%							
Test for overall effect: Z=1.24(P=0.21)							

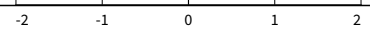
Analysis 3.6. Comparison 3 SMP versus information only, Outcome 6 Function—performance.

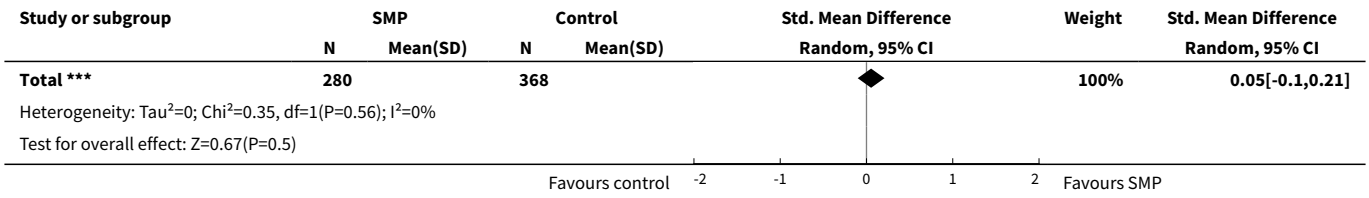
Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Hughes 2004	58	1281.5 (502.9)	32	1106.3 (484.1)		175.24[-36.62,387.1]



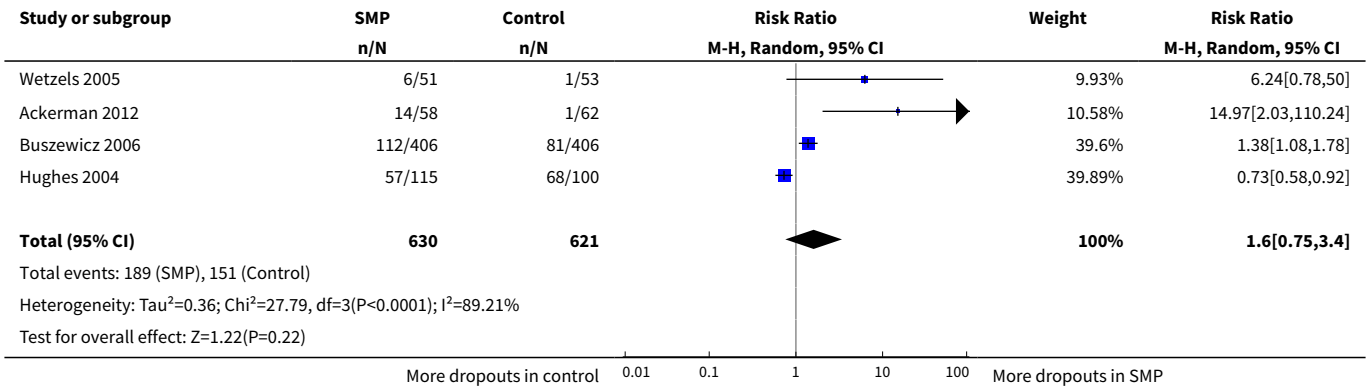
Analysis 3.7. Comparison 3 SMP versus information only, Outcome 7 Quality of life.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Ackerman 2012	38	0.6 (0.3)	56	0.6 (0.2)		14.25%	-0.06[-0.47,0.35]
Buszewicz 2006	242	0.6 (0.3)	312	0.6 (0.3)		85.75%	0.07[-0.1,0.24]

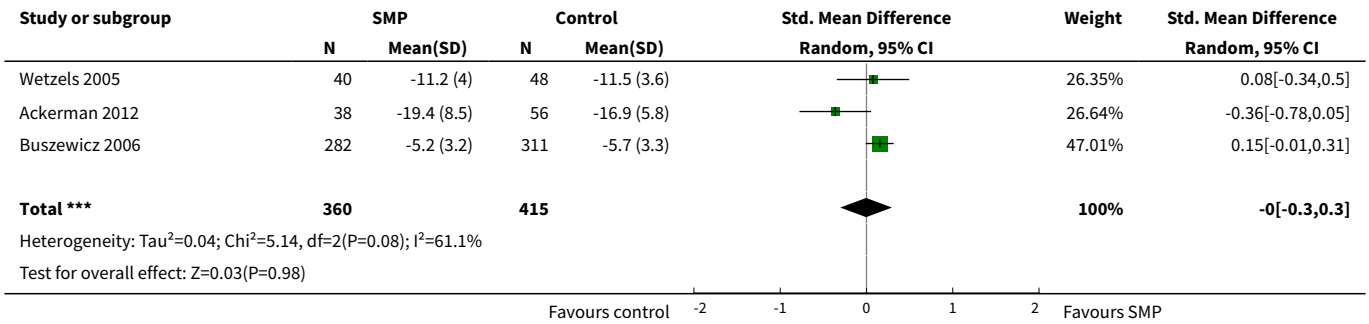




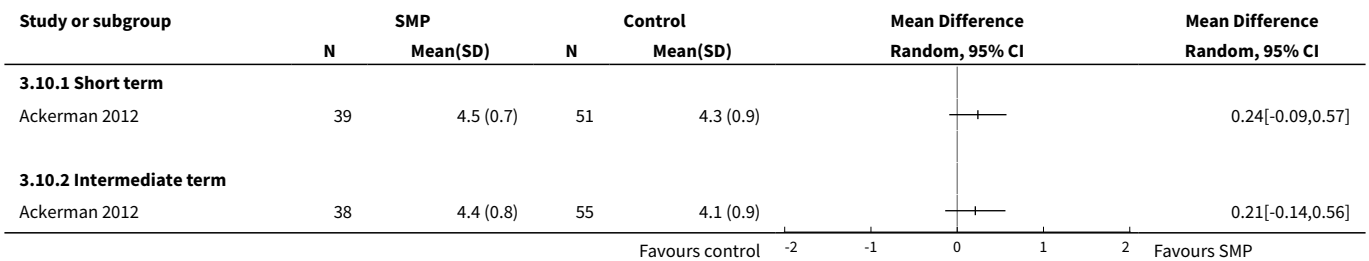
Analysis 3.8. Comparison 3 SMP versus information only, Outcome 8 Withdrawals.



Analysis 3.9. Comparison 3 SMP versus information only, Outcome 9 Emotional distress.

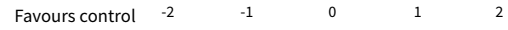


Analysis 3.10. Comparison 3 SMP versus information only, Outcome 10 Health-directed activity.



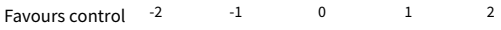
Analysis 3.11. Comparison 3 SMP versus information only, Outcome 11 Social integration and support.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
3.11.1 Short term							
Ackerman 2012	39	4.6 (0.6)	51	4.7 (0.8)		100%	-0.02[-0.44,0.4]
Subtotal ***	**39**		**51**			**100%**	**-0.02[-0.44,0.4]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.09(P=0.93)							
3.11.2 Intermediate term							
Wetzels 2005	40	-11.4 (2.9)	48	-11.9 (2.8)		49.55%	0.17[-0.25,0.59]
Ackerman 2012	38	4.5 (1)	55	4.7 (0.8)		50.45%	-0.21[-0.62,0.21]
Subtotal ***	**78**		**103**			**100%**	**-0.02[-0.39,0.35]**
Heterogeneity: Tau²=0.02; Chi²=1.54, df=1(P=0.21); I²=35.06%							
Test for overall effect: Z=0.11(P=0.91)							



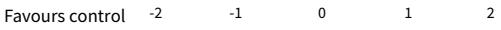
Analysis 3.12. Comparison 3 SMP versus information only, Outcome 12 Health service navigation.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
3.12.1 Short term						
Ackerman 2012	39	4.8 (0.7)	51	4.8 (0.8)		-0.04[-0.33,0.26]
3.12.2 Intermediate term						
Ackerman 2012	38	5 (0.5)	55	4.9 (0.9)		0.14[-0.13,0.41]



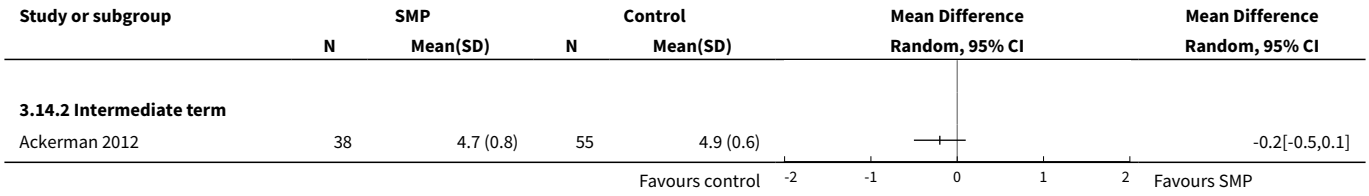
Analysis 3.13. Comparison 3 SMP versus information only, Outcome 13 Skill and technique acquisition.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
3.13.1 Short term						
Ackerman 2012	39	4.5 (0.7)	51	4.4 (0.7)		0.15[-0.13,0.44]
3.13.2 Intermediate term						
Ackerman 2012	38	4.6 (0.7)	55	4.6 (0.7)		-0.06[-0.35,0.23]



Analysis 3.14. Comparison 3 SMP versus information only, Outcome 14 Constructive attitudes and approaches.

Study or subgroup	SMP N	SMP Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
3.14.1 Short term						
Ackerman 2012	39	4.9 (0.6)	51	4.8 (0.8)		0.16[-0.12,0.44]

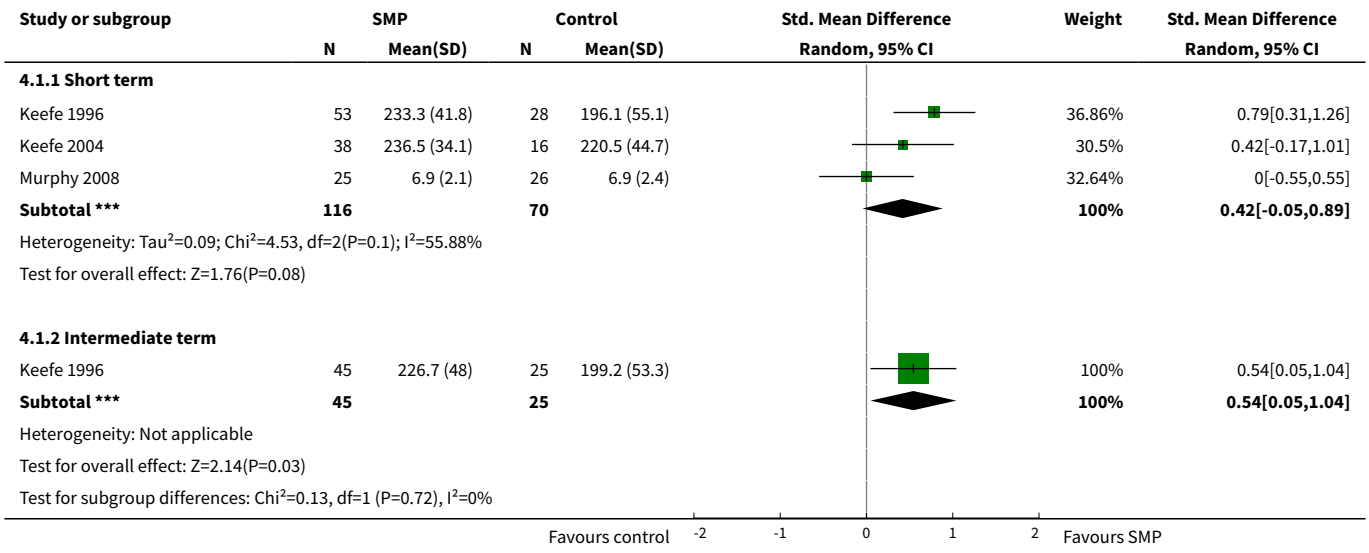


Comparison 4. SMP versus non-SMP intervention

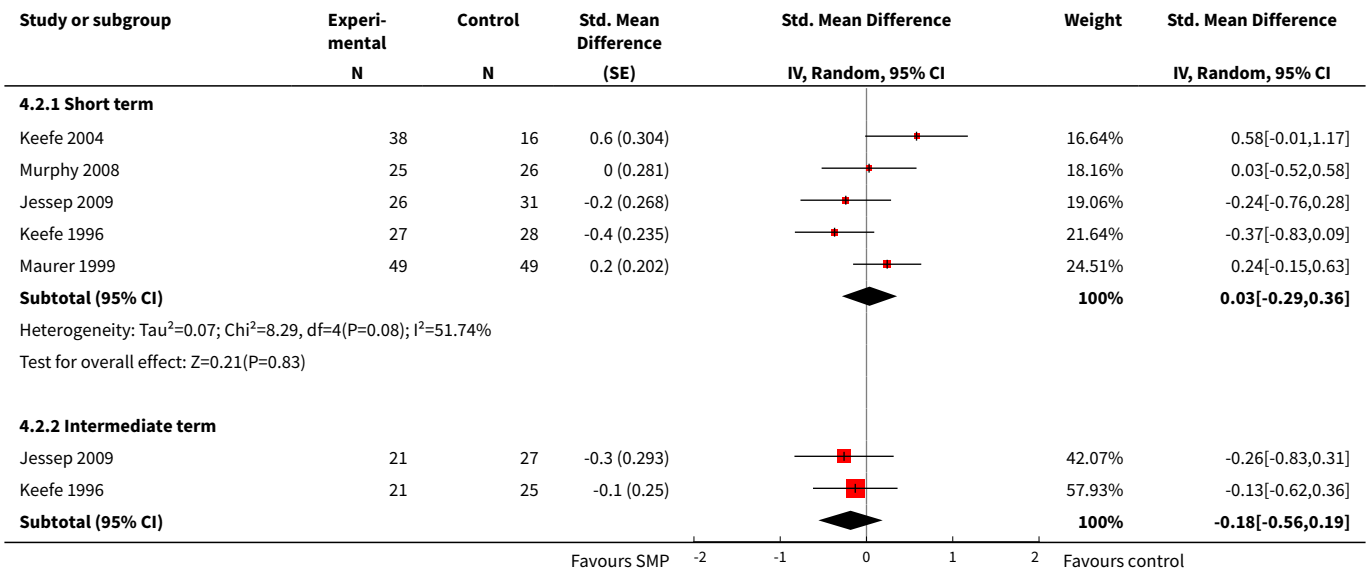
Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Self-management of OA	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Short term	3	186	Std. Mean Difference (IV, Random, 95% CI)	0.42 [-0.05, 0.89]
1.2 Intermediate term	1	70	Std. Mean Difference (IV, Random, 95% CI)	0.54 [0.05, 1.04]
2 Pain	5		Std. Mean Difference (Random, 95% CI)	Subtotals only
2.1 Short term	5		Std. Mean Difference (Random, 95% CI)	0.03 [-0.29, 0.36]
2.2 Intermediate term	2		Std. Mean Difference (Random, 95% CI)	-0.18 [-0.56, 0.19]
3 Global OA scores	1		Std. Mean Difference (Random, 95% CI)	Totals not selected
3.1 Short term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
4 Function—self-reported	3		Std. Mean Difference (Random, 95% CI)	Subtotals only
4.1 Short term	3		Std. Mean Difference (Random, 95% CI)	0.23 [-0.03, 0.48]
4.2 Intermediate term	2		Std. Mean Difference (Random, 95% CI)	-0.17 [-0.54, 0.20]
5 Function—performance	1		Std. Mean Difference (Random, 95% CI)	Totals not selected
6 Quality of life	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
6.1 Short term	1	57	Std. Mean Difference (IV, Random, 95% CI)	0.24 [-0.28, 0.76]
6.2 Intermediate term	2	243	Std. Mean Difference (IV, Random, 95% CI)	-0.01 [-0.28, 0.26]
6.3 Long term	1	178	Std. Mean Difference (IV, Random, 95% CI)	-0.23 [-0.55, 0.10]
7 Withdrawals	7	919	Risk Ratio (M-H, Random, 95% CI)	0.86 [0.69, 1.09]
8 Emotional distress	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
8.1 Short term	3	192	Std. Mean Difference (IV, Random, 95% CI)	0.13 [-0.28, 0.55]
8.2 Intermediate term	2	118	Std. Mean Difference (IV, Random, 95% CI)	0.18 [-0.19, 0.55]

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
9 Constructive atti-tudes and approaches	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
9.1 Short term	2	135	Std. Mean Difference (IV, Random, 95% CI)	0.92 [0.49, 1.34]
9.2 Intermediate term	1	70	Std. Mean Difference (IV, Random, 95% CI)	0.62 [0.12, 1.12]

Analysis 4.1. Comparison 4 SMP versus non-SMP intervention, Outcome 1 Self-management of OA.



Analysis 4.2. Comparison 4 SMP versus non-SMP intervention, Outcome 2 Pain.



Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
Heterogeneity: Tau²=0; Chi²=0.11, df=1(P=0.74); I²=0%						
Test for overall effect: Z=0.97(P=0.33)						
Test for subgroup differences: Chi²=0.77, df=1 (P=0.38), I²=0%						

Favours SMP -2 -1 0 1 2 Favours control

Analysis 4.3. Comparison 4 SMP versus non-SMP intervention, Outcome 3 Global OA scores.

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
4.3.1 Short term					
Maurer 1999	49	49	0.3 (0.204)		0.27[-0.13,0.67]

Favours SMP -2 -1 0 1 2 Favours control

Analysis 4.4. Comparison 4 SMP versus non-SMP intervention, Outcome 4 Function—self-reported.

Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
4.4.1 Short term						
Jessep 2009	26	31	0 (0.265)		24.93%	0.02[-0.5,0.54]
Keefe 1996	27	28	0.2 (0.235)		31.86%	0.23[-0.23,0.69]
Maurer 1999	49	49	0.3 (0.202)		43.21%	0.34[-0.05,0.73]
Subtotal (95% CI)					**100%**	**0.23[-0.03,0.48]**
Heterogeneity: Tau²=0; Chi²=0.92, df=2(P=0.63); I²=0%						
Test for overall effect: Z=1.7(P=0.09)						
4.4.2 Intermediate term						
Jessep 2009	21	27	-0 (0.291)		42.5%	-0.05[-0.62,0.52]
Keefe 1996	21	25	-0.3 (0.25)		57.5%	-0.26[-0.75,0.23]
Subtotal (95% CI)					**100%**	**-0.17[-0.54,0.2]**
Heterogeneity: Tau²=0; Chi²=0.3, df=1(P=0.58); I²=0%						
Test for overall effect: Z=0.9(P=0.37)						
Test for subgroup differences: Chi²=2.93, df=1 (P=0.09), I²=65.88%						

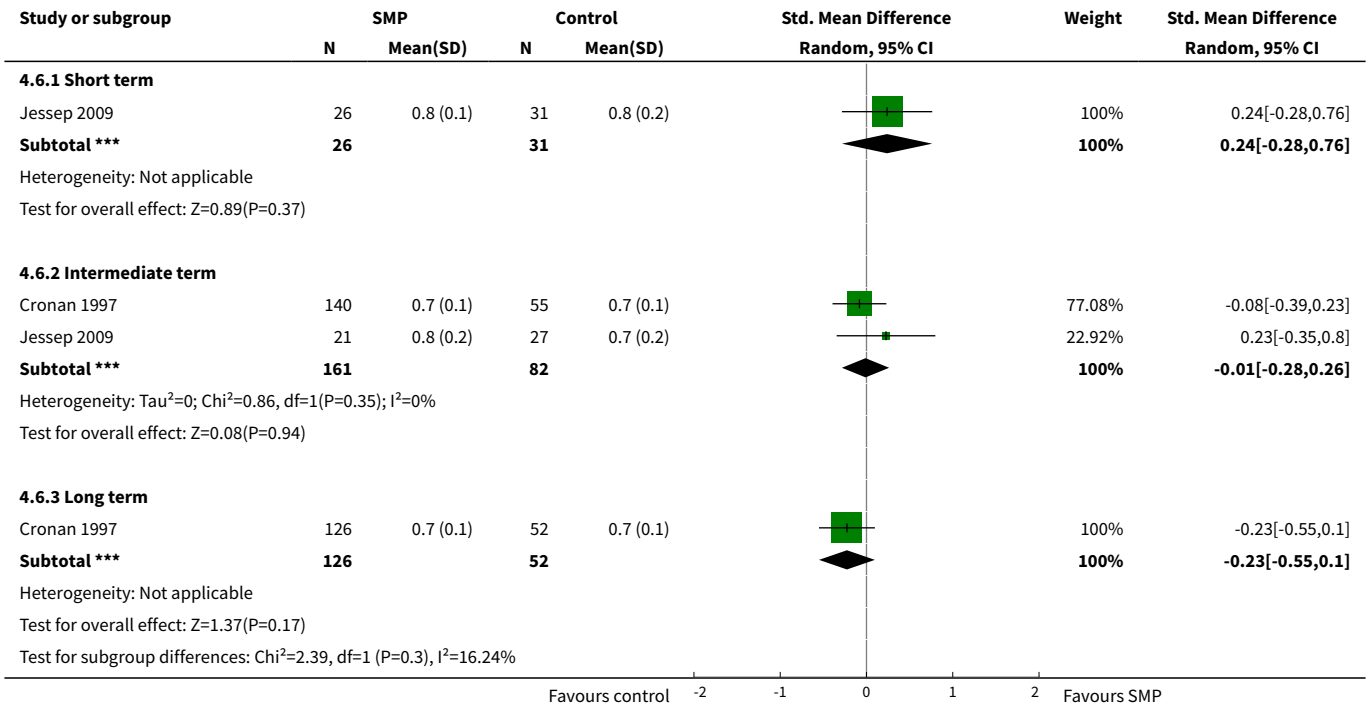
Favours SMP -2 -1 0 1 2 Favours control

Analysis 4.5. Comparison 4 SMP versus non-SMP intervention, Outcome 5 Function—performance.

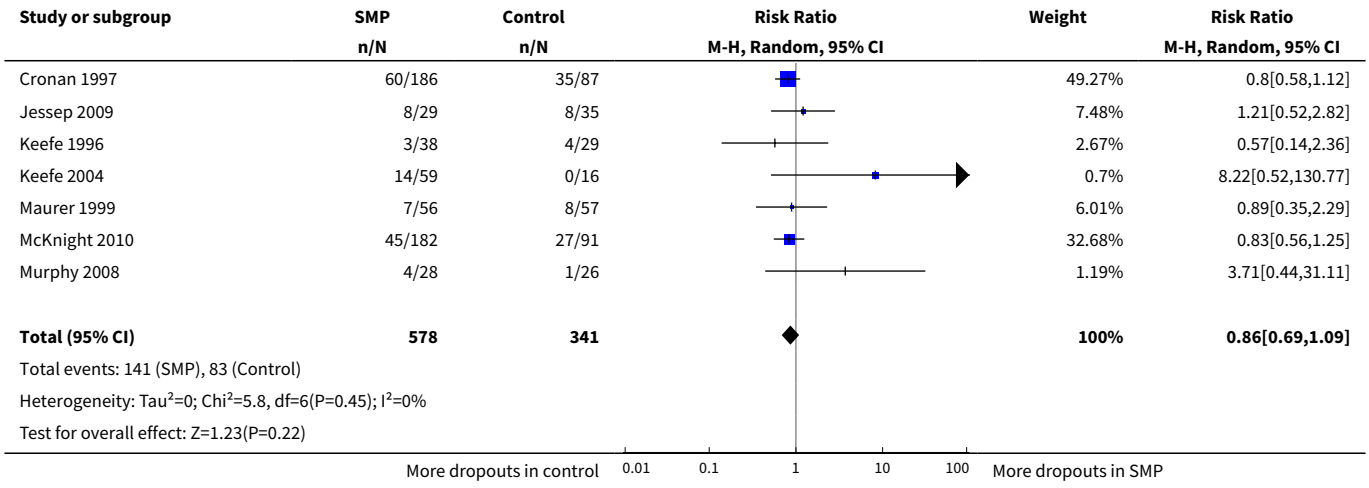
Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
Murphy 2008	25	26	-0.1 (0.281)		-0.09[-0.64,0.46]

Favours control -2 -1 0 1 2 Favours SMP

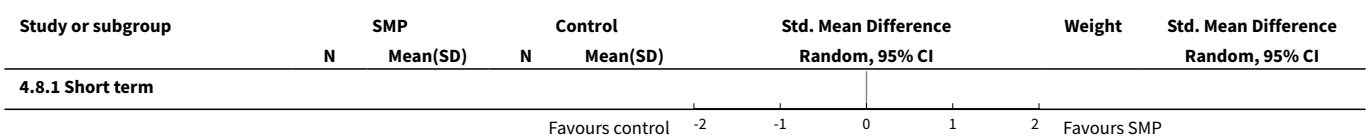
Analysis 4.6. Comparison 4 SMP versus non-SMP intervention, Outcome 6 Quality of life.

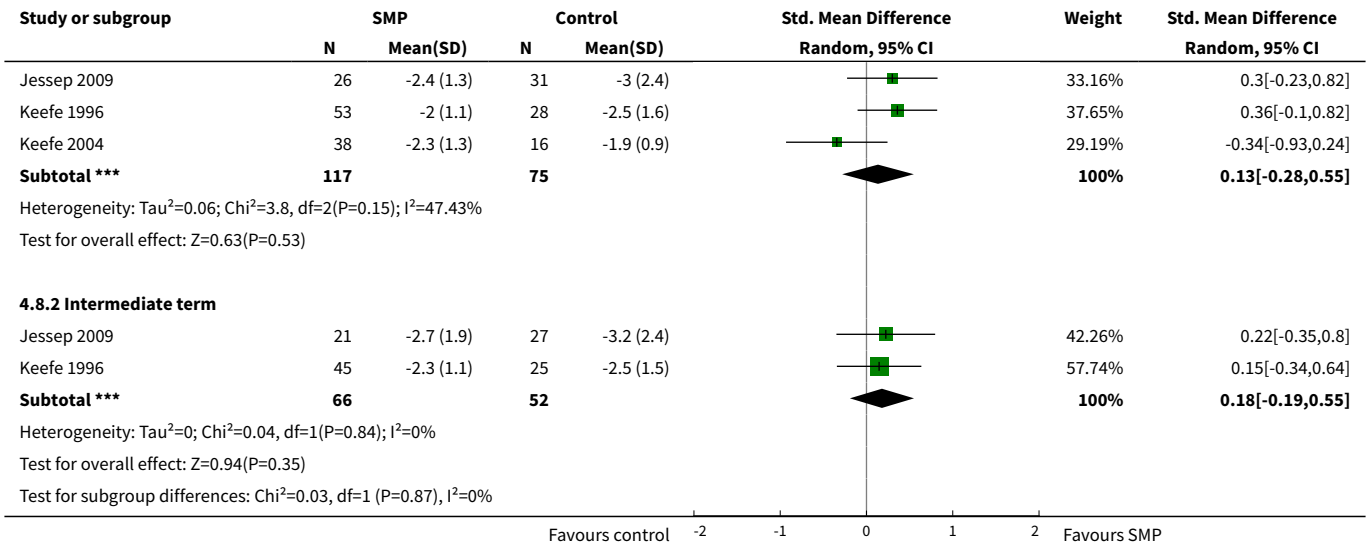


Analysis 4.7. Comparison 4 SMP versus non-SMP intervention, Outcome 7 Withdrawals.

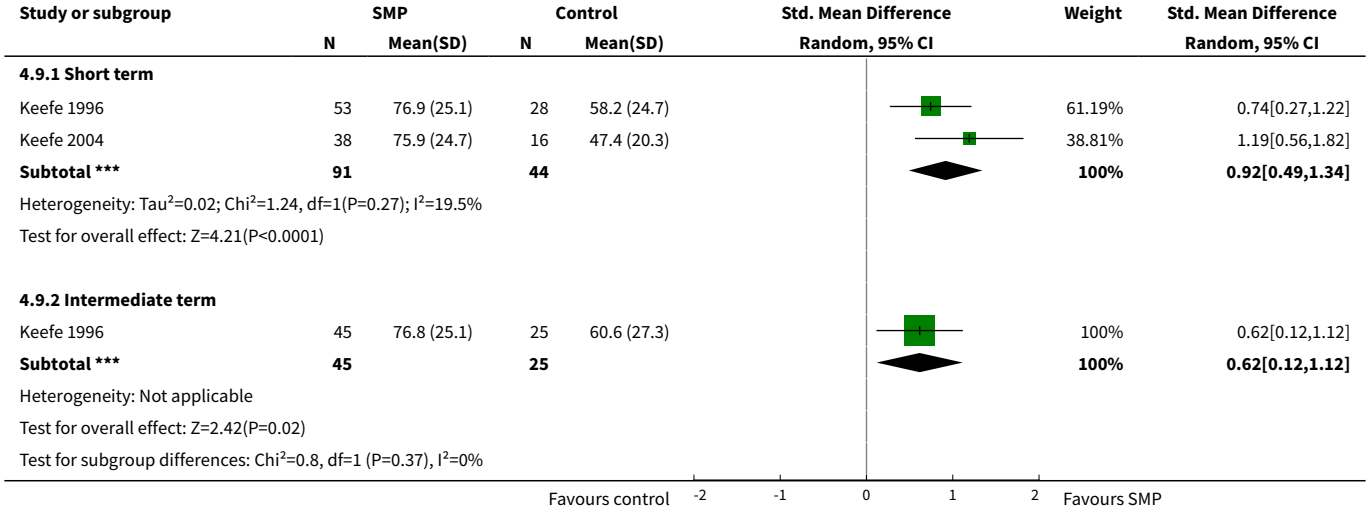


Analysis 4.8. Comparison 4 SMP versus non-SMP intervention, Outcome 8 Emotional distress.





Analysis 4.9. Comparison 4 SMP versus non-SMP intervention, Outcome 9 Constructive attitudes and approaches.



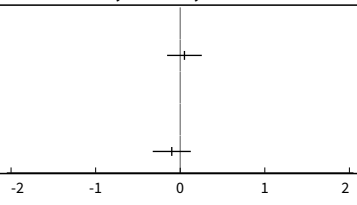
Comparison 5. SMP versus acupuncture

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Global OA scores	1		Std. Mean Difference (Random, 95% CI)	Totals not selected
1.1 Short term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
1.2 Intermediate term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
2 Pain	1		Std. Mean Difference (Random, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Short term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
2.2 Intermediate term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
3 Function self-reported	1		Std. Mean Difference (Random, 95% CI)	Totals not selected
3.1 Short term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
3.2 Intermediate term	1		Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
4 Function performance	1		Std. Mean Difference (Random, 95% CI)	Totals not selected
5 Withdrawals	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

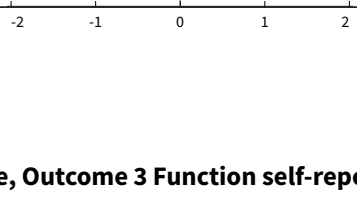
Analysis 5.1. Comparison 5 SMP versus acupuncture, Outcome 1 Global OA scores.

Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
5.1.1 Short term					
Berman 2004	0	0	0.1 (0.105)		0.05[-0.15,0.25]
5.1.2 Intermediate term					
Berman 2004	0	0	-0.1 (0.115)		-0.1[-0.32,0.12]



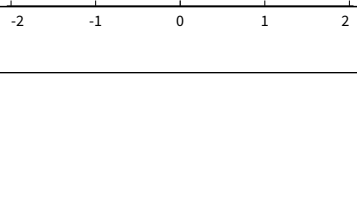
Analysis 5.2. Comparison 5 SMP versus acupuncture, Outcome 2 Pain.

Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
5.2.1 Short term					
Berman 2004	0	0	1 (0.11)		0.95[0.74,1.16]
5.2.2 Intermediate term					
Berman 2004	0	0	1.4 (0.122)		1.37[1.13,1.61]



Analysis 5.3. Comparison 5 SMP versus acupuncture, Outcome 3 Function self-reported.

Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
5.3.1 Short term					
Berman 2004	0	0	1.2 (0.112)		1.22[1,1.44]



Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
5.3.2 Intermediate term					
Berman 2004	0	0	1.5 (0.125)		1.53[1.29,1.77]
			Favours SMP	-2 -1 0 1 2	Favours control

Analysis 5.4. Comparison 5 SMP versus acupuncture, Outcome 4 Function performance.

Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
Berman 2004	0	0	-0.3 (0.115)		-0.3[-0.52,-0.08]
			Favours control	-2 -1 0 1 2	Favours SMP

Analysis 5.5. Comparison 5 SMP versus acupuncture, Outcome 5 Withdrawals.

Study or subgroup	SMP n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Berman 2004	99/189	102/381		1.96[1.58,2.42]
		More dropouts in control	0.01 0.1 1 10 100	More dropouts in SMP

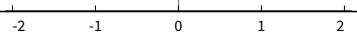
Comparison 6. Subgroup analysis

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Self-management in OA	7		Std. Mean Difference (Random, 95% CI)	0.14 [-0.08, 0.36]
1.1 White, educated, female	3		Std. Mean Difference (Random, 95% CI)	0.29 [0.07, 0.50]
1.2 NOT white, educated, female	4		Std. Mean Difference (Random, 95% CI)	0.03 [-0.29, 0.36]
2 Function self-reported	8		Std. Mean Difference (Random, 95% CI)	-0.12 [-0.23, -0.01]
2.1 White, educated, female	3		Std. Mean Difference (Random, 95% CI)	-0.20 [-0.37, -0.02]
2.2 NOT white, educated, female	5		Std. Mean Difference (Random, 95% CI)	-0.06 [-0.21, 0.08]
3 Pain	9		Std. Mean Difference (Random, 95% CI)	-0.17 [-0.28, -0.06]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 White, educated, female	4		Std. Mean Difference (Random, 95% CI)	-0.11 [-0.30, 0.07]
3.2 NOT white, educated, female	5		Std. Mean Difference (Random, 95% CI)	-0.20 [-0.35, -0.05]
4 Withdrawals	12	3095	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.81, 1.34]
4.1 White, educated, female	7	2129	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.76, 1.45]
4.2 NOT white, educated, female	5	966	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.63, 1.69]

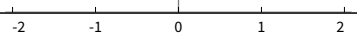
Analysis 6.1. Comparison 6 Subgroup analysis, Outcome 1 Self-management in OA.

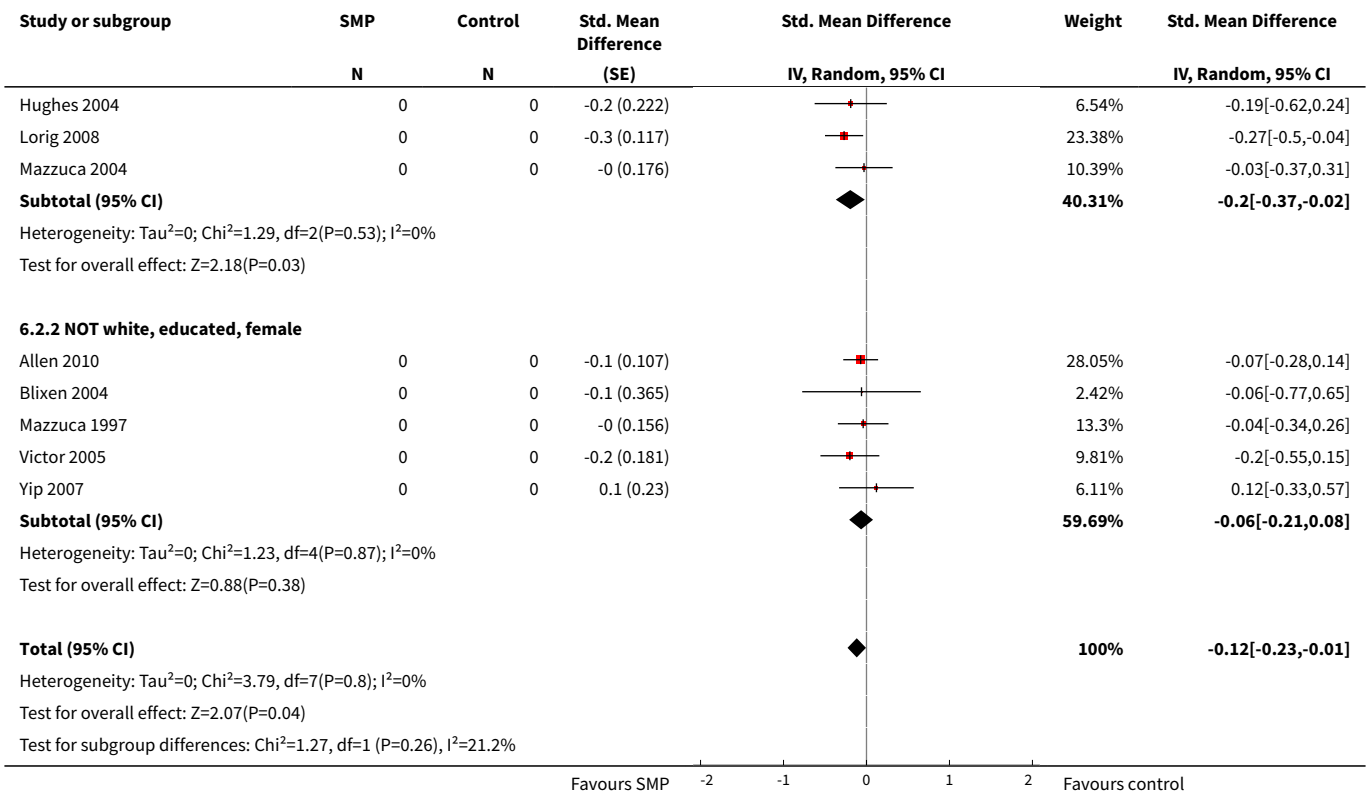
Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
6.1.1 White, educated, female						
Murphy 2008	0	0	0 (0.281)		9.96%	0[-0.55,0.55]
Hughes 2004	0	0	0.5 (0.224)		12.9%	0.53[0.09,0.97]
Lorig 2008	0	0	0.3 (0.12)		20.6%	0.27[0.04,0.5]
Subtotal (95% CI)					**43.46%**	**0.29[0.07,0.5]**
Heterogeneity: Tau^2=0; Chi^2=2.24, df=2(P=0.33); I^2=10.63%						
Test for overall effect: Z=2.61(P=0.01)						
6.1.2 NOT white, educated, female						
Blixen 2004	0	0	0.2 (0.367)		6.87%	0.19[-0.53,0.91]
Yip 2007	0	0	0.4 (0.232)		12.44%	0.39[-0.06,0.84]
Victor 2005	0	0	-0.4 (0.184)		15.59%	-0.4[-0.76,-0.04]
Allen 2010	0	0	0.1 (0.107)		21.64%	0.08[-0.13,0.29]
Subtotal (95% CI)					**56.54%**	**0.03[-0.29,0.36]**
Heterogeneity: Tau^2=0.07; Chi^2=8.29, df=3(P=0.04); I^2=63.82%						
Test for overall effect: Z=0.2(P=0.84)						
Total (95% CI)					**100%**	**0.14[-0.08,0.36]**
Heterogeneity: Tau^2=0.05; Chi^2=14.57, df=6(P=0.02); I^2=58.83%						
Test for overall effect: Z=1.26(P=0.21)						
Test for subgroup differences: Chi^2=1.63, df=1 (P=0.2), I^2=38.71%						



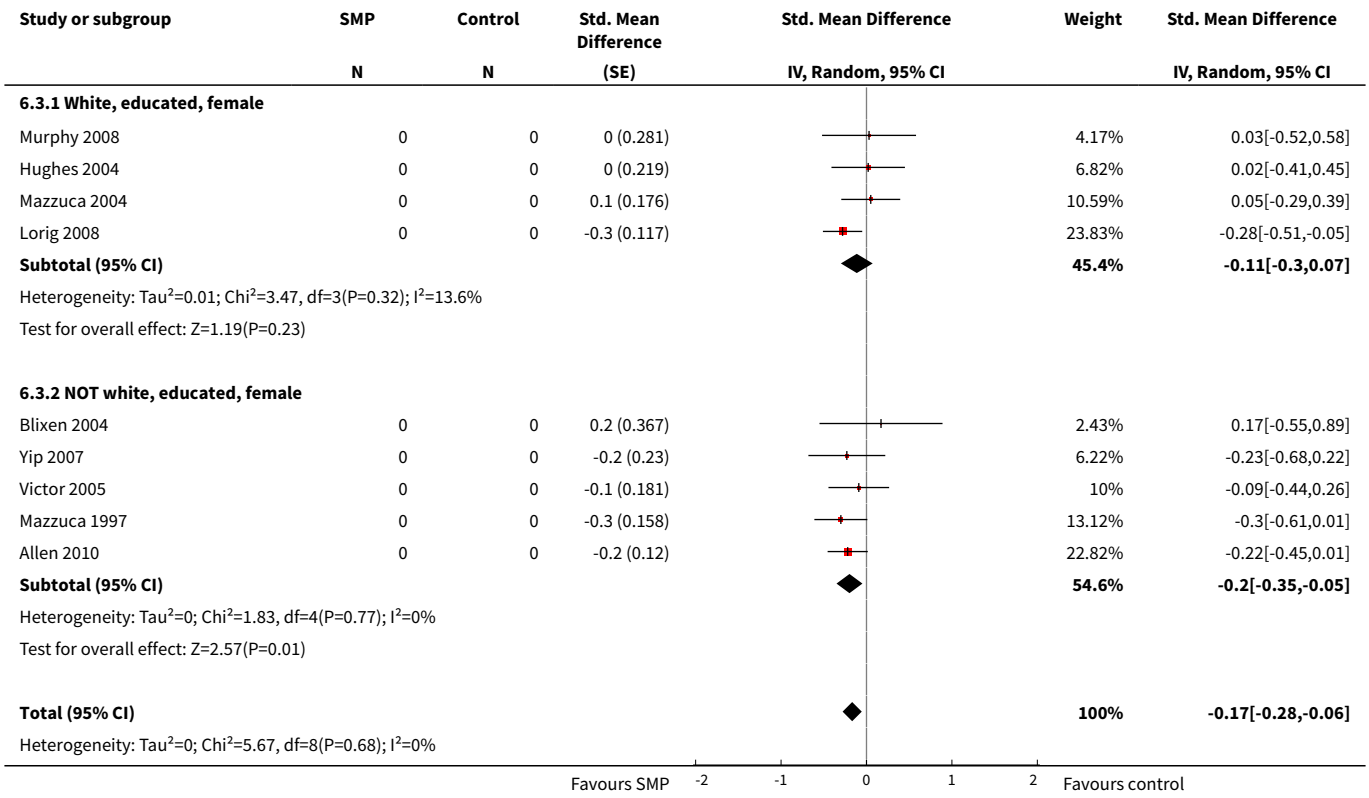
Analysis 6.2. Comparison 6 Subgroup analysis, Outcome 2 Function self-reported.

Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
6.2.1 White, educated, female						

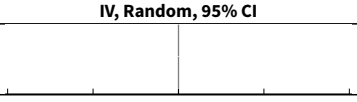




Analysis 6.3. Comparison 6 Subgroup analysis, Outcome 3 Pain.

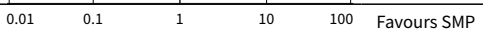


Study or subgroup	SMP N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV, Random, 95% CI	Weight	Std. Mean Difference IV, Random, 95% CI
Test for overall effect: Z=2.93(P=0)						
Test for subgroup differences: Chi²=0.49, df=1 (P=0.49), I²=0%						



Analysis 6.4. Comparison 6 Subgroup analysis, Outcome 4 Withdrawals.

Study or subgroup	SMP n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
6.4.1 White, educated, female					
Murphy 2008	4/28	1/26		1.3%	3.71[0.44,31.11]
Maisiak 1996	7/135	8/135		4.54%	0.88[0.33,2.35]
Mazzuca 2004	31/111	21/75		9.63%	1[0.62,1.6]
McKnight 2010	45/182	27/91		10.52%	0.83[0.56,1.25]
Cronan 1997	60/186	24/90		10.58%	1.21[0.81,1.81]
Lorig 2008	126/433	78/422		12.64%	1.57[1.23,2.02]
Hughes 2004	57/115	68/100		12.88%	0.73[0.58,0.92]
Subtotal (95% CI)	**1190**	**939**		**62.08%**	**1.05[0.76,1.45]**
Total events: 330 (SMP), 227 (Control)					
Heterogeneity: Tau²=0.12; Chi²=24.43, df=6(P=0); I²=75.44%					
Test for overall effect: Z=0.28(P=0.78)					
6.4.2 NOT white, educated, female					
Blixen 2004	1/16	1/16		0.84%	1[0.07,14.64]
Allen 2010	28/174	17/174		8.4%	1.65[0.94,2.9]
Mazzuca 1997	19/105	21/106		8.47%	0.91[0.52,1.6]
Yip 2007	21/88	41/94		10.06%	0.55[0.35,0.85]
Victor 2005	48/120	20/73		10.14%	1.46[0.95,2.25]
Subtotal (95% CI)	**503**	**463**		**37.92%**	**1.04[0.63,1.69]**
Total events: 117 (SMP), 100 (Control)					
Heterogeneity: Tau²=0.19; Chi²=13.32, df=4(P=0.01); I²=69.97%					
Test for overall effect: Z=0.14(P=0.89)					
Total (95% CI)	**1693**	**1402**		**100%**	**1.04[0.81,1.34]**
Total events: 447 (SMP), 327 (Control)					
Heterogeneity: Tau²=0.12; Chi²=37.48, df=11(P<0.0001); I²=70.65%					
Test for overall effect: Z=0.31(P=0.76)					
Test for subgroup differences: Chi²=0, df=1 (P=0.97), I²=0%					



ADDITIONAL TABLES

Table 1. PROGRESS-Plus

Study	P	R	O	Gr	E	S	S	A	D	HL
Allen 2010	US	54% white, 46% non-white, 2% Hispanic	38.3% employed	92.7% male, 7.3% female	33.3% high school education or less	25.7% inadequate income	69.3% married	60.1 (10.4)	32.3% fair or poor health	-
Berman 2004	US	69.1% white, 28.6% African American, 2.3% other	-	35.9% male, 64.1% female	31.1% no college, 68.9% some college	-	-	65.5 (8.6)	75.1% moderate to lesser pain, 24.9% severe to extreme pain	-
Blixen 2004	US	72% white, 28% African American	-	62.5% male, 37.5% female	6% grade 7 to 9, 3% grade 10 and 11, 22% high school, 50% 1 to 4 year college, 9.5% college graduate, 9.5% professional/graduate school	22% < $10,000/y, 13% $10,000 to $19,999/y, 19% $20,000 to $29,999/y, 17.5% $30,000 to $39,999/y, 13.5% $40,000 to $49,999/y, 17.5% > $50,000/y	56.5% married, 25% divorced, 13% single, 25% widowed	70.8 (6.1)	16% joint replacement surgery	Domain social support for health
Buszewicz 2006	UK	> 99% white, < 1% black Caribbean, < 1% black African	-	37% male, 63% female	72.5% no higher education, 27.5% higher education	-	81% house owner, 1% staying with friends or fami-	68.6 (8.4)	-	-

Table 1. PROGRESS-Plus *(Continued)*

							ly, 18.5% rented accommodation			
Calfas 1992	US	97.5% white, 2.5% non-white	82.5% retired, 5% full-time employed, 12.5% part-time employed	27.5% male, 72.5% female	-	-	70% married, 20% divorced, 10% widowed, 5% other	67.0 (7.3)	-	-
Cronan 1997	US	92.3% white, 2.5% African American, 2.2% Hispanic, 1.7% Asian, 1.4% other	-	36.3% male, 64.0% female	31.2% high school, 37.7% some college, 28.7% college degree	27.2% < $20,000, 43.2% $20,000 to $40,000, 19.7% $40,000 to $60,000, 10.0% > $60,000	-	69.2 (5.6)	69.9% other medical conditions present	-
Crotty 2009	Australia	-	11.6% employed, 5.4% home duties, 82.7% retired	39.5% male, 60.5% female	3.5% none to some primary school, 15.0% primary school, 30.6% high school to year 8, 27.2% high school to year 12, 17.4% TAFE/trade, 6.2% university to above	-	48.3% live alone	67.5 (10.8)	32.9% on waiting list for hip replacement	Domains social support for health, navigating health system, actively managing my health
Hansson 2010	Sweden	-	-	-	-	-	-	62.5 (9.4)	-	-
Heuts 2005	The Netherlands	-	40.5% paying job,	40.5% male, 59.5% female	25.5% low, 30.5% middle, 20% high	-	-	51.6 (5.1)	-	-

Table 1. PROGRESS-Plus *(Continued)*

			35.5% no paying job							
Hopman-Rock 2000	The Netherlands	-	-	17% male, 83% female	21.5% primary, 49.5% secondary, 23.5% college/university	-	67.5% living together/married, 27% living alone	65.3 (5.5)	2.5 (1.6) other chronic conditions	-
Hughes 2004	US	72.2% white, 22.1% African American, 2.6% Hispanic, 2.1% Asian-Pacific Islander, 1.1% other	-	16.8% male, 83.3% female	10.5% less than high school, 20.0% high school, 69.6% more than high school	33.1% income less than $20,000	-	73.4	22.4% ARA class I, 64.4% ARA class II, 13.3% ARA class III	-
Hurley 2007	UK	-	-	29.7% male, 70.3% female	-	-	-	67 (range 50 to 91)	-	-
Jessep 2009	UK	-	-	30.5% male, 69.5% female	-	-	-	66.5 (range 51 to 81)	-	-
Keefe 1990	US	-	-	28.3% male, 71.7% female	-	-	-	63.9 (11.5)	-	-
Keefe 1996	US	-	-	40% male, 60% female	-	-	-	62.6	-	-
Keefe 2004	US	-	-	46.9% male, 53.4% female	-	-	-	59.5 (11.1)	-	-

Table 1. PROGRESS-Plus *(Continued)*

Lorig 2008	US	92.3% non-Hispanic white	-	9.9% male, 90.2% female	15.7 (3.1) years of education	-	68.3% married	52.4 (11.6)	-	Domain ability to actively engage with healthcare providers
Maisiak 1996	US	85% white	-	7.7% male, 92.3% female	12.1 years of schooling	-	-	60.4	48% see a specialist	-
Martire 2007	US	-	-	27.3% male, 72.7% female	-	-	-	68.6 (7.6)	-	-
Maurer 1999	US	-	-	58.5% male, 41.5% female	-	-	-	65.4 (8.6)	-	-
Mazzuca 1997	US	69% African American	53% unemployed, 15% employed, 31.5% retired	15% male, 85% female	9.7 (3.1) years of education	96.5% annual income ≤ $20,000	73.5% living alone	62.4 (11.6)	1.6 (1.1) comorbid conditions	-
Mazzuca 2004	US	70.5% white	-	28.5% male, 71.5% female	87.5% 12 or more years of education	-	62% married	61.8 (12.2)	-	-
McKnight 2010	US	91.7% white	-	23.4% male, 77.0% female	63.0% college educated	-	-	52.6 (7.2)	-	-
Murphy 2008	US	91% white	-	11% male, 89% female	66.5% some college to advanced degree	-	22% married	75.3 (7.1)	4.5 (2.1) painful or stiff joints	-
Nunez 2006	Spain	-	84% retired or housewives, 13% permanently disabled,	24.5% male, 75.5% female	-	-	67.5% have family, 30.5% alone, 3%	71.1 (6.7)	33% prior prostheses	-

Table 1. PROGRESS-Plus *(Continued)*

			3% active (sick leave)				living with carer			
Ackerman 2012	Australia	68.5% Australian-born	26% paid employment, 62% retired, 8.5% not working because of medical condition, 3.5% unemployed	40% male, 60% female	12% primary school or less, 46% years 7 to 10, 15.5% years 11 and 12, 14.5% trade/technical education, 12% university	-	63.5% married or living with partner	65.1 (10.9)	-	Domains social support for health, navigating health system, actively managing my health
Victor 2005	UK	64% non-white ethnic group	30% employed, 33.5% professional or managerial job	28% male, 72% female	37% higher education	66% home owner	35.5% living alone, 51% married	63.1 (11.1)	60% OA in both knees, 64% OA in other joints	-
Wetzels 2005	The Netherlands	-	-	24% male, 76% female	52% primary or lower secondary, 48% upper secondary or further	-	-	74.5 (6.4)	-	-
Yip 2007	Hong Kong	100% Asian	26.4% housewife, 8.8% professional and administration, 64.9% service provider and non-professional	16.0% male, 84.1% female	87.4% Form 3 level or below, 12.7% above Form 3	-	69.8% married and living together, 30.3% single	64.8 (10.0)	10.9% one joint OA, 48.9% two joints OA, 23.8% three joints OA, 16.5% four or more joints OA	-

P = Place of residence.
R = Race, ethnicity, culture.
O = Occupation.
Gr = Gender.
E = Education.
S = Socioeconomic status.
S = Social capital.

A = Age (in years).
D = Disability.
HL = Health literacy.
If not stated differently, data are ‘mean (SD)’.
y = year; IQR = interquartile range, - = no information available.
No information was available on 'Religion' and 'Sexual orientation'; therefore these domains are not included in this table.

Table 2. Additional characteristics of included studies

Study	No	Location of OA	BMI	Duration of OA
Allen 2010	523	80% knee, 15% hip, 5% knee and hip	31.8 (6.7)	16.1 (12.2) years
Berman 2004	570	100% knee	-	50.4% < 5 years, 20.7% 6 to 10 years, 28.7% > 10 years
Blixen 2004	32	-	-	8.3 (IQR 2 to 24) years
Buszewicz 2006	812	-	-	-
Calfas 1992	40	-	-	10% < 1 year, 5% 1 to 5 years, 92.5% > 5 years
Cronan 1997	363	-	-	7.0 (5.5) years since diagnosis
Crotty 2009	152	32.9% hip	Height 167.2 (10.1) cm Weight 85.8 (21.0) kg	-
Hansson 2010	114	4.5% hip, 34% knee, 32% hand, 29% more locations	35% BMI 20 to 25, 38.5% BMI 25 to 30, 27% BMI > 30	-
Heuts 2005	297	-	28.1 (5.0)	-
Hopman-Rock 2000	120	-	27.6 (4.3)	3% < 1 year, 20.5% 1 to 3 years, 33.5% 3 to 10 years, 18% 10 to 20 years, 17% > 20 years
Hughes 2004	215	-	-	-
Hurley 2007	431	100% knee	30.2 (range 18 to 51)	6 (IQR 3 to 13) years
Jessep 2009	64	100% knee	29.5 (range 1 to 47)	12.5 (range 0.5 to 55) years
Keefe 1990	99	100% knee	24.2 (23.6)% above ideal weight	-
Keefe 1996	88	100% knee	-	-
Keefe 2004	72	100% knee	-	-

Table 2. Additional characteristics of included studies *(Continued)*

Lorig 2008	551	-	-	-
Maisiak 1996	405	-	-	16.0 years
Martire 2007	242	Hip and knee	-	15.1 (10.9) years
Maurer 1999	113	100% knee	Weight 187.1 (33.9) lb	11.4 (10.5) years
Mazzuca 1997	211	100% knee	-	14.0 (15.9) years
Mazzuca 2004	186	100% knee	-	-
McKnight 2010	273	100% knee	27.7 (4.2)	-
Murphy 2008	54	67% knee, 11% hip, 22% hip and knee	30.1 (5.7)	-
Nunez 2006	100	100% knee	-	11.9 (10.6) months
Ackerman 2012	120	31% hip, 62.5% knee, 6.5% hip and knee	24.5 (IQR 25 to 35)	-
Victor 2005	193	100% knee	-	55% > 3 years
Wetzels 2005	104	53.8% knee, 20.1% hip, 26% hip and knee	-	-
Yip 2007	182	100% knee	-	8.1 (6.8) years

Mean (SD), unless indicated otherwise.
No = Number of participants randomly assigned (total of all groups).
y = year(s).
- = no information available.
IQR = interquartile range.
BMI in kg/m^2.

Table 3. Summary of comparisons

Study *Country*	**Setting**	**Intervention (N)** *Mode/Personnel/Delivery method/Duration*	**Attention control (N)** *Mode/Personnel/Delivery method/Duration*
Allen 2010 US	Primary care	Self-management intervention (174) M: individual P: health professionals De: telephone Du: 12 calls in 12 months	Health education (175) M: individual P: health professionals De: telephone Du: 12 calls in 12 months
Calfas 1992 US	Outpatients	Cognitive-behaviour modification (20) M: group	Traditional education intervention (20) M: group

Table 3. Summary of comparisons *(Continued)*

		P: trained facilitator De: face-to-face Du: 10 weekly sessions	P: health specialists De: face-to-face Du: several lectures
Keefe 1990 US	Outpatients	Pain coping skills training (32) M: group P: nurse and psychologist De: face-to-face Du: 10 sessions in 10 weeks (1.5 hours) and three phone calls	Arthritis education (36) M: group P: nurse and psychologist De: lectures and telephone Du: 10 sessions in 10 weeks (1.5 hours) and three phone calls
Maisiak 1996 US	Primary care and outpatients	Treatment counselling (135) M: individual P: trained counsellors De: telephone Du: 11 sessions in nine months	Symptom monitoring (135) M: individual P: college students trained for two hours De: telephone Du: 11 sessions in nine months
Mazzuca 1997 US	Primary care	Education (105) M: individual P: arthritis nurse educator De: face-to-face/telephone Du: one interview/two calls in one month	Attention control (106) M: group and individual P: not specified De: audiovisual presentation and telephone Du: 20-minute presentation, two phone calls (five to 10 minutes)
Study	**Setting**	**Intervention (N)**	**Information only (N)**
Buszewicz 2006 UK	Primary care	Self-management programme (406) M: group P: trained volunteer De: face-to-face Du: six sessions in six weeks	Education (406) M: individual P: none De: education booklet Du: -
Hughes 2004 US	General population	Fit & Strong (115) M: group P: physical therapist De: face-to-face Du: 24 sessions in eight weeks	Education (100) M: individual P: none De: *Arthritis Helpbook* and handouts Du: -

Table 3. Summary of comparisons *(Continued)*

Ackerman 2012 Australia	Outpatients	Arthritis self-management programme (58) M: group P: peer leader and health professional De: face-to-face Du: six sessions in six weeks	Education only (62) M: individual P: none De: *Arthritis Helpbook* Du: -
Wetzels 2005 The Netherlands	Primary care	Self-management intervention (51) M: individual P: nurse De: face-to-face/telephone Du: one session/one call	Education (53) M: individual P: none De: educational booklet Du: -
Study	**Setting**	**Intervention (N)**	**Usual care/Waiting list/No treatment (N)**
Allen 2010 US	Primary care	Self-management intervention (174) M: individual P: health professionals De: telephone Du: 12 calls in 12 months	Usual care (174)
Blixen 2004 US	Outpatients	Telephone health education strategy (16) M: individual P: advanced practice nurse De: telephone Du: six sessions in six weeks	Usual care (16)
Cronan 1997 US	General population	Education group (97) M: group P: professional health educator De: face-to-face Du: 10 weekly sessions and 10 monthly sessions	No treatment (90)
Cronan 1997 US	General population	Combination education + social support (89) M: group P: professional health educator (first hour) De: face-to-face Du: 10 weekly sessions and 10 monthly sessions	No treatment (90)
Crotty 2009	Outpatients	Patient education (75)	Usual care (77)

Table 3. Summary of comparisons *(Continued)*

Australia		M: individual P: peer support volunteer/research nurse De: telephone/face-to-face Du: six calls in six months/one interview	
Hansson 2010 Sweden	Primary care	PEPOA (61) M: group P: health professionals De: face-to-face Du: five sessions in five weeks	Usual care (53)
Heuts 2005 The Netherlands	Primary care	Self-management programme (149) M: group P: trained physiotherapists De: face-to-face Du: six sessions	Usual care (148)
Hopman-Rock 2000 The Netherlands	General population	Health educational and exercise programme (60) M: group P: peer educator and health professionals De: face-to-face Du: six sessions in six weeks	No treatment (60)
Hurley 2007 UK	Primary care	Individual rehabilitation (146) M: individual P: physiotherapist De: face-to-face Du: 12 sessions in six weeks	Usual care (140)
Hurley 2007 UK	Primary care	Group rehabilitation (132) M: group P: physiotherapist De: face-to-face Du: 12 sessions in six weeks	Usual care (140)
Keefe 1990 US	Outpatients	Pain coping skills training (32) M: group P: nurse and psychologist De: face-to-face	Usual care (31)

Table 3. Summary of comparisons *(Continued)*

		Du: 10 sessions in 10 weeks (1.5 hours) and three phone calls	
Keefe 2004 US	Outpatients and general population	Spouse-assisted coping skills training (18) M: group P: trained psychologist De: face-to-face Du: 12 sessions in 12 weeks	Usual care (18)
Keefe 2004 US	Outpatients and general population	Spouse-assisted coping skills training + exercise (20) M: group P: trained psychologist/exercise trainer De: face-to-face Du: 38 sessions in 12 weeks	Usual care (18)
Lorig 2008 US	General population	Internet-based arthritis self-management programme (433) M: individual P: peer moderators De: Internet Du: 18 sessions in six weeks (at least)	Usual care (422)
Maisiak 1996 US	Primary care and outpatients	Treatment counselling (135) M: individual P: trained counsellors De: telephone Du: 11 sessions in nine months	Usual care (135)
Martire 2007 US	Outpatients	Patient-oriented education and support (89) M: group P: trained facilitator De: face-to-face Du: six sessions in six weeks	Usual care (54)
Martire 2007 US	Outpatients	Couple-oriented education and support (99) M: group P: trained facilitator De: face-to-face Du: six sessions in six weeks	Usual care (54)

Table 3. Summary of comparisons *(Continued)*

Mazzuca 2004 US	Primary care	Nurse-directed intervention (111) M: individual P: arthritis nurse educator De: face-to-face/telephone Du: one interview/five to nine calls inthree to five months	Waiting list (75)
Nunez 2006 US	Outpatients	Self-management programme (51) M: group/individual P: trained health educator De: face-to-face Du: four sessions in three months	Usual care (49)
Victor 2005 UK	Primary care	Self-management (120) M: group P: research nurse De: face-to-face Du: four sessions	Waiting list (73)
Yip 2007 China	Outpatients	Arthritis self-management programme (88) M: group P: nurses De: face-to-face Du: six sessions in six weeks	Usual care (94)
Study	**Setting**	**Intervention (N)**	**Alternate intervention (N)**
Berman 2004 US	General population	Education (189) M: group P: patient education specialist De: face-to-face Du: six sessions in 12 weeks	True/sham acupuncture (190 + 191) M: individual P: acupuncturists De: face-to-face Du: 23 treatments in 26 weeks
Cronan 1997 US	General population	Education group (97) M: group P: professional health educator De: face-to-face Du: 10 weekly sessions and 10 monthly sessions (two hours)	Social support group (87) M: group P: none De: face-to-face Du: 10 weekly sessions and 10 monthly sessions (two hours)

Table 3. Summary of comparisons *(Continued)*

Cronan 1997 US	General population	Combination education + social support (89) M: group P: professional health educator (first hour) De: face-to-face Du: 10 weekly sessions and 10 monthly sessions (two hours)	Social support group (87) M: group P: none De: face-to-face Du: 10 weekly sessions and 10 monthly sessions (two hours)
Jessep 2009 UK	Outpatients	ESCAPE-knee pain (29) M: group P: physiotherapist De: face-to-face Du: 10 sessions in five weeks	Physiotherapy (35) M: not specified P: physiotherapist De: face-to-face Du: maximum of 10 sessions
Keefe 1996 US	Not specified	Coping skills training (29) M: group P: nurse and psychologist De: face-to-face Du: 10 sessions in 10 weeks (two hours)	Arthritis education—spousal support (29) M: group P: nurse and psychologist De: discussions and educational material Du: one time a week (two hours) for 10 weeks
Keefe 1996 US	Not specified	Spouse-assisted coping skills training (30) M: group P: nurse and psychologist De: face-to-face Du: 10 sessions in 10 weeks (two hours)	Arthritis education—spousal support (29) M: group P: nurse and psychologist De: discussions and educational material Du: one time a week (two hours) for 10 weeks
Keefe 2004 US	Outpatients and general population	Spouse-assisted coping skills training (18) M: group P: trained psychologist De: face-to-face Du: 12 sessions in 12 weeks	Exercise (16) M: group P: exercise physiologists (BA level or above) De: face-to-face Du: three times a week (one hour) for 12 weeks
Keefe 2004 US	Outpatients and general population	Spouse-assisted coping skills training + exercise (20) M: group	Exercise (16) M: group

Table 3. Summary of comparisons *(Continued)*

		P: trained psychologist/exercise trainer De: face-to-face Du: 38 sessions in 12 weeks	P: exercise physiologists (BA level or above) De: face-to-face Du: three times a week (one hour) for 12 weeks
Maurer 1999 US	Outpatients	Education (56) M: group P: healthcare professionals De: face-to-face Du: four sessions in eight weeks	Exercise (57) M: not specified P: not specified De: face-to-face Du: three times a week for eight weeks
McKnight 2010 US	General population	Self-management group (87) M: group/individual P: healthcare professionals De: face-to-face/telephone Du: 12 sessions in 12 weeks/staggering calls	Strength training (91) M: not specified P: expert physical trainers De: face-to-face Du: three weekly sessions for nine months (phase 1), then contact every two weeks in first six weeks, then monthly for a total of 15 months (phase 2)
McKnight 2010 US	General population	Combination group (95) M: group/individual P: healthcare professionals and physiotherapists De: face-to-face/telephone Du: 48 sessions in 12 weeks/staggering calls	Strength training (91) M: not specified P: expert physical trainers De: face-to-face, telephone Du: three weekly sessions for nine months (phase 1), then contact every two weeks in first six weeks, then monthly for a total of 15 months (phase 2)
Murphy 2008 US	Senior housing facilities	Exercise + activity strategy training (28) M: group + one individual session P: occupational therapists De: face-to-face Du: eight sessions (1.5 hours) in four weeks	Exercise + education (26) M: group P: health education interventionists De: educational materials Du: two sessions per week (1.5 hours) for four weeks

M = Mode, P = Personnel, De = Delivery method, Du = Duration.

Table 4. heiQ-components addressed in interventions

Intervention	Health-directed activity	Positive and active engagement in life	Emotional distress	Self-monitoring and insight	Constructive attitudes and approaches	Skill and technique acquisition	Social integration and support	Health service navigation
TOTAL + (%)	**29/34**	**9/34**	**13/34**	**27/34**	**15/34**	**32/34**	**4/34**	**11/34**
	85%	**26%**	**38%**	**79%**	**44%**	**94%**	**12%**	**32%**
Allen 2010	+	-	-	+	+	+	-	+
Berman 2004	+	U	+	-	+	+	-	+
Blixen 2004	+	-	+	+	-	+	-	+
Buszewicz 2006	+	+	+	+	+	+	U	+
Calfas 1992	+	+	+	+	+	+	-	-
Cronan 1997 (education)	+	-	-	+	+	+	-	+
Cronan 1997 (combination)	+	-	-	+	+	+	-	+
Crotty 2009	-	-	-	+	+	-	-	-
Hansson 2010	+	-	-	+	-	+	-	-
Heuts 2005	+	-	+	+	-	+	U	+
Hopman-Rock 2000	+	-	-	+	+	+	-	-
Hughes 2004	+	-	-	-	-	+	-	-
Hurley 2007 (Indiv Rehab)	+	+	+	+	+	+	-	-
Hurley 2007 (Group Rehab)	+	+	+	+	+	+	-	-
Jessep 2009	+	+	+	+	+	+	-	-
Keefe 2004	-	-	-	+	+	+	+	-
Keefe 1996 (CST)	-	-	-	+	+	+	-	-

Table 4. heiQ-components addressed in interventions *(Continued)*

Keefe 1996 (SA-CST)	-	-	-	+	+	+	+	-
Keefe 1990	-	-	-	+	+	+	-	-
Lorig 2008	+	+	+	+	+	+	-	+
Maisiak 1996	+	-	-	+	-	+	-	+
Martire 2007 (PES)	+	+	+	+	+	+	-	-
Martire 2007 (CES)	+	+	+	+	+	+	+	-
Maurer 1999	+	-	-	+	-	+	-	-
Mazzuca 1997	+	-	-	-	-	+	-	-
Mazzuca 2004	+	-	-	-	-	+	-	-
McKnight 2010 (SMP)	+	-	-	+	+	+	-	-
McKnight 2010 (combination)	+	-	-	+	+	+	-	-
Murphy 2008	+	-	-	-	+	+	-	-
Nunez 2006	+	-	-	-	-	+	-	-
Ackerman 2012	+	-	+	+	+	+	-	+
Victor 2005	+	-	-	-	+	+	+	-
Wetzels 2005	+	+	-	+	+	-	-	-
Yip 2007	+	-	+	+	+	+	-	+

+ = heiQ component addressed, - = heiQ component not addressed, U = Unclear whether heiQ component was addressed

APPENDICES

Appendix 1. MEDLINE search strategy

MEDLINE (Ovid) 2005 to January Week 2, 2013

1. exp osteoarthritis/

2. osteoarthr$.tw.

3. (degenerative adj2 arthritis).tw.

4. or/1-3

5. exp Self Care/

6. ((self or symptom$) adj (care or help or manag$ or directed or monitor$ or efficacy or admin$)).tw.

7. Patient Education as Topic/

8. ((health or patient$) adj2 (educat$ or information)).tw.

9. exp Consumer Participation/

10. ((patient$ or consumer$) adj part$).tw.

11. "power (psychology)"/

12. empower$.tw.

13. Holistic Health/

14. (holistic or wholistic).tw.

15. exp Rehabilitation/

16. rehab$.tw.

17. "Activities of Daily Living"/

18. (activit$ adj2 daily adj living).tw.

19. social support/

20. (social adj (support or network$)).tw.

21. (support adj system$).tw.

22. exp Adaptation, Psychological/

23. (psychologic$ adj (adjust$ or adapt$)).tw.

24. (cope or copes or coping).tw.

25. exp Behavior Therapy/

26. (adapt$ adj behav$).tw.

27. (behav$ adj (therap$ or intervention$)).tw.

28. or/5-27

29. 4 and 28

30. randomized controlled trial.pt.

31. controlled clinical trial.pt.

32. randomized.ab.

33. placebo.ab.

34. drug therapy.fs.

35. randomly.ab.

36. trial.ab.

37. groups.ab.

38. or/30-37

39. (animals not (humans and animals)).sh.

40. 38 not 39

41. 29 and 40

Appendix 2. EMBASE search strategy

EMBASE (Ovid) 2010 to 2013 Week 2

1. exp osteoarthritis/

2. osteoarthr$.tw.

3. (degenerative adj2 arthritis).tw.

4. or/1-3

5. exp Self Care/

6. ((self or symptom$) adj (care or help or manag$ or directed or monitor$ or efficacy or admin$)).tw.

7. exp Patient Education/

8. ((health or patient$) adj2 (educat$ or information)).tw.

9. exp Patient Participation/

10. exp Consumer/

11. ((patient$ or consumer$) adj part$).tw.

12. exp EMPOWERMENT/

13. empower$.tw.

14. (holistic or wholistic).tw.

15. exp REHABILITATION/

16. rehab$.tw.

17. exp Daily Life Activity/

18. (activit$ adj2 daily adj living).tw.

19. exp Social Support/

20. (social adj (support or network$)).tw.

21. (support adj system$).tw.

22. exp Coping Behavior/

23. (psychologic$ adj (adjust$ or adapt$)).tw.

24. (cope or copes or coping).tw.

25. exp Behavior Therapy/

26. (adapt$ adj behav$).tw.

27. (behav$ adj (therap$ or intervention$)).tw.

28. or/5-27

29. 4 and 28

30. (random$ or placebo$).ti,ab.

31. ((single$ or double$ or triple$ or treble$) and (blind$ or mask$)).ti,ab.

32. controlled clinical trial$.ti,ab.

33. RETRACTED ARTICLE/

34. or/30-33

35. (animal$ not human$).sh,hw.

36. 34 not 35

37. 29 and 36

Appendix 3. *Cochrane Library* search strategy

The Cochrane Library **(Wiley InterScience) January 17, 2013**

#1 MeSH descriptor Osteoarthritis explode all trees

#2 osteoarthr*:ti,ab

#3 degenerative Near/2 arthritis:ti,ab

#4 #1 OR #2 OR #3

#5 MeSH descriptor Self Care explode all trees

#6 ((self or symptom*) next (care or help or manag* or directed or monitor* or efficacy or admin*)):ti,ab

#7 MeSH descriptor Patient Education as Topicexplode all trees

#8 ((health or patient*) near/2 (educat* or information)):ti,ab

#9 MeSH descriptor Patient Participation explode all trees

#10 MeSH descriptor Consumer Participation explode all trees

#11 ((patient* or consumer*) next part*):ti,ab

#12 MeSH descriptor Power (Psychology) explode all trees

#13 empower*:ti,ab

#14 MeSH descriptor Holistic Health explode all trees

#15 (holistic or wholistic):ti,ab

#16 MeSH descriptor Rehabilitation explode all trees

#17 rehab*:ti,ab

#18 MeSH descriptor Activities of Daily Living explode all trees

#19 (activit* near/2 daily next living):ti,ab

#20 MeSH descriptor Social Support explode all trees

#21 (social next (support or network*)):ti,ab

#22 (support next system*):ti,ab

#23 MeSH descriptor Adaptation, Psychological explode all trees

#24 (psychologic* next (adjust* or adapt*)):ti,ab

#25 cope or copes or coping:ti,ab

#26 MeSH descriptor Behavior Therapy explode all trees

#27 adapt* next behav*:ti,ab

#28 (behav* next (therap* or intervention*)):ti,ab

#29 #5 OR #6 OR #7 OR #8 OR #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 OR #26 OR #27 OR #28

#30 #4 AND #29

Appendix 4. CINAHL search strategy

CINAHL (EBSCOhost) 1982 to January 17, 2013

S1 (MH "Osteoarthritis+")
S2 ti osteoarthr* or ab osteoarthr*
S3 ti degenerative N2 arthritis or ab degenerative N2 arthritis
S4 S1 or S2 or S3
S5 (MH "Self Care+")
S6 ti self care or ti self help or ti self manag* or ti self directed or ti self monitor* or ti self efficacy or ti self admin*
S7 ab self care or ab self help or ab self manag* or ab self directed or ab self monitor* or ab self efficacy or ab self admin*
S8 ti symptom* care or ti symptom* help or ti symptom* manag* or ti symptom* directed or ti symptom* monitor* or ti symptom* efficacy or ti symptom* admin*
S9 ab symptom* care or ab symptom* help or ab symptom* manag* or ab symptom* directed or ab symptom* monitor* or ab symptom* efficacy or ab symptom* admin*
S10 ti health N2 educat* or ti health N2 information or ab health N2 educat* or ab health N2 information
S11 ti patient* N2 educat* or ti patient* N2 information or ab patient* N2 educat* or ab patient* N2 information
S12 (MH "Consumer Participation")
S13 ti patient* participat* or ab patient* participat* or ti consumer* participat* or ab consumer* participat*
S14 (MH "Empowerment")
S15 ti empower* or ab empower*

(Continued)

S16 (MH "Holistic Health")
S17 ti holistic or ti wholistic or ab holistic or ab wholistic
S18 (MH "Rehabilitation+")
S19 ti rehab* or ab rehab*
S20 (MH "Activities of Daily Living+")
S21 ti activit* N2 daily living or ab activit* N2 daily living
S22 (MH "Support, Psychosocial+")
S23 ti social support or ab social support or ti social network* or ab social network*
S24 ti support system* or ab support system*
S25 (MH "Coping+")
S26 ti psychologic* adjust* or ti psychologic* adapt*
S27 ab psychologic* adjust* or ab psychologic* adapt*
S28 ti cope or ti copes or coping or ab cope or ab copes or ab coping
S29 (MH "Behavior Therapy+")
S30 ti adapt* behav* or ab adapt* behav*
S31 ti behav* therap* or ti behav* intervention* or ab behav* therap* or ab behav* intervention*
S32 S5 or S6 or S7 or 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 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30 or S31
S33 S4 and S32
S34 (MH "Clinical Trials+")
S35 PT clinical trial
S36 TI clinical* trial* or AB clinical* trial*
S37 TI singl* blind* or TI singl* mask* or TI doub* blind* or TI doubl* mask* or TI trebl* blind* or TI trebl* mask* or TI tripl* blind* or TI tripl* mask*
S38 AB singl* blind* or AB singl* mask* or AB doub* blind* or AB doubl* mask* or AB trebl* blind* or AB trebl* mask* or AB tripl* blind* or AB tripl* mask*
S39 TI Randomi?ed control* trial* or AB Randomi?ed control* trial*
S40 (MH "Random Assignment")
S41 TI Random* allocat* or AB Random* allocat*
S42 TI Placebo* or AB Placebo*

(Continued)

S43 (MH "Placebos")
S44 (MH "Quantitative Studies")
S45 TI Allocat* random* or AB Allocat* random*
S46 S34 or S35 or S36 or S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45
S47 S33 and S46

Appendix 5. PsycINFO search strategy

PsycINFO (Ovid) 1806 to January Week 2 2013

1. arthritis/
2. osteoarthr$.tw.
3. (degenerative adj2 arthritis).tw.
4. or/1-3
5. exp self help techniques/
6. ((self or symptom$) adj (care or help or manag$ or directed or monitor$ or efficacy or admin$)).tw.
7. client education/
8. ((health or patient$) adj2 (educat$ or information)).tw.
9. client participation/
10. ((patient$ or consumer$) adj part$).tw.
11. empowerment/
12. empower$.tw.
13. holistic health/
14. (holistic or wholistic).tw.
15. exp rehabilitation/
16. rehab$.tw.
17. "activities of daily living"/
18. (activit$ adj2 daily adj living).tw.
19. social support/
20. (social adj (support or network$)).tw.
21. (support adj system$).tw.
22. coping behavior/
23. (psychologic$ adj (adjust$ or adapt$)).tw.
24. (cope or copes or coping).tw.
25. exp behavior therapy/

26. (adapt$ adj behav$).tw.

27. (behav$ adj (therap$ or intervention$)).tw.

28. or/5-27

29. 4 and 28

30. limit 29 to "2000 treatment outcome/randomized clinical trial"

Appendix 6. Dissertation abstracts search strategy

Dissertation and Theses January 17, 2013

1. (osteoarthr* OR degenerative arthritis) in abs

2. (self OR symptom*) in abs

3. (care OR manag* OR directed OR monitor* OR efficacy OR admin*) in abs

4. 1 AND 2 AND 3

Appendix 7. Scopus search strategy

Scopus January 17, 2013

#1 (osteoarthr* OR degenerative arthritis) in TITLE-ABS-KEY AND

#2 (self OR symptom*) in TITLE-ABS-KEY AND

#3 (care OR manag* OR directed OR monitor* OR efficacy OR admin*) in TITLE-ABS-KEY

#4 (#1 AND #2 AND 3)

Search limited to Conference Paper

Appendix 8. WHO search strategy

World Health Organization International Clinical Trials Registry Platform April 5, 2011

osteoarthr* OR degenerative arthritis in Condition AND

self* in Intervention

HISTORY

Protocol first published: Issue 1, 2011
Review first published: Issue 1, 2014

Date	Event	Description
11 June 2008	Amended	CMSG ID: C149-P

CONTRIBUTIONS OF AUTHORS

FPBK contributed to study selection, data extraction and management, data analysis and writing of the review.

LRAB contributed to study selection, data extraction and management, data analysis and writing of the review.

RB contributed to development of the protocol, study selection, data analysis, interpretation of data and writing of the review. She is the overall guarantor for the review.

RHO contributed to to development of the protocol, interpretation of data and writing of the review.

RVJ contributed to data analysis, the Summary of findings tables and writing of the review.

VP contributed to development of the protocol, study selection, data analysis, interpretation of data and writing of the review.

DECLARATIONS OF INTEREST

RB and RHO were investigators for one of the included trials but had no part in deciding on inclusion, assessment of risk of bias, data extraction or interpretation of the results (Ackerman 2012).

SOURCES OF SUPPORT

Internal sources

- Cabrini Hospital, Australia.

  In-kind support
- Monash University, Australia.

  In-kind support

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We initially stated that we would search both SCOPUS and the Web of Science database; however, we decided to search only SCOPUS to reduce duplication in the records retrieved.

Our major outcome three was "adverse events or withdrawals" (where reported reasons for withdrawal are related to the intervention). However, none of the included trials reported this outcome. We therefore chose instead to report and compare withdrawals (defined as dropouts both related and unrelated to the study intervention and those lost to follow-up).

Although we stated in the protocol that we would assess healthcare services and medication use as a part of health service navigation, we later decided not to extract these as specific outcomes, as they can be both positive and negative outcomes, depending upon the circumstances.

We decided to split the secondary outcome function into self-reported function (i.e. measured on the WOMAC function subscale) and performance measures of function (i.e. measured on the six-minute walking distance test or the timed up-and-go test). It was judged that these two components of function were so different that they could not be combined, and that both were equally important to be assessed.

We used the PROGRESS-Plus framework instead of the PROGRESS-framework, as this version was published after publication of the protocol.

We used the Health Literacy Questionnaire (HLQ) to assess health literacy in the study population.

Subgroup analysis to explore whether a relationship exists between any of the component domains addressed in the self-management education programmes and participant outcomes was not performed, as available data were insufficient for the primary outcomes of the review. Subgroup analyses to explore the potential impact of age, stage of disease and comorbidities on participant outcomes were also deemed not possible.

A hierarchy for extraction of data was added for the following outcomes: self-management in OA and global OA scores.

We decided that if data on a quality of life scale were provided in both a multi-questionnaire format and a visual analogue scale format, authors would include the multi-questionnaire format data as the first preference.

A good level of reporting was noted across studies for our outcomes of interest; therefore we deemed it unnecessary to include a table with incomplete data.

INDEX TERMS

Medical Subject Headings (MeSH)

Health Literacy; Osteoarthritis [*therapy]; Pain Management [methods]; Patient Education as Topic [*methods]; Program Evaluation; Randomized Controlled Trials as Topic; Self Care [*methods]

MeSH check words

Female; Humans; Male