Cochrane Library

Cochrane Database of Systematic Reviews

Exercise therapy for chronic low back pain (Review)

Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW

Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW.
Exercise therapy for chronic low back pain.
Cochrane Database of Systematic Reviews 2021, Issue 9. Art. No.: CD009790.
DOI: 10.1002/14651858.CD009790.pub2.

www.cochranelibrary.com

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 6
OBJECTIVES 7
METHODS 7
RESULTS 11
Figure 1. 12
Figure 2. 13
Figure 3. 18
DISCUSSION 21
AUTHORS' CONCLUSIONS 25
ACKNOWLEDGEMENTS 26
REFERENCES 27
CHARACTERISTICS OF STUDIES 59
DATA AND ANALYSES 464
Analysis 1.1. Comparison 1: PRIMARY ANALYSES, Outcome 1: Pain (/100): Earliest follow-up 466
Analysis 1.2. Comparison 1: PRIMARY ANALYSES, Outcome 2: Pain (/100): Short-term follow-up (6-12 weeks) 469
Analysis 1.3. Comparison 1: PRIMARY ANALYSES, Outcome 3: Pain (/100): Medium-term follow-up (~6 months) 471
Analysis 1.4. Comparison 1: PRIMARY ANALYSES, Outcome 4: Pain (/100): Long-term follow-up (> 12 months) 473
Analysis 1.5. Comparison 1: PRIMARY ANALYSES, Outcome 5: Function (/100): Earliest follow-up 474
Analysis 1.6. Comparison 1: PRIMARY ANALYSES, Outcome 6: Function (/100): Short-term follow-up (6-12 weeks) 477
Analysis 1.7. Comparison 1: PRIMARY ANALYSES, Outcome 7: Function (/100): Medium-term follow-up (~6 months) 479
Analysis 1.8. Comparison 1: PRIMARY ANALYSES, Outcome 8: Function (/100): Long-term follow-up (> 12 months) 481
Analysis 2.1. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 1: Pain (/100); Earliest follow-up; Exercise vs. no treatment or usual care 483
Analysis 2.2. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 2: Pain (/100); Earliest follow-up; Exercise vs. other conservative treatment 485
Analysis 2.3. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 3: Function (/100); Earliest follow-up; Exercise vs. no treatment or usual care 487
Analysis 2.4. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 4: Function (/100); Earliest follow-up; Exercise vs. other conservative treatment 489
Analysis 3.1. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 1: Pain (/100): Earliest follow-up; Exercise vs. no treatment or usual care 492
Analysis 3.2. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 2: Pain (/100): Earliest follow-up; Exercise vs. other conservative treatment 494
Analysis 3.3. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 3: Function (/100): Earliest follow-up; Exercise vs. no treatment or usual care 496
Analysis 3.4. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 4: Function (/100): Earliest follow-up; Exercise vs. other conservative treatment 498
Analysis 4.1. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 1: Pain (/100): Earliest follow-up; Exercise vs. no treatment or usual care 501
Analysis 4.2. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 2: Pain (/100): Earliest follow-up; Exercise vs. other conservative treatment 503
Analysis 4.3. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 3: Function (/100): Earliest follow-up; Exercise vs. no treatment or usual care 505
Analysis 4.4. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 4: Function (/100): Earliest follow-up; Exercise vs. other conservative treatment 507
Analysis 5.1. Comparison 5: SUBGROUP ANALYSES (Conservative comparisons), Outcome 1: Pain (/100): Earliest follow-up, Exercise vs. other conservative treatments 510
Analysis 5.2. Comparison 5: SUBGROUP ANALYSES (Conservative comparisons), Outcome 2: Function (/100): Earliest follow-up, Exercise vs. other conservative treatments 513

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 6.1. Comparison 6: SENSITIVITY ANALYSES (No assumptions about data), Outcome 1: Pain (/100): Earliest follow-up, no imputation of SD 516
Analysis 6.2. Comparison 6: SENSITIVITY ANALYSES (No assumptions about data), Outcome 2: Function (/100): Earliest follow-up, no imputation of SD 518
Analysis 6.3. Comparison 6: SENSITIVITY ANALYSES (No assumptions about data), Outcome 3: Function (/100): Earliest follow-up, RMDQ measurement only 520
Analysis 7.1. Comparison 7: SENSITIVITY ANALYSES (Low risk of bias), Outcome 1: Pain (/100): Earliest follow-up 522
Analysis 7.2. Comparison 7: SENSITIVITY ANALYSES (Low risk of bias), Outcome 2: Function (/100): Earliest follow-up 524
Analysis 8.1. Comparison 8: SENSITIVITY ANALYSES (Excluding outlying means), Outcome 1: Pain (/100): Earliest follow-up 527
Analysis 8.2. Comparison 8: SENSITIVITY ANALYSES (Excluding outlying means), Outcome 2: Function (/100): Earliest follow-up 530
Analysis 9.1. Comparison 9: SENSITIVITY ANALYSES (Placebo only comparisons), Outcome 1: Pain (/100): Earliest follow-up 532
Analysis 9.2. Comparison 9: SENSITIVITY ANALYSES (Placebo only comparisons), Outcome 2: Function (/100): Earliest follow-up 532
ADDITIONAL TABLES 532
APPENDICES 537
WHAT'S NEW 548
HISTORY 549
CONTRIBUTIONS OF AUTHORS 549
DECLARATIONS OF INTEREST 549
SOURCES OF SUPPORT 549
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 550
INDEX TERMS 550

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

[Intervention Review]

Exercise therapy for chronic low back pain

Jill A Hayden[1], Jenna Ellis[1], Rachel Ogilvie[1], Antti Malmivaara[2], Maurits W van Tulder[3]

[1]Department of Community Health & Epidemiology, Dalhousie University, Halifax, Canada. [2]Centre for Health and Social Economics (CHESS), National Institute for Health and Welfare (THL), Helsinki, Finland. [3]Department of Health Sciences, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, Netherlands

Contact: Jill A Hayden, jhayden@dal.ca.

Editorial group: Cochrane Back and Neck Group.
Publication status and date: Edited (no change to conclusions), published in Issue 10, 2021.

Citation: Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews* 2021, Issue 9. Art. No.: CD009790. DOI: 10.1002/14651858.CD009790.pub2.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Low back pain has been the leading cause of disability globally for at least the past three decades and results in enormous direct healthcare and lost productivity costs.

Objectives

The primary objective of this systematic review is to assess the impact of exercise treatment on pain and functional limitations in adults with chronic non-specific low back pain compared to no treatment, usual care, placebo and other conservative treatments.

Search methods

We searched CENTRAL (which includes the Cochrane Back and Neck trials register), MEDLINE, Embase, CINAHL, PsycINFO, PEDro, SPORTDiscus, and trials registries (ClinicalTrials.gov and World Health Organization International Clinical Trials Registry Platform), and conducted citation searching of relevant systematic reviews to identify additional studies. The review includes data for trials identified in searches up to 27 April 2018. All eligible trials have been identified through searches to 7 December 2020, but have not yet been extracted; these trials will be integrated in the next update.

Selection criteria

We included randomised controlled trials that assessed exercise treatment compared to no treatment, usual care, placebo or other conservative treatment on the outcomes of pain or functional limitations for a population of adult participants with chronic non-specific low back pain of more than 12 weeks' duration.

Data collection and analysis

Two authors screened and assessed studies independently, with consensus. We extracted outcome data using electronic databases; pain and functional limitations outcomes were re-scaled to 0 to 100 points for meta-analyses where 0 is no pain or functional limitations. We assessed risk of bias using the Cochrane risk of bias (RoB) tool and used GRADE to evaluate the overall certainty of the evidence. When required, we contacted study authors to obtain missing data. To interpret meta-analysis results, we considered a 15-point difference in pain and a 10-point difference in functional limitations outcomes to be clinically important for the primary comparison of exercise versus no treatment, usual care or placebo.

Main results

We included 249 trials of exercise treatment, including studies conducted in Europe (122 studies), Asia (38 studies), North America (33 studies), and the Middle East (24 studies). Sixty-one per cent of studies (151 trials) examined the effectiveness of two or more different types of exercise treatment, and 57% (142 trials) compared exercise treatment to a non-exercise comparison treatment. Study participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

had a mean age of 43.7 years and, on average, 59% of study populations were female. Most of the trials were judged to be at risk of bias, including 79% at risk of performance bias due to difficulty blinding exercise treatments.

We found moderate-certainty evidence that exercise treatment is more effective for treatment of chronic low back pain compared to no treatment, usual care or placebo comparisons for pain outcomes at earliest follow-up (MD -15.2, 95% CI -18.3 to -12.2), a clinically important difference. Certainty of evidence was downgraded mainly due to heterogeneity. For the same comparison, there was moderate-certainty evidence for functional limitations outcomes (MD -6.8 (95% CI -8.3 to -5.3); this finding did not meet our prespecified threshold for minimal clinically important difference. Certainty of evidence was downgraded mainly due to some evidence of publication bias.

Compared to all other investigated conservative treatments, exercise treatment was found to have improved pain (MD -9.1, 95% CI -12.6 to -5.6) and functional limitations outcomes (MD -4.1, 95% CI -6.0 to -2.2). These effects did not meet our prespecified threshold for clinically important difference. Subgroup analysis of pain outcomes suggested that exercise treatment is probably more effective than education alone (MD -12.2, 95% CI -19.4 to -5.0) or non-exercise physical therapy (MD -10.4, 95% CI -15.2 to -5.6), but with no differences observed for manual therapy (MD 1.0, 95% CI -3.1 to 5.1).

In studies that reported adverse effects (86 studies), one or more adverse effects were reported in 37 of 112 exercise groups (33%) and 12 of 42 comparison groups (29%). Twelve included studies reported measuring adverse effects in a systematic way, with a median of 0.14 (IQR 0.01 to 0.57) per participant in the exercise groups (mostly minor harms, e.g. muscle soreness), and 0.12 (IQR 0.02 to 0.32) in comparison groups.

Authors' conclusions

We found moderate-certainty evidence that exercise is probably effective for treatment of chronic low back pain compared to no treatment, usual care or placebo for pain. The observed treatment effect for the exercise compared to no treatment, usual care or placebo comparisons is small for functional limitations, not meeting our threshold for minimal clinically important difference. We also found exercise to have improved pain (low-certainty evidence) and functional limitations outcomes (moderate-certainty evidence) compared to other conservative treatments; however, these effects were small and not clinically important when considering all comparisons together. Subgroup analysis suggested that exercise treatment is probably more effective than advice or education alone, or electrotherapy, but with no differences observed for manual therapy treatments.

PLAIN LANGUAGE SUMMARY

Exercise for treatment of chronic low back pain

Is exercise an effective therapy to treat long-lasting low back pain?

Key messages

- Exercise probably reduces pain compared to no treatment, usual care or placebo in people with long-lasting (chronic) low back pain.

- Exercise may reduce pain and improve disability compared to common treatments such as electrotherapy or education.

- There is a lot of research in this field but we need bigger and better designed studies to allow us to draw firm conclusions.

How might exercise help people with long-lasting low back pain?

Long-lasting (chronic) low back pain is a common cause of disability across the world and is expensive in terms of healthcare costs and lost working hours. Exercise therapy aims to increase muscle and joint strength, and improve muscle function and range of motion. This should reduce pain and disability, and speed recovery and return to usual activities. Exercise therapies are designed or prescribed by health professionals and cover a range of exercise types, durations, and delivery methods. Examples of exercise therapies include general physical fitness programmes delivered in a group setting, aerobic exercise in the form of walking programmes, and strengthening of specific muscles or groups of muscles to increase core stability.

What did we want to find out?

We wanted to know whether exercise improves pain and disability for people with chronic low back pain more than no treatment, usual care, placebo or other common treatments. In our review, chronic low back pain is pain that lasts three months or longer or that goes away but returns more than twice in one year. It does not have a specific cause such as a tumour or injury. Examples of common treatments are spinal manipulation, or psychological therapy. ‘Usual care’ is care provided by a family physician.

What did we do?

We searched for studies that assessed the effects of exercise therapy on pain or disability compared to no treatment, usual care, placebo or other common treatments. People in the studies had to be adults with chronic low back pain.

We compared and summarised the results of the studies and rated our confidence in the evidence, based on factors such as study methods and sizes.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

What did we find?

We found 249 studies with a total of 24,486 people. Most studies took place in Europe (122 studies); other common study locations were Asia, North America, and the Middle East. Study participants' average age was 43.7 years; 59% were women. Participants' average pain intensity at the start of the studies was 51 points on a 100-point scale, where 100 is the most pain. They had back pain for 12 weeks to 3 years (78 studies) or longer than 3 years (72 studies); 99 studies did not report how long their participants had low back pain.

Sixty-one per cent of studies (151 studies) examined the effectiveness of two or more different types of exercise, and 57% (142 studies) compared exercise therapy to a non-exercise treatment. The most common types of exercises were core strengthening (127 study groups), mixed exercises (>2 types) (109 study groups), Pilates (29 study groups), general strengthening exercises (52 study groups), and aerobic exercise (30 study groups). Exercise sessions were one-on-one with a healthcare provider (163 study groups) or in a group exercise class (162 study groups). More than half of studies included another treatment alongside exercise (247 study groups), including education or advice (137 study groups), electrotherapy (46 study groups), or manual therapy (21 study groups).

Most studies measured pain (223 studies) and disability (223 studies). Only 12 studies reported data that we could use on unwanted effects of treatments. Studies followed people in the short term (6 to 12 weeks; 184 studies); medium term (13 to 47 weeks; 121 studies) and long-term (48 weeks or more, 69 studies).

We also identified 172 more recent studies that we will add to the next version of our review.

Main results

People receiving exercise therapy rated their pain on average 15 points better and their disability 7 points better, on a scale of 0 to 100, three months after the start of treatment compared to people who had no treatment, usual care or placebo. Exercise is probably more effective for pain (35 studies, 2746 people) and probably slightly more effective for disability (38 studies, 2942 people) than no treatment, usual care or placebo at all follow-up periods.

Exercise may be more effective for pain (64 studies, 6295 people) and is probably more effective for disability (52 studies, 6004 people) than common treatments in the short and medium term.

Few studies reported mostly minor unwanted effects of exercise, most commonly increased low back pain and muscle soreness. However, the non-exercise groups reported similar types and numbers of unwanted effects.

What are the limitations of the evidence?

Our confidence in the evidence is limited. The studies used exercise therapy in different ways and so reported different results from each other. Some studies were very small – the average number of participants was just 98. It is possible that the design of some studies may have made the benefits of exercise seem larger than they are.

How up to date is this evidence?

The evidence is up to date to 28 April 2018.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings: Exercise treatment compared with no treatment, usual care or placebo for chronic low back pain

Outcomes	Anticipated absolute effects*		Number of participants (studies; study groups)	Certainty of the evidence (GRADE)[a]
	Without exercise (no treatment, usual care or placebo[b])	Difference with exercise		
Pain intensity (0-100; 0 = no pain): Earliest follow-up (time point closest to 3 months)	The mean pain intensity outcome at follow-up ranged across the no treatment, usual care or placebo comparison groups from 21 to 81.	The mean pain intensity in the exercise treatment groups was 15 points better than the no treatment, usual care or placebo comparison groups at follow-up (12 to 18 points better).	Participants = 2746; studies = 35; study groups = 47	⊕⊕⊕⊖ **moderate**[c] due to risk of inconsistency
Functional limitations (0-100; 0 = no functional limitations): Earliest follow-up (time point closest to 3 months)	The mean functional limitations outcome at follow-up ranged across the no treatment, usual care or placebo comparison groups from 7 to 58.	The mean functional limitations in the exercise treatment groups was 7 points better than the no treatment, usual care or placebo comparison groups at follow-up (5 to 8 points better).	Participants = 2942; studies = 38; study groups = 50	⊕⊕⊕⊖ **moderate**[d] due to potential publication bias

*The anticipated absolute effects without exercise come from the range of outcomes at follow-up in the no treatment, usual care or comparison groups.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of effect.
Moderate certainty: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of effect.
Very low certainty: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

[a] For all comparisons and outcomes - certainty may be raised to high given that further similar research is very unlikely to change our confidence in the estimate of effect.
[b] Observed effects were consistent when we performed a post hoc sensitivity analysis restricting to only placebo-controlled trials.
[c] Seven studies (10 groups; 526 participants) were judged to have high risk of bias (19% of participant data). Exclusion of these studies in sensitivity analysis did not change conclusions. Serious unexplained inconsistency (substantial heterogeneity I^2 = 75%, point estimates and confidence intervals varied considerably). No concerns about indirectness or imprecision. No evidence of publication bias (Egger's test, P = 0.30).
[d] Nine studies (13 groups; 495 participants) were judged to have high risk of bias (17% of participant data). Exclusion of these studies in sensitivity analysis did not change conclusions. Some unexplained inconsistency (moderate heterogeneity I^2 = 38%, point estimates and confidence intervals varied). No concerns about indirectness or imprecision. Some evidence of publication bias (Egger's test, P = 0.005).

Summary of findings 2. Summary of findings: Exercise treatment compared with other conservative treatments for chronic low back pain

Outcomes	Anticipated absolute effects*	Number of participants	Certainty of the evidence (GRADE)[a]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

	Without exercise (other conservative treatment comparison)	Difference with exercise	(studies; study groups)	
Pain intensity (0-100 points; 0 = no pain): Earliest follow-up (time point closest to 3 months)	The mean pain intensity outcome at follow-up ranged across the other conservative treatment comparison groups from 8 to 73 points.	The mean pain intensity in the exercise treatment groups was 9 points better than the other conservative treatment comparison groups at follow-up (6 to 13 points better).	Participants = 6295; studies = 64; study groups = 85	⊕⊕⊖⊖ **low**[b] due to risk of inconsistency and publication bias
Functional limitations (0-100 points; 0 = no functional limitations): Earliest follow-up (time point closest to 3 months)	The mean functional limitations outcome at follow-up ranged across the other conservative treatment comparison groups from 8 to 60 points.	The mean functional limitations in the exercise treatment groups was 4 points better than the other conservative treatment comparison groups at follow-up (2 to 6 points better).	Participants = 6004; studies = 52; study groups = 69	⊕⊕⊕⊖ **moderate**[c] due to risk of inconsistency

*The anticipated absolute effects without exercise come from the range of outcomes at follow-up in the other conservative treatment comparison groups.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of effect.
Moderate certainty: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of effect.
Very low certainty: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

[a] For all comparisons and outcomes - certainty may be raised one level given that further similar research is very unlikely to change our confidence in the estimate of effect.
[b] Eleven studies (14 groups; 835 participants) were judged to have high risk of bias (13% of participant data). Exclusion of these studies in sensitivity analysis did not change conclusions. Serious inconsistency (considerable heterogeneity I^2 = 81%, point estimates and confidence intervals varied considerably), partially explained by exploration of effect in specific comparison subgroups (I^2 < 75% for five of eight comparisons). No concerns about indirectness or imprecision. Some evidence of potential publication bias (Egger's test, P = 0.001).
[c] Seven studies (9 groups; 571 participants) were judged to have high risk of bias (10% of participant data). Exclusion of these studies in sensitivity analysis did not change conclusions. Serious inconsistency (considerable heterogeneity I^2 = 82%, point estimates and confidence intervals varied considerably), partially explained by exploration of effect in specific comparison subgroups (I^2 < 75% for five of eight comparisons). No concerns about indirectness or imprecision. No evidence of publication bias (Egger's test, P = 0.40).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

BACKGROUND

Low back pain has been the leading cause of disability globally for at least the past three decades (GBD Collaborators 2018) and results in enormous direct healthcare and lost productivity costs (Dagenais 2008; GBD Collaborators 2018; Hayden 2009; Lim 2012; Martin 2008). The majority of the social and economic costs associated with low back pain are attributable to people who have prolonged disability – experiencing chronic and/or recurrent low back pain (Hartvigsen 2018). Despite a large amount of research in the field, there remains uncertainty about the best treatment approach for this population.

Back pain is managed across every level of healthcare systems – from emergent care to community care. Pooled prevalence rates indicate that healthcare utilisation for low back pain varies internationally (Beyera 2019). Recent work suggests that management of low back pain with exercise is associated with lower healthcare system costs and improvements in quality adjusted life years when compared to usual care for acute and chronic patients (Miyamoto 2019).

Description of the condition

Low back pain is defined as pain, muscle tension, or stiffness localised below the costal margin and above the inferior gluteal folds, with or without pain referred to the leg(s). Most patients who experience low back pain are labelled as having 'non-specific low back pain', which is defined as symptoms not attributable to a recognisable, known, specific pathology (for example, fracture, ankylosing spondylitis, spondyloarthritis, infection, neoplasm, or metastasis). Several different structures of the back have been implicated in symptoms of non-specific low back pain, including the musculature, joints, and discs; but also psychosocial factors including maladaptive pain coping behaviours, high baseline functional impairment, presence of psychiatric comorbidities, and low general health status (Chou 2010). These psychosocial factors have been associated with persistent, disabling low back pain.

Most patients with a new episode of acute low back pain have a favourable prognosis, with 75% to 90% recovering in terms of pain and disability (Coste 2004; Grotle 2005), or returning to work within three to four weeks (Henschke 2009). In this review, we consider the effectiveness of exercise treatment for individuals whose symptoms persist beyond the acute stage. Previously, research studies and practice guidelines have defined individuals with low back pain episodes persisting beyond the acute stage as subacute and chronic. In this review, we focus on chronic low back pain which we have defined as pain, muscle tension, or stiffness lasting longer than 12 weeks or recurrent low back pain defined as two episodes in a year, lasting more than 24 hours, with more than 30 days pain-free between. The two back pain populations included in our study (chronic and recurrent) are often described as 'persistent', but for the purpose of this review, they will be referred to as chronic.

Individuals who progress to chronic low back pain often continue to experience long-standing pain. A systematic review of cohort studies found that 33% of patients with low back pain recover in the first three months, but 65% still report pain after one year (Itz 2013). Many factors including biophysical, psychological, social and genetic factors, and comorbidities can contribute to disabling, chronic low back pain. Factors associated with delayed recovery of chronic low back pain include previous sick leave due to low back pain, high disability or pain levels at the onset of chronicity, low levels of education, perceived risk of persistent pain, and psychological factors such as depression, anxiety, catastrophising and self-efficacy (Costa 2009b; Hartvigsen 2018). Specific predictors of poor outcome in patients with chronic low back pain have been inconsistent across prognosis studies. Chronic low back pain appears to be the result of nociceptive input combined with multiple complex factors. (Hayden 2009; Heitz 2009).

Description of the intervention

Exercise treatment is a management strategy that is widely used in low back pain. Exercise treatment encompasses a diverse set of treatments prescribed or planned by a health professional that include conducting specific activities, postures, or movements (or all). Exercise treatments are characterised by several interacting components. They are heterogeneous in treatment design (e.g. standard, individualised), dose (duration, frequency, intensity), delivery format (e.g. clinician supervised, group), specific exercise types (e.g. strengthening, stretching), and may be combined with other conservative treatments. Examples of exercise treatments include general physical fitness programmes delivered in a group setting, aerobic exercise in the form of walking programmes, and strengthening of specific muscles or groups of muscles to increase core stability.

How the intervention might work

Various types of exercise treatment are expected to be associated with different hypothesised mechanisms of effect. Exercise treatment may provide benefits to patients with chronic low back pain through the voluntary contraction of specific muscle groups, movement of the whole body, activities that improve postural musculature, stabilisation, and neuro-co-ordination, or a combination (Powell 2011). These physical therapies aim to increase muscle and joint strength, and improve muscle function and range of motion, with the goals to reduce pain, improve function, and speed patient recovery and return to usual activities. However, recent work has called on the back pain trialist community to prespecify their treatment target to improve the match between intervention and desired outcome (Wood 2019; Wood 2020a). In addition to the physical benefits of exercise treatment, there may be emotional and psychological benefits, which may lead to decreased pain and improved function.

Why it is important to do this review

Exercise is a common approach to the treatment of chronic low back pain. Clinical practice guidelines, including the American College of Physicians (ACP) (Qaseem 2017) and the UK National Institute for Health and Care Excellence (NICE) (National Guideline Centre 2016), recommend exercise as the first line of care for chronic low back pain and recommend exercise programmes that take individual needs, preferences, and capabilities into account. However, the most effective components of exercise interventions for patients with chronic low back pain have yet to be identified. In 2005, our team published the broad Cochrane review investigating the effectiveness of exercise for treating low back pain amongst acute, subacute and chronic populations; we included 43 randomised controlled trials (RCTs) on chronic low back pain in the review (Hayden 2005a; Hayden 2005b; Hayden 2005c). We

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

found evidence of effectiveness in chronic populations relative to comparisons at all follow-up periods; pooled mean improvement was 7.3 points (95% CI 3.7 to 10.9) for pain (out of 100), 2.5 points (1.0 to 3.9) for functional limitations (out of 100) at earliest follow-up. Analyses related to the 2005 Cochrane review found that exercise design and delivery characteristics were statistically associated with improved effectiveness for pain and functional limitations outcomes, more so than specific exercise types (Hayden 2005c). These analyses and results, however, were limited by the small number of trials available at the time.

In 2012, we split the 2005 Cochrane review into two separate Cochrane protocols, one of which that focusses on the chronic low back pain population (Hayden 2012) (presented here), and the other that focusses on the acute low back pain population (Jzelenberg 2011).

Over the past 15 years, many new RCTs have been published investigating the effectiveness of exercise treatment for chronic low back pain. The methodology for conducting and interpreting systematic reviews of interventions is also constantly evolving and several methodologies have been updated since the 2005 review was published (Higgins 2019).

OBJECTIVES

The primary objective of this systematic review is to assess the impact of exercise treatment on pain and functional limitations in adults with chronic non-specific low back pain compared to no treatment, usual care, placebo and other conservative treatments. Related investigations will explore exercise intervention characteristics that decrease pain and improve functional limitations and compare the effectiveness of different types of exercise treatment in adults with chronic non-specific low back pain.

METHODS

Criteria for considering studies for this review

Types of studies

We have included published reports of completed RCTs, with no language restrictions. We included only RCTs as they are the gold standard to assess the effectiveness of health-related interventions and there is a considerable volume of trials for synthesis.

Types of participants

We included studies involving adult participants with chronic nonspecific low back pain of more than 12 weeks' duration (defined as meaning back pain duration of the study group was greater than or equal to 12 weeks). We excluded studies that involved individuals with low back pain caused by specific pathologies (including fracture, ankylosing spondylitis, spondyloarthritis, infection, neoplasm, or metastasis) or conditions (for example, pregnancy), and studies that focussed exclusively on acute exacerbations of chronic low back pain. Studies of non-specific low back pain often include participants with a mixture of individual and symptom characteristics. Patients with leg pain, symptoms or signs consistent with radiculopathy (irritation of the nerve root) were included if back pain was their main complaint. For syntheses, we separately considered the effectiveness of exercise treatment for study populations that allowed study participants with leg pain or neurological symptoms to be included, and populations with moderate (12 weeks to three years) and long (greater than three years) symptom durations of chronic low back pain at baseline (Dunn 2006).

We included participants who were recruited from any setting, including healthcare, occupational, general and mixed populations.

Types of interventions

Exercise is defined as "a series of specific movements with the aim of training or developing the body by a routine practice or as physical training to promote good physical health" (Abenhaim 2000). This encompasses a heterogeneous set of treatments prescribed or planned by a health professional that include conducting specific activities, postures and/or movements with a goal to reduce pain and functional limitations.

Eligible exercise treatment types included muscle strengthening, stretching, core strengthening, flexibility and mobilising exercises, aerobic exercises, functional restoration, McKenzie therapy, and yoga (Table 1). Exercise treatments including mixed types of exercises were included, as were those exercises that fell into an 'other' category (e.g. Tai Chi, rock climbing). Exercise treatments were assigned to one or two main types, or designated as 'mixed' (exercise treatments incorporating three or more types of exercise in which one did not clearly predominate).

We characterised exercise treatment groups by the type of exercise delivered (listed above), exercise treatment design (individually-designed, partially individually-designed, standard), delivery type (independent, independent with follow-up, group-supervised, individually-supervised), dose (low dose, high dose; considering intensity and duration), and inclusion of additional treatments. We defined the design of the exercise treatment as "individually-designed," when the treating therapist completed a clinical history and physical examination and delivered an exercise treatment specifically designed for the individual participant; "partially individually-designed", when the exercise treatment included the same type of exercises for all participants but varied in intensity, duration, or "standard design", when a fixed exercise treatment was delivered to all participants (Hayden 2005c). We categorised delivery type as: "independent" when the participant met with the treating therapist once and then participated in the treatment at home without supervision; "independent with follow-up" when the participant met with the treating therapist initially, participated in the treatment at home, then received follow-up with the therapist at least every six weeks; "group-supervised" if the participant attended supervised group therapy sessions with two or more participants; and "individual" when the participant attended one-on-one sessions with the treating therapist. We categorised dose as "low dose" when the participants received fewer than 20 hours of intervention and "high dose" when the participant received 20 or more hours of intervention. Dose was calculated by multiplying the number of hours per session by the number of sessions. Finally, we indicated all non-exercise co-interventions that were included in the exercise treatment.

We included studies that compared exercise treatment to a) no treatment, usual care or placebo, b) other conservative therapy, or c) another exercise group. We categorised, and hereafter refer to, a comparison group as "no treatment" when participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

received placebo, no or minimal treatment or the same standard medical care they would have without being recruited to the study. We assumed a similar effect of placebo treatments (detuned electrotherapy in included trials) to the most common treatment type in this category, guideline recommended 'usual care', due to few available studies and difficulty of identifying an adequate placebo that blinds providers and participants. We categorised a comparison as "other conservative therapy" when participants received any non-exercise interventions (categorised as education, manual therapy, electrotherapy, psychological therapy, non-exercise physical therapy (multi-modal), back school, relaxation, anti-inflammatories); and "another exercise group" when the comparison included another eligible exercise intervention. We excluded studies that compared exercise treatment with spinal surgery.

Types of outcome measures

Proposed outcome measures for this study were fully described in our published protocol (Hayden 2012). Global improvement was shifted from a primary outcome to a secondary outcome, and satisfaction with treatment and reduction in frequency of analgesic use were removed as secondary outcomes due to limited and inconsistent reporting in early data extraction.

Primary outcomes

We fully reported the following primary outcomes:

- Pain intensity, measured by a pain scale (for example, visual analogue scale (VAS), numerical rating scale (NRS), or McGill pain score (Melzack 1975))
- Functional limitations, measured by a back pain-specific scale (for example, the Roland-Morris Disability Questionnaire (RMDQ) (Roland 2000), or the Oswestry Disability Index (ODI) (Fairbank 1980))

Secondary outcomes

We have also described measurement for the following secondary outcomes (secondary outcomes were not synthesised):

- Return to work/absenteeism
- Health-related quality of life (for example, SF-36 (Ware 1992) (as measured by the general health subscale), EuroQol (Euroquol Research Foundation 2019), general health (for example, as measured on a VAS scale) or similarly validated index)
- Global improvement or perceived recovery
- Adverse events

Outcome assessment data for primary outcomes were extracted for all reported time periods and were grouped into three follow-up periods for the purposes of analyses: short-term (post-treatment assessment closest to three months after randomisation (6-12 weeks)), moderate-term (post-treatment assessment closest to six months after randomisation (13-47 weeks)), and long-term follow-up (post-treatment assessment closest to 12 months after randomisation (48 weeks or more)). For primary analyses, we used the post-treatment time period closest to short-term, omitting studies that only reported time periods less than four weeks as this immediate follow-up was judged to be not meaningful.

Search methods for identification of studies

Electronic and reference searches for this systematic review were conducted periodically (approximately annually) between 2012 and 2020. In August 2019, we re-ran the full electronic search strategy to reconcile the previous, overlapping search returns and to facilitate reporting, then updated the electronic search in December 2020.

Electronic searches

We searched the following databases with no language restrictions to 7 December 2020:

- Cochrane Back and Neck Trials Register (Cochrane Central Register of Controlled Trials (CENTRAL) and the Cochrane Register of Studies (CRS)), searched using CRS Web (7 August 2019) and Wiley CENTRAL (7 December 2020)
- MEDLINE (Epub Ahead of Print, In-Process & Other Non-indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE) (OvidSP, 1946 to 7 December 2020)
- Embase (OvidSP, 1980 to 6 August 2019; Embase.com to 7 December 2020)
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO, 1981 to 7 December 2020)
- PsycINFO (OvidSP, 2002 to July Week 5 2019; EBSCO to 7 December 2020)
- Physiotherapy Evidence Database (PEDro) (6 August 2019)
- SPORTDiscus (EBSCO to 7 December 2020)
- Trial registry web sites: ClinicalTrials.gov and World Health Organization International Clinical Trials Registry Platform (ICTRP) (6 August 2019)

Citations were managed using EndNote X8 software. Search methods are consistent with the Cochrane Back and Neck methods guidelines (Furlan 2015) and the Cochrane Handbook (Higgins 2011). See Appendix 1 for the current strategy. A similar search strategy was used for the Cochrane Back and Neck's acute, chronic and older adult exercise reviews.

We have included studies up to 27 April 2018 in the analyses in this review and have identified all additional eligible studies up to 7 December 2020. Updated inclusion of studies will be conducted as part of a Network Systematic Review pilot project.

Searching other resources

A large number of systematic reviews on the effectiveness of exercise treatment for low back pain have been published. We screened cited references of all identified previously published systematic reviews and conducted citation searches of the previous version of this Cochrane review (Hayden 2005a) (searches conducted in August 2019).

Data collection and analysis

We followed a standard protocol for study selection and data extraction as recommended by the Cochrane Back and Neck methods guidelines (Furlan 2015).

Selection of studies

Two independent reviewers screened citations identified in the literature search based on the title, title and abstract, and

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

subsequently full text, for inclusion in the review. Consensus and, if necessary, a third reviewer were used to resolve disagreements. Conference proceedings, theses, opinion pieces, correspondence and stand-alone abstracts were excluded. Studies published in languages other than English were assessed for inclusion and included in the review whenever feasible using English language abstracts, translation tools and/or review by co-authors and colleagues familiar with the language of publication.

Data extraction and management

We extracted population characteristics (patient population source and setting, study inclusion criteria, mean duration of pain episode, symptom characteristics, mean age and sex of patient populations), intervention characteristics (description and types of exercise treatment, duration and number of treatment sessions, intervention delivery type, and any additional interventions), comparison group characteristics (category and main type for the other conservative treatments), and outcome data. Detailed description of the exercise characteristics extracted is provided in Table 1. Results for primary outcomes pain and functional limitations were extracted as final value scores for inclusion in meta-analyses.

Extracted data were recorded on pre-developed and tested forms using web-based electronic systematic review software (Distiller SR), Evidence Partners Incorporated, Ottawa, Canada) (2017 search updates, onward), and a specially designed Microsoft Access database (2012-2016 search updates). A single reviewer extracted study information and at least one other author checked data extractions from the original study publications. Reviewers were not blinded to authors, institution, or journal of publication as this was not feasible and because they were familiar with the literature. Trial authors were contacted to request any missing study characteristics, data points or risk of bias information, when required. For the sets of studies included prior to 2017, corresponding authors were sent an email with the clarifying questions indicated in the body of the email and sent their responses back via email. For the set of studies identified in the 2017-2018 updates, authors were sent an email with a link to a REDCap (Harris 2019) data capture form, where both extracted data and missing fields were clearly displayed. Authors completed the missing fields, and had the opportunity to correct any incorrect extracted data, then submitted the survey through REDCap.

Assessment of risk of bias in included studies

The risk of bias assessment for RCTs was conducted using the criteria recommended by Cochrane (Higgins 2011) and Cochrane Back and Neck (Furlan 2015). Risk of bias assessments were carried out independently by two authors, with consensus. Any disagreements that could not be resolved through discussion were referred to a third author. Inter-author reliability related to risk of bias was calculated as overall agreement and Kappa scores based on assessments before consensus judgements were reached.

We assessed potential bias related to: selection bias (randomisation, treatment allocation concealment, similarity at baseline), performance bias (blinding of participants and care provider, intention-to-treat), detection bias (blinding of outcome assessors, similar timing of outcome assessment), attrition (missing outcome data), reporting bias (selective outcome reporting), and other biases to assess adherence (avoidance of co-interventions, compliance) (see Table 2 and Table 3). Individual criteria were scored as "high risk", "low risk" or "unclear risk" and reported in a Risk of Bias table. A study with an overall low risk of bias was defined as fulfilling six or more of the twelve criteria items, and with no other fatal flaws; otherwise, a study was considered to have high potential risk. We used sensitivity analyses to assess the robustness of our findings by excluding from the syntheses studies considered to be at high risk of bias. For the most recent set of 64 studies extracted (2017-2018 search updates), we assessed items of the new Cochrane RoB 2 tool (Sterne 2019), which we used to recreate the Cochrane Back and Neck Group criteria (subsequently referred to as RoB 1 tool) presented here for consistency.

In addition to risk of bias, we also comprehensively assessed several characteristics of research integrity for the included studies. Research integrity was assessed using four criteria: publication in a predatory journal, presence of significant plagiarism, high risk of bias, and inadequate reporting of minimal basic CONSORT items. Publication in a predatory journal was assessed by two reviewers with consensus. A study was judged to be published in a predatory journal if they were not on the list of reputable publishers or organisations (e.g. Elsevier, Wiley-Blackwell, Springer, Taylor & Francis, Sage, BMJ, Informa, IOS Press, Oxford University Press, PLOS, ACP, American Medical Association, Canadian Medical Association, World Health Organization), were not on Cabell's white list, and were not listed on the Directory of Open Access Journals but were listed on either Beall's List or on www.PredatoryJournals.com.

Plagiarism was assessed by two independent reviewers with consensus using iThenticate 2020 plagiarism detection software. Our assessment was limited to the introduction, results, and discussion sections. We defined a case of plagiarism using a two-stage process. If a publication failed both stages of the assessment, they were deemed to include plagiarism. In the first stage, we assessed publications to determine if 3% or more of the publication was duplicated from a single source (other than the authors own publications). For those studies with 3% or more duplication from a single source, in the second stage, we assessed if the publication included more than 80% duplication of at least two full sentences (we ignored common phrases or expressions). For both stages, we excluded the text from abstracts, methods, and references lists from assessment. This manual approach, supported by algorithmic plagiarism software, was informed by methods of Lykkesfeldt 2016 and Higgins 2016, and supported by a recent commentary (Weber-Wulff 2019). We chose a liberal approach to cautiously classify a publication as including plagiarism. Given the lack of evidence and clear guidance in the plagiarism field, we feel this approach was appropriate. Twelve studies could not be assessed using iThenticate due to incompatibility issues with the study file format (Bentsen 1997; Bronfort 1996; Dalichau 2000; Gao 2006; Hildebrandt 2000; Kankaanpaa 1999; Kendall 1968; Li 2008; Lie 1999; Muharram 2011; Petersen 2002; Shen 2009).

A study was judged to be inadequately reported if the study failed to report one or more minimal baseline study population characteristics (age, sex, pain intensity, duration of low back pain) and also failed to include or describe one or more key reporting components (study flow chart, study eligibility criteria, method of randomisation, study population description).

A study was judged to be at high risk of bias by two independent reviewers using the RoB 1 tool, as described above.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Studies were excluded from the review due to potential research integrity concerns if they were judged to be either plagiarised or published in a predatory journal in addition to at least one other research integrity concern (high risk of bias, inadequate reporting). For more information regarding the process for developing the research integrity criteria, further description of integrity characteristics and their relationship to other trial methods and reporting characteristics, please see our articles on this topic (Hayden 2020a; Hayden 2021a).

Measures of treatment effect

In the low back pain literature, several outcome measures are used to assess the constructs of pain intensity (for example, 10 mm or 100 mm visual analogue scales (VAS), or 0 to 10 numerical rating scales (NRS)) (see reviews by Chiarotto 2019; Von Korff 2000) and back-related functioning (for example, the 24-point Roland-Morris Disability Questionnaire, or the Oswestry Disability Index scored out of 100) (see reviews by Chiarotto 2016; Kopec 2000). Our primary meta-analyses assessed pain (measured as VAS or NRS) and back-related functioning (measured as Roland-Morris Disability Questionnaire or Oswestry Disability Index, preferentially selecting the Roland-Morris Disability Questionnaire) treatment effects (core outcomes reliably reported in eligible RCTs). Research has established agreement of the VAS and NRS pain scales and this same research established that either can be used interchangeably to assess back pain intensity (Shafshak 2020). With respect to functional limitations, there is evidence that the Roland-Morris Disability Questionnaire and the Oswestry Disability Index are highly correlated and similarly responsive enough for meta-analysis (Chiarotto 2016). The individual trial outcomes for pain and functioning were re-scaled to 0 to 100 points for meta-analyses (for example, a VAS pain score (standard deviation) of 5.1 (2.3) out of 10 would be re-scaled to 51 (23) out of 100), where positive mean effect sizes indicates improvement (i.e. decreased pain and decreased functional limitations). Re-scaling is accepted and common in the back pain field (Kopec 2000) and facilitates comparison and interpretability of the syntheses.

We interpreted a clinically important difference in results for exercise treatment compared to no treatment as a difference in pain of 15 points out of 100, and difference in function of 10 points, calculated as the smallest worthwhile effects based on a 30% reduction, the estimated upper interquartile range for patient-reported smallest worthwhile effect with physiotherapy treatment compared to without (Ferreira 2013), from the average baseline pain (50.9, 95% CI 49.1 to 52.8), and average baseline functional limitations (38.9, 95% CI 35.8 to 42.0), respectively, for included studies. We interpreted smaller differences in effectiveness of exercise treatment compared to other conservative treatments as 'probably meaningful' when the 95% confidence interval was entirely on one side of the no effect line. This is relevant given similar inconveniences and adverse effects for comparison treatments considered in this review (Qaseem 2017). Differences were considered statistically significant at the 5% level.

Unit of analysis issues

Opportunities for unit of analysis issues in this systematic review were due to: 1. repeated observations on participants, 2. studies with more than two intervention groups, 3. non-parallel study designs (cross-over and cluster-randomised trials).

Repeated observations on participants

We assessed available data from multiple follow-up periods of the same treatment groups by defining different outcomes based on different periods of follow-up: short-term (post-treatment assessment closest to three months after randomisation (6-12 weeks)), moderate-term (post-treatment assessment closest to six months after randomisation (13-47 weeks)), and long-term follow-up (post-treatment assessment closest to 12 months after randomisation (48 weeks or more)).

Studies with more than two relevant intervention groups

Exercise treatment groups from included trials were included in the meta-analyses if they had an independent no treatment or other conservative treatment comparison group. This requirement meant that studies with no comparison group (i.e. trials that contrasted multiple exercise treatment groups only) were not included in the primary meta-analyses, and comparison groups were not "double counted" in the meta-analyses. This latter criterion was necessary to avoid correlation in effect sizes resulting from the use of repeated comparison data. We selected groups to include in analyses based on relevance to the review selection criteria. If multiple relevant comparison groups existed in a trial, we split the 'shared' group into two or more groups with smaller sample sizes to allow inclusion in meta-analyses; this allowed investigation of heterogeneity across treatment arms (Higgins 2011).

Our additional network meta-analysis and meta-regression analyses have been reported in a related peer-reviewed publication (Hayden 2021b). These analyses use available data from all treatment groups including indirect and direct comparisons and allow inclusion of data from studies with only multiple exercise group comparisons. We briefly reported these related study results and referred to them in the discussion.

Non-parallel study designs

For any eligible cross-over trials, we conservatively extracted data from the pre-cross-over and post-cross-over periods as if the trial was a parallel trial. For cluster-randomised trials, we assessed the appropriateness of analyses to control for correlation of individuals within clusters. If clustering was not accounted for in a trial, we would reduce the sample size to an estimate of the effective sample size for continuous outcome data in this review.

Dealing with missing data

Any missing data were first requested from individual study authors. In the event that the original authors could not be contacted, or were unable to supply missing data, missing variance scores were imputed using the mean variance from studies with similar populations of low back pain. Where data in study papers were reported as a median and interquartile range (IQR), for studies with moderate to large sample sizes ($n > 25$), we used the median to estimate the mean; for studies with small sample sizes, we planned to use the formula proposed by Hozo and colleagues (Hozo 2005). The standard deviation was calculated with the width of the IQR equivalent to 1.35 times the standard deviation (Higgins 2002). In studies where a range was presented along with the median instead of an IQR, the standard deviation was estimated to be one-quarter of the range (Higgins 2011). We conducted sensitivity analyses to

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

explore the impact of assumptions about missing or incomplete data.

Assessment of heterogeneity

We pooled data if studies were clinically homogeneous with regard to study population, intervention, and outcomes. Statistical heterogeneity was assessed using I^2 statistics and confidence intervals (Higgins 2011). The I^2 statistic describes how much of the total variation across included studies can be attributed to heterogeneity rather than to chance. We interpreted the I^2 statistic according to rough guidance provided in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011):

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

These overlapping intervals reflect that the interpretation of the I^2 statistic depends on the value as well as the size and direction of the treatment effect and variance of the I^2 estimate. The assessment of heterogeneity informed our appraisal of the certainty of evidence available using the GRADE framework.

Assessment of reporting biases

We evaluated between-study publication bias with Egger's test and funnel plots (Egger 1997). We also assessed within-study publication bias by considering whether all expected outcomes were reported for each follow-up period of interest. If we were suspicious of reporting bias for a particular study, we would have contacted study authors for additional information and attempted to locate the study protocol to determine if there were differences between the protocol and publication; this was not required.

Data synthesis

Forest plots were constructed in Review Manager 5.4 (Review Manager 2020) in order to give a pictorial overview of all study results. We used quantitative meta-analyses, which were conducted for continuous outcome measures by pooling weighted mean differences with random-effects models and data from at least three studies (DerSimonian 1986). Prediction intervals were calculated (Riley 2011) to estimate the likely effect of exercise treatment in individual study settings. Stata 14 (StataCorp 2015) (for descriptive analyses and data checking), Review Manager 5.4 (Review Manager 2020), and R packages (R Core Team 2013) (for meta-analyses) were used.

Subgroup analysis and investigation of heterogeneity

We used subgroup analyses to explore heterogeneity due to study-level variables, such as population source and characteristics. We characterised the population sources as healthcare (primary, secondary or tertiary care centres), occupational (patients presenting to occupational healthcare facilities or personnel in compensatory situations), or from a general or mixed population (e.g. including individuals recruited by newspaper advertisements) to differentiate the studies of patients in typical treatment settings (healthcare and occupational) from those including individuals with low back pain who may not normally present for treatment. We separately considered the effect of exercise treatment for populations with moderate (12 weeks to three years) and long (greater than three years) duration of chronic low back pain at baseline (Dunn 2006). We also considered subgroup analyses of study populations with radicular symptoms (leg pain and/or neurological symptoms); no participants with radicular symptoms, including participants with leg pain and/or neurological symptoms; or radicular symptoms not specified. An additional subgroup analysis, not described in the original protocol, explored specific types of conservative treatment comparisons, including advice or education alone, manual therapy, electrotherapy, psychological therapy, back school, relaxation, anti-inflammatory agents or analgesics, or non-exercise physical therapy interventions. We compared primary outcomes for subgroups of studies (Song 2003).

Sensitivity analysis

-We conducted sensitivity analyses in order to explain possible sources of heterogeneity between studies and to determine the robustness of our original analyses. We conducted sensitivity analyses excluding studies with high risk of bias (and exploring the impact of RoB 1 and RoB 2 assessment), to test assumptions about imputing data for studies that did not adequately present variance scores, or where median values were presented, and including only studies with the most homogeneous outcome measures for functional limitations (Roland-Morris Disability Questionnaire only).

We conducted sensitivity analyses that omitted study outcome data from meta-analyses that were judged to be outliers to explore the impact of extreme study results on review conclusions. A study was judged to have an improbable or outlying mean outcome if the absolute difference between any exercise group and any comparison group over all available follow-ups was greater than a predetermined threshold of 30/100 for pain and 20/100 for functional limitations, selected based on clinical judgement.

Summary of findings and assessment of the certainty of the evidence

We have presented results for our primary outcomes of pain and functional limitations in the summary of findings tables, produced using the GRADE framework (GRADEpro). These tables include a summary of the number of studies included in the review, continuous outcome measure results reflecting absolute treatment effects, and a statement regarding the overall quality of the evidence available (see Appendix 2 for description).

RESULTS

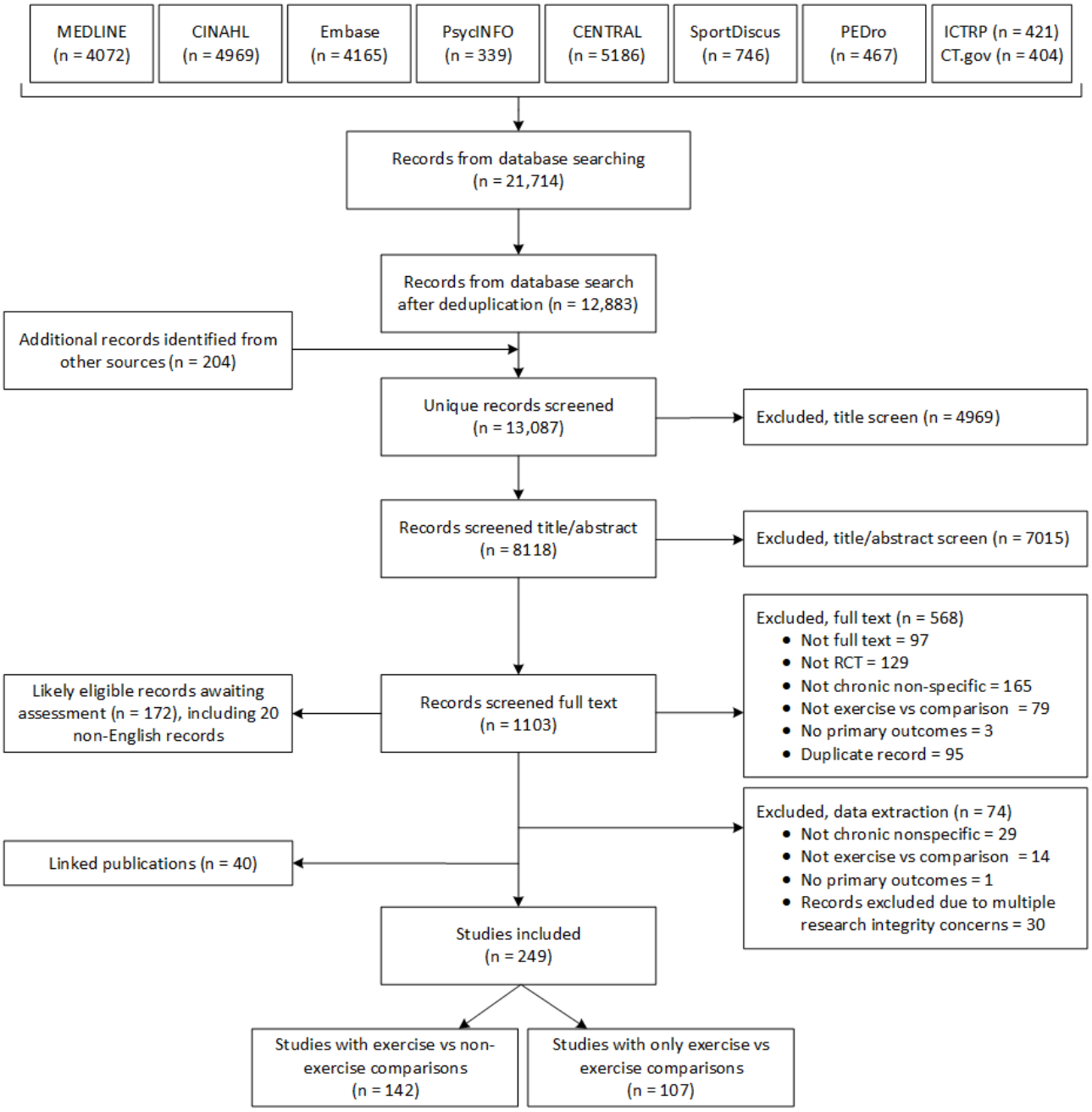
Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

Figure 1 provides detail about the flow of citations and studies through our search and selection process.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 1. Study flow diagram.



Results of the search

Our electronic database searches identified 12,883 unique citations up to December 2020. We screened 1103 full-text publications, identified 279 as potentially eligible, and included 249 RCTs of exercise treatment for chronic low back pain in this review. A further 172 RCTs have been identified as likely to be eligible and are listed in Characteristics of studies awaiting classification.

Included studies

A descriptive summary of the characteristics of the 249 studies (24,486 included participants) is in Figure 2. A complete description of each included study is presented in the Characteristics of included studies and citations are available: References: Included studies.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 2. Summary characteristics of included studies.

Characteristic		All studies (249 trials; 24,486 participants)	Studies with data available, any comparisons (203 trials; 19,633 participants)	Studies with data available, with non-exercise comparison (117 trials; 12,872 participants)
Population source	Healthcare	144 (57.8%)	118 (58.1%)	60 (51.3%)
	Occupational	23 (9.2%)	20 (9.9%)	14 (12.0%)
	General or mixed	71 (28.5%)	58 (28.6%)	39 (33.3%)
	Other	1 (0.4%)	1 (0.5%)	1 (0.9%)
	Not specified	10 (4.0%)	6 (3.0%)	3 (2.6%)
Median # participants (IQR; range)		70 (42, 125; 10, 722)	70 (43, 120; 10, 722)	74 (45, 146; 14, 722)
Mean age (95% CI)[a]		43.7 (42.5, 44.8)	43.5 (42.3, 44.7)	43.7 (42.1, 45.4)
Mean proportion male (95% CI)[b]		41.4 (38.5, 44.4)	40.8 (37.5, 44.1)	40.3 (35.8, 44.9)
LBP category	Chronic	213 (85.5%)	176 (86.7%)	100 (85.5%)
	Mixed Chronic	32 (12.9%)	26 (12.8%)	17 (14.5%)
	Not specified	4 (1.6%)	1 (0.5%)	0
Recurrences included	Yes	19 (7.6%)	16 (7.9%)	9 (7.7%)
	No	230 (92.4%)	187 (92.1%)	108 (92.3%)
Symptom duration, mean	< 3 years	78 (31.3%)	65 (32.0%)	30 (25.6%)
	≥ 3 years	72 (28.9%)	64 (31.5%)	36 (30.8%)
	Not specified	99 (39.8%)	74 (36.5%)	51 (43.6%)
Pain severity at baseline, study mean (95% CI)[c]		50.9 (49.1, 52.8)	51.2 (49.3, 53.1)	51.2 (48.6, 53.8)
Functional limitations at baseline, study mean (95% CI)[d]		38.9 (35.8, 42.0)	38.0 (34.7, 41.4)	40.0 (34.5, 45.6)
Leg pain or neurologic symptoms	All	0	0	0
	Some included	100 (40.2%)	69 (34.0%)	40 (34.2%)
	None included	85 (34.1%)	84 (41.4%)	47 (40.2%)
	Not specified	64 (25.7%)	50 (24.6%)	30 (25.6%)
Total # exercise groups		426 (72.8%)	346 (73.2%)	159 (55.6%)
Total # comparison groups		159 (37.3%)	127 (36.7%)	127 (79.9%)
Exercise groups per study	One	98 (39.4%)	80 (39.4%)	80 (68.4%)
	Two	130 (52.2%)	107 (52.7%)	33 (28.2%)
	Three	16 (6.4%)	12 (5.9%)	3 (2.6%)
	Four	5 (2.0%)	4 (2.0%)	1 (0.9%)
Comparison groups per study	One	127 (51.0%)	107 (52.7%)	107 (91.5%)
	Two	14 (5.6%)	10 (4.9%)	10 (8.6%)
	Three	0	0	0
	Four	1 (0.4%)	0	0
Outcomes assessed	Pain	223 (89.6%)	178 (87.7%)	100 (85.5%)
	Functional limitations	223 (89.6%)	161 (79.3%)	94 (80.3%)
	Work status	36 (14.5%)	27 (13.3%)	13 (11.1%)
	HRQoL	71 (28.5%)	64 (31.5%)	40 (34.2%)
	Global Perceived Recovery	47 (18.9%)	38 (18.7%)	23 (19.7%)
	Adverse Outcomes[e]	86 (36.4%)	74 (37.8%)	51 (45.9%)
Follow-up periods available	Immediate (< 6 weeks)	48 (19.3%)	30 (14.8%)	20 (17.1%)
	Short term (6-12 weeks)	184 (73.9%)	159 (78.3%)	88 (75.2%)
	Moderate (13-47 weeks)	121 (48.6%)	110 (54.2%)	70 (59.8%)
	Long term (48+ weeks)	69 (27.7%)	57 (28.1%)	35 (29.9%)
Funding source reported[f]	Funding received	134 (53.8%)	112 (56.0%)	65 (57.0%)
	No funding	28 (11.2%)	22 (11.0%)	9 (7.9%)
	No statement	80 (32.1%)	66 (33.0%)	40 (35.1%)

[a] Sample available: n=233, n=196, n=114; [b] Sample available: n=222, n=181, n=108; [c] Sample available: n=211, n=187, n=110; [d] Sample available: n=203, n=179, n=99; [e] Sample available: n=236, n=196, n=111; [f] Sample available: n=242, n=200, n=114; IQR=interquartile range; CI=confidence interval; LBP=low back pain; HRQoL=health-related quality of life.

Study design and setting

Studies included in this review were conducted in geographically diverse settings. The trials were conducted in Europe (49%, 122 studies), Asia (15%, 38 studies), North America (13%, 33 studies), the Middle East (10%, 24 studies), South America (6%, 15 studies), Oceania (5%, 12 studies), and Africa (2%, 4 studies); there was also one multi-regional trial. Most of the trials included in this review have been published since 2011 (56%, 139) with an increasing trend since the earliest included trial, published in 1968. Over half (58%) of included studies were conducted in healthcare settings, with

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

29% and 9% of studies conducted in general/mixed population settings, or in occupational settings, respectively (4% of studies did not describe their population source). Sixty-one per cent of studies (151 trials) examined the effectiveness of two or more different types of exercise treatment, and 57% (142 trials) compared exercise to a non-exercise comparison treatment (no treatment, or other conservative therapy). Each study included a median of two groups with a range of two to eight. Figure 2 describes the distribution of exercise and comparison treatment groups across the included studies.

The distribution of the number of subjects participating in the included studies was skewed to the right, with mean of 98 (95% CI: 87 to 109), a median of 70 participants (IQR: 42 to 125), and a range of 10 to 722 study participants. The average number of subjects per treatment group was 42 (95% CI: 38 to 46). Only 19% of included trials (39 of 203 studies that adequately reported data) had sufficient sample size to estimate a statistically significant realistic effect size. Trial registration or prior protocol publication was reported for only 29% of included trials (69 trials). Most study publications included an ethics statement (76%, 185 of 242 studies assessed) and a conflict of interest statement (81%, 195 of 242 assessed). Work status, health-related quality of life, and global perceived recovery were measured in 14%, 29%, and 19% of studies, respectively. Only 44 included studies (18% of studies) included a measure of each of the recommended core outcomes for low back pain trials: pain intensity, functional limitations and health-related quality of life.

Participant characteristics

The study participants in the 249 included trials had a mean age of 43.7 years (95% CI: 42.5 to 44.8); on average, 59% of study populations were female. The participants reported a mean pain intensity at baseline of 51 points on a 100-point scale (95% CI: 49.1 to 52.8), a mean functional limitations score at baseline of 38.0 (35.9 to 40.1), and 100 included studies (40%) had some participants with leg pain or neurologic symptoms in addition to their main complaint of chronic low back pain. The study samples in 78 trials (31%) included participants with chronic low back pain of moderate symptom duration (mean duration 12 weeks to three years) and 72 trials (29%) included participants with longer symptom duration of greater than three years; 99 trials did not specify the mean duration of chronic low back pain of participants. Nineteen trials (8%) included participants with chronic recurrent low back pain symptoms.

Interventions

The 249 included studies had a total of 585 treatment groups: 426 exercise groups and 159 comparison groups. Most exercise treatment groups included mixed types of exercises (i.e. greater than two types) (23%, 110 study groups). The most common specific types of exercises investigated were core strengthening (27%, 131 study groups) and Pilates (6%, 29 study groups), general strengthening exercises (12%, 57 study groups), stretching (11%, 51 study groups), and aerobic exercise (8%, 41 study groups). Other specific types of exercises were investigated in 3% of study groups (16 study groups). We classified 46% of exercise treatments as back-specific (196 study groups), and 29% as whole body exercises (125 study groups). The exercise treatments were classified as standardised for 46% of exercise groups (195 study groups), and were most commonly delivered in an individually-supervised (38%, 163 study groups), or group-supervised setting (38%, 162 study groups), and were of low dose/intensity (< 20 hours) (68%, 288 study groups). More than half of studies included additional treatment with the primary exercise intervention (58%, 247 study groups), including education or advice (32%, 137 study groups), electrotherapy (11%, 46 study groups), or manual therapy (5%, 21 study groups). Psychological interventions (e.g. cognitive behaviour therapy, behavioural support, or psychiatry sessions) were included with the exercise intervention for 14% (58 study groups) of exercise treatment groups.

Non-exercise treatment comparison interventions in the included studies were: no treatment (39%, 62 study groups, including four studies with placebo groups), and other conservative treatments (61%, 97 study groups). Other conservative treatment comparison groups included education (32%, 31 study groups), physical therapy (multi-modal non-exercise) (30%, 29 study groups), manual therapy (15%, 15 study groups), electrotherapy (6%, 6 study groups), or psychological therapy (6%, 6 study groups).

Outcomes

Pain intensity outcomes were reported in 90% of included studies (223 studies). Measurement scales used were most commonly the Visual Analogue Scale (60%, 133 studies) or the Numeric Rating Scale (22%, 50 studies). No other pain outcome measures were reported in 10 or more studies. Our meta-analyses of pain intensity considered only studies that measured with this outcome with the Visual Analogue Scale or the Numeric Rating Scale and adequately reported outcome data (178 studies).

Functional limitations were reported in 90% of included studies (223 studies). The Roland-Morris Disability Questionnaire and the Oswestry Disability Index were the most commonly reported outcome measures, reported in 42% (93 studies), and 37% (83 studies), respectively. No other functional limitations outcome measures were reported in 10 or more studies. Our meta-analyses of functional limitations considered only studies that measured this outcome with the Roland-Morris Disability Questionnaire or the Oswestry Disability Index and adequately reported outcome data (161 studies).

Other outcomes were not consistently reported in the included studies. Global perceived recovery was reported in 19% of included studies (47 studies), with the Global Perceived Effect measure reported by 15 included studies. Health-related quality of life was reported in 29% of included studies (71 studies). The SF-36 was the most common health-related quality of life scale used, reported in 63% (45 of 71 studies). The EuroQol-5D and SF-12 were reported in 17% (12 of 71) and 14% of included studies (10 of 71), respectively. Work status after treatment, for example, return to work for people off work due to a workers' compensation claim, was reported in 14% of included studies (36 studies). We excluded one study that did not report any patient-reported outcomes, according to our protocol (described in the Characteristics of excluded studies).

Adverse effects were measured in 86 studies reporting on 120 exercise groups and 58 comparison groups. There was no mention of measurement of adverse effects in 150 studies (60%) (13 were non-English publications - 5%). Measurement was variable in the 86 studies that mentioned adverse effects; only presence or absence was reported for 74 studies with 112 exercise groups and 42 comparison groups. In these studies, one or more adverse events

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

was reported for 33% of exercise groups (none = 71; one or more = 37), and 29% of comparison groups (none = 30; one or more = 12). Only 12 included studies appeared to measure adverse effects in a systematic way (for all participants at regular time points as part of the study protocol rather than 'as reported'). The twelve studies (19 exercise groups; 7 comparison groups) that systematically measured adverse events reported a range of between zero and an average of 3.2 adverse events (350 reported for 111 participants) per participant in the exercise groups (median 0.14; IQR: 0.01 to 0.57), and an average of 0.02 (1 reported for 64 participants) to 0.6 per person (25 reported for 43 participants) in comparison groups (median 0.12; IQR: 0.02 to 0.32). In the 15 exercise groups reporting adverse events, 11 reported minor only (most commonly increased low back pain and muscle soreness), and four reported minor and major (including disc herniation).

The available follow-up time periods reported in the included studies ranged from immediate to two years. Twelve studies reported only time periods less than four weeks so were not included in any syntheses. One hundred and eighty-four studies reported short-term follow-up. Moderate follow-up outcomes were available for 49% of included studies (121 studies), and long-term follow-up was available for 28% of included studies (69 studies). The median follow-up duration for studies included in the earliest follow-up time period for our review primary analyses was 12 weeks (IQR: 8 to 12 weeks), including 18 studies that reported follow-up time periods between four- and six-weeks post-treatment.

Study data available for analyses

In total, 261 trial authors of the 279 potentially eligible trials were contacted to request missing information. We received responses that resulted in changes to study data from 80 authors (45/186 using email data request - 24% response rate; and 35/75 using REDCap data request - 47% response rate). Study mean outcome data for meta-analyses and related network meta-analyses remained incomplete for 23 studies following author contact (9% of studies available for quantitative analyses). Standard deviations were missing or considered to be improbable for 39 (26%) of the studies available for meta-analyses (22 studies with SD missing before imputation; 10 studies with low outlying or improbable SD values); these studies were included in our primary analyses with data imputed from studies similar on two characteristics (population and chronicity), and subsequently excluded in sensitivity analyses.

Sufficient data for quantitative analyses were available from 203 included studies (19,633 participants) for pain or functional limitations outcomes. The characteristics of these studies were not substantially different from the 46 studies that did not adequately report outcome data (i.e. studies missing data for analyses or pre-defined homogeneous measures of pain and functional limitations), except that they were less likely to report pain (92% versus 79%) and functional limitations outcome measures (91% versus 81%) (Figure 2). We included data from 117 studies (12,872 participants) that had a non-exercise comparison treatment group in at least one of our primary meta-analyses. We also described the characteristics of this subset of studies in Figure 2.

Excluded studies

In total, 491 studies were excluded at the full-text level; a further 74 studies were excluded at the data extraction stage. Of the studies excluded at the full-text stage, 117 studies were not RCTs, 141 studies did not have a majority chronic population, 82 studies were not full-text publications, 76 studies were exact duplicates of previously assessed studies (i.e. publications that inappropriately passed through the duplicate citation screen and were identified as a duplicate at the full-text or data extraction stage), 72 studies did not have an eligible comparison group, and three studies did not include a relevant outcome. Of the 74 studies excluded at the data extraction stage, 29 studies did not have a majority chronic population, 14 studies did not have an eligible comparison group, one study did not include a relevant outcome and 30 studies were excluded due to research integrity concerns.

Of the 30 studies excluded from the review for integrity reasons (1218 participants), 20 studies were excluded as they were judged to be published in a predatory journal in addition to at least one other integrity concern (high risk of bias = 10; inadequate reporting = 19). Eight studies were excluded as they were judged to have substantial plagiarism in addition to at least one other integrity concern (high risk of bias = 3; inadequate reporting = 8). Two studies were excluded as they were judged to be both published in a predatory journal and to have substantial plagiarism; both included an additional research integrity concern (inadequate reporting = 2). Two unrelated trials (different author groups, different countries of conduct) were excluded for research integrity reasons as they presented identical data which could not be explained, despite attempts to contact both author groups, their institutions, and the publishing journals.

The exclusion of studies due to multiple research integrity concerns resulted in 0 to 13% of studies (0 to 6% of participants) being excluded from the primary review meta-analysis comparisons of pain outcomes, and 0 to 11% of studies (0 to 3% of participants) being excluded from primary meta-analysis comparisons of functional limitations outcomes. This is lower than the overall numbers excluded because these studies often did not report data that was usable. There was no substantial impact on the primary meta-analysis effect sizes for any outcomes or follow-up periods with these exclusions.

A complete description of reasons for exclusion for the 74 studies excluded at the data extraction stage can be found in the Characteristics of excluded studies table.

Studies awaiting classification

In total, 172 studies likely to be eligible for inclusion in this review are awaiting classification. These studies were identified in the most recent search update (152 studies), or are subject to translation to English language (20 studies). They have been screened through title abstract and full-text screening and are awaiting data extraction.

Ongoing studies

At each search update, we screened trial registries to identify published versions of completed trials. We did not identify additional ongoing trials due to the high volume of completed trials to include in the review at each update.

Risk of bias in included studies

Of the 249 included trials, 129 (52%) were at risk of selection bias, 195 (78%) were at risk of performance bias, 182 (73%) were at risk of detection bias, 49 (20%) were at risk of attrition bias, 22 (9%) were at risk of reporting bias, and 173 (59%) were susceptible to

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

other potential sources of bias. Although the majority of trials were susceptible to some level of bias, three studies had low risk of bias in all domains assessed (Dougherty 2014; Garcia 2017; Mannion 1999).

Allocation

In total, 120 trials (48%) were deemed to have an appropriate method of randomisation, sufficient allocation concealment, and balanced baseline variables, and thus were at a low risk of selection bias. Sixty-three trials had unclear reporting of the method of sequence generation and 59 had unclear reporting of the method of allocation concealment while 19 did not present any baseline characteristics. Generic statements such as "participants were randomised" without any further information were considered to be of unclear risk of bias.

Blinding

While blinding is often unfeasible due to the nature of the interventions in question, a trial could still be considered at a low risk for performance bias if the risks of biasing effects were adequately mitigated. For example, equally plausible treatment options would lead participants to believe that treatments were comparable in effectiveness, reducing the risk of bias due to lack of blinding of participants. Further, if distinct personnel provided care for a group in their area of expertise, the risk of non-blinded personnel may also be reduced.

Performance bias

Eighty-two trials were considered to have adequately blinded patients, 65 trials had adequately blinded personnel with 59 studies being at a low risk of bias in both domains. Inadequate information was provided to make a decision about participant blinding in 12 studies and personnel blinding in 10 studies; these studies were deemed to have an unclear risk of bias. Sixty-eight trials did not conduct an intent-to-treat analysis and a further 21 did not clearly state whether they did or not.

Of the 185 studies assessed using RoB 1, 183 (99%) were deemed to be at high risk of performance bias. Conversely, within the 64 studies assessed using RoB 2, only 12 (19%) were classified as high risk. None of these 64 trials attempted to blind both patients and personnel.

Detection bias

Our primary outcomes, pain and functional limitations, are patient-reported subjective outcomes. In total, 68 trials were found to have low risk of bias due to outcome assessor blinding. Forty-one studies had inadequate information to make a decision and were deemed to have an unclear risk of bias. Three trials did not have similar timing of assessments across groups and five were missing adequate information to make a judgement.

Incomplete outcome data

Two hundred trials had losses to follow-up that were minimal or that were unlikely to bias the outcome. Dropouts that were unlikely to bias results included those that were balanced between groups or were caused by reasons deemed unrelated to outcomes. Studies could also be rated as low risk for attrition bias if they had evidence of a lack of biasing effect like the results of a sensitivity analysis. Fifteen studies had insufficient information on missing data and were classified as having unclear risk.

Selective reporting

In total, 227 trials were judged to have a low risk of selective reporting bias. When protocols, statistical analysis plans, or trial registrations were not available, a study could still be rated as low risk if there were no omissions of results based on the within-paper's methods section and no obvious omissions such as a lack of pain outcomes or if outcomes were missing from certain follow-up. Seventeen trials did not have adequate reporting and were deemed to have unclear risk.

Other potential sources of bias

In total, 173 trials were deemed to have at least one other potential source of bias, either via a lack of compliance or unbalanced co-interventions. Fifty-seven trials reported poor study adherence rates and a further 75 did not report on compliance. Thirty-nine trials had an imbalance in the external co-interventions sought by patients and a further 82 did not collect any information or attempt to restrict participants' external activities. A study could be rated at low risk for bias if, despite not collecting co-intervention information, patients were instructed to refrain from certain non-exercise treatments.

Intra-rater agreement

Out of 2528 total items, reviewers agreed on 2270 for an overall agreement of 90% and a Kappa value of 0.84 (95% CI 0.82 to 0.85), indicating a high level of agreement. Trials that were assessed using the Cochrane RoB 1 tool had higher agreement (96%, Kappa 0.93, 95% CI 0.92 to 0.94) than those that were assessed using the Cochrane RoB 2 tool (75%, Kappa 0.45, 95% CI 0.42 to 0.46).

Comparisons of RoB 1 and RoB 2

For the most recent set of 64 studies included in this review (included and extracted since the publication of the RoB 2 tool), we used the RoB 2 items to recreate a RoB 1 assessment for consistency across the review. This assessment often included incorporating clarifying questions included in RoB 2 to prompt a more fulsome response. In these domains – blinding of patients, blinding of personnel, missing data, and selective reporting – there were fewer trials classified as having an unclear risk of bias when using RoB 2 (7% of items vs 2% of items classified as unclear). This was not true, however, for the domain of outcome assessor blinding which had a much higher proportion of studies judged to be of unclear risk when using RoB 2.

The domains looking at the blinding of patients and personnel were also more frequently deemed to be low risk of bias when using RoB 2 – 94% of studies were low risk in these domains in RoB 2 vs 7% for RoB 1. Although this could, in part, be due to systematic differences between the sets of trials assessed, in exercise trials blinding is often unfeasible and RoB 2 allowed for a more in-depth assessment of the risk of bias that resulted in trials not being classified as high risk in blinding domains.

Post hoc analysis comparing RoB 1 and RoB 2 assessment tools for trials with specific comparison groups found the differences between the tool judgements (i.e. RoB 2 with fewer trials judged as 'unclear' risk of bias) were greater in trials with no treatment than those with other comparisons. Similar to other comparison types,

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

no treatment comparisons were more likely to be judged low risk of bias for blinding of patients and personnel with RoB 2 than with RoB 1.

Effects of interventions

See: **Summary of findings 1** Summary of findings: Exercise treatment compared with no treatment, usual care or placebo for chronic low back pain; **Summary of findings 2** Summary of findings: Exercise treatment compared with other conservative treatments for chronic low back pain

Complete meta-analysis data, including forest plots and full results are provided in the 'Data and analyses' section. Overall, there is moderate-certainty evidence of effectiveness of exercise treatment for improving pain and functional limitations outcomes (Summary of findings 1; Summary of findings 2).

Pain intensity outcomes

Ninety-nine studies (132 study groups) provided data on 9041 participants for earliest pain outcomes comparing exercise treatment to other comparisons (40% of all included studies, 37% of all included participants) (Analysis 1.1).

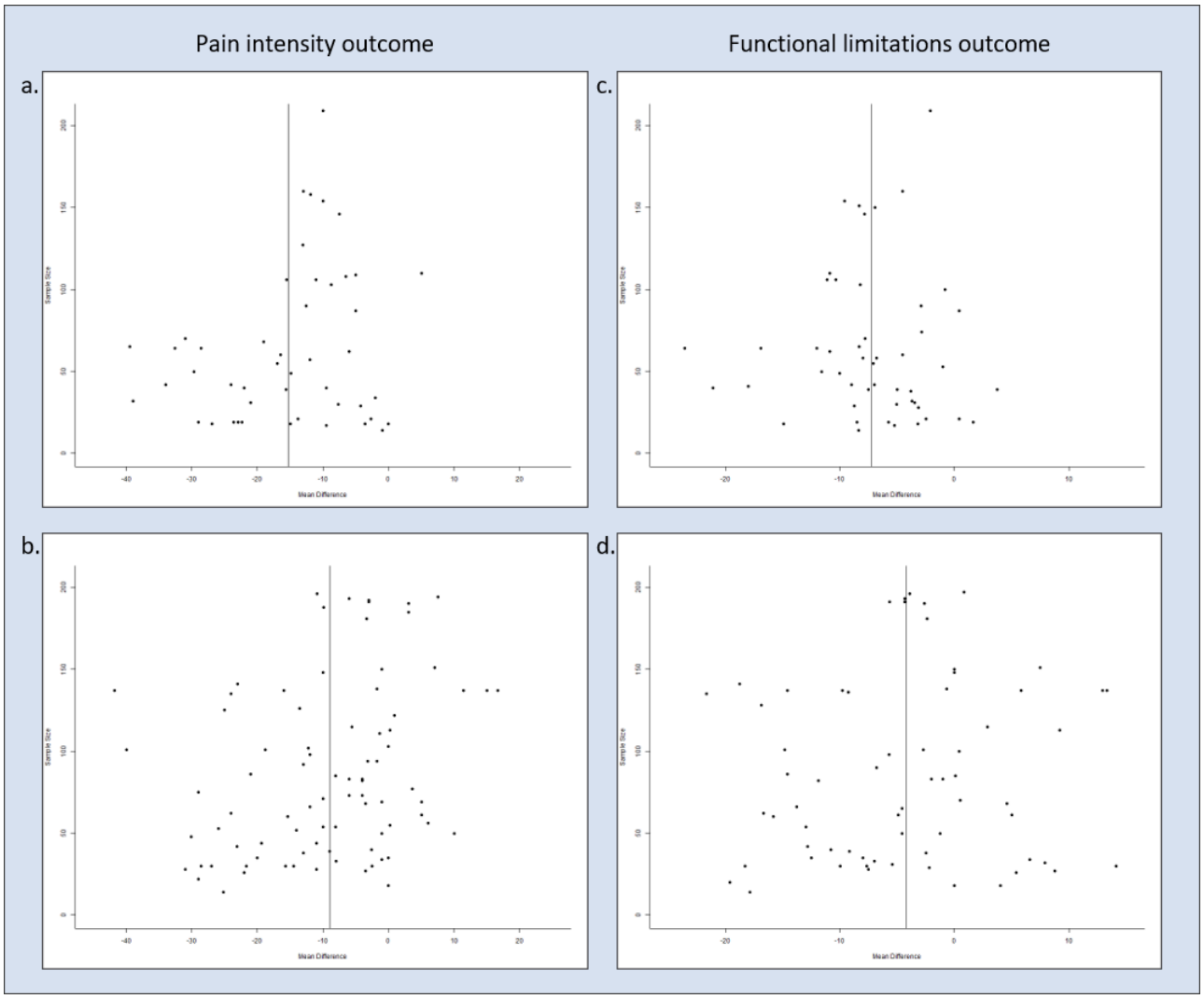
Exercise treatment versus no treatment, usual care or placebo

Thirty-five studies (47 study groups) provided data on 2746 participants comparing exercise treatment to no treatment comparisons (including usual care and placebo) on pain outcomes. The pooled mean difference in pain for exercise treatment compared to no treatment comparisons at the earliest follow-up point was -15.2 (95% CI -18.3 to -12.2; I^2 = 75%). This effect is compatible with a clinically important improvement in pain outcomes for exercise compared to no treatment. The prediction interval describing the expected range of true effect in similar studies is -33.0 to 2.5.

This evidence was judged to be moderate certainty for the comparison of exercise treatment to no treatment. The certainty of the evidence was downgraded due to inconsistency (substantial heterogeneity). There were no serious concerns due to indirectness, imprecision, or publication bias. Seven studies (10 groups; 526 participants) were judged to have high risk of bias (19% of participant data), however, exclusion of these studies in sensitivity analysis did not change conclusions so the certainty of evidence was not downgraded further (Summary of findings 1; Figure 3).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 3. Funnel plots for primary analyses (x axis = mean difference). Earliest follow-up study results for: 1. Pain outcome: a. exercise compared to placebo, no treatment or usual care, b. exercise compared to other conservative treatments; and 2. Functional limitations: c. exercise compared to no treatment, usual care or placebo, d. exercise compared to other conservative treatments.



Separate post hoc analysis of placebo treatments compared to exercise treatments in studies providing pain outcome data (2 studies; 300 participants) found a pooled mean difference of -8.9 (95% CI -15.1 to -2.6; I^2 = 0%) (Analysis 9.1).

Exercise treatment versus other conservative treatments

Sixty-four studies (85 study groups) provided data on 6295 participants comparing exercise treatment to other conservative treatment comparisons. The pooled mean difference in pain for exercise treatment compared to all other conservative treatments investigated at the earliest follow-up point was -9.1 (95% CI -12.6 to -5.6; I^2 = 93%). This improvement did not meet our prespecified threshold for clinically important difference in pain outcomes, but is probably meaningful. The prediction interval describing the expected range of true effect in similar studies is -40.1 to 21.9.

This evidence was judged to be low certainty for the comparison of exercise treatment to other conservative treatments. The certainty of the evidence was downgraded due to inconsistency (considerable heterogeneity), and some evidence of publication bias, however, heterogeneity was partially explained by exploration of effect in specific comparison subgroups. There were no serious concerns due to indirectness or imprecision. Eleven studies (14 groups; 835 participants) were judged to have high risk of bias (13% of participant data), however, exclusion of these studies in sensitivity analysis did not change conclusions so the certainty of evidence was not downgraded further (Summary of findings 2; Figure 3).

Short, medium and long-term follow-up periods

Data about pain outcomes for the short-term follow-up period (approximately three months) were available for exercise treatment compared to no treatment comparisons (MD -16.4, 95% CI -20.3 to -12.4; participants = 2247; studies = 26; study groups = 37; I^2 = 78%) (Analysis 1.2). This mean effect of exercise treatment is most compatible with a clinically important difference. The

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

observed mean effect differences did not meet this threshold of clinically important difference at medium-term (approximately six months) (MD -10.6, 95% CI -15.0 to -6.1; participants = 1491; studies = 17; study groups = 20; I^2 = 79%) (Analysis 1.3) or long-term follow-up (approximately 12 months) (MD -11.8, 95% CI -22.0 to -1.6; participants = 498; studies = 5; study groups = 5; I^2 = 88%) (Analysis 1.4).

For exercise treatment compared to all other conservative treatments investigated at specific follow-up periods, data at short-term (MD -8.6, 95% CI -13.1 to -4.1; participants = 4874; studies = 47; study groups = 60; I^2 = 95%) (Analysis 1.2) and medium-term follow-up periods (MD -8.1, 95% CI -12.3 to -3.9; participants = 4078; studies = 38; study groups = 52; I^2 = 90%) (Analysis 1.3) were similar and probably meaningful. For the long-term follow-up period, smaller non-significant differences were observed (MD -5.2, 95% CI -11.3 to 0.8; participants = 2490; studies = 19; study groups = 28; I^2 = 92%) (Analysis 1.4).

Subgroup analyses

Subgroup analyses explored the impact of population source, symptom duration, study participants with leg pain or neurologic symptoms, and specific type of conservative treatment comparison.

Population source

Exploring population source, the pooled mean difference in pain at the earliest follow-up period for exercise treatments compared to no treatment groups were: healthcare study populations (MD -16.9, 95% CI -20.9 to -13.0; participants = 1454; studies = 18; study groups = 23; I^2 = 68%), occupational populations (MD -22.9, 95% CI -39.7 to -6.1; participants = 252; studies = 3; study groups = 3; I^2 = 90%), and general or mixed populations (MD -12.2, 95% CI -17.1 to -7.3; participants = 1010; studies = 13; study groups = 20; I^2 = 76%) (Analysis 2.1).

For exercise treatment compared to other conservative treatments, the pooled mean difference in pain at the earliest follow-up period were: healthcare study populations (MD -8.0, 95% CI -12.2 to -3.9; participants = 2602; studies = 33; study groups = 46; I^2 = 88%), occupational study populations (MD -14.6, 95% CI -28.4 to -0.8; participants = 935; studies = 9; study groups = 11; I^2 = 97%), and general or mixed populations (MD -8.4, 95% CI -12.8 to -3.9; participants = 2698; studies = 21; study groups = 27; I^2 = 87%) (Analysis 2.2).

Symptom duration

Exploring pain outcome results at the earliest follow-up period according to the symptom duration of the study population, the pooled mean difference for exercise treatments compared to no treatment groups were: moderate symptom duration studies of 12 weeks to three years (MD -16.4, 95% CI -22.6 to -10.3; participants = 523; studies = 8; study groups = 12; I^2 = 77%), long symptom duration > three years (MD -8.5, 95% CI -11.1 to -6.0; participants = 1318; studies = 13; study groups = 17; I^2 = 0%), and for studies that did not specify symptom duration (MD -21.6, 95% CI -26.6 to -16.5; participants = 905; studies = 14; study groups = 18; I^2 = 71%) (Analysis 3.1).

The pooled mean difference for exercise treatments compared to other conservative treatment groups were: moderate symptom duration studies of 12 weeks to three years (MD -10.6, 95% CI -18.1 to -3.0; participants = 1555; studies = 17; study groups = 25; I^2 = 95%), long symptom duration > three years (MD -6.8, 95% CI -11.7 to -1.8; participants = 2094; studies = 17; study groups = 22; I^2 = 83%) and for studies that did not specify symptom duration (MD -9.5, 95% CI -14.2 to -4.7; participants = 2646; studies = 30; study groups = 38; I^2 = 91%) (Analysis 3.2).

Study participants with leg pain or neurologic symptoms

Exploring pain outcome results at earliest follow-up period according to the inclusion of study participants with leg pain or neurologic symptoms, the pooled mean difference for exercise treatments compared to no treatment groups were: studies that excluded participants with leg pain or neurologic symptoms (MD -17.0, 95% CI -22.6 to -11.3; participants = 816; studies = 14; study groups = 20; I^2 = 83%), studies that included participants with leg pain or neurologic symptoms (MD -15.9, 95% CI -20.9 to -10.9; participants = 1382; studies = 13; study groups = 15; I^2 = 74%), and studies that did not specify (MD -11.2, 95% CI -14.7 to -7.8; participants = 548; studies = 8; study groups = 12; I^2 = 15%) (Analysis 4.1).

The pooled mean difference in pain outcomes at earliest follow-up for exercise treatment interventions compared to other conservative treatment groups were: studies that excluded participants with leg pain or neurologic symptoms (MD -9.5, 95% CI -14.6 to -4.4; participants = 1657; studies = 19; study groups = 23; I^2 = 88%), studies that included participants with leg pain or neurologic symptoms (MD -4.9, 95% CI -9.1 to -0.7; participants = 2766; studies = 27; study groups = 35; I^2 = 85%), and studies that did not specify (MD -13.7, 95% CI -20.0 to -7.4; participants = 1872; studies = 18; study groups = 27; I^2 = 93%) (Analysis 4.2).

Type of conservative treatment comparison

The pooled mean difference in pain at earliest follow-up period for exercise treatments compared to other conservative treatments were: education comparisons (MD -12.2, 95% CI -19.4 to -5.0; participants = 2039; studies = 20; study groups = 26, I^2 = 94%), manual therapy comparisons (MD 1.0, 95% CI -3.1 to 5.1; participants = 1632; studies = 12; study groups = 18; I^2 = 74%), electrotherapy comparisons (MD -10.3, 95% CI -23.9 to 3.2; participants = 200; studies = 3; study groups = 6; I^2 = 91%), non-exercise physical therapy comparisons (MD -10.4, 95% CI -15.2 to -5.6; participants = 1869; studies = 21; study groups = 26; I^2 = 89%) and back school (MD -14.0, 95% CI -22.7 to -5.6; participants = 241; studies = 3; study groups = 3; I^2 = 66%). Fewer than three studies were available for each of the other comparison treatments: psychological therapy, relaxation, anti-inflammatories (Analysis 5.1).

Functional limitations outcomes

Eighty-eight studies provided data on 8946 participants for earliest functional limitations outcomes comparing exercise treatment to other comparisons (35% of studies, 37% of participants) (Analysis 1.5).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Exercise treatment versus no treatment, usual care or placebo

Thirty-eight studies (50 study groups) provided data on 2942 participants comparing exercise treatment to no treatment comparisons (including usual care and placebo) on functional limitations outcomes. The pooled mean difference in functional limitations with exercise treatment compared to no treatment comparisons at the earliest follow-up point was -6.8 (95% CI -8.3 to -5.3; I^2 = 38%). This effect did not meet our prespecified threshold for minimal clinically important difference in functional limitations outcomes. The prediction interval describing the expected range of true effect in similar studies is -13.2 to -0.4.

This evidence was judged to be moderate certainty for the comparison of exercise treatment to no treatment comparisons. The certainty of the evidence was downgraded due to potential publication bias. There were no serious concerns due to inconsistency, indirectness, or imprecision. Nine studies (13 groups; 495 participants) were judged to have high risk of bias (17% of participant data), however, exclusion of these studies in sensitivity analysis did not change conclusions so the certainty of evidence was not downgraded further (Summary of findings 1; Figure 3).

Separate post hoc analysis of placebo treatments compared to exercise treatments in studies providing functional limitations outcome data (3 studies; 400 participants) found a pooled mean difference of -5.6 (95% CI -11.2 to -0.01; I^2 = 33%) (Analysis 9.2).

Exercise treatment versus other conservative treatments

Fifty-two studies (69 study groups) provided data on 6004 participants comparing exercise treatment to other conservative treatment comparisons. The pooled mean difference in functional limitations with exercise treatment compared to other conservative treatments at the earliest follow-up point was -4.1 (95% CI -6.0 to -2.2; I^2 = 82%). This improvement did not meet our prespecified threshold for clinically important difference in functional limitations outcomes, but is probably meaningful. The prediction interval describing the expected range of true effect in similar studies is -17.6 to 9.3.

This evidence was judged to be moderate certainty for the comparison of exercise treatment to other conservative treatment. The certainty of the evidence was downgraded due to inconsistency (substantial heterogeneity), however, this was partially explained by exploration of the effect in specific comparison subgroups. There were no serious concerns due to indirectness, imprecision, or publication bias. Seven studies (9 groups; 571 participants) were judged to have high risk of bias (10% of participant data), however, exclusion of these studies in sensitivity analysis did not change conclusions so the certainty of evidence was not downgraded further (Summary of findings 2; Figure 3).

Short, medium and long-term follow-up periods

Data about functional limitations outcomes for the short-term follow-up period (approximately three months) were available for exercise treatment compared to no treatment comparisons (MD -7.4, 95% CI -9.2 to -5.6; participants = 2555; studies = 30; study groups = 41; I^2 = 43%) (Analysis 1.6). This mean effect did not meet our prespecified threshold for clinically important difference. Similar results were observed for medium-term (approximately six months) (MD -6.6, 95% CI -8.3 to -4.8; participants = 1831; studies = 22; study groups = 25; I^2 = 21%) (Analysis 1.7) and for long-term follow-up periods (approximately 12 months) (MD -4.4, 95% CI -7.3 to -1.5; participants = 550; studies = 6; study groups = 6; I^2 = 0%) (Analysis 1.8).

For exercise treatment compared to all other conservative treatments investigated at specific follow-up periods, data at short-term follow-up is probably meaningful (MD -4.0, 95% CI -6.0 to -1.9; participants = 5541; studies = 44; study groups = 58; I^2 = 82%) (Analysis 1.6). For medium-term follow-up (MD -2.2, 95% CI -4.5 to 0.1; participants = 3531; studies = 29; study groups = 40; I^2 = 77%) (Analysis 1.7) and long-term follow-up periods (MD 0.3, 95% CI -2.2 to 2.7; participants = 2816; studies = 15; study groups = 21; I^2 = 71%), smaller non-significant differences were observed (Analysis 1.8).

Subgroup analyses

Subgroup analyses explored the impact of population source, symptom duration, study participants with leg pain or neurologic symptoms, and specific type of conservative treatment comparison.

Population source

Exploring population source, the pooled mean difference in functional limitations at the earliest follow-up period for exercise treatments compared to no treatment groups were: healthcare study populations (MD -7.9, 95% CI -10.0 to -5.7; participants = 1900; studies = 22; study groups = 28; I^2 = 50%), occupational populations (MD -3.4, 95% CI -7.1 to 0.4; participants = 237; studies = 4; study groups = 5; I^2 = 11%), and general or mixed populations (MD -5.7, 95% CI -8.4 to -2.9; participants = 651; studies = 9; study groups = 13; I^2 = 14%) (Analysis 2.3).

For exercise treatment compared to other conservative treatments, the pooled mean difference in functional limitations at the earliest follow-up period were: healthcare study populations (MD -3.3, 95% CI -6.2 to -0.4; participants = 2488; studies = 28; study groups = 35; I^2 = 86%), occupational study populations (MD -2.1, 95% CI -7.0 to 2.9; participants = 583; studies = 4; study groups = 6; I^2 = 81%), and general or mixed populations (MD -5.4, 95% CI -8.1 to -2.7; participants = 2873; studies = 19; study groups = 27; I^2 = 72%) (Analysis 2.4).

Symptom duration

Exploring functional limitations outcome results at the earliest follow-up period according to the symptom duration of the study population, the pooled mean difference for exercise treatments compared to no treatment groups were: moderate symptom duration studies of 12 weeks to three years (MD -4.7, 95% CI -7.1 to -2.4; participants = 543; studies = 8; study groups = 11; I^2 = 3%), long symptom duration > three years (MD -5.6, 95% CI -7.4 to -3.7; participants = 1295; studies = 13; study groups = 17; I^2 = 0%), and for studies that did not specify symptom duration (MD -8.5, 95% CI -11.4 to -5.6; participants = 1104; studies = 17; study groups = 22; I^2 = 62%) (Analysis 3.3).

The pooled mean difference for exercise treatments compared to other conservative treatment groups were: moderate symptom duration studies of 12 weeks to three years (MD -4.8, 95% CI -7.8 to -1.8; participants = 1736; studies = 15; study groups = 21; I^2 = 77%), long symptom duration > three years (MD -4.9, 95% CI -8.5 to -1.2;

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

participants = 2076; studies = 15; study groups = 21; I^2 = 82%) and for studies that did not specify symptom duration (MD -3.0, 95% CI -6.2 to 0.2; participants = 2192; studies = 22; study groups = 27; I^2 = 85%) (Analysis 3.4).

Study participants with leg pain or neurologic symptoms

Exploring functional limitations outcome results at the earliest follow-up period according to the inclusion of study participants with leg pain or neurologic symptoms, the pooled mean difference for exercise treatments compared to no treatment groups were: studies that excluded participants with leg pain or neurologic symptoms (MD -6.0, 95% CI -8.1 to -3.9; participants = 983; studies = 15; study groups = 21; I^2 = 24%), studies that included participants with leg pain or neurologic symptoms (MD -6.6, 95% CI -8.4 to -4.9; participants = 1676; studies = 17; study groups = 20; I^2 = 7%), and studies that did not specify (MD -7.3, 95% CI -13.5 to -1.2; participants = 283; studies = 6; study groups = 9; I^2 = 73%) (Analysis 4.3).

The pooled mean difference in functional limitations outcomes at the earliest follow-up for exercise treatment interventions compared to other conservative treatment groups were: studies that excluded participants with leg pain or neurologic symptoms (MD -5.9, 95% CI -8.9 to -2.9; participants = 2022; studies = 19; study groups = 25; I^2 = 83%), studies that included participants with leg pain or neurologic symptoms (MD -0.9, 95% CI -3.50 to 1.71; participants = 2789; studies = 20; study groups = 27; I^2 = 77%), and studies that did not specify (MD -6.7, 95% CI -10.5 to -2.9; participants = 1193; studies = 13; study groups = 17; I^2 = 73%) (Analysis 4.4).

Type of conservative treatment comparison

The pooled mean difference in functional limitations at the earliest follow-up period for exercise treatments compared to other conservative treatments were: education comparisons (MD -8.0, 95% CI -10.8 to -5.3; participants = 2598; studies = 19; study groups = 26; I^2 = 78%), manual therapy comparisons (MD 1.5, 95% CI -1.5 to 4.4; participants = 1637; studies = 12; study groups = 18; I^2 = 64%), non-exercise physical therapy comparisons (MD -4.3, 95% CI -9.0 to 0.4; participants = 1074; studies = 12; study groups = 13; I^2 = 89%), and electrotherapy comparisons (MD -3.8, 95% CI -10.1 to 2.4; participants = 248; studies = 4; study groups = 6; I^2 = 71%). Fewer than three studies were available for each of the other comparison treatments: psychological therapy, back school, relaxation, and anti-inflammatories (Analysis 5.2).

Sensitivity analyses

The results were robust to assumptions about data, functional limitations outcome measures, risk of bias, and excluding trial mean difference data that were judged to be improbable or outlying. To test assumptions about data, we omitted studies that required imputation of data to represent the follow-up outcomes standard deviations, or where assumptions about the normality of the data were questionable; this omitted 28 studies (28%; 43 groups; 2696 participants) from the primary pain outcome meta-analysis, and 21 studies (23%; 30 groups; 2127 participants) from the primary functional limitations outcome meta-analysis, however, interpretation of results did not change. This sensitivity analysis resulted in pooled estimates of mean differences for pain and functional limitations outcomes for exercise treatment compared to any other intervention at the earliest follow-up of MD -12.1 (95% CI -15.5 to -8.7; I^2 = 93%) (Analysis 6.1) and MD -5.5 (95% CI -7.1 to -3.9; I^2 = 77%) (Analysis 6.2), respectively.

We omitted 35 studies (39%; 45 groups; 2718 participants omitted) to synthesise outcome results for only studies that measured functional limitations using the Roland-Morris Disability Questionnaire. This sensitivity analysis resulted in a pooled estimate of mean difference for functional limitations outcomes for exercise treatment compared to any other intervention at earliest follow-up of MD -6.3 (95% CI -8.3 to -4.3; I^2 = 78%) with similar effect size and heterogeneity observed (Analysis 6.3).

We omitted studies that were judged to have high risk of bias; this omitted 18 studies (18%; 24 groups; 1361 participants) from the primary pain outcome meta-analysis, and 16 studies (18%; 22 groups; 1066 participants) from the primary functional limitations outcome meta-analysis. This sensitivity analysis resulted in pooled estimates of mean differences for pain and functional limitations outcomes for exercise treatment compared to any other intervention at earliest follow-up of MD -11.2 (95% CI -14.1 to -8.3; I^2 = 91%) (Analysis 7.1) and MD -5.6 (95% CI -7.1 to -4.1; I^2 = 77%) (Analysis 7.2), respectively. We conducted a post hoc sensitivity analysis to explore the impact of RoB 1 potentially underestimating bias due to blinding (performance bias and detection bias), and RoB 2 potentially overestimating bias due to outcome measure blinding. These results did not change our results or conclusions (results not presented).

To test robustness of our results, we conducted primary analyses excluding trial data that were judged to be improbable or outlying typical results in the field. These sensitivity analyses resulted in exclusion of data from 12 studies in primary pain outcome meta-analysis comparisons (12%; 17 groups; 1024 participants) (Analysis 8.1), and four studies were excluded in primary functional limitations outcome meta-analysis comparisons (4%; 8 groups; 516 participants) (Analysis 8.2). These results did not change any conclusions, however average effect sizes and heterogeneity observed were smaller for pain and functional limitations outcomes when outliers were omitted. Results were less heterogeneous, and still showed strong effectiveness of exercise treatment, although not meeting our criteria for a clinically important difference compared to no treatment for pain outcome (MD -11.5, 95% CI -13.8 to -9.2; I^2 = 47%).

Importance of population and exercise intervention characteristics

The effects of specific types of exercise treatments and delivery characteristics on pain and functional limitations outcomes were investigated in a companion publication. Detailed network meta-analysis methods and results are presented in Hayden 2021b.

DISCUSSION

This review is an updated assessment of the effectiveness of exercise for treatment of chronic non-specific low back pain. A tremendous number of RCTs have been published since the last publication of the Cochrane review (more than 200 additional studies included in this review). We identified potential research integrity concerns and excluded 30 trials that otherwise met selection criteria. We included 249 RCTs (24,486 participants), with 203 trials providing sufficient data for meta-

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

analyses (19,633 participants): 142 RCTs compared exercise to non-exercise comparisons (117 with data for meta-analyses), and 151 compared different types of exercise treatments. Despite considering study-level sources of clinical heterogeneity, the studies within meta-analyses were heterogeneous. However, we found consistent results across primary, subgroup and sensitivity analyses to draw conclusions about the effectiveness of exercise for treatment of chronic low back pain described below.

Summary of main results

This systematic review provides useful information for primary care clinicians to help guide their patient management and referral practices.

There is moderate-certainty evidence that exercise treatment is probably more effective than no treatment, usual care or placebo for pain intensity and functional limitations outcomes. For pain intensity outcomes, our findings are most compatible with a clinically important difference of 15 points compared to no treatment; for functional limitations outcomes, the mean effect did not meet our prespecified threshold for minimal clinically important difference. The observed effect sizes would likely be smaller for comparison to placebo only.

There is low to moderate-certainty evidence that there is a small difference in pain and functional limitations for exercise treatment compared with other conservative treatments. The effect size did not meet our prespecified threshold of clinically important group-level difference, but is larger than mean effects reported for other treatments for low back pain. For example, a recent systematic review of spinal manipulation for chronic low back pain found that a mean difference of spinal manipulation compared with non-recommended therapies was -7.5 (95% CI -11.5 to -3.5) for short-term pain relief (Rubinstein 2019); a systematic review of behavioural treatment reported an overall mean difference of -5.2 (95% CI -9.8 to -0.6) compared to usual care for short-term pain relief (Henschke 2010); and another review found improvement in pain outcomes with back school compared to no treatment to be a mean difference of -6.1 (95% CI -10.2 to -2.0) (Parreira 2017).

Data from 64 trials comparing exercise treatment to other conservative treatments found exercise to be more effective, although the effect size was small and not clinically important overall (9 points improvement in pain; 4 points improvement in functional limitations); comparisons with some specific other conservative treatments (e.g. electrotherapy, education alone) showed exercise treatments to have larger improvements compatible with a clinically important difference.

Adverse effects of exercise treatment were rarely reported. When they were reported, those likely to be related to the intervention mainly included increased low back pain and muscle soreness.

Overall completeness and applicability of evidence

This review included numerous studies across geographic locations and settings. Most studies were conducted in healthcare settings where exercise treatments are often delivered, making this evidence applicable to clinical practice. However, fewer studies were available in community settings. Only nine studies conducted in occupational settings contributed data to primary meta-analyses, limiting the strength of this evidence. Subgroup analyses suggested that the results are applicable to healthcare, occupational and general/mixed populations.

A broad range of exercise treatments were investigated in the included studies. We classified treatment groups according to exercise types; the most common types included were core strengthening and mixed exercise types. We also classified exercises by design and delivery characteristics. The most common design was standardised (e.g. an aerobics class delivered to all participants), delivered in either a group or individually-supervised format, with a median of 12 hours of exercise time delivered over a median of eight weeks. These most common exercise characteristics are in line with current clinical practice in most regions, suggesting applicability of our findings. We may have misclassified studies based on the (often) brief descriptions available in study reports. By design, we aimed to include studies that met a broad definition of exercise treatment. By using a broad definition, the evidence remains relevant to many stakeholders, and allows assessment of the impact of delivery and design characteristics, which our team found to be important in previous analyses. This broad inclusion likely contributed to the heterogeneity across studies.

Comparison interventions for exercise treatment in the included studies were no treatment, usual care, placebo and various other conservative treatments. There are few studies comparing exercise treatment to several conservative treatments that are described as effective for treating chronic low back pain in the NICE (National Guideline Centre 2016) or ACP (Qaseem 2017) practice guidelines, including NSAIDS (NICE and ACP recommended), psychological therapies (NICE and ACP recommended), acupuncture (ACP recommended), and tramadol or duloxetine (described as second-line treatment options in the ACP guidelines). If any, future trials on exercise treatment should evaluate effectiveness and cost-effectiveness compared with other effective therapies, as this will guide clinicians in their choice for the best treatment.

In a related investigation, we found a dearth of complete reporting of study methods, participant characteristics and results. Over 30% of studies did not report data comparable to a CONSORT flow chart, 25% of studies did not fully report their method of randomisation, and 33% of studies did not adequately report allocation concealment. Half of the included studies failed to report a set of minimal baseline characteristics (age, sex, baseline pain intensity, duration of low back pain). These reporting gaps raise questions about methods and potential biases. Additionally, a substantial proportion of the eligible studies, almost 20%, that measured our primary outcomes did not report data completely or in a form usable in our meta-analyses. If we had not imputed missing values, 38% of studies would have been omitted from analysis as only 62% reported complete usable data. There is a dire need for better reporting of trials. We suggest trialists use the CONSORT tool when reporting trials. However, there is also a need for journal editors to more strictly check whether trial manuscripts meet the recommendations of these tools.

Many of the included studies, particularly more recently published studies, had very small sample sizes (median = 70, IQR: 42 to 125 participants per study; median = 30 per group). Only 26% of included trials had sufficient sample sizes for greater than or equal to 80% power. Fewer than 20% had sufficient sample size to find a statistically significant difference in the treatment effect for a planned realistic observed effect. Exercise is by far the most studied

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

intervention for low back pain. Additional small trials will not add anything to the current evidence.

Despite these gaps in reporting, we think that it is unlikely that inclusion of data from excluded studies would change the conclusions reached in our review.

Quality of the evidence

Overall, there was moderate-quality evidence for the effectiveness of exercise treatment compared to no treatment, usual care or placebo, and low to moderate compared to other conservative treatments for pain and functional limitations outcomes. Evidence was downgraded mainly due to heterogeneity across primary analyses, which could not be explained by study and population-level characteristics assessed in subgroup and sensitivity analyses. Considerable heterogeneity in comparisons of exercise with other conservative treatments could be partially explained by specific comparison treatment subgroups.

We recorded and explored the impact of several sources of heterogeneity, including setting, population, exercise and comparison interventions, outcomes and study design characteristics. Statistical heterogeneity as assessed with the I-squared measure, was substantial in most primary analyses (I^2 = 60-90%) and remained moderate to substantial in most subgroup and sensitivity analyses.

There are several additional important sources of heterogeneity when considering the effectiveness of exercise treatment, including heterogeneity in the exercise treatments, comparison interventions, and individual patient level differences. Our primary analyses considered all types of exercise treatment together. We recognise that these represent a very heterogeneous set of interventions. However, there is limited evidence that one specific type of exercises is more effective than others (Foster 2018). Clinical guidelines recommend exercise programmes that take individual needs, preferences, and capabilities into account. We think that there is value to look broadly at exercise treatment as an intervention since there are likely to be common mechanisms of effect. Further research is needed to better define and assess the most effective components. Exploration of individual level patient characteristics and their relationship with the effectiveness of exercise was beyond the scope of this review, however, is important to consider in future research (Malmivaara 2019).

The included studies broadly represented the heterogeneous chronic low back pain population, so we did not rate down the quality of evidence available for indirectness. Precision was high for the estimates of treatment effect for both pain and functional limitations outcomes due to the large number of trials, and thus participants, available for all primary analysis comparisons. There was evidence of potential publication bias in the analysis of exercise versus other conservative treatment for pain data, and for functional limitations outcomes for exercise versus no treatment, usual care or placebo; we downgraded the certainty of the evidence in these analyses due to this potential publication bias. No evidence of publication bias (Egger's test for funnel plot asymmetry) was found for other analyses.

Potential biases in the review process

Our systematic review, and the evidence presented within, has numerous strengths afforded by our structured processes, comprehensive search, assessment, and the large number of studies available. Limitations in our review are mainly related to potential selection and publication biases. A further source of both the limitations and strengths of this review is the sheer volume of included studies and, as a result, the duration of time to complete the review.

In the interest of completing this review, we were not able to extract data and incorporate results from 152 trials from our most recent electronic search update. We also identified 20 non-English language publications, however, we were not able to translate these in time for submission. These 172 studies are 'awaiting assessment'. Considering the limited change in the conclusions of this systematic review since the last publication in 2005, and the limited impact of multiple subgroup and sensitivity analyses on interpretation of results, we do not expect that the addition of more study data will result in a change of the overall conclusions.

At least two reviewers independently assessed trials for inclusion, extracted and checked data, and assessed risk of bias, with a third reviewer adjudicating any discrepancies. Because of the long duration to complete this review, the personnel involved in the review changed. At each search update (generally yearly) all extracted data were cleaned and re-checked to ensure consistency. This impacted the efficiency of review completion; however, we do not believe that this negatively impacted the reliability of the data. Indeed, it resulted in a very transparent and well documented process of electronic data capture and data validation as all data extraction was moved to Distiller SR.

Additional limitations of our systematic review methods related to length of time to complete this review were due to changes in systematic review methods recommendations since approval of our protocol in 2012. Recommendations for risk of bias assessment changed over the course of our review. We used the Cochrane Back and Neck group operationalisation of the Cochrane RoB 1 tool with high risk of bias defined by a threshold cut-point of the 12 tool items assessed based on the primary pain and functional limitations outcomes (both self-report). This threshold may have misclassified studies as high or low risk of bias. We do not think that this introduced systematic bias; sensitivity analyses modifying the assessment of blinding did not change our conclusions.

Despite feasibility limitations, there are benefits to our large systematic review and overarching review question. We were able to observe patterns in the study reports and data that may be missed with smaller focussed systematic reviews. A comprehensive investigation of research integrity characteristics was prompted by identification of exact duplication of data in two eligible studies (identified during a data checking stage prior to primary analyses). We defined integrity characteristics, planned and implemented additional data extraction without consideration of study results, and conducted analyses with studies blinded (identified by a study ID). This assessment was not described in our original protocol.

We have adopted a conservative estimate of clinical importance based on the upper interquartile range estimate of smallest worthwhile effect from a study of physiotherapy treatment for chronic low back pain (Ferreira 2013), representing a difference that 75% of participants reported as being worthwhile compared to no treatment. Two recent systematic reviews of drug interventions for low back pain have used a lower smallest worthwhile effect of 10 points for pain and for function. With a higher

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

threshold for important difference in pain outcome, we may have underestimated the clinical importance of the difference for exercise treatments. It is also important to consider that the smallest worthwhile effect estimates are specific for treatment and comparison interventions. The available research (Christiansen 2018; Ferreira 2013) estimated the smallest worthwhile effect values for physiotherapy treatment compared to no treatment. One might expect different estimates compared to other conservative treatment comparisons, as we have included in this review.

Although we undertook a comprehensive electronic search strategy with no restrictions and searched the reference lists of included studies and other published reviews, we may have missed relevant studies in the search or screening process (our funnel plot analysis of exercise versus other conservative treatments suggests potential publication bias). However, we consider it highly unlikely that we have missed any important, large studies.

We contacted authors for missing information but had responses from only 31% of those contacted. We did improve response rates (from 25% to 47%) by moving information requests from email to REDCap electronic survey. The overall low response may have introduced bias in our results if author response was selective.

Our definition of research integrity concerns resulting in exclusion of studies from our review was based on a combination of multiple conduct or reporting issues, including serious plagiarism and/or predatory journal publication. We also excluded two trials where there was unexplained exact duplication of study data (unresolved after report to the author institutions and publishing journal editors). Similar criteria have not previously been used to exclude studies from Cochrane reviews. However, we feel strongly that such exclusions are necessary and justified to increase confidence in the evidence base. We may have misclassified and inappropriately excluded useful study data; however, these exclusions did not change our results or conclusions.

Future updates of this review will assess evidence for health-related quality of life and patient-perceived recovery outcomes, as available. We also plan to use RoB 2 consistently across all studies; although there are challenges with the use of the tool, and lower agreement scores (Minozzi 2020), this approach has face validity for study/outcome level assessment of potential biases.

Agreements and disagreements with other studies or reviews

Effectiveness of exercise treatment for low back pain is a very active research field. Hundreds of systematic reviews on focussed populations or exercise types have been published in the past 10 years. While there has been an abundance of systematic reviews published on the effectiveness of exercise for low back pain, it is interesting to note that Almeida 2020 found that 74% of the 28 included reviews of “exercise for low back pain” interventions had an overall confidence rating of critically low. All four reviews rated as high (3) or moderate (1) were prospectively registered Cochrane reviews, including the first publication of this review in 2005. This review presents the most comprehensive identification and evaluation of the evidence about the effectiveness of exercise treatment, broadly defined, for treatment of chronic low back pain. In this systematic review, we have drawn similar conclusions to the previous publication of this Cochrane review (Hayden 2005a), with more confidence, supported by a much larger number of available trials and participants, more robust methods, and several additional analyses. Our observed effect sizes are larger than previously observed. Sensitivity analyses omitting study results that were outlying and considered improbable show smaller effects, however, similar interpretation of effectiveness of exercise treatment. We used this strategy in our primary analysis to present a more conservative and in our opinion more reliable assessment of the evidence.

Large reviews of all treatments for low back pain have made similar conclusions about the effectiveness of exercise treatment. A recent update of an Agency for Healthcare Research and Quality (AHRQ) systematic review of non-pharmacological interventions for chronic musculoskeletal conditions reported moderate-quality evidence of effectiveness of exercise treatment for chronic low back pain at short-term follow-up for function, and low-quality evidence for pain (all time points) and at medium and long-term time points for function (Skelly 2020).

Overlapping focussed reviews of specific types of exercise therapies include published Cochrane reviews (Saragiotto 2016; Schaafsma 2013; Wieland 2017; Yamato 2015), Cochrane protocols (Liu 2013; Machado 2012a; Machado 2012b) and many recent systematic reviews that have been published outside of the Cochrane Library (Alhakami 2019; Namnaqani 2019; Niederer 2020; Shi 2018; Vanti 2019; Wewege 2018; Zhang 2019a; Zou 2019b). These reviews have included five to 29 studies on treatments for chronic low back pain - a small proportion of the studies included here with more focussed review questions and selection criteria, and some differences in methods. Similar to our findings, the Cochrane reviews reported low to moderate-quality evidence that the specific exercise type investigated produced small to moderate improvements in outcome compared to minimal intervention; they similarly reported limited evidence of important differences compared to other types of conservative treatments. Non-Cochrane reviews, often on focussed topics with a small number of included studies, usually reported significant favourable results, often not meeting minimal clinically important outcomes.

We believe that the most effective components and approach of exercise treatments for low back pain patients have yet to be confirmed. Effectiveness studies have only recently raised the importance of considering likely mechanisms of effect and few available trials appear to have planned their interventions considering potentially effective mechanisms (Wood 2019; Wood 2021). Exercise treatments investigated have been heterogeneous in treatment design, dose, delivery format, specific exercise types, and are often combined with other conservative treatments (Ferreira 2010). Even with the large number of trials included in this review, remaining heterogeneity, incomplete descriptions and our potential misclassification makes this difficult to tease apart. Our prior work found that exercise design and delivery characteristics were statistically associated with improved effectiveness for pain and functional limitations outcomes, more so than specific exercise types (Hayden 2005a; Hayden 2005b; Hayden 2005c). Our related network meta-analysis study (Hayden 2021b) found Pilates, McKenzie therapy and functional restoration exercises were more effective than other exercise types for reducing pain and functional limitations. The results in this network meta-analysis were consistent for the benefits of other conservative treatments added to an exercise treatment, and greater benefits of higher dose.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

AUTHORS' CONCLUSIONS

Implications for practice

Clinicians often turn to systematic reviews and the resulting guideline recommendations to inform their treatment of patients. This review fills an identified gap, the need for updated, high-quality review of general exercise (Almeida 2020). We found moderate evidence that exercise probably provides a small benefit for pain outcomes in treatment of chronic low pain compared to no treatment comparisons (including usual care and placebo). The observed treatment effect for functional limitations outcomes was small and was not considered clinically important. Exercise treatment was found to have improved pain and functional limitations outcomes compared to other conservative treatments, however, these effects are small and not considered clinically important. We are not able to make recommendations about specific exercise types based on the results of this review, nor on the work of others to date; however, a related publication by this team will address this topic.

Due to insufficient reporting of adverse events in included trials, we are not able to confirm the safety or harms related to exercise treatment for chronic low back pain; however, the trials that did measure adverse events report few and mostly minor adverse effects, such as muscle pain. Based on available evidence, exercise is likely a good option to manage chronic low back pain. However, when determining if exercise is right for their patient, clinicians can take into consideration a wide range of factors including patient preference, suitability, access, and costs.

Implications for research

We have identified 451 RCTs on exercise treatment for chronic low back pain (279 assessed, 172 awaiting assessment). Low back pain researchers need to avoid further research waste by ensuring future trials contribute to the evidence base by thoughtful planning, robust conduct, large sample size, complete protocol and manuscript reporting, sharing trial data for validation and future meta-syntheses, and publication to properly archive results.

One of the primary benefits of systematic review and meta-analysis is combining studies of small sample sizes. However, the field is now saturated with small exercise trials many of which suffer from poor planning, conduct, and reporting due to limited resources. Efforts and research funding may be better spent on the co-ordination of multi-site studies, and the development of supports for good conduct and reporting.

It is our position that there needs to be more co-ordinated and thoughtful planning of trials in the field. This includes the identification of existing gaps and the planning of exercise treatments and assessment of outcomes based on the proposed mechanisms of effect (Bird 2020; Wood 2020a).

Future trials should assess the recommended core outcome set (Chiarotto 2018), including health-related quality of life and perceived effect; we found these important outcomes were only available for a small subset of studies. Trials should include the systematic measurement of harms, such as adverse events. There is also a need for careful selection of comparison groups to best contribute to evidence about the effective management of low back pain.

In addition, a more comprehensive measurement of baseline participant characteristics will allow for future investigation of treatment effect modifiers and mediating characteristics (even when single trials are insufficiently powered for such analyses). Currently, inconsistent availability and measurement of patient baseline characteristics limit the ability to assess the generalisability of the trials and the overall body of evidence. Furthermore, this limits the usefulness of the trial to contribute to future important IPD meta-analyses. IPD meta-analyses are the most feasible way to investigate potential treatment effect modifiers. Identifying potential treatment effect modifiers and the phenotypes of patients who are likely to benefit more from a specific treatment is an important goal and a long-term back pain research priority (Costa 2009b; Henschke 2007).

We support and encourage initiatives to make research data accessible for testing and further analyses. There are several benefits of this, including opportunities to double-check data (before, during or after publication), which provides a collegial way to ensure the best quality of evidence is disseminated. Data sharing has the additional benefit of enabling analyses of IPD data to identify treatment effect modifiers in a more robust way, without the limitations of ecological bias from study level data.

Finally, we believe there needs to be an increased emphasis on publishing trials in reputable journals with peer review and editorial processes to confirm the robustness of methods, assure complete reporting, and to require the proper archive of results to be included in future evidence reviews. If authors choose to publish their work in predatory journals, there does need to be a transparent system to determine if their trials should be eligible for inclusion in systematic reviews.

Future systematic review methods recommendations

One of the key benefits of a large overarching review, such as this one, is the ability to identify relationships that are not able to be investigated in smaller focussed reviews (e.g. characteristics of delivery, populations). Overarching reviews also have the potential to decrease the duplication of effort that often happens when multiple authors groups extract the same data for the same studies for inclusion in slightly different reviews. Overarching reviews enable the inclusion of trials that do not otherwise meet inclusion criteria for focussed reviews such as mixed exercise treatments (which represented 23% of all exercise groups in this review). However, feasibility of such large reviews is a barrier – they require tremendous time and resources for one team to complete and funding opportunities for systematic reviews are very limited. This presents an opportunity for more co-ordination of efforts within the limits of the current flawed academic credit and funding system.

This review has highlighted additional methods considerations for future systematic reviews, including attention to issues of research integrity in RCTs. We need further study to understand how to appropriately define and measure problematic characteristics, and to understand the relationships between study integrity characteristics and the reported outcomes. Issues of poor conduct, incomplete reporting, insecure archiving, and outright fabrication or fraud cannot be ignored and may be more problematic in some fields. These issues are important for systematic review authors and warrant investigation and consideration in the selection of studies and in the analyses conducted to summarise a body of evidence.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

We particularly draw attention to the issue of outlying and improbable results presented in RCTs. We defined a high threshold to consider results improbable or outlying and conducted sensitivity analyses to exclude these study data. In this review, the results and conclusions did not change in direction with the exclusion of improbable or outlying effects, adding robustness to the conclusions. However, the exclusion of these studies from analyses did impact the interpretation of the clinical importance of results; with a smaller number of studies, individual study results could have more weight in analyses and provide incorrect results and conclusions.

We posit that we should stop doing the same type of trials about exercise treatment. Additional trials are unlikely to add anything to our knowledge, unless they are based on a new and plausible hypothesis for the cause of low back pain and a potential working mechanism of the intervention. Future studies should evaluate whether exercise treatment is cost-effective compared to other interventions with similar effects. This will help patients, clinicians and policy-makers to decide which intervention to choose. Implementation studies that evaluate how to best introduce exercise treatment in healthcare settings, and how to increase compliance seem important as well. Additionally, using modern technologies to support exercise treatments, monitor compliance, give positive feedback, and monitor outcomes may be the best way forward.

Finally, progress in the field is unlikely to be feasible with small individual studies, so it will need to be facilitated by increased international collaboration, prospective planning of multi-site and multi-country trials, standardising measurement of baseline characteristics, and potentially by sharing of data through accessible repositories. Future prospective co-ordination and collaboration for more consistent data collection will help researchers identify treatment effect modifiers. This will further advance a personalised management approach for chronic low back pain.

ACKNOWLEDGEMENTS

We would like to thank Shireen Harbin, Managing Editor of the Cochrane Back and Neck Group, Jordi Pardo Pardo, Managing Editor of the Musculoskeletal Group, Rachel Couban, Trials Search Co-ordinator of the Cochrane Back and Neck Review Group, and Leah Boulos, Evidence Synthesis Coordinator for the Maritime SPOR SUPPORT Unit for their assistance and support conducting literature searches.

We would like to thank Cochrane editors and reviewers, including Bruno T Saragiotto, PhD, Universidade Cidade de São Paulo, Brazil and Timothy Carey MD MPH, Research Professor of Medicine, University of North Carolina at Chapel Hill, NC, USA for their comprehensive peer review. We would also like to thank Kathryn Armitstead for her careful review as consumer reviewer.

We would like to thank Dr. Samuel Stewart, Assistant Professor, Dalhousie University, for his consultation on our analyses. We would like to acknowledge Jenny Cartwright for her contributions to the development of the protocol, and early study screening and data extraction. We would like to thank Sally Miller, Andrea Smith, and Maria Wilson for their contributions to study screening, data extraction and data preparation, and Selena Glover for study screening, data extraction and assistance making judgements about exercise types. We would like to thank Somayyeh Mohammadi for her assessment of Persian abstracts.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

REFERENCES

References to studies included in this review

Adamczyk 2009 *{published data only}*

Adamczyk A, Kiebzak W, Wilk-Franczuk M, Sliwinski Z. Effectiveness of holistic physiotherapy for low back pain. *Ortopedia Traumatologia Rehabilitacja* 2009;**11**(6):562-76.

Akhtar 2017 *{published data only}*

Akhtar MW, Karimi H, Gilani SA. Effectiveness of core stabilization exercises and routine exercise therapy in management of pain in chronic nonspecific low back pain: a randomized controlled clinical trial. *Pakistan Journal of Medical Sciences* 2017;**33**(4):1002-6.

Akodu 2017 *{published data only}*

Akodu AK, Akinbo SRA, Omootunde AS. Comparative effects of muscle energy technique and core stability exercise in the management of patients with non-specific chronic low back pain. *Sports Medicine* 2017;**13**(1):2860-7.

Alaranta 1994 *{published data only}*

Alaranta H, Rytokoski U, Rissanen A, Talo S, Ronnemaa T, Puukka P, et al. Intensive physical and psychosocial training program for patients with chronic low back pain: a controlled clinical trial. *Spine* 1994;**19**(12):1339.

Alayat 2014 *{published data only}*

Alayat MSM, Atya AM, Ali MME, Shosha TM. Long-term effect of high-intensity laser therapy in the treatment of patients with chronic low back pain: a randomized blinded placebo-controlled trial. *Lasers in Medical Science* 2014;**29**:1065-73.

Albaladejo 2010 *{published data only}*

Albaladejo C, Kovacs FM, Royuela A, Pino R, Zamora J. The efficacy of a short education program and a short physiotherapy program for treating low back pain in primary care: a cluster randomized trial. *Spine* 2010;**35**(5):483-96.

Alexandre 2001 *{published data only}*

Alexandre NM, De Moraes MA, Correa Filho HR, Jorge SA. Evaluation of a program to reduce back pain in nursing personnel. *Revista de Saude Publica* 2001;**35**(4):356-61.

Alp 2014 *{published data only}*

Alp A, Mengi G, Avsaroglu AH, Mert M, Sigirli D. Efficacy of core-stabilization exercise and its comparison with home-based conventional exercise in low back pain patients. *Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi* 2014;**60**(Suppl 1):S36-S42.

Andrusaitis 2011 *{published data only}*

Andrusaitis SF, Brech GC, Vitale GF, Greve JM. Trunk stabilization among women with chronic lower back pain: a randomized, controlled, and blinded pilot study. *Clinics* 2011;**66**(9):1645-50.

Arab 2016 *{published data only}*

Arab AM, Saadati H, Sheikhhoseini R. The effect of harmonic technique vs end range loading exercises on pain and disability in patients with non-specific chronic low back pain: a preliminary study. *Journal of Chiropractic Medicine* 2016;**15**(1):3-8.

Arampatzis 2017 *{published data only}*

Arampatzis A, Schroll A, Catala MM, Laube G, Schuler S, Dreinhofer K. A random-perturbation therapy in chronic non-specific low-back pain patients: a randomised controlled trial. *European Journal of Applied Physiology* 2017;**117**(12):2547-60.

Areeudomwong 2017 *{published data only}*

Areeudomwong P, Wongrat W, Neammesri N, Thongsakul T. A randomized controlled trial on the long-term effects of proprioceptive neuromuscular facilitation training, on pain-related outcomes and back muscle activity, in patients with chronic low back pain. *Musculoskeletal Care* 2017;**15**(3):218-29.

Atalay 2017 *{published data only}*

Atalay E, Akova B, Gur H, Sekir U. Effect of upper-extremity strengthening exercises on the lumbar strength, disability and pain of patients with chronic low back pain: a randomized controlled study. *Journal of Sports Science and Medicine* 2017;**16**(4):595-603.

Aure 2003 *{published data only}*

Aure O, Nilsen J, Vasseljen O. Manual therapy and exercise therapy in patients with chronic low back pain. A randomized, controlled trial with 1-year follow-up. *Spine* 2003;**28**(6):525-32.

Azevedo 2017 *{published data only}*

Azevedo DC, Ferreira PH, Santos HO, Oliveira DR, De Souza JVL, Costa LOP. Movement system impairment-based classification treatment versus general exercises for chronic low back pain: randomized controlled trial. *Physical Therapy* 2017;**98**(1):28-39.

Bade 2017 *{published data only}*

Bade M, Cobo-Estevez M, Neeley D, Pandya J, Gunderson T, Cook C. Effects of manual therapy and exercise targeting the hips in patients with low-back pain: a randomized controlled trial. *Journal of Evaluation in Clinical Practice* 2017;**23**(4):734-40.

Barberini 2011 *{published data only}*

Barberini A, Bertozzi L, Nardi A, Ricci B. Hydrotherapy on workers in the poultry industry with low back pain. A randomized controlled trial with 3-months follow-up. *Scienza Riabilitativa* 2011;**13**(4):5-11.

Bellido-Fernandez 2018 *{published data only}*

Bellido-Fernandez L, Jimenez-Rejano JJ, Chillon-Martinez R, Gomez-Benitez MA, De-La-Casa-Almeida M, Rebollo-Salas M. Effectiveness of massage therapy and abdominal hypopressive gymnastics in nonspecific chronic low back pain: a randomized controlled pilot study. *Evidence-Based Complementary and Alternative Medicine* 2018;**0**:1-9.

Bendix 1995 *{published data only}*

Bendix AF, Bendix T, Labriola M, Boekgaard P. Functional restoration for chronic low back pain: two-year follow-up of two randomized clinical trials. *Spine* 1998;**23**(6):717-25.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bendix AF, Bendix T, Lund C, Kirkbak S, Ostenfeld S, Kirking BC. Comparison of three intensive programs for chronic low back pain patients: a prospective, randomized, observer-blinded study with one-year follow-up. *Scandinavian Journal of Rehabilitation Medicine* 1997;**29**(2):81-9.

* Bendix AF, Bendix T, Ostenfeld S, Bush E, Andersen A. Active treatment programs for patients with chronic low back pain: a prospective randomized, observer-blinded study. *European Spine Journal* 1995;**4**(3):148-52.

Bendix 2000 *{published data only}*

Bendix T, Bendix A, Labriola M, Haestrup C, Ebbehoj N. Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study. *Spine* 2000;**25**(19):2494-500.

Ben-Salah 2009 *{published data only}*

Ben-Salah FZ, Fendri Y, Jellad A, Boudoukhane S, Rejeb N. Efficacy and treatment compliance of a home-based rehabilitation programme for chronic low back pain: a randomized, controlled study. *Annals of Physical and Rehabilitation Medicine* 2009;**52**(6):485-96.

Bentsen 1997 *{published data only}*

Bentsen H, Lindgärde F, Manthorpe R. The effect of dynamic strength back exercise and/or a home training program in 57-year old women with chronic low back pain: results of a prospective randomized study with a 3-year follow-up period. *Spine* 1997;**22**(13):1494-500.

Bhadauria 2017 *{published data only}*

Bhadauria EA, Gurudut P. Comparative effectiveness of lumbar stabilization, dynamic strengthening, and Pilates on chronic low back pain: randomized clinical trial. *Journal of Exercise Rehabilitation* 2017;**13**(4):477-85.

Bi 2013 *{published data only}*

Bi X, Zhao J, Zhao L, Liu Z, Zhang J, Sun D, et al. Pelvic floor muscle exercise for chronic low back pain. *Journal of International Medical Research* 2013;**41**(1):146-52.

Bid 2017 *{published data only}*

Bid DD, Soni NC, Yadav AS, Rathod PV. A study on central sensitization in chronic non-specific low back pain. *Indian Journal of Physiotherapy & Occupational Therapy* 2017;**11**(4):165-75.

Blodt 2015 *{published data only}*

Blodt S, Pach D, Kaster T, Ludtke R, Icke K, Reisshauer A, et al. Qigong versus exercise therapy for chronic low back pain in adults: a randomized controlled non-inferiority trial. *European Journal of Pain* 2015;**19**(1):123-31.

Bramberg 2017 *{published data only}*

Bramberg EB, Bergstrom G, Jensen I, Hagberg J, Kwak L. Effects of yoga, strength training and advice on back pain: a randomized controlled trial. *BMC Musculoskeletal Disorders* 2017;**18**(1):132-43.

Bronfort 1996 *{published data only}*

Bronfort G, Goldsmith CH, Nelson C, Boline PD, Anderson AV. Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomized, observer-blinded clinical trial. *Journal of Manipulative & Physiological Therapeutics* 1996;**19**(9):570-82.

Bronfort 2011 *{published data only}*

Bronfort G, Maiers MJ, Evans RL, Schulz CA, Bracha Y, Svendsen KH, et al. Supervised exercise, spinal manipulation, and home exercise for chronic low back pain: a randomized clinical trial. *Spine* 2011;**11**(7):585-98.

Cabak 2017 *{published data only}*

Cabak A, Rudnicka A, Kulej L, Tomaszewski W. Biopsychosocial rehabilitation programme for patients with chronic back pain: pilot study. *Ortopedia Traumatologia Rehabilitacja* 2017;**19**(2):165-74.

Cai 2017 *{published data only}*

Cai C, Yang Y, Kong PW. Comparison of lower limb and back exercises for runners with chronic low back pain. *Medicine and Science in Sports & Exercise* 2017;**49**(12):2374-84.

Cairns 2006 *{published data only}*

Cairns MC, Foster NE, Wright C. Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. *Spine* 2006;**31**(19):E670-81.

Calmels 2004 *{published data only}*

Calmels P, Jacob JF, Fayolle-Minon I, Charles C, Bouchet JP, Rimaud D, et al. Use of isokinetic techniques vs standard physiotherapy in patients with chronic low back pain. Preliminary results. *Annales de Readaptation et de Medecine Physique* 2004;**47**(1):20-7.

Cambron 2006 *{published data only}*

* Cambron JA, Gudavalli MR, Hedeker D, McGregor M, Jedlicka J, Keenum M, et al. One-year follow-up of a randomized clinical trial comparing flexion distraction with an exercise program for chronic low-back pain. *Journal of Alternative and Complementary Medicine* 2006;**12**(7):659-68.

Gudavalli MR, Cambron JA, McGregor M, Jedlicka J, Keenum M, Ghanayem AJ, et al. A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for chronic low back pain. *European Spine Journal* 2006;**15**(7):1070-82.

Carr 2005 *{published data only}*

Carr JL, Klaber Moffett JA, Howarth E, Richmond SJ, Torgerson DJ, Jackson DA, et al. A randomized trial comparing a group exercise programme for back pain patients with individual physiotherapy in a severely deprived area. *Disability & Rehabilitation* 2005;**272**(16):929-37.

Cecchi 2010 *{published data only}*

Cecchi F, Molino-Lova R, Chiti M, Pasquini G, Paperini A, Conti AA, et al. Spinal manipulation compared with back school and with individually delivered physiotherapy for the treatment

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

of chronic low back pain: a randomized trial with one-year follow-up. *Clinical Rehabilitation* 2010;**24**(1):26-36.

Chan 2011 *{published data only}*

Chan CW, Mok NW, Yeung EW. Aerobic exercise training in addition to conventional physiotherapy for chronic low back pain: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* 2011;**92**(10):1681-5.

Chan 2017 *{published data only}*

Chan AYP, Ford JJ, Surkitt LD, Richards MC, Slater SL, Davidson M, et al. Individualised functional restoration plus guideline-based advice vs advice alone for non-reducible discogenic low back pain: a randomised controlled trial. *Physiotherapy* 2017;**103**(2):121-30.

Chatzitheodorou 2007 *{published data only}*

Chatzitheodorou D, Kabitsis C, Malliou P, Mougios V. A pilot study of the effects of high-intensity aerobic exercise versus passive interventions on pain, disability, psychological strain, and serum cortisol concentrations in people with chronic low back pain. *Physical Therapy* 2007;**87**(3):304-12.

Chen 2014 *{published data only}*

Chen HM, Wang HH, Chen CH, Hu HM. Effectiveness of a stretching exercise program on low back pain and exercise self-efficacy among nurses in Taiwan: a randomized clinical trial. *Pain Management Nursing* 2014;**15**(1):283-91.

Chown 2008 *{published data only}*

Chown M, Whittamore L, Rush M, Allan S, Stott D, Archer M. A prospective study of patients with chronic back pain randomised to group exercise, physiotherapy or osteopathy. *Physiotherapy* 2008;**94**(1):21-8.

Chung 2018 *{published data only}*

Chung SH, You YY, Lee HJ, Sim SH. Effects of stabilization exercise using flexi-bar on functional disability and transverse abdominis thickness in patients with chronic low back pain. *Journal of Physical Therapy Science* 2018;**30**(3):400-4.

Cortell-Tormo 2018 *{published data only}*

Cortell-Tormo JM, Sanchez PT, Chulvi-Medrano I, Tortosa-Martinez J, Manchado-Lopez C, Llana-Belloch S, et al. Effects of functional resistance training on fitness and quality of life in females with chronic nonspecific low-back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2018;**31**(1):95-105.

Costa 2009b *{published data only}*

Costa LO, Maher CG, Latimer J, Hodges PW, Herbert RD, Refshauge KM, et al. Motor control exercise for chronic low back pain: a randomized placebo-controlled trial. *Physical Therapy* 2009;**89**(12):1275-86.

Costantino 2014 *{published data only}*

Costantino C, Romiti D. Effectiveness of Back School program versus hydrotherapy in elderly patients with chronic non-specific low back pain: a randomized clinical trial. *Acta BioMedica* 2014;**85**(3):52-61.

Critchley 2007 *{published data only}*

Critchley DJ, Ratcliffe J, Noonan S, Jones RH, Hurley MV. Effectiveness and cost-effectiveness of three types of physiotherapy used to reduce chronic low back pain disability: a pragmatic randomized trial with economic evaluation. *Spine* 2007;**32**(14):1474-81.

Cruz-Diaz 2016 *{published data only}*

Cruz-Diaz D, Martinez-Amat A, Osuna-Pacrez MC, Torre-Cruz MJ, Hita-Contreras F. Short- and long-term effects of a six-week clinical Pilates program in addition to physical therapy on postmenopausal women with chronic low back pain: a randomized controlled trial. *Disability & Rehabilitation* 2016;**38**(13):1300-8.

Cruz-Diaz 2017 *{published data only}*

Cruz-Diaz D, Bergamin M, Gobbo S, Martinez-Amat A, Hita-Contreras F. Comparative effects of 12 weeks of equipment based and mat Pilates in patients with chronic low back pain on pain, function and transversus abdominis activation. A randomized controlled trial. *Complementary Therapies in Medicine* 2017;**33**:72-7.

Cruz-Diaz 2018 *{published data only}*

Cruz-Diaz D, Romeu M, Velasco-Gonzalez C, Martinez-Amat A, Hita-Contreras F. The effectiveness of 12 weeks of Pilates intervention on disability, pain and kinesiophobia in patients with chronic low back pain: a randomized controlled trial. *Clinical Rehabilitation* 2018;**32**(9):1249-57.

Cuesta-Vargas 2012 *{published data only}*

Cuesta-Vargas AI, Adams N, Salazar JA, Belles A, Hazanas S, Arroyo-Morales M. Deep water running and general practice in primary care for non-specific low back pain versus general practice alone: randomized controlled trial. *Clinical Rheumatology* 2012;**31**(7):1073-8.

Cuesta-Vargas 2015 *{published data only}*

Cuesta-Vargas AI, White M, Gonzalez-Sanchez M, Kuisma R. The optimal frequency of aquatic physiotherapy for individuals with chronic musculoskeletal pain: a randomised controlled trial. *Disability Rehabilitation* 2015;**37**(4):311-8.

Curnow 2009 *{published data only}*

Curnow D, Cobbin D, Wyndham J, Boris Choy ST. Altered motor control, posture and the Pilates method of exercise prescription. *Journal of Bodywork and Movement Therapies* 2009;**13**(1):104-11.

Dalichau 2000 *{published data only}*

Dalichau S, Scheele K. Effects of elastic lumbar belts on the effect of a muscle training program for patients with chronic back pain. *Zeitschrift fur Orthopadie und Ihre Grenzgebiete* 2000;**138**(1):8-16.

Da Luz 2014 *{published data only}*

Da Luz MA, Costa LO, Fuhro FF, Manzoni AC, Oliveira NT, Cabral CM. Effectiveness of mat Pilates or equipment-based Pilates exercises in patients with chronic nonspecific low back pain: a randomized controlled trial. *Physical Therapy* 2014;**94**(5):623-31.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Devasahayam 2014 *{published data only}*

Devasahayam AJ, Lim Siang CK, Goh MR, You Lim JP, Pua PY. Delivering a back school programme with a cognitive behavioural modification: a randomised pilot trial on patients with chronic non-specific low back pain and functional disability. *Proceedings of Singapore Healthcare* 2014;**23**(3):218-25.

Deyo 1990 *{published data only}*

Deyo RA, Walsh NE, Martin DC, Schoenfeld LS, Ramamurthy S. A controlled trial of transcutaneous electrical nerve stimulation (TENS) and exercise for chronic low back pain. *New England Journal of Medicine* 1990;**322**(23):1627-34.

Diaz-Arribas 2009 *{published data only}*

Diaz-Arribas MJ, Sanchez MR, Hervas PP, Chicharro JL, Carrere TA, Molina PO, et al. Effectiveness of the physical therapy Godelive Denys-Struyf method for nonspecific low back pain primary care randomized control trial. *Spine* 2009;**34**(15):1529-38.

Djavid 2007 *{published data only}*

Djavid GE, Mehrdad R, Ghasemi M, Hasan-Zadeh H, Sotoodeh-Manesh A, Pouryaghoub G. In chronic low back pain, low level laser therapy combined with exercise is more beneficial than exercise alone in the long term: a randomised trial. *Australian Journal of Physiotherapy* 2007;**53**(3):155-60.

Donzelli 2006 *{published data only}*

Donzelli S, Di Domenica E, Cova AM, Galletti R, Giunta N. Two different techniques in the rehabilitation treatment of low back pain: a randomized controlled trial. *Europa Medicophysica* 2006;**42**(3):205-10.

Dougherty 2014 *{published data only}*

Dougherty PE, Karuza J, Savino D, Katz P. Evaluation of a modified clinical prediction rule for use with spinal manipulative therapy in patients with chronic low back pain: a randomized clinical trial. *Chiropractic & Manual Therapies* 2014;**22**(1):41-53.

Dufour 2010 *{published data only}*

Dufour N, Thamsborg G, Oefeldt A, Lundsgaard C, Stender S. Treatment of chronic low back pain: a randomized, clinical trial comparing group-based multidisciplinary biopsychosocial rehabilitation and intensive individual therapist-assisted back muscle strengthening exercises. *Spine* 2010;**35**(5):469-76.

Elnaggar 1991 *{published data only}*

Elnaggar IM, Nordin M, Sheikhzadeh A, Parnianpour M, Kahanovitz N. Effects of spinal flexion and extension exercises on low-back pain and spinal mobility in chronic mechanical low-back pain patients. *Spine* 1991;**16**(8):967-72.

Engbert 2011 *{published data only}*

Engbert K, Weber M. Effects of therapeutic climbing in patients with chronic low back pain a randomized controlled study. *Spine* 2011;**36**(11):842-9.

Farajzadeh 2017 *{published data only}*

* Farajzadeh F, Ghaderi F, Asghari Jafarabadi M, Azghani MR, Eteraf Oskoui MA, Rezaie M. Effects of McGill stabilization exercise on pain and disability, range of motion and dynamic balance indices in patients with chronic nonspecific low back pain. *Journal of Babol University of Medical Sciences* 2017;**19**(10):21-7.

Ghorbanpour A, Azghani MR, Taghipour M, Salahzadeh Z, Ghaderi F, Oskouei AE. Effects of McGill stabilization exercises and conventional physiotherapy on pain, functional disability and active back range of motion in patients with chronic non-specific low back pain. *Journal of Physical Therapy Science* 2018;**30**(4):481-5.

Feldwieser 2018 *{published data only}*

Feldwieser F, Kiselev J, Hardy S, Garcia-Agundez A, Eicher C, Steinhagen-Thiessen E, et al. Evaluation of biofeedback based bridging exercises on older adults with low back pain: a randomized controlled trial. *Physiotherapy Practice and Research* 2018;**39**(1):15-25.

Ferreira 2007 *{published data only}*

* Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, et al. Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. *Pain* 2007;**131**(1-2):31-7.

Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Maher C, Refshauge K. Relationship between spinal stiffness and outcome in patients with chronic low back pain. *Manual Therapy* 2009;**14**(1):61-7.

Franca 2012 *{published data only}*

* Franca FR, Burke TN, Caffaro RR, Ramos LA, Marques AP. Effects of muscular stretching and segmental stabilization on functional disability and pain in patients with chronic low back pain: a randomized, controlled trial. *Journal of Manipulative & Physiological Therapeutics* 2012;**35**(4):279-85.

Franca FR, Burke TN, Hanada ES, Marques AP. Segmental stabilization and muscular strengthening in chronic low back pain: a comparative study. *Clinics* 2010;**65**(10):1013-7.

Fransoo 2006 *{published data only}*

Fransoo P. Comparison of three types of treatment for lower back pain. *Kinesitherapie La Revue* 2006;**49**:31-5.

Frost 1995 *{published data only}*

* Frost H, Klaber Moffett JA, Moser JS, Fairbank JCT. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. *British Medical Journal* 1995;**310**(6973):151-4.

Frost H, Lamb SE, Klaber Moffett JA, Fairbank JCT, Moser JS. A fitness programme for patients with chronic low back pain: 2-year follow-up of a randomised controlled trial. *Pain* 1998;**75**(2-3):273-9.

Frost 2004 *{published data only}*

* Frost H, Lamb SE, Doll HA, Taffe Carver P, Stewart-Brown S. Randomized controlled trial of physiotherapy compared

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

with advice for low back pain. *British Medical Journal* 2004;**329**:708-11.

Rivero-Arias O, Gray A, Frost H, Lamb SE, Stewart-Brown S. Cost-utility analysis of physiotherapy treatment compared with physiotherapy advice in low back pain. *Spine* 2006;**31**(12):1381-7.

Garcia 2013 *{published data only}*

* Garcia AN, Costa LC, Da Silva TM, Gondo FL, Cyrillo FN, Costa RA, et al. Effectiveness of back school versus McKenzie exercises in patients with chronic nonspecific low back pain: a randomized controlled trial. *Physical Therapy* 2013;**93**(6):729-47.

Garcia AN, Gondo FL, Costa RA, Cyrillo FN, Costa LO. Effects of two physical therapy interventions in patients with chronic non-specific low back pain: feasibility of a randomized controlled trial. *Revista Brasileira de Fisioterapia* 2011;**15**(5):420-7.

Hlobil H, Staal JB, Twisk J, Koke A, Ariens G, Smid T, et al. The effects of a graded activity intervention for low back pain in occupational health on sick leave, functional status and pain: 12-month results of a randomized controlled trial. *Journal of Occupational Rehabilitation* 2005;**15**(4):569-80.

Garcia 2017 *{published data only}*

Garcia AN, Costa L, Hancock MJ, Souza FS, Gomes G, Almeida MO, et al. McKenzie method of mechanical diagnosis and therapy was slightly more effective than placebo for pain, but not for disability, in patients with chronic non-specific low back pain: a randomised placebo controlled trial with short and longer term follow-up. *British Journal of Sports Medicine* 2017;**52**(9):594-600.

Gatti 2011 *{published data only}*

Gatti R, Faccendini S, Tettamanti A, Barbero M, Balestri A, Calori G. Efficacy of trunk balance exercises for individuals with chronic low back pain: a randomized clinical trial. *Journal of Orthopaedic & Sports Physical Therapy* 2011;**41**(8):542-52.

Geisser 2005 *{published data only}*

Geisser ME, Wiggert EA, Haig AJ, Colwell MO. A randomized, controlled trial of manual therapy and specific adjuvant exercise for chronic low back pain. *Clinical Journal of Pain* 2005;**21**(6):463-70.

Ghaderi 2016 *{published data only}*

Ghaderi F, Mohammadi K, Amir Sasan R, Niko Kheslat S, Oskouei AE. Effects of stabilization exercises focusing on pelvic floor muscles on low back pain and urinary incontinence in women. *Urology* 2016;**93**:50-4.

Gladwell 2006 *{published data only}*

Gladwell V, Head S, Haggar M, Beneke R. Does a program of Pilates improve chronic non-specific low back pain? *Journal of Sport Rehabilitation* 2006;**15**(4):338-50.

Goldby 2006 *{published data only}*

Goldby LJ, Moore AP, Doust J, Trew ME. A randomized controlled trial investigating the efficiency of musculoskeletal physiotherapy on chronic low back disorder. *Spine* 2006;**31**(10):1083-93.

Goode 2018 *{published data only}*

Goode AP, Stark Taylor S, Hastings SN, Stanwyck C, Coffman CJ, Allen KD. Effects of a home-based telephone-supported physical activity program for older adult veterans with chronic low back pain. *Physical Therapy* 2018;**98**(5):369-80.

Groessl 2017 *{published data only}*

Groessl EJ, Liu L, Chang DG, Wetherell JL, Bormann JE, Atkinson J, et al. Yoga for military veterans with chronic low back pain: a randomized clinical trial. *American Journal of Preventive Medicine* 2017;**53**(5):599-608.

Gunay 2014 *{published data only}*

Gunay S, Yildirim Y, Karadibak D. The effect of the muscle endurance training on the chronic low back pain. *Fizyoterapi Rehabilitasyon* 2014;**25**(1):28-34.

Gur 2003 *{published data only}*

Gur A, Karakoc M, Cevik R, Nas K, Sarac AJ, Karakoc M. Efficacy of low power laser therapy and exercise on pain and functions in chronic low back pain. *Lasers in Surgery and Medicine* 2003;**32**(3):233-8.

Hagen 2010 *{published data only}*

Hagen EM, Odelien KH, Lie SA, Eriksen HR. Adding a physical exercise programme to brief intervention for low back pain patients did not increase return to work. *Scandinavian Journal of Public Health* 2010;**38**(7):731-8.

Hall 2011 *{published data only}*

Hall AM, Maher CG, Lam P, Ferreira M, Latimer J. Tai chi exercise for treatment of pain and disability in people with persistent low back pain: a randomized controlled trial. *Arthritis Care Research* 2011;**63**(11):1576-83.

Halliday 2016 *{published data only}*

Halliday MH, Pappas E, Hancock MJ, Clare HA, Pinto RZ, Robertson G, et al. A randomized clinical trial comparing the McKenzie method and motor control exercises in people with chronic low back pain and a directional preference: 1-year follow-up. *Physiotherapy* 2019;**46**(7):514-22.

* Halliday MH, Pappas E, Hancock MJ, Clare HA, Pinto RZ, Robertson G, et al. A randomized controlled trial comparing the McKenzie method to motor control exercises in people with chronic low back pain and a directional preference. *Journal of Orthopaedic & Sports Physical Therapy* 2016;**46**(7):514-22.

Hansen 1993 *{published data only}*

Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L, et al. Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low back pain. *Spine* 1993;**18**(1):98-107.

Harris 2017 *{published data only}*

Harris A, Moe TF, Eriksen HR, Tangen T, Lie SA, Tveito TH, et al. Brief intervention, physical exercise and cognitive behavioural

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

group therapy for patients with chronic low back pain (the CINS trial). *European Journal of Pain* 2017;**21**(8):1397-407.

Harts 2008 *{published data only}*

Harts CC, Helmhout PH, De Bie RA, Staal JB. A high-intensity lumbar extensor strengthening program is little better than a low-intensity program or a waiting list control group for chronic low back pain: a randomised clinical trial. *Australian Journal of Physiotherapy* 2008;**54**(1):23-31.

Hartvigsen 2010 *{published data only}*

Hartvigsen J, Morso L, Bendix T, Manniche C. Supervised and non-supervised Nordic walking in the treatment of chronic low back pain: a single blind randomized clinical trial. *BMC Musculoskeletal Disorders* 2010;**11**:30-9.

Hasanpour-Dehkordi 2017 *{published data only}*

Hasanpour-Dehkordi A, Dehghani A, Solati K. A comparison of the effects of Pilates and McKenzie training on pain and general health in men with chronic low back pain: a randomized trial. *Indian Journal of Palliative Care* 2017;**23**(1):36-40.

Helmhout 2004 *{published data only}*

Helmhout PH, Hart CC, Staal JB, Da Bie RA. Rationale and design of a multicenter randomized controlled trial on a 'minimal intervention' in Dutch army personnel with nonspecific low back pain. *BMC Musculoskeletal Disorders* 2004;**5**(1):40-9.

* Helmhout PH, Harts CC, Staal JB, Candel MJ, De Bie RA. Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. *European Spine Journal* 2004;**13**(6):537-47.

Hemmilä 1997 *{published data only}*

* Hemmilä HM, Keinanen-Kiukaanniemi SM, Levoska S, Puska P. Does folk medicine work? A randomized clinical trial on patients with prolonged back pain. *Archives of Physical Medicine and Rehabilitation* 1997;**78**(6):571-7.

Hemmila H, Keinanen-Kiukaanniemi SM, Levoska S, Puska P. Long-term effectiveness of bone-setting, light exercise therapy, and physiotherapy for prolonged back pain: a randomized controlled trial. *Journal of Manipulative & Physiological Therapeutics* 2002;**25**:99-104.

Henchoz 2010 *{published data only}*

Henchoz Y, De Goumoens P, Norberg M, Paillex R, So AKL. Role of physical exercise in low back pain rehabilitation: a randomized controlled trial of a three-month exercise program in patients who have completed multidisciplinary rehabilitation. *Spine* 2010;**35**(12):1192-9.

Henry 2014 *{published data only}*

Henry SM, Van Dillen LR, Ouellette-Morton RH, Hitt JR, Lomond KV, DeSarno MJ, et al. Outcomes are not different for patient-matched versus nonmatched treatment in subjects with chronic recurrent low back pain: a randomized clinical trial. *Spine* 2014;**14**(12):2799-810.

Highland 2018 *{published data only}*

Highland KB, Schoomaker A, Rojas W, Suen J, Ahmed A, Zhang Z, et al. Benefits of the restorative exercise and strength training for operational resilience and excellence yoga program for chronic lower back pain in service members: a pilot randomized control trial. *Archives of Physical Medicine and Rehabilitation* 2018;**99**(1):91-8.

Hildebrandt 2000 *{published data only}*

Hildebrandt VH, Proper KI, Van den Berg R, Douwes M, Van den Heuvel SG, Van Buuren S. Cesar therapy is temporarily more effective in patients with chronic low back pain than the standard treatment by family practitioner: randomized, controlled and blinded clinical trial with 1 year follow-up. *Nederlands Tijdschrift voor Geneeskunde* 2000;**144**:2258-64.

Iversen 2018 *{published data only}*

Iversen VM, Vasseljen O, Mork PJ, Gismervik S, Bertheussen GF, Salvesen O, et al. Resistance band training or general exercise in multidisciplinary rehabilitation of low back pain? A randomized trial. *Scandinavian Journal of Medicine & Science in Sports* 2018;**28**(9):2074-83.

Jans 2006 *{published data only}*

Jans MP, De Korte EM, Heinrich J, Hildebrandt VH. Intermittent follow-up treatment with Cesar exercise therapy in patients with subacute or chronic aspecific low back pain: results of a randomized, controlled trial with a 1.5-year follow-up. *Nederlands Tijdschrift Voor Fysiotherapie* 2006;**116**(5):111-6.

Jaromi 2012 *{published data only}*

Jaromi M, Nemeth A, Kranicz J, Laczko T, Betlehem J. Treatment and ergonomics training of work-related lower back pain and body posture problems for nurses. *Journal of Clinical Nursing* 2012;**21**(11/12):1776-84.

Jaromi 2018 *{published data only}*

Jaromi M, Kukla A, Szilagyi B, Simon-Ugron A, Bobaly VK, Makai A, et al. Back School programme for nurses has reduced low back pain levels: a randomised controlled trial. *Journal of Clinical Nursing* 2018;**27**(5-6):e895-902.

Javadian 2012 *{published data only}*

Javadian Y, Behtash H, Akbari M, Taghipour-Darzi M, Zekavat H. The effects of stabilizing exercises on pain and disability of patients with lumbar segmental instability. *Journal of Back and Musculoskeletal Rehabilitation* 2012;**25**(3):149-55.

Jensen 2012 *{published data only}*

Jensen RK, Leboeuf-Yde C, Wedderkopp N, Sorensen JS, Manniche C. Rest versus exercise as treatment for patients with low back pain and modic changes. A randomized controlled clinical trial. *BMC Medicine* 2012;**10**:22-35.

Johanssen 1995 *{published data only}*

Johanssen F, Remvig L, Kryger P, Beck P, Warming S, Lybeck K, et al. Exercises for chronic low back pain: a clinical trial. *Journal of Orthopaedic & Sports Physical Therapy* 1995;**22**(2):52-9.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jousset 2004 *{published data only}*

Jousset N, Fanello S, Bontoux L, Dubus V, Billabert C, Vielle B, et al. Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. *Spine* 2004;**29**(5):487-93.

Kaapa 2006 *{published data only}*

Kaapa EH, Frantsi K, Sarna S, Malmivaara A. Multidisciplinary group rehabilitation versus individual physiotherapy for chronic nonspecific low back pain: a randomized trial. *Spine* 2006;**31**(4):371-6.

Kader 2012 *{published data only}*

Kader D, Radha S, Smith F, Wardlaw D, Scott N, Rege A, et al. Evaluation of perifacet injections and paraspinal muscle rehabilitation in treatment of low back pain. A randomised controlled trial. *Ortopedia Traumatologia Rehabilitacja* 2012;**14**(3):251-9.

Kankaanpaa 1999 *{published data only}*

Kankaanpaa M, Taimela S, Airaksinen O, Hanninen O. The efficacy of active rehabilitation in chronic low back pain. Effect on pain intensity, self-experienced disability, and lumbar fatigability. *Spine* 1999;**24**(10):1034-42.

Keane 2017 *{published data only}*

Keane LG. Comparing aquastretch with supervised land based stretching for chronic lower back pain. *Journal of Bodywork and Movement Therapies* 2017;**21**(2):297-305.

Kell 2009 *{published data only}*

Kell RT, Asmundson GJ. A comparison of two forms of periodized exercise rehabilitation programs in the management of chronic nonspecific low-back pain. *Journal of Strength & Conditioning Research* 2009;**23**(2):513-23.

Kendall 1968 *{published data only}*

Kendall PH, Jenkins JM. Exercises for backache: a double-blind controlled trial. *Physiotherapy* 1968;**54**:154-7.

Kendall 2015 *{published data only}*

Kendall KD, Emery CA, Wiley JP, Ferber R. The effect of the addition of hip strengthening exercises to a lumbopelvic exercise programme for the treatment of non-specific low back pain: a randomized controlled trial. *Journal of Science and Medicine in Sport* 2015;**18**(6):626-31.

Kim 2015a *{published data only}*

Kim TH, Kim EH, Cho H. The effects of the CORE programme on pain at rest, movement-induced and secondary pain, active range of motion, and proprioception in female office workers with chronic low back pain: a randomized controlled trial. *Clinical Rehabilitation* 2015;**29**(7):653-62.

Kim 2017 *{published data only}*

Kim BR, Lee HJ. Effects of proprioceptive neuromuscular facilitation-based abdominal muscle strengthening training on pulmonary function, pain, and functional disability index in chronic low back pain patients. *Journal of Exercise Rehabilitation* 2017;**13**(4):486-90.

Kim 2018a *{published data only}*

Kim YW, Kim NY, Chang WH, Lee SC. Comparison of the therapeutic effects of a sling exercise and a traditional stabilizing exercise for clinical lumbar spinal instability. *Journal of Sport Rehabilitation* 2018;**27**(1):47-54.

Kladny 2003 *{published data only}*

Kladny B, Fischer FC, Haase I. Evaluation of specific stablizing exercise in the treatment of low back pain and lumbar disk disease in outpatient rehabilitation. *Zeitschrift fur Orthopadie und Ihre Grenzgbiete* 2003;**141**:401-5.

Kofotolis 2006 *{published data only}*

Kofotolis N, Kellis E. Effects of two 4-week proprioceptive neuromuscular facilitation programs on muscle endurance, flexibility, and functional performance in women with chronic low back pain. *Physical Therapy* 2006;**86**(7):1001-12.

Kofotolis 2016 *{published data only}*

Kofotolis N, Kellis E, Vlachopoulos SP, Gouitas I, Theodorakis Y. Effects of Pilates and trunk strengthening exercises on health-related quality of life in women with chronic low back pain. *Journal of Back & Musculoskeletal Rehabilitation* 2016;**29**(4):649-59.

Koldas 2008 *{published data only}*

Koldas DS, Sonel TB, Kurtais Y, Atay MB. Comparison of three different approaches in the treatment of chronic low back pain. *Clinical Rheumatology* 2008;**27**(7):873-81.

Kool 2007 *{published data only}*

* Kool J, Bachmann S, Oesch P, Knuesel O, Ambergen T, De Bie R, et al. Function-centered rehabilitation increases work days in patients with nonacute nonspecific low back pain: 1-year results from a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* 2007;**88**(9):1089-94.

Kool JP, Oesch PR, Bachmann S, Knuesel O, Dierkes JG, Russo M, et al. Increasing days at work using function-centered rehabilitation in nonacute nonspecific low back pain: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* 2005;**86**(5):857-64.

Koumantakis 2005 *{published data only}*

Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only: randomized controlled trial of patients with recurrent low back pain. *Physical Therapy* 2005;**85**(3):209-25.

Krein 2013 *{published data only}*

Krein SL, Kadri R, Hughes M, Kerr EA, Piette JD, Holleman R, et al. Pedometer-based internet-mediated intervention for adults with chronic low back pain: randomized controlled trial. *Journal of Medical Internet Research* 2013;**15**(8):e181-e208. [PMID: 23969029]

Kumar 2009a *{published data only}*

Kumar S, Negi MP, Sharma VP, Shukla R, Dev R, Mishra UK. Efficacy of two multimodal treatments on physical strength of occupationally subgrouped males with low back pain. *Journal of Back & Musculoskeletal Rehabilitation* 2009;**22**(3):179-88.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kumar 2009b *{published data only}*

Kumar S, Sharma VP, Negi MP. Efficacy of dynamic muscular stabilization techniques (DMST) over conventional techniques in rehabilitation of chronic low back pain. *Journal of Strength & Conditioning Research* 2009;**23**(9):2651-9.

Kumar 2010 *{published data only}*

Kumar S, Sharma VP, Shukla R, Dev R. Comparative efficacy of two multimodal treatments on male and female sub-groups with low back pain (part II). *Journal of Back & Musculoskeletal Rehabilitation* 2010;**23**(1):1-9.

Kuukkanen 2000 *{published data only}*

Kuukkanen T, Malkia E, Kautiainen H, Pohjolainen T. Effectiveness of a home exercise programme in low back pain: a randomized five-year follow-up study. *Physiotherapy Research International* 2007;**12**(4):213-24.

* Kuukkanen TM, Malkia EA. An experimental controlled study on postural sway and therapeutic exercise in subjects with low back pain. *Clinical Rehabilitation* 2000;**14**(2):192-202.

Lee 2016a *{published data only}*

Lee JS. The effects of strength exercise and walking on lumbar function, pain level, and body composition in chronic back pain patients. *Journal of Exercise Rehabilitation* 2016;**12**(5):463-70.

Lewis 2005 *{published data only}*

Lewis JS, Hewitt JS, Billington L, Cole S, Byng J, Karayiannis S. A randomized clinical trial comparing two physiotherapy interventions for chronic low back pain. *Spine* 2005;**30**(7):711-21.

Li 2008 *{published data only}*

Li Y, Wang Y, Zhang Y. Rehabilitation nursing effect of carrying out backbone deep muscles stability training for patients with chronic low back pain. *Chinese Nursing Research* 2008;**22**(5A):1166-7.

Lie 1999 *{published data only}*

Lie H, Frey S. Mobilizing or stabilizing exercise in degenerative disk disease in the lumbar region? *Tidsskrift for Den Norske Laegeforening* 1999;**119**:2051-3.

Limke 2008 *{published data only}*

Limke JC, Rainville J, Pena E, Childs L. Randomized trial comparing the effects of one set versus two sets of resistance exercises for outpatients with chronic low back pain and leg pain. *European Journal of Physical and Rehabilitation Medicine* 2008;**44**(4):399-405.

Little 2008 *{published data only}*

Little P, Lewith G, Webley F, Evans M, Beattie A, Middleton K, et al. Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain. *British Medical Journal* 2008;**337**:a884-92.

Little 2014 *{published data only}*

Little P, Stuart B, Stokes M, Nicholls C, Roberts L, Preece S, et al. Alexander technique and supervised physiotherapy exercises in back pain (aspen) feasibility trial. *Efficacy and Mechanism Evaluation* 2014;**1**(2):1-106.

Lomond 2014 *{published data only}*

Lomond KV, Jacobs JV, Hitt JR, DeSarno MJ, Bunn JY, Henry SM. Effects of low back pain stabilization or movement system impairment treatments on voluntary postural adjustments: a randomized controlled trial. *Spine Journal* 2014;**15**(4):596-606.

Long 2004 *{published data only}*

Long A, Donelson R, Fung T. Does it matter which exercise? A randomized control trial of exercise for low back pain. *Spine* 2004;**29**(23):2593-602.

Longo 2016 *{published data only}*

Longo CA, Ferrari S, Lanza A, Vanti C, Stradiotti P, Ausenda CD. Effects of the "standing posture with flexion of the trunk" on the lumbar multifidus tropism in patients with chronic low back pain. A randomized, controlled pilot study. *Scienza Riabilitativa* 2016;**18**(2):5-12.

Lopes 2017 *{published data only}*

Lopes S, Correia C, Felix G, Lopes M, Cruz A, Ribeiro F. Immediate effects of Pilates based therapeutic exercise on postural control of young individuals with non-specific low back pain: a randomized controlled trial. *Complementary Therapies in Medicine* 2017;**34**:104-10.

Macedo 2012 *{published data only}*

* Macedo LG, Latimer J, Maher CG, Hodges PW, McAuley JH, Nicholas MK, et al. Effect of motor control exercises versus graded activity in patients with chronic nonspecific low back pain: a randomized controlled trial. *Physical Therapy* 2012;**92**(3):363-77.

Macedo LG, Maher CG, Hancock MJ, Kamper SJ, McAuley JH, Stanton TR, et al. Predicting response to motor control exercises and graded activity for patients with low back pain: preplanned secondary analysis of a randomized controlled trial. *Physical Therapy* 2014;**94**(11):1543-54.

Machado 2007 *{published data only}*

Machado LA, Azevedo DC, Capanema MB, Neto TN, Cerceau DM. Client-centered therapy vs exercise therapy for chronic low back pain: a pilot randomized controlled trial in Brazil. *Pain Medicine* 2007;**8**(3):251-8.

Magalhaes 2017 *{published data only}*

Magalhaes MO, Comachio J, Ferreira PH, Pappas E, Marques AP. Effectiveness of graded activity versus physiotherapy in patients with chronic nonspecific low back pain: midterm follow up results of a randomized controlled trial. *Brazilian Journal of Physical Therapy* 2017;**22**(1):82-91.

Magnussen 2005 *{published data only}*

Magnussen L, Rognsvag T, Tveito TH, Eriksen HR. Effect of a brief cognitive training programme in patients with long-lasting back pain evaluated as unfit for surgery. *Journal of Health Psychology* 2005;**10**(2):233-43.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Manniche 1988 *{published data only}*

* Manniche C, Hesselsoe G, Bentzen L, Christensen I, Lundberg E. Clinical trial of intensive muscle training for chronic low back pain. *Lancet* 1988;**2**(8626-7):1473-6.

Manniche C, Lundberg E, Christensen I, Bentzen L, Hesselsoe G. Intensive dynamic back exercises for chronic low back pain: a clinical trial. *Pain* 1991;**47**(1):53-63.

Mannion 1999 *{published data only}*

* Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. *Spine* 1999;**24**(23):2435-48.

Mannion AF, Muntener M, Taimela S, Dvorak J. Comparison of three active therapies for chronic low back pain: results of a randomized clinical trial with one-year follow-up. *Rheumatology* 2001;**40**(7):772-8.

Mannion AF, Taimela S, Muntener M, Dvorak J. Active therapy for chronic low back pain part 1. Effects on back muscle activation, fatigability, and strength. *Spine* 2001;**26**(8):897-908.

Marshall 2008 *{published data only}*

Marshall P, Murphy B. Self-report measures best explain changes in disability compared with physical measures after exercise rehabilitation for chronic low back pain. *Spine* 2008;**33**(3):326-38.

Marshall 2013 *{published data only}*

Brooks C, Kennedy S, Marshall PWM. Specific trunk and general exercise elicit similar changes in anticipatory postural adjustments in patients with chronic low back pain: a randomized controlled trial. *Spine* 2012;**37**(25):E1543-50.

* Marshall PWM, Kennedy S, Brooks C, Lonsdale C. Pilates exercise or stationary cycling for chronic nonspecific low back pain: does it matter? a randomized controlled trial with 6-month follow-up. *Spine* 2013;**38**(15):E952-9.

Masharawi 2013 *{published data only}*

Masharawi Y, Nadaf N. The effect of non-weight bearing group-exercising on females with non-specific chronic low back pain: a randomized single blind controlled pilot study. *Journal of Back and Musculoskeletal Rehabilitation* 2013;**26**(4):353-9.

Masse-Alarie 2016 *{published data only}*

Masse-Alarie H, Beaulieu LD, Preuss R, Schneider C. Influence of paravertebral muscles training on brain plasticity and postural control in chronic low back pain. *Scandinavian Journal of Pain* 2016;**12**:74-83.

Maul 2005 *{published data only}*

Maul I, Laubli T, Oliveri M, Krueger H. Long-term effects of supervised physical training in secondary prevention of low back pain. *European Spine Journal* 2005;**14**(6):599-611.

Mazloum 2018 *{published data only}*

Mazloum V, Sahebozamani M, Barati A, Nakhaee N, Rabiei P. The effects of selective Pilates versus extension-based exercises on rehabilitation of low back pain. *Journal of Bodywork and Movement Therapies* 2018;**22**(4):999-1003.

McCaskey 2018 *{published data only}*

McCaskey MA, Wirth B, Schuster-Amft C, De Bruin ED. Postural sensorimotor training versus sham exercise in physiotherapy of patients with chronic non-specific low back pain: an exploratory randomised controlled trial. *PLOS One* 2018;**13**(3):e0193358.

McDonough 2013 *{published data only}*

McDonough SM, Tully MA, Boyd A, O'Connor SR, Kerr DP, O'Neill SM, et al. Pedometer-driven walking for chronic low back pain: a feasibility randomized controlled trial. *Clinical Journal of Pain* 2013;**29**(11):972-81.

Mehling 2005 *{published data only}*

Mehling WE, Hamel KA, Acree M, Byl N, Hecht FM. Randomized, controlled trial of breath therapy for patients with chronic low-back pain. *Alternative Therapies in Health & Medicine* 2005;**11**(4):44-52.

Michaelson 2016 *{published data only}*

Aasa P, Berglund B, Michaelson L, Aasa U. Individualized low-load motor control exercises and education versus a high-load lifting exercise and education to improve activity, pain intensity, and physical performance in patients with low back pain: a randomized controlled trial. *Journal of Orthopaedic & Sports Physical Therapy* 2015;**45**(2):77-85.

Berglund L, Aasa B, Michaelson P, Aasa U. Effects of low-load motor control exercises and a high-load lifting exercise on lumbar multifidus thickness: a randomized controlled trial. *Spine* 2016;**42**(15):e876-82.

* Michaelson P, Holmberg D, Aasa B, Aasa U. High load lifting exercise and low load motor control exercises as interventions for patients with mechanical low back pain: a randomized controlled trial with 24-month follow-up. *Journal of Rehabilitation Medicine* 2016;**48**(5):456-63.

Michalsen 2016 *{published data only}*

Michalsen A, Kunz N, Jeitler M, Brunnhuber M, Meier L, Ludtke R, et al. Effectiveness of focused meditation for patients with chronic low back pain - a randomized controlled clinical trial. *Complementary Therapies in Medicine* 2016;**26**:79-84.

Miller 2005 *{published data only}*

Miller ER, Schenk RJ, Kames JL, Rousselle JG. A comparison of the Mckenzie approach to a specific spine stabilization program for chronic low back pain. *Journal of Manual and Manipulative Therapy* 2005;**13**(2):103-12.

Miyamoto 2013 *{published data only}*

Miyamoto GC, Oliveira Pena CL, Galvanin T, Nunes Cabral CM. Efficacy of the addition of modified Pilates exercises to a minimal intervention in patients with chronic low back pain: a randomized controlled trial. *Physical Therapy* 2013;**93**(3):310-20.

Miyamoto 2018 *{published data only}*

Miyamoto GC, Franco KF, Van Dongen JM, Franco Y, De Oliveira NT, Amaral DD, et al. Different doses of Pilates-based exercise therapy for chronic low back pain: a randomised

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

controlled trial with economic evaluation. *British Journal of Sports Medicine* 2018;**52**(13):859-68.

Mohseni-Bandpei 2011 *{published data only}*

Mohseni-Bandpei MA, Rahmani N, Behtash H, Karimloo M. The effect of pelvic floor muscle exercise on women with chronic non-specific low back pain. *Journal of Bodywork and Movement Therapies* 2011;**15**(1):75-81.

Monticone 2014 *{published data only}*

* Monticone M, Ambrosini E, Rocca B, Magni S, Brivio F, Ferrante S. A multidisciplinary rehabilitation programme improves disability, kinesiophobia and walking ability in subjects with chronic low back pain: results of a randomised controlled pilot study. *European Spine Journal* 2014;**23**(10):2105-13.

Monticone M. Group-based task-oriented exercises aimed at managing kinesiophobia improved disability in chronic low back pain with consumer summary. *European Journal of Pain* 2016;**20**(4):541-51.

Moon 2013 *{published data only}*

Moon HJ, Choi KH, Kim DH, Kim HJ, Cho YK, Lee KH, et al. Effect of lumbar stabilization and dynamic lumbar strengthening exercises in patients with chronic low back pain. *Annals of Rehabilitation Medicine* 2013;**37**(1):110-7.

Moreschi-Guastala 2016 *{published data only}*

Moreschi-Guastala FA, Guerini MH, Klein PF, Leite VC, Cappellazzo R, Facci LM. Effect of global postural re-education and isostretching in patients with nonspecific chronic low back pain: a randomized clinical trial. *Fisioterapia em Movimento* 2016;**29**(3):515-25.

Morone 2011 *{published data only}*

* Morone G, Paolucci T, Alcuri MR, Vulpiani MC, Matano A, Bureca I, et al. Quality of life improved by multidisciplinary back school program in patients with chronic non-specific low back pain: a single blind randomized controlled trial. *European Journal of Physical & Rehabilitation Medicine* 2011;**47**(4):533-41.

Paolucci T, Morone G, Iosa M, Fusco A, Alcuri R, Matano A, et al. Psychological features and outcomes of the Back School treatment in patients with chronic non-specific low back pain. A randomized controlled study. *European Journal of Physical and Rehabilitation Medicine* 2012;**48**(2):245-53.

Morone 2012 *{published data only}*

Morone G, Iosa M, Paolucci T, Fusco A, Alcuri R, Spadini E, et al. Efficacy of perceptive rehabilitation in the treatment of chronic nonspecific low back pain through a new tool: a randomized clinical study. *Clinical Rehabilitation* 2012;**26**(4):339-50.

Moseley 2002 *{published data only}*

Moseley L. Combined physiotherapy and education is efficacious for chronic low back pain. *Australian Journal of Physiotherapy* 2002;**48**(4):297-302.

Moussouli 2014 *{published data only}*

Moussouli M, Vlachopoulos SP, Kofotolis ND, Theodorakis Y, Malliou P, Kellis E. Effects of stabilization exercises on health-related quality of life in women with chronic low back pain. *Journal of Physical Activity and Health* 2014;**11**:1295-303.

Muharram 2011 *{published data only}*

Muharram A, Liu W, Wang Z, Sun L, Wu W. Shadow boxing for relief of chronic low back pain. *International Journal of Athletic Therapy & Training* 2011;**16**(6):29-33.

Murtezani 2011 *{published data only}*

Murtezani A, Govori V, Meka VS, Ibraimi Z, Rrecaj S, Gashi S. A comparison of Mckenzie therapy with electrophysical agents for the treatment of work related low back pain: a randomized controlled trial. *Journal of Back & Musculoskeletal Rehabilitation* 2014;**28**(2):247-53.

* Murtezani A, Hundozi H, Orovcanec N, Sllamniku S, Osmani T. A comparison of high intensity aerobic exercise and passive modalities for the treatment of workers with chronic low back pain: a randomized, controlled trial. *European Journal of Physical and Rehabilitation Medicine* 2011;**47**(3):359-66.

Nabavi 2018 *{published data only}*

Nabavi N, Mohseni Bandpei MA, Mosallanezhad Z, Rahgozar M, Jaberzadeh S. The effect of 2 different exercise programs on pain intensity and muscle dimensions in patients with chronic low back pain: a randomized controlled trial. *Journal of Manipulative & Physiological Therapeutics* 2018;**41**(2):102-10.

Nagrale 2012 *{published data only}*

Nagrale AV, Patil SP, Gandhi RA, Learman K. Effect of slump stretching versus lumbar mobilization with exercise in subjects with non-radicular low back pain: a randomized clinical trial. *Journal of Manual and Manipulative Therapy* 2012;**20**(1):35-42.

Nassif 2011 *{published data only}*

Nassif H, Brosset N, Guillaume M, Delore-Milles E, Tafflet M, Buchholz F, et al. Evaluation of a randomized controlled trial in the management of chronic lower back pain in a French automotive industry: an observational study. *Archives of Physical Medicine & Rehabilitation* 2011;**92**(12):1927-36.

Natour 2015 *{published data only}*

Natour J, De Araujo Cazotti L, Ribeiro LH, Baptista AS, Jones A. Pilates improves pain, function and quality of life in patients with chronic low back pain: a randomized controlled trial. *Clinical Rehabilitation* 2015;**29**(1):59-68.

Nemcic 2013 *{published data only}*

Nemcic T, Budisin V, Vrabec-Matkovic D, Grazio S. Comparison of the effects of land-based and water-based therapeutic exercises on the range of motion and physical disability in patients with chronic low-back pain: single-blinded randomized study. *Acta Clinica Croatica* 2013;**52**(3):321-7.

Nguyen 2017 *{published data only}*

Nguyen C, Boutron I, Rein C, Baron G, Sanchez K, Palazzo C, et al. Intensive spa and exercise therapy program for returning to work for low back pain patients: a randomized controlled trial. *Scientific Reports* 2017;**7**(1):1-10.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Niemisto 2003 *{published data only}*

* Niemisto L, Lahtinen-Suopanki T, Rissanen P, Lindgren KA, Sarna S, Hurri H. A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. *Spine* 2003;**28**(19):2185-91.

Niemisto L, Rissanen P, Sarna S, Lahtinen-Suopanki T, Lindgren KA, Hurri H. Cost-effectiveness of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain: a prospective randomized trial with 2-year follow-up. *Spine* 2005;**30**(10):1109-15.

Riipinen M, Niemisto L, Lindgren KA, Hurri H. Psychosocial differences as predictors for recovery from chronic low back pain following manipulation, stabilizing exercises and physician consultation or physician consultation alone. *Journal of Rehabilitation Medicine* 2005;**37**(3):152-8.

Noori 2011 *{published data only}*

Noori S, Ghasemi G, Khayambashi K, Karimi A, Minasian V, Alizamani S. Effect of exercise therapy and physiotherapy on patients with chronic low back pain. *Journal of Isfahan Medical School* 2011;**29**(151):1-7.

Oh 2014 *{published data only}*

Oh HW, Lee MG, Jang JY, Jin JJ, Cha JY, Jin YY, et al. Time-effects of horse simulator exercise on psychophysiological responses in men with chronic low back pain. *Isokinetics and Exercise Science* 2014;**22**(2):153-63.

Ojoawo 2017 *{published data only}*

Ojoawo AO, Hassan MA, Olaogun MOB, Johnson EO, Mbada CE. Comparative effectiveness of two stabilization exercise positions on pain and functional disability of patients with low back pain. *Journal of Exercise Rehabilitation* 2017;**13**(3):363-71.

Okafor 2012 *{published data only}*

Okafor UAC, Solanke TA, Akinbo SRA, Odebiyi DO. Effect of aerobic dance on pain, functional disability and quality of life on patients with chronic low back pain. *South African Journal of Physiotherapy* 2012;**68**(3):11-4.

Olivier 2008 *{published data only}*

Olivier N, Lepretre A, Caby I, Dupuis MA, Prieur F. Does exercise therapy for chronic lower-back pain require daily isokinetic reinforcement of the trunk muscles? *Annales de R'adaptation et de Medecine Physique* 2008;**51**(4):284-91.

Paolucci 2012 *{published data only}*

Paolucci T, Fusco A, Iosa M, Grasso MR, Spadini E, Paolucci S, et al. The efficacy of a perceptive rehabilitation on postural control in patients with chronic nonspecific low back pain. *International Journal of Rehabilitation Research* 2012;**35**(4):360-6.

Paolucci 2017 *{published data only}*

Paolucci T, Zangrando F, Iosa M, De Angelis S, Marzoli C, Piccinini G, et al. Improved interoceptive awareness in chronic low back pain: a comparison of Back school versus Feldenkrais method. *Disability & Rehabilitation* 2017;**39**(10):994-1001.

Patti 2016 *{published data only}*

Patti A, Bianco A, Paoli A, Messina G, Montalto MA, Bellafiore M, et al. Pain perception and stabilometric parameters in people with chronic low back pain after a Pilates exercise program: a randomized controlled trial. *Medicine* 2016;**95**(2):e2414-21.

Paungmali 2017 *{published data only}*

Paungmali A, Joseph LH, Sitilertpisan P, Pirunsan U, Uthaikhup S. Lumbopelvic core stabilization exercise and pain modulation among individuals with chronic nonspecific low back pain. *Pain Practice* 2017;**17**(8):1008-14.

Petersen 2002 *{published data only}*

* Petersen T, Kryger P, Ekdahl C, Olsen S, Jacobsen S. The effect of McKenzie therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic low back pain: a randomized controlled trial. *Spine* 2002;**27**(16):1702-9.

Petersen T, Larsen K, Jacobsen S. One-year follow-up comparison of the effectiveness of McKenzie treatment and strengthening training for patients with chronic low back pain: outcome and prognostic factors. *Spine* 2007;**32**(26):2948-56.

Petersen 2011 *{published data only}*

Petersen T, Larsen K, Nordsteen J, Olsen S, Fournier G, Jacobsen S. The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: a randomized controlled trial. *Spine* 2011;**36**(24):1999-2010.

Phattharasupharerk 2019a *{published data only}*

Phattharasupharerk S, Purepong N, Eksakulkla S, Siriphorn A. Effects of Qigong practice in office workers with chronic non-specific low back pain: a randomized control trial. *Journal of Bodywork and Movement Therapies* 2019;**23**(2):375-81.

Puntumetakul 2013 *{published data only}*

Puntumetakul R, Areeudomwong P, Emasithi A, Yamauchi J. Effect of 10-week core stabilization exercise training and detraining on pain-related outcomes in patients with clinical lumbar instability. *Patient Preference and Adherence* 2013;**7**:1189-99.

Puppin 2011 *{published data only}*

Puppin MA, Marques AP, Da Silva AG, Neto HD. Stretching in nonspecific chronic low back pain: a strategy of the GDS method. *Fisioterapia e Pesquisa* 2011;**18**(2):116-21.

Quinn 2011 *{published data only}*

Quinn K, Barry S, Barry L. Do patients with chronic low back pain benefit from attending Pilates classes after completing conventional physiotherapy treatment? *Physiotherapy Ireland* 2011;**32**(1):5-12.

Rantonen 2012 *{published data only}*

Rantonen J, Karppinen J, Vehtari A, Luoto S, Viikari-Juntura E, Hupli M, et al. Effectiveness of three interventions for secondary prevention of low back pain in the occupational health setting:

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

a randomised controlled trial with a natural course control. *BMC Public Health* 2018;**18**(1):598-611.

* Rantonen J, Luoto S, Vehtari A, Hupli M, Karppinen J, Malmivaara A, et al. The effectiveness of two active interventions compared to self-care advice in employees with non-acute low back symptoms: a randomised, controlled trial with a 4-year follow-up in the occupational health setting. *Occupational & Environmental Medicine* 2012;**69**(1):12-20.

Rasmussen-Barr 2003 *{published data only}*

Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I. Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. *Manual Therapy* 2003;**8**(4):233-41.

Rasmussen-Barr 2009 *{published data only}*

Rasmussen-Barr E, Ang B, Arvidsson I, Nilsson-Wikmar L. Graded exercise for recurrent low-back pain: a randomized, controlled trial with 6-, 12-, and 36-month follow-ups. *Spine* 2009;**34**(3):221-8.

Rhee 2012 *{published data only}*

Rhee HS, Kim YH, Sung PS. A randomized controlled trial to determine the effect of spinal stabilization exercise intervention based on pain level and standing balance differences in patients with low back pain. *Medical Science Monitor* 2012;**18**(3):174-81.

Risch 1993 *{published data only}*

Risch SV, Norvell NK, Pollock ML, Risch ED, Langer H, Fulton M, et al. Lumbar strengthening in chronic low back pain patients: physiologic and psychological benefits. *Spine* 1993;**18**(2):232-8.

Rittweger 2002 *{published data only}*

Rittweger J, Just K, Kautzsch K, Reeg P, Felsenberg D. Treatment of chronic lower back pain with lumbar extension and whole-body vibration exercise: a randomized controlled trial. *Spine* 2002;**27**(17):1829-34.

Ritvanen 2007 *{published data only}*

Ritvanen T, Zaproudina N, Nissen M, Leinonen V, Hanninen O. Dynamic surface electromyographic responses in chronic low back pain treated by traditional bone setting and conventional physical therapy. *Journal of Manipulative & Physiological Therapeutics* 2007;**30**(1):31-7.

Roche-Leboucher 2011 *{published data only}*

Roche-Leboucher G, Petit-Lemanac'h A, Bontoux L, Dubus-Bausiere V, Parot-Shinkel E, Fanello S, et al. Multidisciplinary intensive functional restoration versus outpatient active physiotherapy in chronic low back pain: a randomized controlled trial. *Spine* 2011;**36**(26):2235-42.

Ryan 2010 *{published data only}*

Ryan CG, Gray HG, Newton M, Granat MH. Pain biology education and exercise classes compared to pain biology education alone for individuals with chronic low back pain: a pilot randomised controlled trial. *Manual Therapy* 2010;**15**(4):382-7.

Rydeard 2006a *{published data only}*

Rydeard R, Leger A, Smith D. Back pain and functional disability: a randomized controlled trial. *Journal of Orthopaedic & Sports Physical Therapy* 2006;**36**(7):472-84.

Salamat 2017 *{published data only}*

Salamat S, Talebian S, Bagheri H, Maroufi N, Jafar Shaterzadeh M, Kalbasi G, et al. Effect of movement control and stabilization exercises in people with extension related non-specific low back pain: a pilot study. *Journal of Bodywork & Movement Therapies* 2017;**21**(4):860-5.

Salavati 2016 *{published data only}*

Salavati M, Akhbari B, Takamjani IE, Bagheri H, Ezzati K, Kahlaee AH. Effect of spinal stabilization exercise on dynamic postural control and visual dependency in subjects with chronic non-specific low back pain. *Journal of Bodywork & Movement Therapies* 2016;**20**(2):441-8.

Saner 2016 *{published data only}*

Saner J, Sieben JM, Kool J, Luomajoki H, Bastiaenen CH, De Bie RA. A tailored exercise program versus general exercise for a subgroup of patients with low back pain and movement control impairment: short-term results of a randomised controlled trial. *Journal of Bodywork & Movement Therapies* 2016;**20**(1):189-202.

Saper 2009 *{published data only}*

Saper RB, Sherman KJ, Cullum-Dugan D, Davis RB, Phillips RS, Culpepper L. Yoga for chronic low back pain in a predominantly minority population: a pilot randomized controlled trial. *Alternative Therapies in Health & Medicine* 2009;**15**(6):18-27.

Saper 2013 *{published data only}*

Saper RB, Boah AR, Keosaian J, Cerrada C, Weinberg J, Sherman KJ. Comparing once-versus twice-weekly yoga classes for chronic low back pain in predominantly low income minorities: a randomized dosing trial. *Evidence-based Complementary and Alternative Medicine* 2013;**2013**:1-3.

Saper 2017 *{published data only}*

Saper RB, Lemaster C, Delitto A, Sherman KJ, Herman PM, Sadikova E, et al. Yoga, physical therapy, or education for chronic low back pain: a randomized noninferiority trial. *Annals of Internal Medicine* 2017;**167**(2):85-94.

Schenkman 2009 *{published data only}*

Schenkman ML, Jordan S, Akuthota V, Roman M, Kohrt WM, Hearty T, et al. Functional movement training for recurrent low back pain: lessons from a pilot randomized controlled trial. *PM&R: Journal of Injury, Function, & Rehabilitation* 2009;**1**(2):137-46.

Schinhan 2016 *{published data only}*

Schinhan M, Neubauer B, Pieber K, Gruber M, Kainberger F, Castellucci C, et al. Climbing has a positive impact on low back pain: a prospective randomized controlled trial. *Clinical Journal of Sport Medicine* 2016;**26**(3):199-205.

Segal-Snir 2016 *{published data only}*

Segal-Snir Y, Lubetzky VA, Masharawi Y. Rotation exercise classes did not improve function in women with non-specific

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

chronic low back pain: a randomized single blind controlled study. *Journal of Back & Musculoskeletal Rehabilitation* 2016;**29**(3):467-75.

Sertpoyraz 2009 *{published data only}*

Sertpoyraz F, Eyigor S, Karapolat H, Capaci K, Kirazli Y. Comparison of isokinetic exercise versus standard exercise training in patients with chronic low back pain: a randomized controlled study. *Clinical Rehabilitation* 2009;**23**(3):238-47.

Shankar 2011 *{published data only}*

Shankar N, Thakur M, Tandon OP, Saxena AK, Arora S, Bhattacharya N. Autonomic status and pain profile in patients of chronic low back pain and following electro acupuncture therapy: a randomized control trial. *Indian Journal of Physiology & Pharmacology* 2011;**55**(1):25-36.

Shaughnessy 2004 *{published data only}*

Shaughnessy M, Caulfield B. A pilot study to investigate the effect of lumbar stabilisation exercise training on functional ability and quality of life in patients with chronic low back pain. *International Journal of Rehabilitation Research* 2004;**27**(4):297-301.

Shen 2009 *{published data only}*

Shen YC, Shen YX. Effect of eperisone hydrochloride on the paraspinal muscle blood flow in patients with chronic low back pain: a randomized controlled observation. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2009;**13**(7):1293-6.

Sherman 2005 *{published data only}*

Sherman KJ, Cherkin DC, Erro J, Miglioretti DL, Deyo RA. Comparing yoga, exercise, and a self-care book for chronic low back pain: a randomized, controlled trial. *Annals of Internal Medicine* 2005;**12**:849-56.

Sherman 2011 *{published data only}*

Sherman KJ, Cherkin DC, Wellman RD, Cook AJ, Hawkes RJ, Delaney K, et al. A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. *Archives of Internal Medicine* 2011;**171**(22):2019-26.

Shirado 2010 *{published data only}*

Shirado O, Doi T, Akai M, Hoshino Y, Fujino K, Hayashi K, et al. Multicenter randomized controlled trial to evaluate the effect of home-based exercise on patients with chronic low back pain: the Japan low back pain exercise therapy study. *Spine* 2010;**35**(17):E811-9.

Shnayderman 2013 *{published data only}*

Shnayderman I, Katz-Leurer M. An aerobic walking programme versus muscle strengthening programme for chronic low back pain: a randomized controlled trial. *Clinical Rehabilitation* 2013;**27**(3):207-14.

Sjogren 2006a *{published data only}*

Sjogren T, Nissinen KJ, Jarvenpaa SK, Ojanen MT, Vanharanta H, Malkia EA. Effects of a workplace physical exercise intervention on the intensity of low back symptoms in office workers: a cluster randomized controlled cross-over design. *Journal of Back and Musculoskeletal Rehabilitation* 2006;**19**(1):13-24.

Smeets 2008 *{published data only}*

* Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, Van der Heijden GJ, Knottnerus JA. Chronic low back pain: physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain* 2008;**134**(3):263-76.

Smeets RJ, Vlaeyen JWS, Hidding A, Kester ADM, Van der Heijden GJM, Van Geel ACM. Active rehabilitation for chronic low back pain: cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial. *BMC Musculoskeletal Disorders* 2006;**7**:1-16.

Smith 2011 *{published data only}*

Bruce-Low S, Smith D, Burnet S, Fisher J, Bissell G, Webster L. One lumbar extension training session per week is sufficient for strength gains and reductions in pain in patients with chronic low back pain ergonomics. *Ergonomics* 2012;**55**(4):500-7.

* Smith D, Bissell G, Bruce-Low S, Wakefield C. The effect of lumbar extension training with and without pelvic stabilization on lumbar strength and low back pain. *Journal of Back & Musculoskeletal Rehabilitation* 2011;**24**(4):241-9.

Sorensen 2010 *{published data only}*

Sorensen PH, Bendix T, Manniche C, Korsholm L, Lemvigh D, Indahl A. An educational approach based on a non-injury model compared with individual symptom-based physical training in chronic LBP. A pragmatic, randomised trial with a one-year follow-up. *BMC Musculoskeletal Disorders* 2010;**11**:212-25.

Soukup 1999 *{published data only}*

Glomsrod B, Lonn JH, Soukup MG, Bo K, Larsen S. 'Active back school', prophylactic management for low back pain: three-year follow-up of a randomized, controlled trial. *Journal of Rehabilitation Medicine* 2001;**33**:26-30.

Lonn JH, Glomsrod B, Soukup MG, Bo K, Larsen S. Active back school: prophylactic management for low back pain. A randomized controlled, 1-year follow-up study. *Spine* 1999;**24**:865-71.

Soukup M, Lonn J, Glomsrod B, Bo K, Larsen S. Exercises and education as secondary prevention for recurrent low back pain. *Physiotherapy Research International* 2001;**6**(1):27-39.

* Soukup MG, Glomsrod B, Lonn JH, Bo K, Larsen S, Fordyce WE. The effect of a Mensendieck exercise program as secondary prophylaxis for recurrent low back pain: a randomized, controlled trial with 12-month follow-up. *Spine* 1999;**24**(15):1585-92.

Soundararajan 2016 *{published data only}*

Soundararajan LR, Thankappan SM. Efficacy of the multifidus retraining program in computer professionals with chronic low back pain. *Asian Spine Journal* 2016;**10**(3):450-6.

Sousa 2009 *{published data only}*

Sousa KS, Orfale AG, Meireles SM, Leite JR, Natour J. Assessment of a biofeedback program to treat chronic low back pain. *Journal of Musculoskeletal Pain* 2009;**17**(4):369-77.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Steele 2013 *{published data only}*

Steele J, Bruce-Low S, Smith D, Jessop D, Osborne N. A randomized controlled trial of limited range of motion lumbar extension exercise in chronic low back pain. *Spine* 2013;**38**(15):1245-52.

Stephan 2011 *{published data only}*

Stephan A, Goebel S, Schmidtbleicher D. Effects of machine-based strength training in the therapy of chronic back pain. *Deutsche Zeitschrift Fur Sportmedizin* 2011;**62**(3):69-74.

Sung 2013 *{published data only}*

Sung PS. Disability and back muscle fatigability changes following two therapeutic exercise interventions in participants with recurrent low back pain. *Medical Science Monitor* 2013;**19**(1):40-8.

Tekur 2012 *{published data only}*

* Tekur P, Nagarathna R, Chametcha S, Hankey A, Nagendra HR. A comprehensive yoga programs improves pain, anxiety and depression in chronic low back pain patients more than exercise: an RCT. *Complementary Therapies in Medicine* 2012;**20**(3):107-18.

Tekur P, Singphow C, Nagendra HR, Raghuram N. Effect of short-term intensive yoga program on pain, functional disability and spinal flexibility in chronic low back pain: a randomized control study. *Journal of Alternative & Complementary Medicine* 2008;**14**(6):637-44.

Teut 2016 *{published data only}*

Teut M, Knilli J, Daus D, Roll S, Witt CM. Qigong or yoga versus no intervention in older adults with chronic low back pain - a randomized controlled trial. *Journal of Pain* 2016;**17**(7):796-805.

Tilbrook 2011 *{published data only}*

Chuang LH, Soares MO, Tilbrook H, Cox H, Hewitt CE, Aplin J, et al. A pragmatic multicentered randomized controlled trial of yoga for chronic low back pain: economic evaluation. *Spine* 2012;**37**(18):1593-1601.

* Tilbrook HE, Cox H, Hewitt CE, Kangombe AR, Chuang LH, Jayakody S, et al. Yoga for chronic low back pain: a randomized trial. *Annals of Internal Medicine* 2011;**155**(9):569-78.

Torstensen 1998 *{published data only}*

Torstensen TA, Ljunggren AE, Meen HD, Odland E, Mowinckel P, Geijerstam S. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain: a pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. *Spine* 1998;**23**(23):2616-24.

Tritilanunt 2001 *{published data only}*

Tritilanunt T, Wajanavisit W. The efficacy of an aerobic exercise and health education program for treatment of chronic low back pain. *Journal of the Medical Association of Thailand* 2001;**84**(Suppl 2):S528-33.

Turner 1990 *{published data only}*

Turner JA, Clancy S, McQuade KJ, Cardenas DD. Effectiveness of behavioral therapy for chronic low back pain: a component analysis. *Journal of Consulting and Clinical Psychology* 1990;**58**(5):573-9.

Ulger 2017 *{published data only}*

Ulger O. The effect of manual therapy and exercise in patients with chronic low back pain: double blind randomized controlled trial. *Journal of Back and Musculoskeletal Rehabilitation* 2017;**30**(6):1303-9.

Underwood 2004 *{published data only}*

Underwood M. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. *British Medical Journal* 2004;**329**(7479):1377-81.

Unsgaard-Tondel 2010 *{published data only}*

* Unsgaard-Tondel M, Fladmark AM, Salvesen O, Vasseljen O. Motor control exercises, sling exercises, and general exercises for patients with chronic low back pain: a randomized controlled trial with 1-year follow-up. *Physical Therapy* 2010;**90**(10):1426-40.

Vasseljen O, Unsgaard-Tondel M, Westad C, Mork PJ. Effect of core stability exercises on feed-forward activation of deep abdominal muscles in chronic low back pain: a randomized controlled trial. *Spine* 2012;**37**(13):1101-8.

Valenza 2017 *{published data only}*

Valenza M, Rodriguez-Torres J, Cabrera-Martos I, Diaz-Pelegrina A, Aguilar-Ferrandiz M, Castellote-Caballero Y. Results of a Pilates exercise program in patients with chronic non-specific low back pain: a randomized controlled trial. *Clinical Rehabilitation* 2017;**31**(6):753-60.

Van der Roer 2008 *{published data only}*

Van der Roer N, Van Tulder M, Barendse J, Knol D, Van Mechelen W, De Vet H. Intensive group training protocol versus guideline physiotherapy for patients with chronic low back pain: a randomised controlled trial. *European Spine Journal* 2008;**17**(9):1193-200.

Vibe Fersum 2013 *{published data only}*

* Vibe Fersum K, O'Sullivan P, Skouen JS, Smith A, Kvale A. Efficacy of classification-based cognitive functional therapy in patients with non-specific chronic low back pain: a randomized controlled trial. *European Journal of Pain* 2013;**17**(6):916-28.

Vibe Fersum K, Smith A, Kvale A, Skouen JS, O'Sullivan P. Cognitive functional therapy in patients with non-specific chronic low back pain: a randomized controlled trial 3-year follow-up. *European Journal of Pain* 2019;**23**(8):1416-24.

Vincent 2014a *{published data only}*

* Vincent HK, George SZ, Seay AN, Vincent KR, Hurley RW. Resistance exercise, disability, and pain catastrophizing in obese adults with back pain. *Medicine & Science in Sports & Exercise* 2014;**46**(9):1693-701.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Vincent HK, Vincent KR, Seay AN, Conrad BP, Hurley RW, George SZ. Back strength predicts walking improvement in obese, older adults with chronic low back pain. *PM&R: Journal of Injury, Function, & Rehabilitation* 2014;**6**(5):418-26.

Vollenbroek-Hutten 2004 *{published data only}*

Vollenbroek-Hutten MM, Hermens HJ, Wever D, Gorter M, Rinket J, Ijzerman MJ. Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clinical Rehabilitation* 2004;**18**(5):566-79.

Wajswelner 2012 *{published data only}*

Wajswelner H, Metcalf B, Bennell K. Clinical Pilates versus general exercise for chronic low back pain: randomized trial. *Medicine & Science in Sports & Exercise* 2012;**44**(7):1197-205.

Weifen 2013 *{published data only}*

Weifen W, Muheremu A, Chaohui C, Md LW, Lei S. Effectiveness of tai chi practice for non-specific chronic low back pain on retired athletes: a randomized controlled study. *Journal of Musculoskeletal Pain* 2013;**21**(1):37-45.

Weiner 2008 *{published data only}*

Weiner DK, Perera S, Rudy TE, Glick RM, Shenoy S, Delitto A. Efficacy of percutaneous electrical nerve stimulation and therapeutic exercise for older adults with chronic low back pain: a randomized controlled trial. *Pain* 2008;**140**(2):344-57.

Williams 2005 *{published data only}*

Williams KA, Petronis J, Smith D, Goodrich D, Wu J, Ravi N, et al. Effect of Iyengar yoga therapy for chronic low back pain. *Pain* 2005;**115**(1-2):107-17.

Williams 2009 *{published data only}*

Williams K, Abildso C, Steinberg L, Doyle E, Epstein B, Smith D, et al. Evaluation of the effectiveness and efficacy of Iyengar yoga therapy on chronic low back pain. *Spine* 2009;**34**(19):2066-76.

Woods 2008 *{published data only}*

Woods MP, Asmundson GJ. Evaluating the efficacy of graded in vivo exposure for the treatment of fear in patients with chronic back pain: a randomized controlled clinical trial. *Pain* 2008;**136**(3):271-80.

Xueqiang 2012 *{published data only}*

Xueqiang W, Jiejiao Z, Xia B, Jing L. Effect of core stability training on patients with chronic low back pain. *HealthMED* 2012;**6**(3):754-9.

Yelland 2004 *{published data only}*

Yelland MJ, Glasziou PP, Bogduk N, Schluter PJ, McKernon M. Prolotherapy injections, saline injections, and exercises for chronic low-back pain: a randomized trial. *Spine* 2004;**29**(1):9-16.

Yeung 2003 *{published data only}*

Yeung CK, Leung MC, Chow DH. The use of electro-acupuncture in conjunction with exercise for the treatment of chronic low-back pain. *Journal of Alternative & Complementary Medicine* 2003;**9**(4):479-90.

You 2014 *{published data only}*

You JH, Kim SY, Oh DW, Chon SC. The effect of a novel core stabilization technique on managing patients with chronic low back pain: a randomized, controlled, experimenter-blinded study. *Clinical Rehabilitation* 2014;**28**(5):460-9.

Yozbatiran 2004 *{published data only}*

Yozbatiran N, Yildirim Y, Parlak B. Effects of fitness and aquafitness exercises on physical fitness in patients with chronic low back pain. *Pain Clinic* 2004;**16**(1):35-42.

References to studies excluded from this review

Akbari 2008 *{published data only}*

Akbari A, Khorashadizadeh S, Abdi G. The effect of motor control exercise versus general exercise on lumbar local stabilizing muscles thickness: randomized controlled trial of patients with chronic low back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2008;**21**(2):105-12.

Aleksiev 2014 *{published data only}*

Aleksiev AR. Ten-year follow-up of strengthening versus flexibility exercises with or without abdominal bracing in recurrent low back pain. *Spine* 2014;**39**(13):997-1003.

Ali 2013 *{published data only}*

Ali S, Ali SM, Memon KN. Effectiveness of core stabilization exercises versus McKenzie's exercises in chronic lower back pain. *Medical Forum Monthly* 2013;**24**(12):82-5.

Anema 2007 *{published data only}*

Anema JR, Steenstra IA, Bongers PM, De Vet HC, Knol DL, Loisel P, et al. Multidisciplinary rehabilitation for subacute low back pain: graded activity or workplace intervention or both? A randomized controlled trial. *Spine* 2007;**32**(3):291-8.

Bergstrom 2012 *{published data only}*

Bergstrom C, Jensen I, Hagberg J, Busch H, Bergstrom G. Effectiveness of different interventions using a psychosocial subgroup assignment in chronic neck and back pain patients: a 10-year follow-up. *Disability & Rehabilitation* 2012;**34**(2):110-8.

Bhatnagar 2017 *{published data only}*

Bhatnagar G, Sahu M. Comparison of pelvic floor exercises and conventional regimen in patients with chronic low back pain. *Indian Journal of Physiotherapy & Occupational Therapy* 2017;**11**(3):38-42.

Browder 2007 *{published data only}*

Browder DA, Childs JD, Cleland JA, Fritz JM. Effectiveness of an extension-oriented treatment approach in a subgroup of subjects with low back pain: a randomized clinical trial. *Physical Therapy* 2007;**87**(12):1608-18.

Celestini 2005 *{published data only}*

Celestini M, Marchese A, Serenelli A, Graziani G. A randomized controlled trial on the efficacy of physical exercise in patients

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

braced for instability of the lumbar spine. *Europa Medicophysica* 2005;**41**(3):223-31.

Chatzitheodorou 2008 *{published data only}*

Chatzitheodorou D, Mavromoustakos S, Milioti S. The effect of exercise on adrenocortical responsiveness of patients with chronic low back pain, controlled for psychological strain. *Clinical Rehabilitation* 2008;**22**(4):319-28.

Cho 2014 *{published data only}*

Cho HY, Kim EH, Kim J. Effects of the CORE exercise program on pain and active range of motion in patients with chronic low back pain. *Journal of Physical Therapy Science* 2014;**26**(8):1237-40.

Chul 2016 *{published data only}*

Chul KI, Myoung HE, Hwang-Yong KI, Eun-Jeong KI. The effects of forced breathing exercise on the lumbar stabilization in chronic low back pain patients. *Journal of Physical Therapy Science* 2016;**28**(12):3380-3.

Demirel 2008 *{published data only}*

Demirel R, Ucok K, Kavuncu V, Gecici O, Evcik D, Dundar U. Effects of balneotherapy with exercise in patients with low back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2008;**21**(4):263-72.

Diaz-Arribas 2015 *{published data only}*

Diaz-Arribas MJ, Kovacs FM, Royuela A, Fernandez-Serrano M, Gutierrez-Fernandez L, San Martin-Pariente O, et al. Effectiveness of the Godelieve Denys-Struyf (GDS) method in patients with low back pain: a cluster randomized controlled trial. *Physical Therapy* 2015;**95**(3):319-36.

Dillen 2016 *{published data only}*

Dillen LR, Norton BJ, Sahrmann SA, Evanoff BA, Harris-Hayes M, Holtzman GW, et al. Efficacy of classification-specific treatment and adherence on outcomes in people with chronic low back pain. A one-year follow-up, prospective, randomized, controlled clinical trial. *Manual Therapy* 2016;**24**:52-64.

Dundar 2009 *{published data only}*

Dundar U, Solak O, Yigit I, Evcik D, Kavuncu V. Clinical effectiveness of aquatic exercise to treat chronic low back pain: a randomized controlled trial. *Spine* 2009;**34**(14):1436-40.

Ferreira 2016 *{published data only}*

Ferreira G, Stieven F, Araujo F, Wiebusch M, Rosa C, Plentz R, et al. Neurodynamic treatment did not improve pain and disability at two weeks in patients with chronic nerve-related leg pain: a randomised trial. *Journal of Physiotherapy* 2016;**62**(4):197-202.

Ford 2016 *{published data only}*

Ford JJ, Hahne AJ, Surkitt LD, Chan AY, Richards MC, Slater SL, et al. Individualised physiotherapy as an adjunct to guideline-based advice for low back disorders in primary care: a randomised controlled trial. *British Journal of Sports Medicine* 2016;**50**(4):237-45.

Froholdt 2012 *{published data only}*

Froholdt A, Reikeraas O, Holm I, Keller A, Brox JI. No difference in 9-year outcome in CLBP patients randomized to lumbar fusion versus cognitive intervention and exercises. *European Spine Journal* 2012;**21**(12):2531-8.

Galantino 2004 *{published data only}*

Galantino ML, Bzdewka TM, Eissler-Russo JL, Holbrook ML, Mogck EP, Geigle P, et al. The impact of modified hatha yoga on chronic low back pain: a pilot study. *Alternative Therapies In Health And Medicine* 2004;**10**(2):56-9.

Gao 2006 *{published data only}*

Gao MX, Liu XY, Zhen P, Wang P, Wang XM, Xue Y. Evaluation of the efficacy of exercise program in the remission of low back pain in manage tank soldiers. *Chinese Journal of Clinical Rehabilitation* 2006;**10**(8):32-4.

Haufe 2017 *{published data only}*

Haufe S, Wiechmann K, Stein L, Kuck M, Smith A, Meineke S, et al. Low-dose, non-supervised, health insurance initiated exercise for the treatment and prevention of chronic low back pain in employees. Results from a randomized controlled trial. *PLOS One* 2017;**12**(6):e0178585.

Hosseinifar 2013 *{published data only}*

Hosseinifar M, Akbari M, Behtash H, Amiri M, Sarrafzadeh J. The effects of stabilization and McKenzie exercises on transverse abdominis and multifidus muscle thickness, pain, and disability: a randomized controlled trial in nonspecific chronic low back pain. *Journal of Physical Therapy Science* 2013;**25**(12):1541-5.

Jeong 2015 *{published data only}*

Jeong UC, Sim JH, Kim CY, Hwang-Bo G, Nam CW. The effects of gluteus muscle strengthening exercise and lumbar stabilization exercise on lumbar muscle strength and balance in chronic low back pain patients. *Journal of Physical Therapy Science* 2015;**27**(12):3813-6.

Jeong 2016 *{published data only}*

Jeong UC, Kim CY, Park YH, Hwang-Bo G, Nam CW. The effects of self-mobilization techniques for the sciatic nerves on physical functions and health of low back pain patients with lower limb radiating pain. *Journal of Physical Therapy Science* 2016;**28**(1):46-50.

Ju 2015 *{published data only}*

Ju T, Choi W, Yang Y, Lee S. Effects of hip mobilization on pain and function for chronic low back pain individuals with limited range of hip joint motion. *Indian Journal of Science and Technology* 2015;**8**(26):345–8.

Ju-Hyun 2017 *{published data only}*

Ju-Hyun L, Tae-Ho K. The treatment effect of hamstring stretching and nerve mobilization for patients with radicular lower back pain. *Journal of Physical Therapy Science* 2017;**29**(9):1578-82.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kachanathu 2012 *{published data only}*

Kachanathu SJ, Zakaria AR, Sahni A, Jaiswal P. Chronic low back pain in fast bowlers a comparative study of core spinal stabilization and conventional exercises. *Journal of Physical Therapy Science* 2012;**24**(9):821-5.

Kaeding 2017 *{published data only}*

Kaeding T, Karch A, Schwarz R, Flor T, Wittke T, Kuck M, et al. Whole-body vibration training as a workplace-based sports activity for employees with chronic low-back pain. *Scandinavian Journal of Medicine & Science in Sports* 2017;**27**(12):2027-39.

Kamali 2014 *{published data only}*

Kamali F, Panahi F, Ebrahimi S, Abbasi L. Comparison between massage and routine physical therapy in women with sub acute and chronic nonspecific low back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2014;**27**(4):475.

Kamali 2018 *{published data only}*

Kamali F, Zamanlou M, Ghanbari A, Alipour A, Bervis S. Comparison of manipulation and stabilization exercises in patients with sacroiliac joint dysfunction patients: a randomized clinical trial. *Journal of Bodywork and Movement Therapies* 2018;**23**(1):177-82.

Kang 2016 *{published data only}*

Jeong-Il K, Dae-Keun J, Hyun C. Effect of exhalation exercise on trunk muscle activity and Oswestry Disability Index of patients with chronic low back pain. *Journal of Physical Therapy Science* 2016;**28**(6):1738-42.

Kim 2013 *{published data only}*

Kim JH, Kim YE, Bae SH, Kim KY. The effect of the neurac sling exercise on postural balance adjustment and muscular response patterns in chronic low back pain patients. *Journal of Physical Therapy Science* 2013;**25**(8):1015-9.

Kim 2014 *{published data only}*

Kim SS, Min WK, Kim JH, Lee BH. The effects of VR-based Wii fit yoga on physical function in middle-aged female LPB patients. *Journal of Physical Therapy Science* 2014;**26**(4):549-52.

Kim 2015b *{published data only}*

Kim JY, Choi WJ, Seo TH. Effect of proprioceptive neuromuscular facilitation integration pattern and swiss ball training on pain and balance in elderly patients with chronic back pain. *Journal of Physical Therapy Science* 2015;**27**(10):3237-40.

Ko 2018 *{published data only}*

Ko KJ, Ha GC, Yook YS, Kang SJ. Effects of 12-week lumbar stabilization exercise and sling exercise on lumbosacral region angle, lumbar muscle strength, and pain scale of patients with chronic low back pain. *Journal of Physical Therapy Science* 2018;**30**(1):18-22.

Konstantinou 2007 *{published data only}*

Konstantinou K, Foster N, Rushton A, Baxter D, Wright C, Breen A. Flexion mobilizations with movement techniques: the immediate effects on range of movement and pain in subjects with low back pain. *Journal of Manipulative & Physiological Therapeutics* 2007;**30**(3):178-85.

Kumar 2011 *{published data only}*

Kumar SP. Efficacy of segmental stabilization exercise for lumbar segmental instability in patients with mechanical low back pain: a randomized placebo controlled crossover study. *North American Journal of Medical Sciences* 2011;**3**(10):461.

Lee 2014a *{published data only}*

Lee C, Hwangbo K, Lee IS. The effects of combination patterns of proprioceptive neuromuscular facilitation and ball exercise on pain and muscle activity of chronic low back pain patients. *Journal of Physical Therapy Science* 2014;**26**(1):93-6.

Lee 2014b *{published data only}*

Lee CW, Hyun J, Kim SG. Influence of Pilates mat and apparatus exercises on pain and balance of business women with chronic low back pain. *Journal of Physical Therapy Science* 2014;**26**(4):475-7.

Lee 2016b *{published data only}*

Lee SB, Cho WJ. The effect of sling exercise on sagittal lumbosacral angle and intervertebral disc area of chronic low back pain patients. *Journal of Exercise Rehabilitation* 2016;**12**(5):471-5.

Lee 2017 *{published data only}*

Lee JH. Effects of the centrifugal contraction exercise of the rectus abdominis accompanied with the isometric exercise of the transversus abdominis on pain and balance of lower back pain patients. *Journal of Physical Therapy Science* 2017;**29**(8):1405-8.

Lee 2018 *{published data only}*

Lee DK, Kim YN, Chi-Bok P, Mi-Sook P. The effect of actively induced vibration using shoulder joint on pain and dysfunction in patients with low back pain. *Journal of Physical Therapy Science* 2018;**30**(1):23-6.

Maciaszek 2016 *{published data only}*

Maciaszek J, Skrypnik D, Ratajczak M, Stemplewski R, Osinski W, Bogdanski P, et al. Two aerobic exercise programs in management of back pain among middle-aged obese women: a randomized controlled study. *Human Movement* 2016;**17**(2):72-9.

Majiwala 2017 *{published data only}*

Majiwala BA, Warude TA, Pawar A. Effect of isometric and isotonic exercise training on core muscle in patients with non-specific low back pain. *Asian Journal of Pharmaceutical and Clinical Research* 2017;**10**(6):308-11.

Masse-Alarie 2017 *{published data only}*

Masse-Alarie H, Beaulieu LD, Preuss R, Schneider C. Repetitive peripheral magnetic neurostimulation of multifidus muscles combined with motor training influences spine motor control and chronic low back pain. *Clinical Neurophysiology* 2017;**128**(3):442-53.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mirovsky 2006 *{published data only}*

Mirovsky Y, Grober A, Blankstein A, Stabholz L. The effect of ambulatory lumbar traction combined with treadmill on patients with chronic low back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2006;**19**(2-3):73-8.

Moon 2017 *{published data only}*

Moon JH, Jung JH, Won YS, Cho HY. Immediate effects of Graston Technique on hamstring muscle extensibility and pain intensity in patients with nonspecific low back pain. *Journal of Physical Therapy Science* 2017;**29**(2):224-7.

Nambi 2014 *{published data only}*

Nambi GS, Inbasekaran D, Khuman R, Devi S, Shanmugananth, Jagannathan K. Changes in pain intensity and health related quality of life with Iyengar yoga in nonspecific chronic low back pain: a randomized controlled study. *International Journal of Yoga* 2014;**7**(1):48-53.

Nazzal 2013 *{published data only}*

Nazzal ME, Saadah MA, Saadah LM, Al-Omari MA, Al-Oudat ZA, Nazzal MS, et al. Management options of chronic low back pain. A randomized blinded clinical trial. *Neurosciences* 2013;**18**(2):152-9.

Park 2013 *{published data only}*

Park JH, Lee SH, Ko DS. The effects of the Nintendo Wii exercise program on chronic work-related low back pain in industrial workers. *Journal of Physical Therapy Science* 2013;**25**(8):985-8.

Park 2019 *{published data only}*

Park SH, Lee MM. Effects of a progressive stabilization exercise program using respiratory resistance for patients with lumbar instability: a randomized controlled trial. *Medical Science Monitor* 2019;**25**:1740-8.

Paungmali 2018 *{published data only}*

Paungmali A, Joseph LH, Punturee K, Sitilertpisan P, Pirunsan U, Uthaikhup S. Immediate effects of core stabilization exercise on β-endorphin and cortisol levels among patients with chronic nonspecific low back pain: a randomized crossover design. *Journal of Manipulative & Physiological Therapeutics* 2018;**41**(3):181-8.

Pengel 2007 *{published data only}*

Pengel LH, Refshauge KM, Maher CG, Nicholas MK, Herbert RD, McNair P. Physiotherapist-directed exercise, advice, or both for subacute low back pain: a randomized trial. *Annals of Internal Medicine* 2007;**146**(11):787-96.

Rasmussen 2015 *{published data only}*

Rasmussen CDN, Holtermann A, Bay H, Sogaard K, Jorgensen MB. A multifaceted workplace intervention for low back pain in nurses' aides: a pragmatic stepped wedge cluster randomised controlled trial. *Pain* 2015;**156**(9):1786-94.

Rhon 2018 *{published data only}*

Rhon DA, Miller RB, Fritz JM. Effectiveness and downstream healthcare utilization for patients that received early physical therapy versus usual care for low back pain: a randomized clinical trial. *Spine* 2018;**43**(19):1313-21.

Schaller 2017 *{published data only}*

Schaller A, Petrowski K, Pfoertner TK, Froboese I. Effectiveness of a theory-based multicomponent intervention (Movement Coaching) on the promotion of total and domain-specific physical activity: a randomised controlled trial in low back pain patients. *BMC Musculoskeletal Disorders* 2017;**18**:1-14.

Schiltenwolf 2006 *{published data only}*

Schiltenwolf M, Buchner M, Heindl B, Von Reumont J, Muller A, Eich W. Comparison of a biopsychosocial therapy (BT) with a conventional biomedical therapy (MT) of subacute low back pain in the first episode of sick leave: a randomized controlled trial. *European Spine Journal* 2006;**15**(7):1083-92.

Sjogren 2006b *{published data only}*

Sjogren T, Nissinen KJ, Jarvenpaa SK, Ojanen MT, Vanharanta H, Malkia EA. Effects of a workplace physical exercise intervention on the intensity of low back symptoms in office workers: a cluster randomized controlled cross-over design. *Journal of Back and Musculoskeletal Rehabilitation* 2006;**19**(1):13-24.

Suni 2006 *{published data only}*

Suni J, Rinne M, Natri A, Statistisian MP, Parkkari J, Alaranta H. Control of the lumbar neutral zone decreases low back pain and improves self-evaluated work ability: a 12-month randomized controlled study. *Spine* 2006;**31**(18):E611-20.

Suni 2017 *{published data only}*

Suni JH, Rinne M, Tokola K, Manttari A, Vasankari T. Effectiveness of a standardised exercise programme for recurrent neck and low back pain: a multicentre, randomised, two-arm, parallel group trial across 34 fitness clubs in Finland. *BMJ Open Sport & Exercise Medicine* 2017;**3**(1):e000233.

Tang 2016 *{published data only}*

Tang S, Qian X, Zhang Y, Liu Y. Treating low back pain resulted from lumbar degenerative instability using Chinese Tuina combined with core stability exercises: a randomized controlled trial. *Complementary Therapies in Medicine* 2016;**25**:45-50.

Telles 2016 *{published data only}*

Telles S, Bhardwaj AK, Gupta RK, Sharma SK, Monro R, Balkrishna A. A randomized controlled trial to assess pain and magnetic resonance imaging-based (MRI-based) structural spine changes in low back pain patients after yoga practice. *Medical Science Monitor* 2016;**22**:3228-47.

Tsauo 2009 *{published data only}*

Tsauo JY, Chen WH, Liang HW, Jang Y. The effectiveness of a functional training programme for patients with chronic low back pain: a pilot study. *Disability and Rehabilitation* 2009;**31**(13):1100-6.

Wand 2004 *{published data only}*

Wand BM, Bird C, McAuley JH, Dore CJ, MacDowell M, De Souza LH. Early intervention for the management of acute low back pain: a single-blind randomized controlled trial of biopsychosocial education, manual therapy, and exercise. *Spine* 2004;**29**(21):2350-6.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Waqqar 2016 *{published data only}*

Waqqar S, Shakil-Ur-Rehman S, Ahmad S. McKenzie treatment versus mulligan sustained natural apophyseal glides for chronic mechanical low back pain. *Pakistan Journal of Medical Sciences* 2016;**32**(2):476-9.

Whitfill 2010 *{published data only}*

Whitfill T, Haggard R, Bierner SM, Pransky G, Hassett RG, Gatchel RJ. Early intervention options for acute low back pain patients: a randomized clinical trial with one-year follow-up outcomes. *Journal of Occupational Rehabilitation* 2010;**20**(2):256-63.

Woo 2016 *{published data only}*

Woo SD, Kim TH. The effects of lumbar stabilization exercise with thoracic extension exercise on lumbosacral alignment and the low back pain disability index in patients with chronic low back pain. *Journal of Physical Therapy Science* 2016;**28**(2):680-4.

Wright 2005 *{published data only}*

Wright A, Lloyd-Davies A, Williams S, Ellis R, Strike P. Individual active treatment combined with group exercise for acute and subacute low back pain. *Spine* 2005;**30**(11):1235-41.

Wu 2004 *{published data only}*

Wu JX, Wang B. Early intervention of aerobic exercise to rehabilitation of non-specific low back pain. *Chinese Journal of Clinical Rehabilitation* 2004;**8**(26):5718-20.

Yilmaz Yelvar 2017 *{published data only}*

Yilmaz Yelvar GD, Cirak Y, Dalkilinc M, Parlak Demir Y, Guner Z, Boydak A. Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial. *European Spine Journal* 2017;**26**(2):538-45.

Yoo 2012 *{published data only}*

Yoo YD, Lee YS. The effect of core stabilization exercises using a sling on pain and muscle strength of patients with chronic low back pain. *Journal of Physical Therapy Science* 2012;**24**(8):671-4.

Yoon 2013 *{published data only}*

Yoon KS, Park SD. The effects of ankle mobilization and active stretching on the difference of weight-bearing distribution, low back pain and flexibility in pronated-foots subjects. *Journal of Exercise Rehabilitation* 2013;**9**(2):292.

Yu 2016 *{published data only}*

Yu SH, Sim YH, Kim MH, Bang JH, Son KH, Kim JW, et al. The effect of abdominal drawing-in exercise and myofascial release on pain, flexibility, and balance of elderly females. *Journal of Physical Therapy Science* 2016;**28**(10):2812-5.

Zahin 2018 *{published data only}*

Zahin N, Karahan AY, Albayrak Z. Effectiveness of physical therapy and exercise on pain and functional status in patients with chronic low back pain: a randomized-controlled trial. *Turkish Journal of Physical Medicine & Rehabilitation* 2018;**64**(1):52-8.

References to studies awaiting assessment

Abadi 2019 *{published data only}*

Abadi FH, Sankaravel M, Zainuddin FF, Elumalai G, Razli AI. The effect of aquatic exercise program on low-back pain disability in obese women. *Journal of Exercise Rehabilitation* 2019;**15**(6):855-60.

Ahmadi 2020 *{published data only}*

Ahmadi H, Adib H, Selk-Ghaffari M, Shafizad M, Moradi S, Madani Z, et al. Comparison of the effects of the Feldenkrais method versus core stability exercise in the management of chronic low back pain: a randomised control trial. *Clinical Rehabilitation* 2020;**34**(12):1449-57.

Ahmadizadeh 2019 *{published data only}*

Ahmadizadeh Z, Ehsani F, Samaei SA, Mirmohamadkhani M. The effect of stabilization exercises along with self-care training on transverse abdominal activity, pain, and disability in mothers with low back pain having children with CP: a RCT. *American Journal of Physical Medicine & Rehabilitation* 2019;**99**(2):156-60.

Ak 2016 *{published data only}*

Ak A, Sra A, Cs O. Comparative efficacy of core stabilization exercise and Pilates exercise on patients with nonspecific chronic low back pain. *Romanian Journal of Physical Therapy* 2016;**21**(38):14-22.

Akodu 2020 *{published data only}*

Akodu AK, Odunfa KS. Effects of core stabilization and McKenzie back extension exercises on pain, disability and insomnia in patients with non-specific chronic low back pain. *Sports Medicine Journal* 2020;**16**(1):3190-6.

Alfuth 2016 *{published data only}*

Alfuth M, Cornely D. Chronic low back pain: comparison of mobilization and core stability exercises. *Orthopade* 2016;**45**(7):579-90.

Alikhajeh 2020 *{published data only}*

Alikhajeh Y, Barabadi E, Mohammad Rahimi GR. A comparison of 6 weeks of aquatic exercise and kinesio taping in patients with chronic nonspecific low back pain. *Journal of Sport Rehabilitation* 2021;**30**:37-42.

Almhdawi 2020 *{published data only}*

Almhdawi KA, Obeidat DS, Kanaan SF, Oteir AO, Mansour ZM, Alrabbaei H. Efficacy of an innovative smartphone application for office workers with chronic non-specific low back pain: a pilot randomized controlled trial. *Clinical Rehabilitation* 2020;**34**(10):1282-91.

Altinbilek 2020 *{published data only}*

Altinbilek T, Murat S. A comparison of application frequency of physical therapy modalities in patients with chronic mechanical low back pain. *Turkish Journal of Physical Medicine and Rehabilitation* 2020;**66**(2):201-9.

Amorim 2019 *{published data only}*

Amorim AB, Pappas E, Simic M, Ferreira ML, Jennings M, Tiedemann A, et al. Integrating mobile-health, health coaching,

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

and physical activity to reduce the burden of chronic low back pain trial (IMPACT): a pilot randomised controlled trial. *BMC Musculoskeletal Disorders* 2019;**20**(1):71-85.

Ansari 2020 *{published data only}*

Ansari S, Elmieh A, Alipour A. The effect of aquatic exercise on functional disability, flexibility and function of trunk muscles in postmenopausal women with chronic non-specific low back pain: randomized controlled trial. *Science and Sports* 2021;**36**(3):e103-10.

Areeudomwong 2018 *{published data only}*

Areeudomwong P, Buttagat V. Proprioceptive neuromuscular facilitation training improves pain-related and balance outcomes in working-age patients with chronic low back pain: a randomized controlled trial. *Brazilian Journal of Physical Therapy* 2018;**23**(5):428-36.

Areeudomwong 2019 *{published data only}*

Areeudomwong P, Buttagat V. Comparison of core stabilisation exercise and proprioceptive neuromuscular facilitation training on pain-related and neuromuscular response outcomes for chronic low back pain: a randomised controlled trial. *Malaysian Journal of Medical Sciences* 2019;**26**(6):77-89.

Barbosa 2013 *{published data only}*

Barbosa Silva DK, Murata E, Freitas CD. Comparação do método Pilates com e sem a estabilização segmentar na dor lombar crónica. *Revista Terapia Manual* 2013;**11**(51):90-4.

Barni 2018 *{published data only}*

Barni L, Calabretta L, Lepori L, Pasquetti P, Gulisano M, Freddolini M. Does an aerobic exercise improve outcomes in older sedentary nonspecific low back pain subjects? A randomized controlled study. *Topics in Geriatric Rehabilitation* 2018;**34**(2):88-94.

Barradas 2015 *{published data only}*

Barradas LPF, Da Silva LFBP, De Sousa RC, Matos LKBL, Carvalho AFM. The effects of the exercises of segmental stabilization in low back pain. *Manual Therapy, Posturology & Rehabilitation Journal* 2015;**13**:1-6.

Batool 2019 *{published data only}*

Batool F, Afzal W, Ahmad A, Gilani SA. Effectiveness of core stability with and without neck extension on pain intensity and functional disability in patients with chronic low back pain. *Rawal Medical Journal* 2019;**44**(4):797-800.

Bello 2018 *{published data only}*

Bello B, Adeniyi AF. Effects of lumbar stabilisation and treadmill exercise on function in patients with chronic mechanical low back pain. *International Journal of Therapy & Rehabilitation* 2018;**25**(9):493-9.

Brodsky 2019 *{published data only}*

Brodsky M, Hansen A, Bjerke W. Randomized pilot trial for a community-based group stretching exercise program for chronic low back pain. *Global Advances in Health & Medicine* 2019;**8**:1-7.

Bruehl 2020 *{published data only}*

Bruehl S, Burns JW, Koltyn K, Gupta R, Buvanendran A, Edwards D, et al. Are endogenous opioid mechanisms involved in the effects of aerobic exercise training on chronic low back pain? A randomized controlled trial. *Pain* 2020;**161**(12):2887-97.

Buttagat 2019 *{published data only}*

Buttagat V, Techakhot P, Wiriya W, Mueller M, Areeudomwong P. Effectiveness of traditional Thai self-massage combined with stretching exercises for the treatment of patients with chronic non-specific low back pain: a single-blinded randomized controlled trial. *Journal of Bodywork and Movement Therapies* 2019;**24**(1):19-24.

Calatayud 2020 *{published data only}*

Calatayud J, Guzman-Gonzalez B, Andersen LL, Cruz-Montecinos C, Morell MT, Roldan R, et al. Effectiveness of a group-based progressive strength training in primary care to improve the recurrence of low back pain exacerbations and function: a randomised trial. *International Journal of Environmental Research and Public Health* 2020;**17**(22):8326-40.

Canaway 2018 *{published data only}*

Canaway A, Pincus T, Underwood M, Shapiro Y, Chodick G, Ben-Ami N. Is an enhanced behaviour change intervention cost-effective compared with physiotherapy for patients with chronic low back pain? Results from a multicentre trial in Israel. *British Medical Journal Open* 2018;**8**(4):1-8.

Cavalcanti 2020 *{published data only}*

Cavalcanti IF, Antonino GB, Do Monte-Silva KK, Guerino MR, Ferreira APDL, De Araújo MDGR. Global postural re-education in non-specific neck and low back pain treatment: a pilot study. *Journal of Back and Musculoskeletal Rehabilitation* 2020;**33**(5):823-8.

Chen 2018 *{published data only}*

Chen J, Wang C. Core stability training improves back muscle endurance and increases peak values of isokinetic strength of back muscles. *Chinese Journal of Tissue Engineering Research* 2018;**22**(36):5797-802.

Chhabra 2018 *{published data only}*

Chhabra HS, Sharma S, Verma S. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. *European Spine Journal* 2018;**27**(11):2862-74.

Cimarras-Otal 2020 *{published data only}*

Cimarras-Otal C, Marcen-Cinca N, Rabal-Pelay J, Lacarcel-Tejero B, Alcazar-Crevillen A, Villalba-Ruete JA, et al. Adapted exercises versus general exercise recommendations on chronic low back pain in industrial workers: a randomized control pilot study. *Work* 2020;**67**(3):733-40.

Cox 2010 *{published data only}*

Cox H, Tilbrook H, Aplin J, Semlyen A, Torgerson D, Trewhela A, et al. A randomised controlled trial of yoga for the treatment of chronic low back pain: results of a pilot study. *Complementary Therapies in Clinical Practice* 2010;**16**(4):187-93.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Das 2019 *{published data only}*

Das H, Jayaseelan V, Manikandanesan S, Sharma D, Rehman T. Effectiveness of multipurpose health-worker-led exercise therapy on pain reduction among patients with chronic nonspecific low backache in primary health-care setting: a randomized control trial. *Journal of Family Medicine & Primary Care* 2019;**8**(1):199-202.

Da Silva 2014 *{published data only}*

Da Silva TM, Silva NN, E Souza Rocha SH, Marques de Oliveira D, Karina Monte-Silva K, Silva Tenório A, et al. Back school program for back pain: education or physical exercise? *Conscientiae Saude* 2014;**13**(4):506-15.

Daulat 2014 *{published data only}*

Daulat A, Goodlad E. A phase II pilot study comparing a home total body strengthening programme plus manual therapy with a standard physiotherapy exercise regimen plus manual therapy in the management of chronic low back pain. *International Musculoskeletal Medicine* 2014;**36**(3):87-95.

Daulat 2016 *{published data only}*

Daulat A. A pragmatic randomized controlled trial to compare a novel group physiotherapy programme with a standard group exercise programme for managing chronic low back pain in primary care. *International Musculoskeletal Medicine* 2016;**38**(3/4):97-108.

De Giogrio 2018 *{published data only}*

De Giorgio A, Padulo J, Kuvazizi G. Effectiveness of yoga combined with back school program on anxiety, kinesiophobia and pain in people with non-specific chronic low back pain: a prospective randomized trial. *Muscles, Ligaments & Tendons Journal* 2018;**8**(1):104-12.

Delicia 2019 *{published data only}*

Delicia Roshini P, Antony Leo Aseer P. Motor control training in chronic low back pain. *Journal of Clinical and Diagnostic Research* 2019;**13**(4):YC01-5.

Del Pozo 2011 *{published data only}*

Del Pozo-Cruz B, Hernandez Mocholi MA, Adsuar JC, Parraca JA, Muro I, Gusi N. Effects of whole body vibration therapy on main outcome measures for chronic non-specific low back pain: a single-blind randomized controlled trial. *Journal of Rehabilitation Medicine* 2011;**43**(8):689-94.

Demirel 2019 *{published data only}*

Demirel A, Oz M, Ozel YA, Cetin H, Ulger O. Stabilization exercise versus yoga exercise in non-specific low back pain: pain, disability, quality of life, performance: a randomized controlled trial. *Complementary Therapies in Clinical Practice* 2019;**35**:102-8.

Dimer 2019 *{published data only}*

Dimer LR, Santos SM, Steffen EA, Matos SL, Daitx RB, Döhnert MB. Neuromuscular electrical stimulation associated with core stability exercises in nonspecific postural low back pain: a randomized clinical trial. *Muscles, Ligaments & Tendons Journal* 2019;**9**(3):446-56.

Dineshkumar 2015 *{published data only}*

Dineshkumar SK, Dibyendunarayan B, Ramalingam AT. Effect of abdominal drawing-in maneuver along with resisted ankle dorsi-flexion to activate transverse abdominis muscle in chronic non-specific low back pain. *Romanian Journal of Physical Therapy* 2015;**21**(36):40-7.

Ehsani 2019 *{published data only}*

Ehsani F, Hedayati R, Bagheri R, Jaberzadeh S. The effects of stabilization exercise on the thickness of lateral abdominal muscles during standing tasks in women with chronic low back pain: a randomized triple-blinded clinical trial study. *Journal of Sport Rehabilitation* 2019;**29**(7):942-51.

Fang 2015 *{published data only}*

Fang L, Yan J. The effect of Wuqinxi Exercise on mechanics characteristic of abdominal and back muscles and pain in patients with chronic nonspecific low back pain. *Shanghai Journal of Traditional Chinese Medicine* 2015;**9**:49-53.

Fatoye 2020 *{published data only}*

Fatoye F, Gebrye T, Fatoye C, Mbada CE, Olaoye MI, Odole AC, et al. The clinical and cost-effectiveness of telerehabilitation for people with nonspecific chronic low back pain: randomized controlled trial. *Journal of Medical Iinternet Research mHealth and uHealth* 2020;**8**(6):e15375.

Ferreira 2017 *{published data only}*

Ferreira M, Melo L, Cabral K, Nascimento L, Guerino M, Araujo M. Maitland in chronic lumbar pain of young adults improves pain and functionality. *Manual Therapy, Posturology & Rehabilitation Journal* 2017;**15**:1-7.

França 2019 *{published data only}*

França FJR, Burke TN, Magalhães MO, Comachio J, Carvalho Silva APM, Marques AP, et al. Motor control training compared with transcutaneous electrical nerve stimulation in patients with disc herniation with associated radiculopathy: a randomized controlled trial. *American Journal of Physical Medicine & Rehabilitation* 2019;**98**(3):207-14.

Galan-Martin 2020 *{published data only}*

Galan-Martin MA, Montero-Cuadrado F, Lluch-Girbes E, Coca-Lopez MC, Mayo-Iscar A, Cuesta-Vargas A. Pain neuroscience education and physical therapeutic exercise for patients with chronic spinal pain in Spanish physiotherapy primary care: a pragmatic randomized controlled trial. *Journal of Clinical Medicine* 2020;**9**(4):1201-23.

Gao 2018 *{published data only}*

Gao F, Yuan S, Zheng S. Effect observation of Baduanjin training on chronic nonspecific low back pain. *Rehabilitation Medicine* 2018;**04**:1-6.

Gardner 2019 *{published data only}*

Gardner T, Refshauge K, McAuley J, Hubscher M, Goodall S, Smith L. Combined education and patient-led goal setting intervention reduced chronic low back pain disability and intensity at 12 months: a randomised controlled trial. *British Journal of Sports Medicine* 2019;**53**(22):1424-31.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ghasemi 2020 *{published data only}*

Ghasemi C, Amiri A, Sarrafzadeh J, Dadgoo M, Maroufi N. Comparison of the effects of craniosacral therapy, muscle energy technique, and sensorimotor training on non-specific chronic low back pain. *Anaesthesia, Pain and Intensive Care* 2020;**24**(5):532-43.

Gholami 2019 *{published data only}*

Gholami Borujeni B, Yalfani A. Reduction of postural sway in athletes with chronic low back pain through eight weeks of inspiratory muscle training: a randomized controlled trial. *Clinical Biomechanics* 2019;**69**:215-20.

Gohil 2018 *{published data only}*

Gohil D, Samuel R. The efficacy of a lumbar strengthening program in lumbar spine derangement syndrome 1. *International Journal of Pharma and Bio Sciences* 2018;**9**(1):B232-7.

Groessl 2020 *{published data only}*

Groessl EJ, Liu L, Schmalzl L, Chang DG, McCarthy A, Chun WI, et al. Secondary outcomes from a randomized controlled trial of yoga for veterans with chronic low-back pain. *International Journal of Yoga Therapy* 2020;**30**(1):69-76.

Gupta 2019 *{published data only}*

Gupta P, Mohanty PP, Pattnaik M. The effectiveness of aerobic exercise program for improving functional performance and quality of life in chronic low back pain. *Indian Journal of Physiotherapy & Occupational Therapy* 2019;**13**(2):155-60.

Harper 2019 *{published data only}*

Harper B, Steinbeck L, Aron A. Fascial manipulation vs. standard physical therapy practice for low back pain diagnoses: a pragmatic study. *Journal of Bodywork & Movement Therapies* 2019;**23**(1):115-21.

Heidari 2018 *{published data only}*

Heidari RS, Sahebozamani M, Karimi AF. Comparison of the effects of 8 weeks of core stability exercise on ball and sling exercise on the quality of life and pain in the female with non-specific chronic low back pain (NSLBP). *Journal of Zanjan University of Medical Sciences and Health Services* 2018;**26**(117):44-56.

Hosseinifar 2017 *{published data only}*

Hosseinifar M, Ghiasi F, Akbari A, Ghorbani M. The effect of stabilization exercises on lumbar lordosis in patients with low back pain. *Annals of Tropical Medicine & Public Health* 2017;**10**(6):1779-84.

Ibrahim 2018 *{published data only}*

Ibrahim AA, Akindele MO, Ganiyu SO. Motor control exercise and patient education program for low resource rural community dwelling adults with chronic low back pain: a pilot randomized clinical trial. *Journal of Exercise Rehabilitation* 2018;**14**(5):851-63.

Jahantiqh 2018 *{published data only}*

Jahantiqh F, Abdollahimohammad A, Firouzkouhi M, Ebrahiminejadm V. Effects of reiki versus physiotherapy on relieving lower back pain and improving activities daily living of patients with intervertebral disc hernia. *Journal of Evidence-based Integrative Medicine* 2018;**23**:1-6.

Jinnouchi 2020 *{published data only}*

Jinnouchi H, Matsudaira K, Kitamura A, Kakihana H, Oka H, Hayama-Terada M, et al. Effects of brief self-exercise education on the management of chronic low back pain: a community-based, randomized, parallel-group pragmatic trial. *Modern Rheumatology* 2021;**31**(4):890-8.

Joseph 2018 *{published data only}*

Joseph LH, Hancharoenkul B, Sitilertpisan P, Pirunsan U, Paungmali A. Comparison of effects between core stability training and sports massage therapy among elite weightlifters with chronic non-specific low back pain: a randomized cross-over study. *Asian Journal of Sports Medicine* 2018;**9**(1):1-8.

Kamali 2019 *{published data only}*

Kamali F, Zamanlou M, Ghanbari A, Alipour A, Bervis S. Comparison of manipulation and stabilization exercises in patients with sacroiliac joint dysfunction patients: a randomized clinical trial. *Journal of Bodywork & Movement Therapies* 2019;**23**(1):177-82.

Kang 2018 *{published data only}*

Kang TW, Lee JH, Park DH, Cynn HS. Effect of 6-week lumbar stabilization exercise performed on stable versus unstable surfaces in automobile assembly workers with mechanical chronic low back pain. *Work* 2018;**60**(3):445-54.

Karimzadeh 2016 *{published data only}*

Karimzadeh F, Letafatkar A, Ghasemi G. The effect of 8 weeks core stabilization exercises on pain and functional disability induced by low back pain in the mothers of children with cerebral palsy. *Scientific Journal of Kurdistan University of Medical Sciences* 2016;**21**(3):34-44.

Kell 2011 *{published data only}*

Kell RT, Risi AD, Barden JM. The response of persons with chronic nonspecific low back pain to three different volumes of periodized musculoskeletal rehabilitation. *Journal of Strength & Conditioning Research* 2011;**25**(4):1052-64.

Khandhar 2020 *{published data only}*

Khandhar RY, Sathya P, Paul J. Comparative effect of trunk balance exercise over conventional back care exercise in patients with chronic mechanical low back pain. *Indian Journal of Public Health Research and Development* 2020;**11**(6):786-91.

Kim 2018b *{published data only}*

Kim H, Kwon BS, Park JW, Lee A, Nam CW, Park T, et al. Effect of whole body horizontal vibration exercise in chronic low back pain patients: vertical versus horizontal vibration exercise. *Annals of Rehabilitation Medicine* 2018;**42**(6):804-13.

Kim 2019 *{published data only}*

Kim T, Lee J, Oh S, Kim S, Yoon BC. Effectiveness of simulated horseback riding for chronic low back pain patients: a randomized controlled trial. *Journal of Sport Rehabilitation* 2019;**29**(2):179-85.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kim 2020 *{published data only}*

Kim B, Yim J. Core stability and hip exercises improve physical function and activity in patients with non-specific low back pain: a randomized controlled trial. *Tohoku Journal of Experimental Medicine* 2020;**251**(3):193-206.

Kim 2020a *{published data only}*

Kim S, Jee Y. Effects of 3D moving platform exercise on physiological parameters and pain in patients with chronic low back pain. *Medicina* 2020;**56**(7):351-66.

Kim 2020b *{published data only}*

Kim SH, Park KN, Kwon OY. Classification-specific treatment improves pain, disability, fear-avoidance beliefs, and erector spinae muscle activity during walking in patients with low back pain exhibiting lumbar extension-rotation pattern: a randomized controlled trial. *Journal of Manipulative and Physiological Therapeutics* 2020;**43**(2):123-33.

Krishna 2020 *{published data only}*

Krishna D, Deepeshwar S, Devi B. Yoga-based relaxation technique facilitates sustained attention in patients with low back pain: a pilot study. *Advances in Mind-Body Medicine* 2020;**34**(3):11-7.

Kumar 2013 *{published data only}*

Kumar A, Gupta M, Katyal T. Effect of trunk muscles stabilization exercises and general exercises on pain in recurrent non specific low back ache. *International Research Journal of Medicine and Medical Sciences* 2013;**1**(1):23-6.

Kuvacic 2018 *{published data only}*

Kuvacic G, Fratini P, Padulo J, Antonio DI, De Giorgio A. Effectiveness of yoga and educational intervention on disability, anxiety, depression, and pain in people with CLBP: a randomized controlled trial. *Complementary Therapies in Clinical Practice* 2018;**31**:262-7.

Kwon 2020 *{published data only}*

Kwon SH, Oh SJ, Kim DH. The effects of lumbar stabilization exercise on transversus abdominis muscle activation capacity and function in low back pain patients. *Isokinetics and Exercise Science* 2020;**28**(2):147-52.

Leeuw 2008 *{published data only}*

Leeuw M, Goossens ME, Van Breukelen GJ, De Jong JR, Heuts PH, Smeets RJ, et al. Exposure in vivo versus operant graded activity in chronic low back pain patients: results of a randomized controlled trial. *Pain* 2008;**138**(1):192-207.

Leonard 2018 *{published data only}*

Joseph LH, Benjamaporn H, Patraporn S, Ubon P, Aatit P. Effects of massage as a combination therapy with lumbopelvic stability exercises as compared to standard massage therapy in low back pain: a randomized cross-over study. *International Journal of Therapeutic Massage & Bodywork* 2018;**11**(4):16-22.

Li 2015 *{published data only}*

Li L, Feng Z. Clinical observation of treating chronic nonspecific low back pain with Baduanjin combining sling exercise therapy. *Rheumatology & Arthritis* 2015;**4**:16-9.

Li 2017 *{published data only}*

Li H, Zhang H, Liu S, Wang Y, Gai D, Lu Q, et al. Rehabilitation effect of exercise with soft tissue manipulation in patients with lumbar muscle strain. *Nigerian Journal of Clinical Practice* 2017;**20**(5):629-33.

Lima 2018 *{published data only}*

Lima VP, Nunes RAM, Da Silva JB, Paz GA, Jesus M, De Castro JPB, et al. Pain perception and low back pain functional disability after a 10-week core and mobility training program: a pilot study. *Journal of Back & Musculoskeletal Rehabilitation* 2018;**31**(4):637-43.

Liu 2018 *{published data only}*

Liu J, Zhao W, Yuan Y. Effects of Tai Chi on the event-related potential of patients with chronic non-specific low back pain. *Chinese Journal of Sports Medicine* 2018;**10**:826-32.

Liu 2019 *{published data only}*

Liu J, Yeung A, Xiao T, Tian X, Kong Z, Zou L, et al. Chen-style tai chi for individuals (aged 50 years old or above) with chronic non-specific low back pain: a randomized controlled trial. *International Journal of Environmental Research & Public Health* 2019;**16**(3):517-26.

Madadi-Shad 2020 *{published data only}*

Madadi-Shad M, Jafarnezhadgero AA, Sheikhalizade H, Dionisio VC. Effect of a corrective exercise program on gait kinetics and muscle activities in older adults with both low back pain and pronated feet: a double-blind, randomized controlled trial. *Gait & Posture* 2020;**76**:339-45.

Mataran-Penarrocha 2020 *{published data only}*

Mataran-Penarrocha GA, Lara Palomo IC, Antequera Soler E, Gil-Martinez E, Fernandez-Sanchez M, Aguilar-Ferrandiz ME, et al. Comparison of efficacy of a supervised versus non-supervised physical therapy exercise program on the pain, functionality and quality of life of patients with non-specific chronic low-back pain: a randomized controlled trial. *Clinical Rehabilitation* 2020;**34**(7):948-59.

Matheve 2020 *{published data only}*

Matheve T, Bogaerts K, Timmermans A. Virtual reality distraction induces hypoalgesia in patients with chronic low back pain: a randomized controlled trial. *Journal of Neuroengineering and Rehabilitation* 2020;**17**(1):55.

Matos 2020 *{published data only}*

Matos FP, Dantas EHM, Oliveira FB, Castro JBP, Conceicao M, Nunes RAM, et al. Analysis of pain symptoms, flexibility and hydroxyproline concentration in individuals with low back pain submitted to Global Postural Re-education and stretching. *Pain Management* 2020;**10**(3):167-77.

Mazloum 2016 *{published data only}*

Mazloum V, Sahebozamani M, Barati A, Nakhaee N. Comparing the effects of Pilates training and McKenzie exercises on core muscles cross-sectional area and strength in patients with chronic non-specific low back pain: a clinical trial. *Journal of Mazandaran University of Medical Sciences* 2016;**26**(143):48-61.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mbada 2014 *{published data only}*

Mbada CE, Ayanniyi O, Ogunlade SO, Orimolade EA, Oladiran AB, Ogundele AO. Influence of Mckenzie protocol and two modes of endurance exercises on health-related quality of life of patients with long-term mechanical low-back pain. *Pan African Medical Journal* 2014;**17**(5):1-7.

Mbada 2019 *{published data only}*

Mbada CE, Olaoye MI, Dada OO, Ayanniyi O, Johnson OE, Odole AC, et al. Comparative efficacy of clinic-based and telerehabilitation application of McKenzie therapy in chronic low-back pain. *International Journal of Telerehabilitation* 2019;**11**(1):41-58.

Moncelon 2015 *{published data only}*

Moncelon S, Otero J. Méthode McKenzie et lombalgiques chroniques avec Préférence Directionnelle. *Kinesitherapie Revue* 2015;**15**(160):31-7.

Nambi 2020 *{published data only}*

Nambi G, Abdelbasset WK, Alqahtani BA, Alrawaili SM, Abodonya AM, Saleh AK. Isokinetic back training is more effective than core stabilization training on pain intensity and sports performances in football players with chronic low back pain: a randomized controlled trial. *Medicine* 2020;**99**(21):e20418.

Nambi 2020a *{published data only}*

Nambi G, Abdelbasset WK, Elsayed SH, Alrawaili SM, Abodonya AM, Saleh AK, et al. Comparative effects of isokinetic training and virtual reality training on sports performances in university football players with chronic low back pain - randomized controlled study. *Evidence-Based Complementary and Alternative Medicine* 2020;**2020**:2981273.

Nambi 2021 *{published data only}*

Nambi G, Abdelbasset WK, Alrawaili SM, Abodonya AM, Saleh AK. Virtual reality or Isokinetic training; its effect on pain, kinesiophobia and serum stress hormones in chronic low back pain: a randomized controlled trial. *Technology and Health Care* 2021;**29**(1):155-66.

Narouei 2020 *{published data only}*

Narouei S, Barati AH, Akuzawa H, Talebian S, Ghiasi F, Akbari A, et al. Effects of core stabilization exercises on thickness and activity of trunk and hip muscles in subjects with nonspecific chronic low back pain. *Journal of Bodywork and Movement Therapies* 2020;**24**(4):138-46.

Natraj 2018 *{published data only}*

Natraj P, Anandaselvashankar A, Manikumar M, Sankara Kumaran P. Efficacy of slump stretching in combinations with conventional therapy in chronic non-radicular low back pain: a randomized control trial. *Biomedicine* 2018;**38**(3):396-403.

Neyaz 2019 *{published data only}*

Neyaz O, Sumila L, Nanda S, Wadhwa S. Effectiveness of hatha yoga versus conventional therapeutic exercises for chronic nonspecific low-back pain. *Journal of Alternative & Complementary Medicine* 2019;**25**(9):938-45.

Ning 2015 *{published data only}*

Ning X, Wu L, Wang T, He X, Yu Y. Clinical study on five mimic-animal boxing combined with core muscular strength exercise for the treatment of nonspecific low back pain. *Journal of Traditional Chinese Orthopedics and Traumatology* 2015;**11**:25-8.

Noormohammadpour 2018 *{published data only}*

Noormohammadpour P, Kordi M, Mansournia MA, Akbari-Fakhrabadi M, Kordi R. The role of a multi-step core stability exercise program in the treatment of nurses with chronic low back pain: a single-blinded randomized controlled trial. *Asian Spine Journal* 2018;**12**(3):490-502.

O'Keeffe 2020 *{published data only}*

O'Keeffe M, O'Sullivan P, Purtill H, Bargary N, O'Sullivan K. Cognitive functional therapy compared with a group-based exercise and education intervention for chronic low back pain: a multicentre randomised controlled trial (RCT). *British Journal of Sports Medicine* 2020;**54**(13):782-9.

Oh 2020 *{published data only}*

Oh YJ, Park SH, Lee MM. Comparison of effects of abdominal draw-in lumbar stabilization exercises with and without respiratory resistance on women with low back pain: a randomized controlled trial. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* 2020;**26**:e921295.

Oka 2018 *{published data only}*

Oka H, Nomura N, Asada F, Takano K, Nitta Y, Uchima Y, et al. The effect of the "One Stretch" exercise on the improvement of low back pain in Japanese nurses: a large-scale, randomized, controlled trial. *Modern Rheumatology* 2018;**29**(5):861-6.

Oliveira 2011 *{published data only}*

Oliveira PD, Blanco PE, Facci LM. Comparing the hydrokinesiotherapy effects with land exercises in non specific low back pain patients: randomized clinical Trial. *Revista Terapia Manual* 2011;**9**(45):470-7.

Oliveira 2020 *{published data only}*

Oliveira Meirelles F, Oliveira Muniz Cunha JC, Silva EB. Osteopathic manipulation treatment versus therapeutic exercises in patients with chronic nonspecific low back pain: a randomized, controlled and double-blind study. *Journal of Back and Musculoskeletal Rehabilitation* 2020;**33**(3):367-77.

Oratsch 2019 *{published data only}*

Oratsch C, Pipam W, Kostenberger M, Apich G, Likar R. Treatment for chronic back pain? Active multimodal, interdisciplinary pain therapy vs. physiotherapy-physical therapy for chronic back pain. *Schmerz* 2019;**33**(4):337-46.

Owen 2020 *{published data only}*

Owen PJ, Miller CT, Rantalainen T, Simson KJ, Connell D, Hahne AJ, et al. Exercise for the intervertebral disc: a 6-month randomised controlled trial in chronic low back pain. *European Spine Journal* 2020;**29**(8):1887-99.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Pacrez-de-la-Cruz 2017 *{published data only}*

Pacrez-de-la-Cruz S, Puentes-Fern S, Rocamora-Pacrez P. Effectiveness of a program of Romana's Pilates for non-specific low back pain. a pilot study. *Revista Internacional de Medicina* 2017;**17**(68):667-76.

Park 2020 *{published data only}*

Park S, Park S, Min S, Kim CJ, Jee YS. A randomized controlled trial investigating the effects of equine simulator riding on low back pain, morphological changes, and trunk musculature in elderly women. *Medicina* 2020;**56**(11):610-24.

Park 2020a *{published data only}*

Park SJ, Kim YM, Yang SR. Effects of lumbar segmental stabilization exercise and respiratory exercise on the vital capacity in patients with chronic back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2020;**33**(5):841-8.

Patil 2018 *{published data only}*

Patil NJ, Nagaratna R, Tekur, P, Manohar PV, Hemant B, Dhanashri P. A randomized trial comparing effect of yoga and exercises on quality of life in among nursing population with chronic low back pain. *International Journal of Yoga* 2018;**11**(3):208-14.

Prado 2019 *{published data only}*

Prado ERA, Meireles SM, Carvalho ACA, Mazoca MF, Motta Neto AM, Barboza Da Silva R, et al. Influence of isostretching on patients with chronic low back pain. A randomized controlled trial. *Physiotherapy Theory & Practice* 2019;**1**:1-8.

Rabiei 2020 *{published data only}*

Rabiei P, Sheikhi B, Letafatkar A. Comparing pain neuroscience education followed by motor control exercises with group-based exercises for chronic low back pain: a randomized controlled trial. *Pain Practice* 2020;**21**(3):333-42.

Rahbar 2018 *{published data only}*

Rahbar M, Salekzamani Y, Jahanjou F, Eslamian F, Niroumand A, Dolatkhah N. Effect of hippotherapy simulator on pain, disability and range of motion of the spinal column in subjects with mechanical low back pain: a randomized single-blind clinical trial. *Journal of Back & Musculoskeletal Rehabilitation* 2018;**31**(6):1183-92.

Rajpal 2008 *{published data only}*

Rajpal N, Arora M, Chauhan V. A study on efficacy of Pilates & Pilates & Mckenzie exercises in postural low back pain: a rehabilitative protocol. *Indian Journal of Physiotherapy & Occupational Therapy* 2008;**1**(1):33-56.

Ramos 2018 *{published data only}*

Ramos LAV, Callegari B, Franca FJR, Magalhaes MO, Burke TN, Silva APMC Carvalho E, et al. Comparison between transcutaneous electrical nerve stimulation and stabilization exercises in fatigue and transversus abdominis activation in patients with lumbar disk herniation: a randomized study. *Journal of Manipulative and Physiological Therapeutics* 2018;**41**(4):323-31.

Ranaalsadat 2018 *{published data only}*

Ranaalsadat A, Sahib MA, Karimi F. Eight weeks of central stability training on a ball with hanging rope training on quality of life and pain in women with chronic nonspecific low back pain: clinical trial research. *Journal of Zanjan University of Medical Sciences & Health Services* 2018;**26**(117):45-57.

Raoul 2019 *{published data only}*

Raoul T, Malferiot J, Barizien N, Lhuissier FJ. Effects of a muscle strengthening program designed for spine extensors in triathletes with chronic back pain: randomized controlled trial in 67 athletes. *Journal de Traumatologie du Sport* 2019;**36**(3):183-93.

Rodriguez-Romero 2019 *{published data only}*

Rodriguez-Romero B, Bello O, Vivas Costa J, Carballo-Costa L. A therapeutic exercise program improves pain and physical dimension of health-related quality of life in young adults: a randomized controlled trial. *American Journal of Physical Medicine & Rehabilitation* 2019;**98**(5):392-8.

Rungthip 2018 *{published data only}*

Rungthip P, Rungthip C, Su Su H, Weerasak T, Pongsatorn S, Boucaut R. The effect of core stabilization exercise on lumbar joint position sense in patients with subacute non-specific low back pain: a randomized controlled trial. *Journal of Physical Therapy Science* 2018;**30**(11):1390-6.

Rydeard 2006b *{published data only}*

Rydeard R, Leger L, Smith D. Pilates-based therapeutic exercise: effect on subjects with nonspecific chronic low back pain and functional disability: a randomized controlled trial. *Journal of Orthopaedic & Sports Physical Therapy* 2006;**36**(7):472-84.

Saiklang 2020 *{published data only}*

Saiklang P, Puntumetakul R, Selfe J, Yeowell G. An evaluation of an innovative exercise to relieve chronic low back pain in sedentary workers. *Human Factors* 2020 [Epub ahead of print]:18720820966082.

Salas 2018 *{published data only}*

Salas E, Hulla R, Vanzzini N, Trowbridge C, Brotto M, Keller D, et al. The relative efficacy of two exercise methods for older adults with chronic low back pain: a preliminary randomized control study. *Journal of Applied Biobehavioral Research* 2018;**24**(1):1-11.

Samir 2016 *{published data only}*

Samir SM, Zaky LA, Soliman MO. Mulligan versus Maitland mobilizations in patients with chronic low back dysfunction. *International Journal of PharmTech Research* 2016;**9**(6):92-9.

Sarker 2019 *{published data only}*

Sarker KK, Sethi J, Mohanty U. Effect of spinal manipulation on pain sensitivity, postural sway, and health-related quality of life among patients with non-specific chronic low back pain: a randomised control trial. *Journal of Clinical and Diagnostic Research* 2019;**13**(2):YC01-5.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sawant 2019 *{published data only}*

Sawant RS, Shinde SB. Effect of hydrotherapy based exercises for chronic nonspecific low back pain. *Indian Journal of Physiotherapy & Occupational Therapy* 2019;**13**(1):133-8.

Schmidt 2020 *{published data only}*

Schmidt AM, Schiøttz-Christensen B, Foster NE, Laurberg TB, Maribo T. The effect of an integrated multidisciplinary rehabilitation programme alternating inpatient interventions with home-based activities for patients with chronic low back pain: a randomized controlled trial. *Clinical Rehabilitation* 2020;**34**(3):382-93.

Schulz 2019 *{published data only}*

Schulz C, Evans R, Maiers M, Schulz K, Leininger B, Bronfort G. Spinal manipulative therapy and exercise for older adults with chronic low back pain: a randomized clinical trial. *Chiropractic & Manual Therapies* 2019;**27**:21-35.

Sedaghati 2017 *{published data only}*

Sedaghati P, Arjmand A, Sedaghati N. Comparison of the effects of different training approaches on dynamic balance and pain intensity in the patients with chronic back pain. *Scientific Journal of Kurdistan University of Medical Sciences* 2017;**22**(4):45-56.

Seo 2019 *{published data only}*

Seo HR, Kim TH. The effects of gyrotonic expansion system exercise and trunk stability exercise on muscle activity and lumbar stability for the subjects with chronic low back pain. *Journal of Exercise Rehabilitation* 2019;**15**(1):129-33.

Shabnam 2011 *{published data only}*

Shabnam H, Reza R, Nooroddin K. Effects of consecutive supervised core stability training on pain and disability in women with nonspecific chronic low back pain. *Koomesh* 2011;**12**(3):244-52.

Shamsi 2019 *{published data only}*

Shamsi M, Mirzaei M, Samadzadeh S, Schuermans J. Comparing the effects of static stretching and strengthening in lengthened position on EMG activity of hamstring muscle in patients with chronic non-specific LBP having shortened muscle: a randomised controlled clinical trial. *European Journal of Physiotherapy* 2019;**0**:1-7.

Shamsi 2020 *{published data only}*

Shamsi M, Mirzaei M, Shahsavari S, Safari A, Saeb M. Modeling the effect of static stretching and strengthening exercise in lengthened position on balance in low back pain subject with shortened hamstring: a randomized controlled clinical trial. *BMC Musculoskeletal Disorders* 2020;**21**(1):1-9.

Shamsi 2020a *{published data only}*

Shamsi M, Shahsavari S, Safari A, Mirzaei M. A randomized clinical trial for the effect of static stretching and strengthening exercise on pelvic tilt angle in LBP patients. *Journal of Bodywork and Movement Therapies* 2020;**24**(3):15-20.

Shariat 2019 *{published data only}*

Shariat A, Alizadeh R, Moradi V, Afsharnia E, Hakakzadeh A, Ansari NN, et al. The impact of modified exercise and relaxation therapy on chronic lower back pain in office workers: a randomized clinical trial. *Journal of Exercise Rehabilitation* 2019;**15**(5):703-8.

Shebib 2019 *{published data only}*

Shebib R, Bailey JF, Smittenaar P, Perez DA, Mecklenburg G, Hunter S. Randomized controlled trial of a 12-week digital care program in improving low back pain. *NPJ Digital Medicine* 2019;**2**:1-12.

Sheeran 2013 *{published data only}*

Sheeran L, Van Deursen R, Caterson B, Sparkes V. Classification-guided versus generalized postural intervention in subgroups of nonspecific chronic low back pain: a pragmatic randomized controlled study. *Spine* 2013;**38**(19):1613-25.

Sherman 2020 *{published data only}*

Sherman KJ, Wellman RD, Hawkes RJ, Phelan EA, Lee T, Turner JA. T'ai Chi for chronic low back pain in older adults: a feasibility trial. *Journal of Alternative and Complementary Medicine* 2020;**26**(3):176-89.

Singh 2019 *{published data only}*

Singh S, Khan N, Agarwal R. Effect of rhythmic stabilization exercise v/s conventional physiotherapy on pain and disability with patients of chronic mechanical low back pain. *Indian Journal of Physiotherapy & Occupational Therapy* 2019;**13**(4):122-6.

Sipaviciene 2020 *{published data only}*

Sipaviciene S, Kliziene I. Effect of different exercise programs on non-specific chronic low back pain and disability in people who perform sedentary work. *Clinical Biomechanics* 2020;**73**:17-27.

Sokhanguei 2017 *{published data only}*

Sokhanguei Y, Sadoughi Noorabadi M, Sadoughi Noorabadi N, Hatami M. The effect of a Pilates program on chronic non-specific lower back pain and stable range of motion in 40 to 60 year old women. *Journal of Zanjan University of Medical Sciences and Health Services* 2017;**25**(108):20-30.

Song 2008 *{published data only}*

Song H, Gao L. A study on effect of Taijiquan on lumbar disc protrusion. *Journal of Beijing Sport University* 2008;**5**:1-5.

Springrova 2020 *{published data only}*

Springrova I, Krejcova A, Bendíkova E, Tomkova Š, Lubkowska W, Mroczek B. Comparison of the impact of two physical therapy methods on pain and disability in patients with non-specific lower back pain: a controlled clinical pilot study. *Family Medicine & Primary Care Review* 2020;**22**(2):146-51.

Srivastav 2018 *{published data only}*

Srivastav N, Joshi S, Kushwah SS. Comparison between effectiveness of lumbar stabilization exercises and conventional physical therapy in the management of mechanical low back pain. *Indian Journal of Physiotherapy & Occupational Therapy* 2018;**12**(4):28-33.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Srivastava 2020 *{published data only}*

Srivastava S, Kumar KUD, Mittal H, Dixit S, Nair A. Short-term effect of muscle energy technique and mechanical diagnosis and therapy in sacroiliac joint dysfunction: a pilot randomized clinical trial. *Journal of Bodywork and Movement Therapies* 2020;**24**(3):63-70.

Suh 2019 *{published data only}*

Suh JH, Kim H, Jung GP, Ko JY, Ryu JS. The effect of lumbar stabilization and walking exercises on chronic low back pain: a randomized controlled trial. *Medicine* 2019;**98**(26):1-10.

Surkitt 2016 *{published data only}*

Surkitt LD, Ford JJ, Chan AYP, Richards MC, Slater SL, Pizzari T, et al. Effects of individualised directional preference management versus advice for reducible discogenic pain: a pre-planned secondary analysis of a randomised controlled trial. *Manual Therapy* 2016;**25**:69-80.

Tagliaferri 2020 *{published data only}*

Tagliaferri SD, Miller CT, Ford JJ, Hahne AJ, Main LC, Rantalainen T, et al. Randomized trial of general strength and conditioning versus motor control and manual therapy for chronic low back pain on physical and self-report outcomes. *Journal of Clinical Medicine* 2020;**9**(6):1726-43.

Takinaci 2019 *{published data only}*

Takinaci ZD, Karagulle M, Issever H, Karagulle MZ. Clinical efficacy of spa therapy (balneotherapy) for chronic low back pain: a randomized single-blind trial. *European Journal of Integrative Medicine* 2019;**29**:1-5.

Teixeira 2016 *{published data only}*

Teixeira AL, Da Silva NF, Da Silva EM. Evaluation of functional disability and pain in patients with chronic low back pain submitted to physiotherapy. *Manual Therapy, Posturology & Rehabilitation Journal* 2016;**14**:1-6.

Teychenne 2019 *{published data only}*

Teychenne M, Lamb KE, Main L, Miller C, Hahne A, Ford J, et al. General strength and conditioning versus motor control with manual therapy for improving depressive symptoms in chronic low back pain: a randomised feasibility trial. *PLOS One* 2019;**14**(8):1-11.

Thomas 2016 *{published data only}*

Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. *Journal of Pain* 2016;**17**(12):1302-17.

Tul 2019 *{published data only}*

Tul Ain SQ, Shakil U, Rehman S, Maryam M, Kiani SK. Effects of sustained natural apophyseal glides with and without thoracic posture correction techniques on mechanical back pain: a randomized control trial. *Journal of the Pakistan Medical Association* 2019;**69**(11):1584-7.

Turner 2018 *{published data only}*

Turner BJ, Liang Y, Simmonds MJ, Rodriguez N, Bobadilla R, Yin Z. Randomized trial of chronic pain self-management program in the community or clinic for low-income primary care patients. *Journal of General Internal Medicine* 2018;**33**(5):668-77.

Van Baal 2020 *{published data only}*

Van Baal K, Allofs J, Ehrenbrusthoff K, Gruneberg C, Hering T, Kopkow C, et al. Effects of a movement control and tactile acuity training in patients with nonspecific chronic low back pain and control impairment: a randomised controlled pilot study. *BMC Musculoskeletal Disorders* 2020;**21**(1):794.

Vasseljen 2010a *{published data only}*

Vasseljen O, Fladmark AM. Abdominal muscle contraction thickness and function after specific and general exercises: a randomized controlled trial in chronic low back pain patients. *Manual Therapy* 2010;**15**(5):482-9.

Verbrugghe 2019 *{published data only}*

Verbrugghe J, Agten A, Stevens S, Hansen D, Demoulin C, Eijnde BO, et al. Exercise intensity matters in chronic nonspecific low back pain rehabilitation. *Medicine & Science in Sports & Exercise* 2019;**51**(12):2434-42.

Verbrugghe 2020 *{published data only}*

Verbrugghe J, Agten A, Stevens S, Hansen D, Demoulin C, Eijnde BO, et al. High intensity training to treat chronic nonspecific low back pain: effectiveness of various exercise modes. *Journal of Clinical Medicine* 2020;**9**(8):2401-15.

Verra 2018 *{published data only}*

Verra ML, Angst F, Brioschi R, Lehmann S, Benz T, Aeschlimann A, et al. Effectiveness of subgroup-specific pain rehabilitation: a randomized controlled trial in patients with chronic back pain. *European Journal of Physical & Rehabilitation Medicine* 2018;**54**(3):358-70.

Vincent 2014b *{published data only}*

Vincent HK, George SZ, Seay AN, Vincent KR, Hurley RW. Resistance exercise, disability, and pain catastrophizing in obese adults with back pain. *Medicine & Science in Sports & Exercise* 2014;**46**(9):1693-701.

Wahyuddin 2020 *{published data only}*

Wahyuddin W, Vongsirinavarat M, Mekhora K, Bovonsunthonchai S, Adisaipoapun R. Immediate effects of muscle energy technique and stabilization exercise in patients with chronic low back pain with suspected facet joint origin: a pilot study. *Hong Kong Physiotherapy Journal* 2020;**40**(2):109-19.

Walsh 2020 *{published data only}*

Walsh N, Jones L, Phillips S, Thomas R, Odondi L, Palmer S, et al. Facilitating activity and self-management for people with arthritic knee, hip or lower back pain (FASA): a cluster randomised controlled trial. *Musculoskeletal Science and Practice* 2020;**50**:102271.

Wang 2019 *{published data only}*

Wang XQ, Gu W, Chen BL, Wang X, Hu HY, Zheng HL, et al. Effects of whole-body vibration exercise for non-specific chronic low back pain: an assessor-blind, randomized controlled trial. *Clinical Rehabilitation* 2019;**33**(9):1445-57.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Wang 2020 *{published data only}*

Wang X, Zhu JS. Effect of Cangguitanxue acupuncture combined with suspension exercise therapy on chronic low back pain. *Chinese Acupuncture & Moxibustion* 2020;**40**(7):739-43.

Waseem 2019 *{published data only}*

Waseem M, Karimi H, Gilani SA, Hassan D. Treatment of disability associated with chronic non-specific low back pain using core stabilization exercises in Pakistani population. *Journal of Back & Musculoskeletal Rehabilitation* 2019;**32**(1):149-54.

Wattamwar 2012 *{published data only}*

Wattamwar RB, Nadkarni K. Effect of conventional occupational therapy and yoga in chronic low back pain. *Indian Journal of Occupational Therapy* 2012;**44**(2):1-8.

Weissenfels 2019 *{published data only}*

Weissenfels A, Wirtz N, Dormann U, Kleinoder H, Donath L, Kohl M, et al. Comparison of whole-body electromyostimulation versus recognized back-strengthening exercise training on chronic nonspecific low back pain: a randomized controlled study. *BioMed Research International* 2019;**2019**:5745409.

Yalfani 2020 *{published data only}*

Yalfani A, Raeisi Z, Koumasian Z. Effects of eight-week water versus mat pilates on female patients with chronic nonspecific low back pain: double-blind randomized clinical trial. *Journal of Bodywork and Movement Therapies* 2020;**24**(4):70-5.

Yang 2015 *{published data only}*

Yang J, Seo D. The effects of whole body vibration on static balance, spinal curvature, pain, and disability of patients with low back pain. *Journal of Physical Therapy Science* 2015;**27**(3):805-8.

Yao 2020 *{published data only}*

Yao C, Li Z, Zhang S, Wu Z, Zhu Q, Fang L. Effects of Wuqinxi in the patients with chronic low back pain: a randomized controlled trial. *Evidence-Based Complementary and Alternative Medicine* 2020;**2020**:1428246.

Yuan 2016 *{published data only}*

Yuan RL, Yan B, Li GZ, Huang Z, Ma J, Liu DZ. Observation on the forward curative effect of back muscle functional exercise in the treatment of lumbar vertebra small joint synovial incarcerated. *Chinese Community Doctors* 2016;**32**(17):186-7.

Zadro 2019 *{published data only}*

Zadro JR, Shirley D, Simic M, Mousavi SJ, Ceprnja D, Maka K, et al. Video-game-based exercises for older people with chronic low back pain: a randomized controlled table trial (GAMEBACK). *Physical Therapy* 2019;**99**(1):14-27.

Zakari 2019 *{published data only}*

Zakari UU, Bello B, Sokumbi GO, Yakasai AM, Danazumi MS. Comparison of the effects of positional release therapy and lumbar stabilization exercises in the management of chronic mechanical low back pain: a randomized controlled trial. *Critical Reviews in Physical and Rehabilitation Medicine* 2019;**31**(4):321-32.

Zarzycka 2018 *{published data only}*

Zarzycka K, Zarzycki M, Kolasa P. Comparative evaluation of rehabilitation systems for patients with chronic lumbar overload syndrome. *Aktualnosci Neurologiczne* 2018;**18**(1):5-13.

Zhang 2019a *{published data only}*

Zhang X, Bi X, Shao J, Sun D, Zhang C, Liu Z. Curative effects on muscle function and proprioception in patients with chronic lumbar disk herniation using isokinetic trunk muscle strength training. *International Journal of Clinical and Experimental Medicine* 2019;**12**(4):4311-20.

Zhongliang 2018 *{published data only}*

Zhongliang Z, Xiaojie Y, Bangzhong L, Min Z, Liya Z, Dongdong Y. Spinal mechanical impulsive manipulation versus conventional physiotherapy for chronic low back pain: differences in range of lumbar motion and trunk muscle strength. *Chinese Journal of Tissue Engineering Research* 2018;**22**(35):5637-41.

Zou 2019a *{published data only}*

Zou L, Zhang Y, Liu Y, Tian X, Xiao T, Liu X, et al. The effects of tai chi chuan versus core stability training on lower-limb neuromuscular function in aging individuals with non-specific chronic lower back pain. *Medicina* 2019;**55**(3):60-70.

Additional references

Abenhaim 2000

Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, et al. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. *Spine* 2000;**25**(4 Suppl):1S-33S.

Alhakami 2019

Alhakami AM, Davis S, Qasheesh M, Shaphe A, Chahal A. Effects of McKenzie and stabilization exercises in reducing pain intensity and functional disability in individuals with nonspecific chronic low back pain: a systematic review. *Journal of Physical Therapy Science* 2019;**31**(7):590-7.

Almeida 2020

Almeida MO, Yamato TP, Parreira P, Costa LO, Kamper S, Saragiotto BT. Overall confidence in the results of systematic reviews on exercise therapy for chronic low back pain: a cross-sectional analysis using the Assessing the Methodological Quality of Systematic Reviews (AMSTAR) 2 tool. *Brazilian Journal of Physical Therapy* 2020;**24**(2):103-17.

Beyera 2019

Beyera GK, O'Brien J, Campbell S. Health-care utilisation for low back pain: a systematic review and meta-analysis of population-based observational studies. *Rheumatology International* 2019;**39**(10):1663-79.

Bird 2020

Bird ML, Mortenson WB, Eng JJ. Evaluation and facilitation of intervention fidelity in community exercise programs through an adaptation of the TIDier framework. *BMC Health Services Research* 2020;**20**(1):68-75.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chiarotto 2016

Chiarotto A, Maxwell LJ, Terwee CB, Wells GA, Tugwell P, Ostelo RW. Roland-Morris Disability Questionnaire and Oswestry Disability Index: which has better measurement properties for measuring physical functioning in nonspecific low back pain? systematic review and meta-analysis. *Physical Therapy* 2016;**96**(10):1620-37.

Chiarotto 2018

Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LO, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain* 2018;**159**(3):481-95.

Chiarotto 2019

Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement properties of visual analogue scale, numeric rating scale, and pain severity subscale of the brief pain inventory in patients with low back pain: a systematic review. *Journal of Pain* 2019;**20**(3):245-63.

Chou 2010

Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *Journal of the American Medical Association* 2010;**303**(13):1295-302.

Christiansen 2018

Christiansen DH, De Vos Anderson NB, Poulsen PH, Ostelo RW. The smallest worthwhile effect of primary care physiotherapy did not differ across musculoskeletal pain sites. *Journal of Clinical Epidemiology* 2018;**101**:44-52.

Coste 2004

Coste J, Lefrancois G, Guillemin F, Pouchot J. Prognosis and quality of life in patients with acute low back pain: insights from a comprehensive inception cohort study. *Arthritis and Rheumatism* 2004;**51**(2):168-76.

Dagenais 2008

Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine Journal* 2008;**8**(1):8-20.

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**(3):177-88.

Distiller SR [Computer program]

Distiller SR. Ottawa, ON: Evidence Partners, 2020.

Dunn 2006

Dunn KM, Croft PR. The importance of symptom duration in determining prognosis. *Pain* 2006;**121**(1-2):126-32.

Egger 1997

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 1997;**315**(7109):629-34.

EndNote X8 [Computer program]

EndNote X8. Clarivate Analytics, 2018.

Euroquol Research Foundation 2019

EuroQol Research Foundation. Q-5D-5L User Guide. euroqol.org/publications/user-guides 2019 (accessed prior to 5 Sept 2021).

Fairbank 1980

Fairbank J, Couper J, Davies J, O'Brien JP. The Oswestry Low Back Pain Questionnaire. *Physiotherapy* 1980;**66**:271-3.

Ferreira 2010

Ferreira ML, Smeets RJ, Kamper SJ, Ferreira PH, Machado LA. Can we explain heterogeneity among randomized clinical trials of exercise for chronic back pain? A meta-regression analysis of randomized controlled trials. *Physical Therapy* 2010;**90**(10):1383-403.

Ferreira 2013

Ferreira ML, Herbert RD, Ferreira PH, Latimer J. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. *Journal of Clinical Epidemiology* 2013;**66**(12):1397-404.

Foster 2018

Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet* 2018;**391**(10137):2368-83.

Furlan 2015

Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. *Spine* 2015;**40**(21):1660-73.

GBD Collaborators 2018

GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease study 2017. *Lancet* 2018;**392**(10159):1789-1858.

GRADEpro [Computer program]

GRADEpro Guideline Development Tool. Evidence Prime, Incorporated, Version accessed prior to 5 Sept 2021. Hamilton (ON): GRADE Working Group, McMaster University, 2015. www.gradepro.org.

Grotle 2005

Grotle M, Brox JI, Veierod MB, Glomsrod B, Lonn JH, Vollestad NK. Clinical course and prognostic factors in acute low back pain: patients consulting primary care for the first time. *Spine* 2005;**30**(8):976-82.

Hahne 2011

Hahne AJ, Ford JJ, Surkitt LD, Richards MC, Chan AY, Thompson SL, Hinman RS, Taylor NF. Specific treatment of problems of the spine (STOPS): design of a randomised controlled trial comparing specific physiotherapy versus advice for people with subacute low back disorders. *BMC Musculoskeletal Disorders* 2011;**12**:104-20.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Harris 2019

Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software partners. *Journal of Biomedical Informatics* 2019;**95**:1-10.

Hartvigsen 2018

Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. *Lancet* 2018;**391**(10137):2356-67.

Hayden 2005c

Hayden JA, Van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Annals of Internal Medicine* 2005;**142**(9):776-85.

Hayden 2009

Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *Journal of Clinical Epidemiology* 2009;**62**(8):781-96.

Hayden 2020a

Hayden JA. Predatory publishing dilutes and distorts evidence in systematic reviews. *Journal of Clinical Epidemiology* 2020;**121**:117-9.

Hayden 2020b

Hayden JA, Wilson MN, Stewart S, Cartwright JL, Smith AO, Riley RD, Van Tulder M, Chronic Low Back Pain IPD Meta-Analysis Group. Exercise treatment effect modifiers in persistent low back pain: an individual participant data meta-analysis of 3514 participants from 27 randomised controlled trials. *British Journal of Sports Medicine* 2020;**54**:1277-93.

Hayden 2021a

Hayden JA, Ellis J, Ogilvie R, Boulos L, Stanojevic S. Meta-epidemiological study of publication integrity, and quality of conduct and reporting of randomized trials included in a systematic review of low back pain. *Journal of Clinical Epidemiology* 2021;**134**:65-78.

Hayden 2021b

Hayden JA, Ellis J, Ogilvie R, Stewart S, Bagg MK, Stanojevic S, Yamato TP, Saragiotto BT. Some types of exercise are more effective than others in people with chronic low back pain: a network meta-analysis. *Journal of Physiotherapy* 2021;**67**(4):252-62.

Heitz 2009

Heitz CA, Hilfiker R, Bachmann LM, Joronen H, Lorenz T, Uebelhart D, et al. Comparison of risk factors predicting return to work between patients with subacute and chronic non-specific low back pain: systematic review. *European Spine Journal* 2009;**18**(12):1829-35.

Henschke 2007

Henschke N, Maher CG, Refshauge KM, Das A, McAuley JH. Low back pain research priorities: a survey of primary care practitioners. *BMC Family Practice* 2007;**8**:40-5.

Henschke 2009

Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Characteristics of patients with acute low back pain presenting to primary care in Australia. *Clinical Journal of Pain* 2009;**25**(1):5-11.

Henschke 2010

Henschke N, Ostelo RW, Van Tulder MW, Vlaeyen JW, Morley S, Assendelft WJ, et al. Behavioural treatment for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 7. Art. No: CD002014. [DOI: 10.1002/14651858.CD002014.pub3]

Higgins 2002

Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**(11):1539-58.

Higgins 2011

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 training.cochrane.org/handbook/archive/v5.1.

Higgins 2016

Higgins JR, Lin FC, Evans JP. Plagiarism in submitted manuscripts: incidence, characteristics and optimization of screening-case study in a major specialty medical journal. *Research Integrity and Peer Review* 2016;**1**:13-21.

Higgins 2019

Higgins J, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Methodological Expectations of Cochrane Intervention Reviews (MECIR): Standards for the conduct and reporting of new Cochrane Intervention Reviews, reporting of protocols and the planning, conduct and reporting of updates. Cochrane, London, 2019. community.cochrane.org/mecir-manual (accessed prior to 5 Sept 2021);(1.07).

Hozo 2005

Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology* 2005;**5**:13.

iThenticate 2020 [Computer program]

iThenticate Plagiarism Detection Software. Turnitin LLC, 2020.

Itz 2013

Itz CJ, Geurts JW, Van Kleef M, Nelemans P. Clinical course of non-specific low back pain: a systematic review of prospective cohort studies set in primary care. *European Journal of Pain* 2013;**17**(1):5-15.

Jzelenberg 2011

Jzelenberg W, Rubinstein SM, Hayden J, Koes BW, Van Tulder MW. Exercise therapy for acute non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2011, Issue 10. Art. No: CD009365. [DOI: 10.1002/14651858.CD009365]

Kopec 2000

Kopec JA. Measuring functional outcomes in persons with back pain: a review of back-specific questionnaires. *Spine* 2000;**25**(24):3110-4.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lim 2012

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease study 2010. *Lancet* 2012;**380**(9859):2224-60.

Liu 2013

Liu H, Yao K, Zhang J, Li L, Wu T, Brox JI, et al. Sling exercise therapy for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 9. Art. No: CD010689. [DOI: 10.1002/14651858.CD010689]

Lykkesfeldt 2016

Lykkesfeldt J. Strategies for using plagiarism software in the screening of incoming journal manuscripts: recommendations based on a recent literature survey. *Basic & Clinical Pharmacology and Toxicology* 2016;**119**(2):161-4.

Machado 2012a

Machado L, Lin CWC, Clare H, Van Tulder MW. The McKenzie method for (sub)acute non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. Art. No: CD009711. [DOI: 10.1002/14651858.CD009711]

Machado 2012b

Machado L, Van Tulder MW, Lin CWC, Clare H, Hayden JA. The McKenzie method for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. Art. No: CD009712. [DOI: 10.1002/14651858.CD009712]

Malmivaara 2019

Malmivaara A. Generalizability of findings from randomized controlled trials is limited in the leading general medical journals. *Journal of Clinical Epidemiology* 2019;**107**:36-41.

Martin 2008

Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, et al. Expenditures and health status among adults with back and neck problems. *Journal of the American Medical Association* 2008;**299**(6):656-64.

Melzack 1975

Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;**1**:277-99.

Minozzi 2020

Minozzi S, Cinquini M, Gianola S, Gonzalez-Lorenzo M, Banzi R. The revised Cochrane risk-of-bias tool for randomised trials (RoB 2) showed low inter-rater reliability and challenges in its application. *Journal of Clinical Epidemiology* 2020;**126**:37-44.

Miyamoto 2019

Miyamoto GC, Lin CC, Cabral CMN, Van Dongen JM, Van Tulder MW. Cost-effectiveness of exercise therapy in the treatment of non-specific neck pain and low back pain: a systematic review with meta-analysis. *British Journal of Sports Medicine* 2019;**53**(3):172-81.

Moffett 2000

Moffett JK, Frost H 86:295e305. Back to fitness programme: the manual for physiotherapists to set up the classes. *Physiotherapy* 2000;**85**:295-305.

Namnaqani 2019

Namnaqani FI, Mashabi AS, Yaseen KM, Alshehri MA. The effectiveness of McKenzie method compared to manual therapy for treating chronic low back pain: a systematic review. *Journal of Musculoskeletal and Neuronal Interactions* 2019;**19**(4):492-9.

National Guideline Centre 2016

National Guideline Centre. National Institute for Health and Care Excellence: Clinical Guidelines. In: Low Back Pain and Sciatica in Over 16s: Assessment and Management. London: National Institute for Health and Care Excellence (UK), 2016.

Niederer 2020

Niederer D, Mueller J. Sustainability effects of motor control stabilisation exercises on pain and function in chronic nonspecific low back pain patients: a systematic review with meta-analysis and meta-regression. *PLOS One* 2020;**15**(1):e0227423.

Parreira 2017

Parreira P, Heymans MW, Van Tulder MW, Esmail R, Koes BW, Poquet N, et al. Back Schools for chronic non-specific low back pain. *Cochrane Database of Systematic Reviews* 2017, Issue 8. Art. No: CD011674. [DOI: 10.1002/14651858.CD011674.pub2]

Powell 2011

Powell KE, Paluch AE, Blair SN. Physical activity for health: What kind? How much? How intense? On top of what? *Annual Review of Public Health* 2011;**32**:349-65.

Qaseem 2017

Qaseem A, Wilt TJ, McLean RM, Forciea MA. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. *Annals of Internal Medicine* 2017;**166**(7):514-30.

R Core Team 2013 [Computer program]

R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2013.

Review Manager 2020 [Computer program]

Review Manager (RevMan). Version 5.4. Copenhagen: Nordic Cochrane Centre: The Cochrane Collaboration, 2020.

Richardson 1999

Richardson C, Jull G, Hodges P. Therapeutic exercise for spinal segmental stabilisation in low back pain: scientific basis and clinical approach. Churchill Livingstone, 1999.

Riley 2011

Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *British Medical Journal* 2011;**342**:d549. [PMID: 21310794]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Roland 2000

Roland M, Fairbank J. The Roland Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 2000;**25**:3115-24.

Rubinstein 2019

Rubinstein SM, De Zoete A, Van Middelkoop M, Assendelft WJJ, De Boer MR, Van Tulder MW. Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. *British Medical Journal* 2019;**364**(1689):1-15.

Saragiotto 2016

Saragiotto BT, Maher CG, Yamato TP, Costa LO, Menezes Costa LC, Ostelo R, et al. Motor control exercise for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2016, Issue 1. Art. No: CD012004. [DOI: 10.1002/14651858.CD012004]

Schaafsma 2013

Schaafsma FG, Whelan K, Van der Beek AJ, Van der Es-Lambeek LC, Ojajärvi A, Verbeek JH. Physical conditioning as part of a return to work strategy to reduce sickness absence for workers with back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No: CD001822. [DOI: 10.1002/14651858.CD001822.pub3]

Shafshak 2020

Shafshak TS, Elnemr R. The visual analogue scale versus numerical rating scale in measuring pain severity and predicting disability in low back pain. *Journal of Clinical Rheumatology* 2020 [Epub ahead of print]. [DOI: 10.1097/RHU.0000000000001320]

Shi 2018

Shi Z, Zhou H, Lu L, Pan B, Wei Z, Yao X, et al. Aquatic exercises in the treatment of low back pain: a systematic review of the literature and meta-analysis of eight studies. *American Journal of Physical Medicine & Rehabilitation* 2018;**97**(2):116-22.

Skelly 2020

Skelly AC, Chou R, Dettori JR, Turner JA, Friedly JL, Rundell SD, et al, Agency for Healthcare Research and Quality (US). Noninvasive nonpharmacological treatment for chronic pain: a systematic review update. effectivehealthcare.ahrq.gov/products/nonpharma-treatment-pain/research-2018 (accessed prior to 5 Sept 2021).

Song 2003

Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *British Medical Journal* 2003;**326**(7387):472.

StataCorp 2015 [Computer program]

Stata Statistical Software. StataCorp, Version 14. College Station, TX: StataCorp LP, 2015.

Sterne 2019

Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *British Medical Journal* 2019;**366**:1-8.

Vanti 2019

Vanti C, Andreatta S, Borghi S, Guccione AA, Pillastrini P, Bertozzi L. The effectiveness of walking versus exercise on pain and function in chronic low back pain: a systematic review and meta-analysis of randomized trials. *Disability & Rehabilitation* 2019;**41**(6):622-32.

Vlaeyen 2002

Vlaeyen JW, de Jong J, Sieben J, Crombez G. Psychological approaches to pain management: graded exposure in vivo for pain-related fear. Guilford, 2002.

Von Korff 2000

Von Korff M, Jensen MP, Karoly P. Assessing global pain severity by self-report in clinical and health services research. *Spine* 2000;**25**(24):3140-51.

Ware 1992

Ware JE, Sherbourne CD. The MOS 36 item short-form health survey (SF-36). 1: conceptual framework and item selection. *Medical Care* 1992;**30**:473-80.

Weber-Wulff 2019

Weber-Wulff D. Plagiarism detectors are a crutch, and a problem. *Nature* 2019;**567**(7749):435.

Wewege 2018

Wewege MA, Booth J, Parmenter BJ. Aerobic vs. resistance exercise for chronic non-specific low back pain: a systematic review and meta-analysis. *Journal of Back and Musculoskeletal Rehabilitation* 2018;**31**(5):889-99.

Wieland 2017

Wieland LS, Skoetz N, Pilkington K, Vempati R, D'Adamo CR, Berman BM. Yoga treatment for chronic non-specific low back pain. *Cochrane Database of Systematic Reviews* 2017, Issue 1. Art. No: CD010671. [DOI: 10.1002/14651858.CD010671]

Wood 2019

Wood L, Foster N, Hayden JA. Workshop WB8: Agreeing on treatment targets for future trials of exercise for back pain. In: International Back and Neck Forum; 2019 July 4; Quebec City, Canada. 2019.

Wood 2020a

Wood L, Ogilvie R, Hayden JA. Specifying the treatment targets of exercise interventions: do we? *British Journal of Sports Medicine* 2020;**54**:1235-36.

Wood 2021

Wood L, Foster NE, Lewis M, Bishop A. Exercise interventions for persistent non-specific low back pain - does matching outcomes to treatment targets make a difference? A systematic review and meta-analysis. *Journal of Pain* 2021;**22**(2):107-26.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yamato 2015

Yamato TP, Maher CG, Saragiotto BT, Hancock MJ, Ostelo R, Cabral CMN, et al. Pilates for low back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 7. Art. No: CD010265. [DOI: 10.1002/14651858.CD010265]

Zou 2019b

Zou L, Zhang Y, Yang L, Loprinzi PD, Yeung AS, Kong J, et al. Are mindful exercises safe and beneficial for treating chronic lower back pain? A systematic review and meta-analysis of randomized controlled trials. *Journal of Clinical Medicine* 2019;**8**(5):628-54.

References to other published versions of this review

Hayden 2005a

Hayden JA, Van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Art. No: CD000335. [DOI: 10.1002/14651858.CD000335.pub2]

Hayden 2005b

Hayden JA, Van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. *Annals of Internal Medicine* 2005;**142**(9):765-75.

Hayden 2012

Hayden JA, Cartwright J, Van Tulder MW, Malmivaara A. Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 4. Art. No: CD009790. [DOI: 10.1002/14651858.CD009790]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Adamczyk 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Poland, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 60 (E1 = 30, E2 = 30) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Strengthening exercises for the abdominal, gluteal, hip, and back muscles, post-isometric relaxation of muscles, active mobilisations, stretches; type = mixed; duration = 2 weeks; dose = low; design = individualised; delivery = individual; additional intervention = lumbar support & other **Exercise Group 2 (E2):** Strengthening exercises for abdominal, back, and thigh muscles and buttocks, stretching exercises for dorsal muscles, hip flexors and ischiotibial muscle, respiratory exercises; type = stretching & strengthening; duration = 2 weeks; dose = low; design = standardised; delivery = group; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 2 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Adamczyk 2009 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Patients were randomly assigned to two groups of 30.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Dropout rate not reported
Participants analysed in group allocated (attrition bias)	Unclear risk	Not described
Selective reporting (reporting bias)	Low risk	Yes
Groups similar at baseline (selection bias)	Unclear risk	Not described
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Yes

Akhtar 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Pakistan, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 120 (E1 = 60, E2 = 60) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Akhtar 2017 *(Continued)*

	Mean age (years): 46 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Core stabilisation exercise: exercises targeting deep muscles of the abdomen, ultrasound and transcutaneous electrical nerve stimulation; type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = individual; additional intervention = electrotherapy **Exercise Group 2 (E2):** Routine physical therapy exercise "Exercise not specifically targeted to core muscle of the spine"; type = strengthening & stretching; duration = 6 weeks; dose = low; design = standardised; delivery = individual; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation method used computer-generated random number table.
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; article stated patients were "blinded form[sic] the treatment they received" but this did not seem possible; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. No information on which therapists provided which treatments; exercise protocols were explicit, and adequately distinct.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Exercise versus exercise; unlikely that either treatment seemed to be better from the patient perspective.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 12 of 120 lost or discontinued
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found; fully reported within this publication
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline except no report of sex distribution or duration of symptoms

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Akhtar 2017 *(Continued)*

Co-interventions avoided or similar (performance bias)	High risk	No information on co-interventions reported in this study
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance, adherence or attendance in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Visual Analogue Scale (for pain) is a well-validated tool in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Akodu 2017

Study characteristics

Methods	**Study design:** RCT **Setting:** Nigeria, healthcare **Exercise groups:** 4 **Comparison groups:** 0
Participants	**Number of participants:** 74 (E1 = 18, E2 = 18, E3 = 20, E4 = 18) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 48 **Sex (female):** 47%
Interventions	**Exercise Group 1 (E1):** Muscle energy technique: undefined posture held against therapist resistance to promote relaxation of the other muscle in the antagonistic pair; type = stretching & strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Core stabilisation; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Exercise Group 3 (E3):** Muscle energy technique, core stabilisation; type = stretching & core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Exercise Group 4 (E4):** Educational booklet and conventional stretches; type = stretching; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was accomplished "via a computer-generated random number sequence".

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Akodu 2017 *(Continued)*

Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments. 2. All 4 arms received some level of activity in their treatment and novel interventions would not be easily accessible outside of the study setting.
Blinding of care provider (performance bias)	High risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Lack of care provider blinding may have caused deviations from intended intervention, as providers had significant contact with all treatment groups, and different expectations of groups' efficacies; 3. If lack of care provider blinding did cause deviations from intended interventions, these deviations would not have been balanced because of different expectations of groups' efficacies; 4. If lack of care provider blinding did cause deviations from intended interventions, we expected overestimation of "better" groups' outcomes and underestimation of control group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for this study were the patients themselves, who could not be blinded to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Muscle energy technique is a relatively novel sounding intervention and some patients may have higher expectations for it than the other interventions.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 7% (5/74).
Participants analysed in group allocated (attrition bias)	Low risk	1. Not explicitly stated but no indication that patients were analysed otherwise; numbers matched up for analysis and allocation.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all outcomes and analyses fully reported; no obvious omissions
Groups similar at baseline (selection bias)	High risk	Baseline similarity between groups on age, body mass index and pain; function was not similar, sex distribution was suspiciously not reported by group, and duration of symptoms was not measured.
Co-interventions avoided or similar (performance bias)	Low risk	No mention of co-interventions in the study
Compliance acceptable in all groups (performance bias)	Low risk	All patients completed the study; seemed indicative of all patients followed up; not all patients completed every session; no information on compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were identical regardless of treatment group 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alaranta 1994

Study characteristics	
Methods	**Study design:** RCT **Setting:** Finland, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 293 (E1 = 141, E2 = 152) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 40 **Sex (female):** 55%
Interventions	**Exercise Group 1 (E1):** Muscle training; pool exercises; indoor and outdoor activities (as part of a multidisciplinary programme); type = mixed; duration = 3 weeks; dose = high; design = standardised; delivery = group; additional intervention = electrotherapy & manual therapy & back school **Exercise Group 2 (E2):** Guided or self-controlled physical exercises including aerobic, strengthening, and stretching exercises; type = mixed; duration = 3 weeks; dose = high; design = individualised; delivery = independent with follow-up; additional intervention = psychological therapy & relaxation
Outcomes	**Core outcomes reported:** Function (Million Index); work (sick-leave days and occupational handicap) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** Financial disclosure/device statement category: 3, 7 **Funding source:** Not reported **Other:** SDs imputed

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	A total of 98% of patients from the baseline examinations participated in the 12-month follow-up.
Participants analysed in group allocated (attrition bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alaranta 1994 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Alayat 2014

Study characteristics

Methods	**Study design:** RCT **Setting:** Saudi Arabia, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 72 (E1 = 28, E2 = 24, C1 = 20) **Chronic LBP duration:** 13.9 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 33 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Exercises for strengthening, stretching, mobilising, co-ordinating, and stabilising abdominal, back, and pelvic muscles; type = core strengthening; duration = 4 weeks; dose = high; design = individualised; delivery = individual; additional intervention = electrotherapy **Exercise Group 2 (E2):** Exercises for strengthening, stretching, mobilising, co-ordinating, and stabilising abdominal, back, and pelvic muscles; type = mixed; duration = 4 weeks; dose = high; design = individualised; delivery = individual; additional intervention = placebo **Comparison Group 1 (C1):** Other conservative treatment (electrotherapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were assigned a specific ID# and randomised using a GraphPad programme.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alayat 2014 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Patients did not know which group they were assigned to or which treatment they would be offered.
Blinding of care provider (performance bias)	High risk	Single-blinded only
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the three treatment groups at baseline.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Family members were responsible to record when the study participant performed the home exercises and all reported no deficiencies.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Albaladejo 2010

Study characteristics	
Methods	**Study design:** Cluster-RCT (NCT00439764) **Setting:** Spain, healthcare **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 348 (E1 = 100, C1 = 109, C2 = 139) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 51 **Sex (female):** 67%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Albaladejo 2010 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Stretching and active exercises for abdominal, lumbar and thoracic back extensors, psoas, ischiotibial, and pelvic muscles; type = core strengthening & stretching; duration = 1 week; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & relaxation **Comparison Group 1 (C1):** Usual care/no treatment (healthy nutrition booklet and 15-minute group talk) **Comparison Group 2 (C2):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (12-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global assessment pain/disability (4 categories)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Regional Government of Castilla-Leon; Kovacs Foundation **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation to the three groups was performed blindly according to a random numbers table.
Allocation concealment (selection bias)	Low risk	Primary care physicians were not informed about which group they had been allocated to.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors said there was blinding, however, we believe this was not possible.
Blinding of care provider (performance bias)	Unclear risk	Primary care physicians were not informed about which group they had been allocated to.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All questionnaires were self-administered and completed by the patients on their own, unaccompanied by healthcare staff or third parties.
Incomplete outcome data (attrition bias) All outcomes	Low risk	There were no losses to follow-up.
Participants analysed in group allocated (attrition bias)	Low risk	There were no losses to follow-up.
Selective reporting (reporting bias)	Low risk	Yes
Groups similar at baseline (selection bias)	Low risk	Groups were very similar
Co-interventions avoided or similar (performance bias)	Low risk	All patients received the usual treatment for low back pain in the Spanish National Health Service as provided by their primary care physician.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Albaladejo 2010 *(Continued)*

Compliance acceptable in all groups (performance bias)	Low risk	In order to reduce the number of losses to follow-up, a treatment co-ordinator blinded to the kind of intervention patients had received, was in charge.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Post-treatment, follow-up the same

Alexandre 2001

Study characteristics	
Methods	**Study design:** RCT **Setting:** Brazil, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 33 (E1 = 14, C1 = 19) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 37 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Multiple components: exercise plus home exercises; type = stretching & strengthening; duration = 16 weeks; dose = high; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 16 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** São Paulo Research Foundation (Process n. 1997/05744-3) **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Support for judgement was not available.
Allocation concealment (selection bias)	Unclear risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias)	Unclear risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alexandre 2001 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Unclear risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Alp 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 48 (E1 = 24, E2 = 24) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 50 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Core stabilisation exercises: stabilisation exercises for the multifidus and transversus abdominis muscles (30 minutes), warm-up (5 minutes), stretching (5 minutes), cool down (5 minutes); type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Conventional exercise: lumbar isometric and lumbar flexion-extension exercises, 1 set of 20 repetitions daily; type = flexibility/mobilising; duration = 6 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 13 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alp 2014 *(Continued)*

Other: Information modified for author contact, SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated table of numbers
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All of the patients completed the programme, and none of the patients had radicular pain.
Participants analysed in group allocated (attrition bias)	Low risk	All of the participants were included in the statistical analysis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	At baseline evaluation, the groups were determined to be homogeneous for patient characteristics and clinical measurements (except 36-Item Short Form Health Survey phydimension)
Co-interventions avoided or similar (performance bias)	Low risk	Patients instructed not to take non-steroidal anti-inflammatory drugs or any analgesic during the intervention period
Compliance acceptable in all groups (performance bias)	Low risk	Two subjects in the core-stabilization exercise group and three subjects in the home-based conventional exercise group missed some of the exercise sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Andrusaitis 2011

Study characteristics

Methods	**Study design:** RCT

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Andrusaitis 2011 *(Continued)*

	Setting: Brazil, not specified **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 10 (E1 = 5, E2 = 5) **Chronic LBP duration:** 7.5 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 46 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Progressive strengthening exercises for abdominal, back, and hip muscles, 3 sets of 10 repetitions; warm-up on bike (10 minutes); type = core strengthening; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = not specified **Exercise Group 2 (E2):** Stabilisation exercises starting with the dorsal decubitus and progressing to the ventral decubitus, in seated, 4-support, and standing positions, warm-up on bike (10 minutes); type = core strengthening; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was performed by means of a draw using opaque envelopes containing folded papers that allocated patients.
Allocation concealment (selection bias)	Low risk	Randomisation was performed by means of a draw using opaque envelopes containing folded papers that allocated patients.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The evaluation protocol was administered one week before the treatment was started and one week after it was terminated by two experienced evaluators.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Andrusaitis 2011 *(Continued)*

Groups similar at baseline (selection bias)	High risk	Patients in Group A presented symptoms of longer duration (a mean of 78 months) than Group B's symptoms (a mean of 7.2 months).
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Arab 2016

Study characteristics	
Methods	**Study design:** RCT (NCT02178202) **Setting:** Iran, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 14 (E1 = 7, E2 = 7) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 37 **Sex (female):** 43%
Interventions	**Exercise Group 1 (E1):** Harmonic mobilisation technique: lower limb lifted by sling, patients asked to move leg side to side; type = flexibility/mobilising; duration = 2 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** End range loading, back extensions; type = core strengthening; duration = 2 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 2 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Used "the blocked randomization method (based on sample size in each group)"
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Arab 2016 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were quite distinct and clearly prescribed.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and function questionnaires are subjective and can be altered by knowledge of intervention; 3. No clear better intervention, both groups receiving active treatment, no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	1. No description of dropout rate; 2. Analysis did not account for missing data; no sensitivity analysis; 3. Increased pain and function could cause missingness; 4. No indication of missing values; 5. No indication of missing values
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. Study analysed and reported as described in methods section; protocol was registered (NCT02178202) but contained no pre-planned outcomes.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline, except did not report duration of symptoms.
Co-interventions avoided or similar (performance bias)	Low risk	"To control the effects of other interventions, the patients were asked to abstain from exercise, medication use, and/or any type of modalities through the period of study".
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Numeric Rating Scale (for pain) and Roland-Morris Disability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Arampatzis 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Germany, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 45 (E1 = 24, C1 = 21) **Chronic LBP duration:** Not specified (not specified)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Arampatzis 2017 *(Continued)*

	Neurological/radicular symptoms: No participants **Mean age (years):** 32 **Sex (female):** 42%
Interventions	**Exercise Group 1 (E1):** Seventy-minute abdominal perturbation training as stabilisation exercises with 10 minutes whole body exercises and 10 minutes cool down exercises; type = core strengthening; duration = 13 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Low Back Pain Rating Scale) **Follow-up time periods available for syntheses:** 13 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** German Federal Institute of Sport Science; MiSpEx National Research Network for Medicine in Spine Exercise (Grant Number 080102A/11-14) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Block random-number procedure
Allocation concealment (selection bias)	Low risk	Treatment allocation was likely concealed; the randomisation procedure was controlled by a member of the same institution, who was not involved in the design and execution of the study.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Exercise therapy required machinery that would not be accessible outside of study context.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Therapists seemed to have had minimal contact with control group; exercise therapy so distinct and specific that no deviation could casually occur.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors in this study were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Likely that lack of patient blinding led to perception that experimental group was better than control, and so biased the outcome assessments accordingly; pain study results supported this.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 5 patients of 45 did not finish the study, 11%.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the treatment group to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this paper all outcomes and analyses fully reported, functional status not measured, but isometric measurements likely took place.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Arampatzis 2017 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar at baseline on age, sex, pain, height and weight, but there was no reporting of duration of symptoms.
Co-interventions avoided or similar (performance bias)	Low risk	Co-interventions were largely avoided, as the study excluded patients with continuous dependency on pain relief medication or physiotherapist treatment.
Compliance acceptable in all groups (performance bias)	Low risk	No information on compliance/adherence/attendance to sessions was reported.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were identical, regardless of treatment group; 2. Visual Analogue Scale (for pain) is a well-validated tool in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Areeudomwong 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Thailand, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 42 (E1 = 21, C1 = 21) **Chronic LBP duration:** 9.6 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 74%
Interventions	**Exercise Group 1 (E1):** Proprioceptive neuromuscular facilitation training focussed on trunk muscles; type = core strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 4 weeks (short); 16 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Mae Fah Luang University, Chiang Rai, Thailand **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	"Block randomization, with block sizes of two, four and six"
Allocation concealment (selection bias)	Low risk	"Allocation results were concealed in sealed and opaque envelopes with consecutive numbering".

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Areeudomwong 2017 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were quite distinct and they had little contact with control group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Possible that lack of patient blinding could have caused bias in outcome assessment because the exercise programmes could be perceived as better than the control; outcome data supported this.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. None of the patients were lost to follow-up at any time point.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: all planned analyses were executed and reported for all primary and secondary outcomes.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	No report of co-interventions in this study
Compliance acceptable in all groups (performance bias)	Low risk	All patients in the experimental intervention completed training for 4 weeks; "all participants participated throughout the study period".
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Numeric Rating Scale (for pain) and Roland-Morris Disability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Atalay 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 20 (E1 = 10, E2 = 10)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Atalay 2017 *(Continued)*

	Chronic LBP duration: 31.2 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 25 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Conventional lumbar exercises (core strengthening) plus exercises of the upper back, neck, and shoulders (strengthening), also static stretching; type = mixed; duration = 6 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none **Exercise Group 2 (E2):** Conventional stretching, isometric back exercises, mobilisation, stabilisation, and back and abdominal concentric strengthening exercises; type = stretching & core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Department of the Scientific Research Projects of Uludag University (Project Number = HDP(T)-2014/4) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Drawing lots
Allocation concealment (selection bias)	Low risk	Treatment allocation was likely concealed; it appeared that patients directly drew their own lot, and so research personnel could not interfere with this process.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from intended interventions, as they were highly structured and required delivery by the trained exercise specialist.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from intended intervention, as the treatments were all distinct, structured, and none was clearly better than others.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were patients themselves, who could not be blinded to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Neither treatment was obviously better than the other; no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate not described; it seemed like data might have been presented for all participants in Table 3; 2. Wilcoxon matched-pair test dropped missing, no sensitivity analyses; 3. Missingness could be caused by increased disability from low back pain; 4. No information on dropout or missing data reported; intervention groups were so similar that it was unlikely that any missing data would be related to treatment/outcome.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that patients were analysed according to the treatment to which they were randomised

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Atalay 2017 *(Continued)*

Selective reporting (reporting bias)	Low risk	1. No statistical analysis plan or protocol found; all planned analyses were executed and reported for all primary and secondary outcomes; pain and function were not fully reported; only graphs given.
Groups similar at baseline (selection bias)	Low risk	Characteristics that were collected and were similar; no indication of problematic randomisation
Co-interventions avoided or similar (performance bias)	Low risk	No mention of co-interventions was reported in the study; given the similarity of interventions in each group and baseline similarity, use of co-interventions was likely similar.
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance/adherence/attendance was reported in this study.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were identical, regardless of treatment group; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Very small trial, under-reported results for pain and function; only graphs given

Aure 2003

Study characteristics	
Methods	**Study design:** RCT **Setting:** Norway, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 49 (E1 = 22, E2 = 27) **Chronic LBP duration:** 13 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 47%
Interventions	**Exercise Group 1 (E1):** Strengthening, stretching, mobilising, co-ordination, stabilising exercises for abdominal, back, and pelvic muscles and lower limbs using equipment; type = mixed; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = not specified **Exercise Group 2 (E2):** Stretching exercises; type = stretching; duration = 8 weeks; dose = low; design = individualised; delivery = independent; additional intervention = manual therapy & anti-inflammatory/analgesics
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); work (self-reported return to work) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Foundation for Education and Research In Physiotherapy, Norway **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Aure 2003 *(Continued)*

Random sequence generation (selection bias)	Low risk	Pre-sealed envelopes provided by external research corporation
Allocation concealment (selection bias)	Low risk	Evaluating two common physiotherapy treatments
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Self-report, except range of motion, which was blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Two/27 in manual therapy group, 1/22 in exercise therapy group dropouts
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	Yes, during study period but not controlled in follow-up period
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Azevedo 2017

Study characteristics	
Methods	**Study design:** RCT (NCT02221609) **Setting:** Brazil, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 148 (E1 = 74, E2 = 74) **Chronic LBP duration:** 85.05 months (long) **Neurological/radicular symptoms:** Some participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Azevedo 2017 *(Continued)*

	Mean age (years): 42 **Sex (female):** 61%
Interventions	**Exercise Group 1 (E1):** Movement system impairment-based classification treatment: involved patient education, analysis and modification of daily activities, and prescription of specific exercises; type = functional restoration; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Strengthening and stretching of the core and lower limbs; type = strengthening & stretching; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 8.7 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Conselho Nacional de Desenvolvimento Cientifico e Tecnologico/Brazil (CNPQ grant number 470273/2013-5); Coordenacao de Aperfeicoamento de Pessoal de Nıvel Superior **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation method entailed "generating a randomization schedule in Excel".
Allocation concealment (selection bias)	Low risk	Allocations were concealed using sealed, opaque, and sequentially numbered envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Protocol and progression of exercises was provided in the appendix; one arm involved movement and alignment information which could have been provided to the other group in casual physiotherapy sessions.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Neither treatment is obviously better than the other; no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Loss to follow-up at each time point was only 2% (3/148).
Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was conducted.
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported according to registered protocol (NCT02221609).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Azevedo 2017 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	High risk	No information on co-interventions reported in this study
Compliance acceptable in all groups (performance bias)	High risk	Mean adherence was 67% and 49.3% in the comparator and experimental intervention, respectively.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Numeric Rating Scale (for pain) and Roland-Morris Disability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	No other sources of bias noticed; generally well-reported

Bade 2017

Study characteristics	
Methods	**Study design:** RCT (NCT01900925) **Setting:** USA and Chile, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 90 (E1 = 47, C1 = 43) **Chronic LBP duration:** 18.4 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 46 **Sex (female):** 41%
Interventions	**Exercise Group 1 (E1):** Usual low back pain physiotherapy care as determined by clinician (manual therapy, trunk strengthening, directional preference, flexion, mobilisation, traction, counselling, endurance) with hip strengthening exercises and hip mobilisations; type = core strengthening; duration = not specified weeks; dose = not reported; design = partially individualised; delivery = individual; additional intervention = advice/education & psychological therapy & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Rating of Change Scale)) **Follow-up time periods available for syntheses:** 2 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation used "the roll of a dice by the treating therapist; for allocation, die numbers of 1, 3, and 5 were allocated to one group; die numbers 2, 4, and 6 were allocated to the other group".

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bade 2017 *(Continued)*

Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Pragmatic part of intervention based on guidelines; explicitly instructed not to perform hip therapy to control group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Possible that lack of patient blinding could have caused bias in outcome assessment because the exercise programmes could be perceived as better than the control; some outcome data supported this.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate at 2 weeks (the only usable time point) was 7% (6/90).
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported according to registered protocol (NCT01900925): explicitly stated that "recovery expectations" was dropped due to lack of cross-cultural applicability.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Not enough information presented to assess co-intervention use; anyone using steroidal medication was excluded.
Compliance acceptable in all groups (performance bias)	High risk	Difficult to assess as discharge was variable and there was no set number of sessions to attend
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function) and Numeric Rating Scale (for pain) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Barberini 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, occupational **Exercise groups:** 1 **Comparison groups:** 1

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Barberini 2011 *(Continued)*

Participants	**Number of participants:** 22 (E1 = 10, C1 = 12) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 47 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Exercising in water and general exercises (as part of Back School); type = mixed; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (back school)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Support for judgement was not available.
Allocation concealment (selection bias)	Unclear risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	High risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Barberini 2011 *(Continued)*

Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.
Other bias	Unclear risk	Not available: non-English publication

Bellido-Fernandez 2018

Study characteristics	
Methods	**Study design:** RCT (NCT02721914) **Setting:** Spain, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 27 (E1 = 9, E2 = 9, C1 = 9) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 29 **Sex (female):** 85%
Interventions	**Exercise Group 1 (E1):** Combined abdominal hypopressive gymnastics and massage therapy: abdominal hypopressive gymnastics are static postural exercises that aim to strengthen the abdominal girdle and stabilise the spine; type = core strengthening; duration = 5 weeks; dose = low; design = standardised; delivery = individual; additional intervention = manual therapy **Exercise Group 2 (E2):** Abdominal hypopressive gymnastics alone, static postural exercises that aim to strengthen the abdominal girdle and stabilise the spine; type = core strengthening; duration = 5 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index); HRQoL (12-Item Short Form Survey (Spanish)) **Follow-up time periods available for syntheses:** 5 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Simple random sampling, "drawing from a hat"
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed, as patients directly selected their random assignment from an opaque container.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to the intervention due to the nature of the difference interventions (massage versus exercise); 2. Both groups getting sufficient treatment, unlikely to be dissatisfied; experimental hard to access outside of study.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bellido-Fernandez 2018 *(Continued)*

Blinding of care provider (performance bias)	Low risk	1. Care provider could not be blinded because a single physiotherapist performed the treatments for all intervention groups; 2. Explicit set of exercises in protocol
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain intensity and disability were the patients themselves, who could not be blinded to intervention due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Likely that massage group was perceived as more effective because it was more direct.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. None of the patients dropped out of the study; all who were randomised were analysed.
Participants analysed in group allocated (attrition bias)	Low risk	1. "The effectiveness of the three applied interventions was examined by the intention-to-treat method".
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistic analysis plan found: within this publication all outcomes and analyses were fully reported; no obvious lapses.
Groups similar at baseline (selection bias)	Low risk	Groups were relatively similar at baseline for age, sex, function and pain intensity; did not report baseline duration of symptoms.
Co-interventions avoided or similar (performance bias)	Low risk	One of the exclusion criteria was being under pharmacological treatment (co-intervention), though it did not specify whether this treatment had to be directly for low back pain.
Compliance acceptable in all groups (performance bias)	Low risk	No direct reporting of compliance, though there seemed to be implication that attendance to treatment sessions was 100%.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All participants were assessed on pain intensity and function identically, regardless of treatment group assignment; 2. Numeric rating scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Bendix 1995

Study characteristics	
Methods	**Study design:** RCT **Setting:** Denmark, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 106 (E1 = 40, E2 = 31, C1 = 35) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** 75%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bendix 1995 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Functional restoration: comprehensive multidisciplinary approach including aerobics, strengthening, stretching; type = functional restoration; duration = 3 weeks; dose = high; design = standardised; delivery = group; additional intervention = psychological therapy **Exercise Group 2 (E2):** Aerobics and strengthening; type = aerobic & strengthening; duration = 6 weeks; dose = high; design = standardised; delivery = group; additional intervention = back school **Comparison Group 1 (C1):** Other conservative treatment (psychological therapy)
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Activities of Daily Living Scale); work (work readiness) **Follow-up time periods available for syntheses:** 17 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Danish Rheumatism Association, Nycomed-DAK; AP Moller and Wife's Foundation; Pensam; Danish Insurance Association; Meyer's Foundation; Minister Ema Hamilton's Foundation; Director Ib Henriksen's Foundation; Research Foundation of Copenhagen University; Hafnia Foundation; Peter Ryholt's Foundation; Ingrid Munkholm's Foundation; Danish Society for Manual Medicine; Lily Benthine Lund's Foundation **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bendix 1995 *(Continued)*

Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Bendix 2000

Study characteristics	
Methods	**Study design:** RCT **Setting:** Denmark, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 138 (E1 = 64, E2 = 74) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 41 **Sex (female):** 65%
Interventions	**Exercise Group 1 (E1):** Functional restoration: comprehensive multidisciplinary approach including aerobics, strengthening, stretching; type = functional restoration; duration = 3 weeks; dose = high; design = standardised; delivery = group; additional intervention = psychological therapy & back school **Exercise Group 2 (E2):** Aerobics and strengthening (machines); type = aerobic & strengthening; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Low Back Pain Rating Scale (Manniche)); work (sick leave days); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global effect on quality of life)) **Follow-up time periods available for syntheses:** 52 weeks (long)
Notes	**Conflicts of interest:** Conflict of interest category: 14 **Funding source:** Danish Rheumatism Association; Gerda and Aage Hensch Foundation; Director Ib Henriksen's Fund; Insurance Company for Industrial Injuries; Lilly Benthine Lunds Fund; DANICA Pension; Municipal Pension Insurance Company Ltd.; Danish Society for Manual Medicine **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bendix 2000 *(Continued)*

Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	High risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Ben-Salah 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Tunisia, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 107 (E1 = 54, E2 = 53) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 75%
Interventions	**Exercise Group 1 (E1):** Eighteen exercises: 4 self-positioning exercises (2 in extension and 2 in flexion); 8 muscle stretching exercises; 4 abdominal and truck strengthening exercises; type = mixed; duration = 4 weeks; dose = high; design = standardised; delivery = independent; additional intervention = none **Exercise Group 2 (E2):** Flexibility training, stretching, proprioception exercises and muscle strengthening exercises (as part of standard rehabilitation therapy); type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = individual; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Quebec Back Pain Disability Scale)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ben-Salah 2009 *(Continued)*

	Follow-up time periods available for syntheses: 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised, prospective clinical trial with two parallel groups
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the two groups in terms of epidemiological factors (age, gender, history of spinal surgery, sports)
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Unclear risk	Compliance with the home-based exercise programme was good: 87.8% at the end of the supervised programme (t4) and 88.6% at the M3 consultation.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bentsen 1997

Study characteristics	
Methods	**Study design:** RCT **Setting:** Sweden, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 74 (E1 = 41, E2 = 33) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 57 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Dynamic strengthening back exercises: 3 months at gym, 9 months at home; type = core strengthening; duration = 26 weeks; dose = high; design = standardised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Home exercises; type = strengthening; duration = 26 weeks; dose = high; design = standardised; delivery = independent; additional intervention = none
Outcomes	**Core outcomes reported:** Function (Million Index); work (sick leave days in previous 12 months) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Device status category: 1 **Funding source:** AMF trygghetsforsakring, Stockholm, Sweden **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bentsen 1997 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Bhadauria 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 45 (E1 = 15, E2 = 15, E3 = 15) **Chronic LBP duration:** 9.8 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 35 **Sex (female):** 33%
Interventions	**Exercise Group 1 (E1):** Pilates: majority of exercises are abdominal or back focussed with warm-up and cool-down stretching exercises; type = Pilates; duration = 1.43 weeks; dose = low; design = standardised; delivery = individual; additional intervention = electrotherapy & heat/ice **Exercise Group 2 (E2):** Lumbar stabilisation: photographs of all exercises performed focussed on trunk-strength, warm-up and cool-down stretching exercises; type = core strengthening; duration = 1.43 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = electrotherapy & heat/ice **Exercise Group 3 (E3):** Lumbar dynamic strengthening focussed on abdominals and lumbar muscles; type = core strengthening; duration = 1.43 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = electrotherapy & heat/ice
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 1.43 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** SDs imputed, sufficient data not available for inclusion in meta-analyses

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Patients were "randomly allocated".

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bhadauria 2017 *(Continued)*

Allocation concealment (selection bias)	Low risk	"Randomly allocated in three groups by envelop method"
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were all quite distinct.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and function questionnaires require judgement which could be susceptible to a lack of blinding; 3. No treatment was obviously better than the others; no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 8 of 44 dropped out, 18%.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all outcomes and analyses were fully reported.
Groups similar at baseline (selection bias)	High risk	Pilates group had significantly longer duration of symptoms (1 year longer than other two groups), and almost significantly higher average body mass index (3-4 points higher).
Co-interventions avoided or similar (performance bias)	High risk	No information presented about co-interventions
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance, adherence or attendance presented in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Bi 2013

Study characteristics	
Methods	**Study design:** RCT **Setting:** China, healthcare **Exercise groups:** 2 **Comparison groups:** 0

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bi 2013 *(Continued)*

Participants	**Number of participants:** 47 (E1 = 23, E2 = 24) **Chronic LBP duration:** 7.9 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 30 **Sex (female):** 45%
Interventions	**Exercise Group 1 (E1):** Lumbar strengthening and pelvic floor exercises; type = strengthening; duration = 24 weeks; dose = high; design = standardised; delivery = group; additional intervention = electrotherapy **Exercise Group 2 (E2):** Lumbar strengthening exercises; type = strengthening; duration = 24 weeks; dose = high; design = standardised; delivery = group; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 24 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Science and Technology Development Fund of Shanghai Pudong (PKJ2008-Y39); Program of Shanghai Pudong Subject Chief Scientist (PWRd2010-06); Science and Technology Development Fund of Shanghai Pudong (PKJ2011-Y05) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No info given other than saying the trial was "double-blind, randomized, controlled"
Allocation concealment (selection bias)	Low risk	Used opaque closed envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	In flow chart
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no statistically significant between-group differences in any characteristic at baseline.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bi 2013 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Bid 2017

Study characteristics

Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 128 (E1 = 64, C1 = 64) **Chronic LBP duration:** 41.58 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 41 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** McKenzie exercise programme with no reported modifications; type = McKenzie; duration = 8 weeks; dose = high; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Rating of Change Scale)) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Hajee A.M. Lockhat & Dr. A.M. Moolla Sarvajanik Hospital, Rampura, Surat **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Method of randomisation was described as a "lottery", which if carried out correctly, would adequately randomise participants.
Allocation concealment (selection bias)	High risk	No information reported as to whether or not treatment allocation was concealed adequately
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Unlikely that patients could be kept blinded to their allocation, given how different the two treatment groups were; 2. Despite the treatment groups being so different, patients likely couldn't have effected a change in intervention protocol, as the interventions depended heavily on facilitation by trained personnel.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bid 2017 *(Continued)*

Blinding of care provider (performance bias)	Low risk	1. Care providers (physiotherapist) could not have been blinded because they delivered either the McKenzie programme or the conventional physiotherapy programme; 2. Explicit set of exercises in protocol
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and disability were the patients, and thus were likely not able to be blinded to their intervention allocation; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Mackenzie is tailored and sounds much better than "conventional" physiotherapy; likely altered expectations of effectiveness and assessment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No description of dropout rate; 2. No evidence analysis corrected for missing data, no sensitivity analyses; 3. Increased low back pain or decreased function can cause missingness; 4. No information
Participants analysed in group allocated (attrition bias)	Low risk	1. Though not explicitly reported, it appeared that patients were analysed according to the treatment group to which they were allocated.
Selective reporting (reporting bias)	Low risk	1. No registered protocol, however all analysis described in the methods were reported.
Groups similar at baseline (selection bias)	Low risk	No significant differences in baseline age, sex, body mass index, pain duration, pain or disability; significantly more patients with central sensitization in experimental group
Co-interventions avoided or similar (performance bias)	Low risk	Study exclusion criteria explicitly excluded patients who were receiving conflicting or ongoing co-interventions (though not specified what counted as a co-intervention).
Compliance acceptable in all groups (performance bias)	Low risk	Compliance (attendance) was not reported in the study, and so not able to determine exact compliance.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Within each outcome, assessments were performed with identical tools and at the same time points; 2. Numeric Pain Rating Scale (for pain) and Roland-Morris Disability Questionnaire (for disability) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Blodt 2015

Study characteristics	
Methods	**Study design:** RCT (NCT00993980) **Setting:** Germany, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 127 (E1 = 64, E2 = 63) **Chronic LBP duration:** 2.95 years (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 47 **Sex (female):** 81%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Blodt 2015 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Qigong: 21 movements for spine and legs (14 in level 1, 7 in level 2); type = other (Tai Chi, Qigong); duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Dynamic gym ball exercises followed by strengthening and stretching exercises; type = mixed; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = relaxation
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Karl and Veronica Carstens Foundation **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Block randomisation with list from the 'ranuni' random number generator of the SAS/Stat software
Allocation concealment (selection bias)	Low risk	Randomised using a button
Blinding of participants and personnel (performance bias) All outcomes	Low risk	In the recruitment and patients' information, both interventions were described as similar useful and plausible for treating low back pain.
Blinding of care provider (performance bias)	High risk	Neither patients nor therapists were blinded to the treatment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	In the recruitment and patients' information, both interventions were described as similarly useful and plausible for treating low back pain.
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1
Participants analysed in group allocated (attrition bias)	Low risk	Analysis was based on available data (full analysis set based on intention-to-treat).
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	High risk	More women in the Qigong group; Qigong group had higher school education; Qigong participants had shorter LBP duration.
Co-interventions avoided or similar (performance bias)	Low risk	Same medication as usual allowed; no complimentary treatments

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Blodt 2015 *(Continued)*

Compliance acceptable in all groups (performance bias)	Low risk	Described as "high"
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Both groups assessed at three, six, and 12 months

Bramberg 2017

Study characteristics	
Methods	**Study design:** RCT (NCT01653782) **Setting:** Sweden, mixed **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 159 (E1 = 52, E2 = 52, C1 = 55) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 46 **Sex (female):** 71%
Interventions	**Exercise Group 1 (E1):** Kundalini Yoga; type = yoga; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Strength training: focusses on core muscle strengthening and stabilisation; type = core strengthening; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Chronic Pain Grade Scale - Pain Subscale); function (Chronic Pain Grade Scale- Disability Subscale); work (sickness absence, sickness presenteeism) **Follow-up time periods available for syntheses:** 24 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Swedish Research Council for Health, Working life and Welfare (2008-0849) **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	"A random allocation sequence was generated by the statistician".
Allocation concealment (selection bias)	Low risk	"For each participant an opaque envelope was opened, in consecutive order, by an external research assistant not involved in the inclusion process".
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients were not blinded to their assigned intervention; 2. Patients did not know the content of the different intervention arms; no report of external activity.
Blinding of care provider (performance bias)	Low risk	1. "The yoga leader and physiotherapist were not blinded"; 2. Lack of care provider blinding probably did not cause deviations from intended interventions because each intervention was clearly defined and distinct.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bramberg 2017 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	1. Outcomes in question were assessed by the patient who was not blinded to the intervention; 2. Patients did not know the content of the different intervention arms; no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 26% (40/152) which introduces a risk of bias; 2. Used last-value-carried-forward method, and also relative frequency imputation; not valid methods, artificially improve precision and do not necessarily remove bias; 3. No significant differences between those who were lost to follow-up and participants in terms of age, sex, or pre-intervention values on neck and back pain
Participants analysed in group allocated (attrition bias)	Low risk	1. "An intention-to-treat analysis was conducted which included all randomized individuals, irrespective of whether they had adhered to the intervention programme or not".
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported according to registered protocol (NCT01653782), with some additional outcomes (pain intensity at 6 and 12 months).
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline except duration of symptoms was not reported.
Co-interventions avoided or similar (performance bias)	High risk	Some element of co-intervention was encouraged in the control group based on educational materials.
Compliance acceptable in all groups (performance bias)	High risk	Proportion of participants who adhered to the recommendations (exercised at least 2 times/week) during the 6-month follow-up was: 54% (yoga), 34% (strength training) and 42% (control group).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Chronic Pain Grade Scale analysed using the von Korff protocol (for pain) is a well-validated tool in the low back pain context.
Other bias	Low risk	No other sources of bias noticed; generally well-reported

Bronfort 1996

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 174 (E1 = 71, E2 = 52, C1 = 51) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 41 **Sex (female):** 47%
Interventions	**Exercise Group 1 (E1):** Dynamic trunk (Manniche) and abdominal strengthening; type = core strengthening; duration = 11 weeks; dose = high; design = standardised; delivery = individual; additional intervention = manual therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bronfort 1996 *(Continued)*

	Exercise Group 2 (E2): Dynamic trunk (Manniche) and abdominal strengthening; type = core strengthening; duration = 11 weeks; dose = high; design = standardised; delivery = individual; additional intervention = anti-inflammatory/analgesics **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 5 weeks (short); 11 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Foundation for Chiropractic Education and Research Award (#91-3-1) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bronfort 1996 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Bronfort 2011

Study characteristics

Methods	**Study design:** RCT (NCT00269347) **Setting:** USA, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 301 (E1 = 100, E2 = 101, C1 = 100) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 61%
Interventions	**Exercise Group 1 (E1):** Dynamic trunk strengthening exercises (trunk extensions and leg extensions), abdominal exercises using low-technology methods; type = core strengthening; duration = 12 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = not specified **Exercise Group 2 (E2):** Simple stretching and strengthening exercises, including lumbar extension, bridging, and abdominal crunches; type = stretching & strengthening; duration = 12 weeks; dose = low; design = individualised; delivery = independent with follow-up; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Low Back Pain Rating Scale (Manniche)); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global improvement (9-point)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** RHG: consulting: Pfizer (B); speaking/teaching arrangements: Merck (B), Takeda (B); Scientific Advisory Board: Pfizer (B); research support (investigator salary): Roche (B); Grants: National Institute on Aging (B). TAG: royalties: MSD (F); fellowship support: Synthes (F), Stryker (F), Abbot (F), MSD (F). EET: royalties: Medtronic (F); consulting: Medtronic (F); speaking/teaching arrangements: Stryker (B); trips/travel: Medtronic (A); Scientific Advisory Board: United Health Care (B); fellowship support: Medtronic (E, paid to institution/employer), Synthes Spine (E, paid to institution/employer), Zimmer Spine (C, paid to institution/employer) **Funding source:** Medtronic; Synthes Spine; Zimmer Spine **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Restricted randomisation using a 1:1:1 allocation ratio was applied using four strata.
Allocation concealment (selection bias)	Low risk	Before the trial, the project statistician generated a randomisation list using randomly mixed permuted blocks of different sizes.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bronfort 2011 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	The active interventions used in this trial made blinding patients and providers to treatment type impossible.
Blinding of care provider (performance bias)	High risk	The active interventions used in this trial made blinding patients and providers to treatment type impossible.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Objective outcome assessment was performed by examiners masked to treatment allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Overall attrition rate was 19%, with no differences among treatment groups
Participants analysed in group allocated (attrition bias)	Low risk	All analyses used the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Randomisation resulted in three groups comparable on most baseline variables.
Co-interventions avoided or similar (performance bias)	Low risk	Author response - documented care provided to ensure it met protocol
Compliance acceptable in all groups (performance bias)	Low risk	Overall, adherence to study interventions was high with 96% of the spinal manipulative therapy group, 86% of the supervised exercise therapy group, and 96% of the home exercise and advice group.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Cabak 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Poland, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 68 (E1 = 22, C1 = 46) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 59 **Sex (female):** 72%
Interventions	**Exercise Group 1 (E1):** Physiotherapist teaches exercises movements and recommends 5 minutes of exercise each day at home; type = stretching; duration = 9 weeks; dose = low; design = standardised;

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cabak 2017 *(Continued)*

	delivery = independent with follow-up; additional intervention = advice/education & psychological therapy & manual therapy & anti-inflammatory/analgesics **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey (Polish)) **Follow-up time periods available for syntheses:** 9 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were "randomly assigned" to groups.
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the treatment groups; 2. Unlikely that lack of patient blinding altered the intended intervention, as the treatments were entirely in control of care providers.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding altered intended intervention, as care providers had very little contact with control group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors in this study were patients, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Reasonably likely that patients saw the experimental group as "better" than the control group (no intervention at all); all outcomes favoured the experimental group.
Incomplete outcome data (attrition bias) All outcomes	High risk	1. No description of dropout rate; 2. No evidence of missing data, much less of any biased analyses; 3. Impossible to guess whether an outcome value's missingness depended on its true value, as there was no information as to whether or not there was any missing data in the first place; 4. No evidence as to whether or not there was any missing data in each group; 5. It was impossible to guess whether an outcome value's missingness depended on its true value, as there was no information as to whether or not there was any missing data in the first place.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the treatment to which they were randomised.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found; within this publication all outcomes and analyses were fully reported.
Groups similar at baseline (selection bias)	High risk	While age and sex were reported and were similar between treatment groups, there was no report of baseline body mass index, pain, function or duration of symptoms.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cabak 2017 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	Intervention group could not use any additional therapies or medical consultations during the study period; control patients received usual care without any limitations.
Compliance acceptable in all groups (performance bias)	Low risk	It appeared that all participants in the intervention group attended all three mandated physiotherapy sessions (100% compliance).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All patients were assessed identically on outcomes, regardless of treatment group; 2. Visual Analogue Scale (for pain) and Roland-Morris Disability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Cai 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Singapore, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 84 (E1 = 28, E2 = 28, E3 = 28) **Chronic LBP duration:** 65.4 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 27 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Lumbar stabilisation exercises focussing on transverse abdominal and lumbar multifidus muscles, with progression to have resistance from upper limb and less supported standing base with warm-up (general stretching and stationary bicycling for 15 minutes); type = core strengthening; duration = 8 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Lumbar extensor exercises with wrist and ankle weights increasing as progression of the exercise intensity; type = core strengthening; duration = 8 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 3 (E3):** Lower limb exercises: knee and hip extensor and hip abductor strengthening exercise, warm-up (general stretching exercises and stationary bicycling for 15 minutes); type = strengthening; duration = 8 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Patient Specific Functional Scale (running)) **Follow-up time periods available for syntheses:** 13 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Institute of Education Academic Research Fund, Singapore **Other:** Information modified for author contact; SDs imputed

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomisation was performed "using a block randomization table", which was likely generated with an adequately random method.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cai 2017 *(Continued)*

Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention; interventions were all quite similar.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were all quite structured with distinct protocols.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. No treatment was obviously better than the others, all getting exercise; no reason to alter patient response; patients were also blinded to their previous answer.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 12% of patients were lost to follow-up (10/84), which did not cause concern over risk of bias.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: all planned analyses were executed and reported for all primary and secondary outcomes.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Study excluded from enrolment if they used pain medications.
Compliance acceptable in all groups (performance bias)	High risk	Participants attended an average of 56 supervised sessions (out of 16 requested = 35%); also only completed an average of 293 (out of 40 requested = 73%) home exercise sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Numeric Rating Scale (for pain) is a well-validated tool in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Cairns 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 2 **Comparison groups:** 0

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cairns 2006 *(Continued)*

Participants	**Number of participants:** 97 (E1 = 47, E2 = 50) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 39 **Sex (female):** 51%
Interventions	**Exercise Group 1 (E1):** Endurance training for deep abdominal and multifidus muscles; type = core strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = electrotherapy & manual therapy **Exercise Group 2 (E2):** Exercise described as 'conventional' treatment; type = mixed; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = electrotherapy & manual therapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	An adaptive stratified randomisation procedure was used incorporating minimization, using laterality of symptoms.
Allocation concealment (selection bias)	Low risk	Patients were naive to allocation.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Double-blinding was not possible, so single-blinding, with a credible alternate treatment was used.
Blinding of care provider (performance bias)	High risk	Double-blinding was not possible.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Double-blinding was not possible, so-single blinding, with a credible alternate treatment, was used.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Of 97 patients, nine patients, evenly spread across both groups, did not complete a course of treatment.
Participants analysed in group allocated (attrition bias)	Low risk	For the intention-to-treat analysis, data of patients who withdrew or failed to respond to follow-up were included until they withdrew.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The clinical and demographic characteristics of the groups were well balanced at randomisation.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cairns 2006 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	No restriction was placed on prescribed or over-the-counter medication but hydrotherapy, back school, or other group therapy was prohibited.
Compliance acceptable in all groups (performance bias)	Unclear risk	Although measured informally, which is representative of much clinical practice, it was a limitation of the study that compliance was not formally measured.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Calmels 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** France, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 17 (E1 = 9, E2 = 8) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Isokinetic strengthening exercises (Cybex machines); type = strengthening; duration = 2 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Physiotherapy exercises: series of 3 groups of exercises (whole body); type = stretching; duration = 2 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = manual therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Quebec Back Pain Disability Scale) **Follow-up time periods available for syntheses:** 2 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Calmels 2004 *(Continued)*

Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Unclear risk	Support for judgement is not available

Cambron 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 235 (E1 = 112, C1 = 123) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 37%
Interventions	**Exercise Group 1 (E1):** Strength, flexibility, and cardiovascular exercises to strengthen muscles and improve trunk flexibility; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cambron 2006 *(Continued)*

	Follow-up time periods available for syntheses: 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Health Resources and Services Administration (Grant #R18 AH 10001) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Subjects were randomised to one of two forms of treatment.
Allocation concealment (selection bias)	Low risk	Used manilla envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	High risk	Overall, 61 subjects withdrew from the study: 38 during the trial intervention (13 flexion distraction and 25 exercise program) and 23 during the follow-up (14 flexion distraction, 9 exercise program).
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Patient demographics and baseline characteristics of the two groups were compared.
Co-interventions avoided or similar (performance bias)	High risk	Subjects were asked not to take any pain-reducing medication prior to data collection at weeks zero and five.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Carr 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 237 (E1 = 118, E2 = 119) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 61%
Interventions	**Exercise Group 1 (E1):** Low impact aerobics, strengthening and stretching exercises for main muscle groups; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = psychological therapy & relaxation **Exercise Group 2 (E2):** One or more of: McKenzie (68%), strengthening (15%), stretching (18%), core (11%), other (12%) (as part of physiotherapy); type = mixed; duration = 4 weeks; dose = low; design = individualised; delivery = individual; additional intervention = other & physiotherapy
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (12-Item Short Form Survey; EuroQol 5D) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A pre-prepared list of random numbers that had been stratified by physiotherapy departments in blocks of six were used for randomisation.
Allocation concealment (selection bias)	Low risk	Patients included in the study were randomised, using sealed envelopes, either to the Back to Fitness programme or to individual physiotherapy.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Physiotherapy assessors were blind to treatment allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Carr 2005 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The clinical and demographic characteristics of the patients in the two groups were fairly well balanced at randomisation.
Co-interventions avoided or similar (performance bias)	Unclear risk	Individual physical therapy included co-interventions but Group ex group did not.
Compliance acceptable in all groups (performance bias)	High risk	Did subanalysis and excluded non-attenders and this did not 'notably' change the results.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Cecchi 2010

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 210 (E1 = 70, E2 = 70, C1 = 70) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 59 **Sex (female):** 67%
Interventions	**Exercise Group 1 (E1):** Postural, respiratory, and back exercises; type = mixed; duration = 3 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & relaxation **Exercise Group 2 (E2):** Active movement exercises selected by physiotherapist; type = mixed; duration = 3 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & other & physiotherapy **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Roland-Morris Disability Questionnaire); work (sick leave (did you lose any working days because of low back pain in the past six months?)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Fondazione Don Gnocchi Foundation, Scientific Institute **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Once enrolled, each patient was given a progressive number and the secretary of the outpatient department assigned patients to their treatment group.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cecchi 2010 *(Continued)*

Allocation concealment (selection bias)	Low risk	The enrolling physician was blind as to which number corresponded to which treatment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Patients could not be blinded.
Blinding of care provider (performance bias)	High risk	Patients could not be blinded.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Baseline and follow-up questionnaires were administered by three independent blinded interviewers at the outpatient department.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Among those assigned to back school, two patients discontinued the treatment, one because of concurrent illness and one because of family problems.
Participants analysed in group allocated (attrition bias)	High risk	Five patients lost to follow-up
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences
Co-interventions avoided or similar (performance bias)	High risk	Table 3 shows low back pain recurrences, low back pain-related use of drugs and request of further treatment for low back pain.
Compliance acceptable in all groups (performance bias)	Low risk	All participants who completed back school and individual physiotherapy attended at least 12 of the 15 sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Chan 2011

Study characteristics	
Methods	**Study design:** RCT (ISRCTN23753357) **Setting:** Hong Kong, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 46 (E1 = 24, E2 = 22) **Chronic LBP duration:** 13 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 47 **Sex (female):** 78%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chan 2011 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Back mobilisation exercises, abdominal stabilisation exercises, aerobic training programme, home exercises; type = mixed; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = other & physiotherapy **Exercise Group 2 (E2):** General exercise; type = mixed; duration = 8 weeks; dose = low; design = individualised; delivery = independent; additional intervention = other & physiotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Aberdeen Back Pain Scale) **Follow-up time periods available for syntheses:** 8 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Department of Rehabilitation Sciences, Hong Kong Polytechnic University; Department of Physiotherapy, David Trench Rehabilitation Centre **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Subjects were randomly allocated.
Allocation concealment (selection bias)	Low risk	Subjects were randomly allocated.
Blinding of participants and personnel (performance bias) All outcomes	High risk	The nature of the aerobic exercise training made it impossible to conceal treatment allocation to the subjects or the investigators.
Blinding of care provider (performance bias)	High risk	The nature of the aerobic exercise training made it impossible to conceal treatment allocation to the subjects or the investigators.
Blinding of outcome assessment (detection bias) All outcomes	High risk	The nature of the aerobic exercise training made it impossible to conceal treatment allocation to the subjects or the investigators.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Four/46 lost to follow-up
Participants analysed in group allocated (attrition bias)	Low risk	Intention-to-treat analysis was carried out for all analyses.
Selective reporting (reporting bias)	Unclear risk	Did not test pain at 12 months
Groups similar at baseline (selection bias)	Low risk	No significant differences were found between the two groups.
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	There was an overall exercise attendance rate of 91.3%.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chan 2011 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Chan 2017

Study characteristics

Methods	**Study design:** RCT (ACTRN12609000834257) **Setting:** Australia, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 96 (E1 = 50, C1 = 46) **Chronic LBP duration:** 14.03 weeks (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 52%
Interventions	**Exercise Group 1 (E1):** Individualised functional restoration involving motor control exercises in muscles around the lumbar spine; type = functional restoration & core strengthening; duration = 10 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index); work (degree of interference at work, low back pain days missed); HRQoL (EuroQol 5D); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Rating of Change Scale)) **Follow-up time periods available for syntheses:** 10 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** JF reported being an employee and equity partner of the primary funding source (LifeCare Health). This funding arrangement was declared in our published trial protocol. No other LifeCare staff or equity partners had any involvement with the conduct of the trial including study design; in the collection, analysis, and interpretation of data; in the writing of this paper; and in the decision to submit the paper for publication. During the trial, five of the authors (AH, LS, AC, MR, SS) were subcontracted to LifeCare Health and were paid for treating participants with low back disorders and other musculoskeletal conditions. In order to minimise any potential for bias, all authors had full access to the study data, while both JF and AC had final responsibility for the decision to submit for publication. Three of the authors (JF, AH and AC) provide practitioner education programmes that cover some of the treatments included in this trial. **Funding source:** LifeCare Health **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Block randomisation schedule (random block sizes), stratified for treatment centre using an internet-based randomisation program
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed using an offsite randomisation service which allocated treatment after volunteer consent and enrolment by the researchers.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chan 2017 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the interventions; 2. Experimental intervention not accessible outside of study, any deviations would be to alternate activity sources which was encouraged for the control group.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of provider blinding caused deviations from intended interventions; control group had only 2 brief contacts with provider, and there was built-in latitude to their content.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcomes in question were assessed by the patients who were not blinded to the intervention; 2. Pain and function questionnaires are subjective and are susceptible to being altered by a lack of blinding; 3. Lack of blinding is likely to alter patient response in an inactive versus active treatment study; active treatment here appeared intricate and likely has high expectations surrounding it.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 7 of 96 were lost to follow-up at the 10-week follow-up.
Participants analysed in group allocated (attrition bias)	Low risk	1. Intent-to-treat analysis was conducted.
Selective reporting (reporting bias)	Low risk	1. Study results were analysed and reported as described in study protocol (Hahne 2011, STOPS Trial Protocol).
Groups similar at baseline (selection bias)	Low risk	Treatment groups were all similar at baseline on all relevant characteristics.
Co-interventions avoided or similar (performance bias)	High risk	Patients in the control group sought significantly more healthcare co-intervention than the experimental group, including chiropractic and massage treatments; medication use was not different.
Compliance acceptable in all groups (performance bias)	Low risk	Participants attended a mean (SD) of 17 (05) (out of 2 max, 85%) sessions in the control group and 83 (21) (out of 10 max, 83%) sessions in the experimental group.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical among all patients, regardless of treatment group; 2. Numeric Rating Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Chatzitheodorou 2007

Study characteristics	
Methods	**Study design:** RCT **Setting:** Greece, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 20 (E1 = 10, C1 = 10) **Chronic LBP duration:** Not specified (not specified)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chatzitheodorou 2007 *(Continued)*

	Neurological/radicular symptoms: Not specified **Mean age (years):** 42 **Sex (female):** 45%
Interventions	**Exercise Group 1 (E1):** Fifteen-minute warm-up (callisthenic exercises of arms and legs), leg stretching exercises, running on treadmill; type = aerobic & stretching; duration = 12 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (electrotherapy)
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Eligible subjects were recruited for the study and were randomly assigned to an exercise group or to a control group by block randomisation.
Allocation concealment (selection bias)	Low risk	As eligible subjects were enrolled in the trial, they were assigned to the group that contained their number of recruitment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	After the end of each subject's programme, they completed the questionnaires under the same circumstances.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences between groups
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chatzitheodorou 2007 *(Continued)*

Compliance acceptable in all groups (performance bias)	Low risk	Adherence to the exercise programme was 98% and adherence in control group 96%.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Chen 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Taiwan, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 127 (E1 = 64, C1 = 63) **Chronic LBP duration:** 4.09 years (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 34 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Stretching exercise programme: warm-up exercise (10 minutes), back pain exercise and core muscle training (30 minutes), and relaxation exercises (10 minutes); type = mixed; duration = 26 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (usual activities for 50 minutes per time, 3 times per week)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Low risk	Each participant was provided with an unmarked and sealed envelope containing a randomly generated allocation number.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chen 2014 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	Eleven participants in the experimental group and 13 in the control group failed to complete the study.
Participants analysed in group allocated (attrition bias)	Low risk	The presented analyses involved using intention-to-treat, with missing data being substituted with the last-observation-carried-forward procedure.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	High risk	Difference in participant age (P = 0.01; Table 2)
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Chown 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** United Kingdom, healthcare **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 239 (E1 = 80, C1 = 80, C2 = 79) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 43 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Circuit training, basic postural setting with transverse abdominis and multifidus muscles, home stretching exercises; type = core strengthening & stretching; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (physical therapy) **Comparison Group 2 (C2):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); work (questions); HRQoL (EuroQol 5D) **Follow-up time periods available for syntheses:** 6 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** St. Albans and Hemel Hempstead National Health Services Trust Research and Development Consortium

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chown 2008 *(Continued)*

Other: None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Patients were assigned at random to one of the three therapy regimens by an independent administrator, using block randomisation methods.
Allocation concealment (selection bias)	Low risk	Where feasible, individuals involved in the conduct and analysis of the study were blind to either group membership and/or baseline assessments.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All follow-up assessments were undertaken by an independent assessor who was blind to baseline measurements and group allocation.
Incomplete outcome data (attrition bias) All outcomes	High risk	Ninety-eight of the 239 (41%) eligible participants provided six-week follow-up data.
Participants analysed in group allocated (attrition bias)	High risk	Given the severe reduction in the number of participants with follow-up measurements at six weeks, the results of statistical analysis have been presented as key summary statistics with 95% confidence limits.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Patient randomisation yielded three groups that were highly comparable in terms of basic demography.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	High risk	Group therapy had the worst attendance, with only 40% of patients completing all therapy sessions, compared with 74% and 80% within other treatments.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Chung 2018

Study characteristics

Methods	**Study design:** RCT **Setting:** South Korea, healthcare

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chung 2018 *(Continued)*

	Exercise groups: 2 **Comparison groups:** 0
Participants	**Number of participants:** 27 (E1 = 14, E2 = 13) **Chronic LBP duration:** 12.77 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 33 **Sex (female):** 63%
Interventions	**Exercise Group 1 (E1):** Lumbar stabilisation with Flexi-Bar: participants conducted the abdominal drawing-in manoeuvre in standing, hook-lying, quadruped, and prone positions by maintaining each motion for 10 seconds; type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Lumbar stabilisation only: abdominal drawing-in manoeuvre in standing, hook-lying, quadruped, and prone positions by maintaining each motion for 10 seconds; type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Only cited as "randomized"
Allocation concealment (selection bias)	High risk	No description of randomisation process to assess concealment of treatment allocation
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their intervention due to the nature of the treatments; 2. Both groups receiving exercise and would be difficult to access the experimental protocol outside of the study context.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviations from protocol, as their role in each treatment group was nearly identical, aside from Flexi-bar inclusion.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For outcomes of interest (pain and disability) the outcome assessors were the participants, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Control group receiving usual exercise whereas the experimental group received a intervention potential considered “trendy” that could alter response.
Incomplete outcome data (attrition bias) All outcomes	High risk	1. No description of dropout rate 2. No evidence of the analysis correcting for bias, excluded those who did not have enough time to attend; 3. Increased low back pain or decreased function could cause missing outcome; 4. No information on missingness; 5. No information reported on those lost to follow-up

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chung 2018 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	1. All participants appeared to have been analysed according to their randomised allocation.
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: within this paper there were no obvious omissions in outcomes or analyses.
Groups similar at baseline (selection bias)	Low risk	All relevant baseline characteristics were measured and were sufficiently similar between treatment groups.
Co-interventions avoided or similar (performance bias)	High risk	No reporting of how study addressed co-interventions
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance/attendance reported
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcomes were assessed at the same time; within each outcome, all participants were assessed using the same tools; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Under reported, no other apparent sources of bias

Cortell-Tormo 2018

Study characteristics	
Methods	**Study design:** RCT **Setting:** Spain, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 24 (E1 = 12, C1 = 12) **Chronic LBP duration:** 19.6 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Whole body strengthening including deadlifts, squats, lunges, rows, pushes, reaches with warm-up and cool-down of cardio and stretching; type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no description)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Spanish Ministry of Economy and Competitiveness (DEP2011-30009-C02-01) **Other:** Information modified for author contact
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cortell-Tormo 2018 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Patients were "randomly assigned" to an intervention.
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because care providers did not interact with control group patients during the study period.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in a exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	High risk	1. Dropout rate was 21% (5/24); 2. Missing data excluded in analysis; no sensitivity analysis; 3. Missingness could be caused by increased low back pain and disability; 4. Dropout rate in exercise group was 8%, and in control group was 33%; 5. No information on reasoning for any dropouts
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found. Within this publication, analyses and outcomes were fully reported.
Groups similar at baseline (selection bias)	Low risk	Groups were similar for all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Subjects were excluded if they were receiving concurrent treatments from another practitioner for their low back pain or were receiving medications other than analgesics and nonsteroidal anti-inflammatory drugs.
Compliance acceptable in all groups (performance bias)	Low risk	Eleven of 12 patients in the exercise group attended more than 95% of the sessions; control group adhered to "no treatment" (other than dropouts).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function) and Visual Analogue Scale (for pain) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Costa 2009b

Study characteristics	
Methods	**Study design:** RCT (ACTRN012605000262606) **Setting:** Australia, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 154 (E1 = 77, C1 = 77) **Chronic LBP duration:** 330 weeks (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 54 **Sex (female):** 60%
Interventions	**Exercise Group 1 (E1):** Two-stage motor control exercises: 1st stage exercises for retraining multifidus and transversus abdominal muscles, pelvic floor muscles and control of spinal posture and movement with biofeedback; 2nd stage functional tasks and home exercise; type = core strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Placebo: detuned shortwave diathermy and detuned ultrasound
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Research & Development Grant, The University of Sydney and the Physiotherapy Research Foundation–Australian Physiotherapy Association **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation sequence was computer-generated by one of the investigators.
Allocation concealment (selection bias)	Low risk	Allocation was concealed in sequentially numbered, sealed, opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The placebo treatment was designed to he structurally equivalent to the active intervention, providing similar contact time with the physical therapist.
Blinding of care provider (performance bias)	High risk	The nature of the interventions precluded blinding of the treatment provider.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcomes obtained by telephone interview by an investigator who was blinded to the treatment allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Of the 154 participants who were randomly assigned to groups, 152 attended the two-monthly follow-up (98.7%) and 145 attended both six- and 12-month follow-up.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Costa 2009b *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	The statistical analysis was performed on an intention-to-treat basis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The characteristics of the participants in the two groups were similar at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Ten patients from the exercise group and one patient from the placebo group reported use of co-interventions during the study period.
Compliance acceptable in all groups (performance bias)	Low risk	Out of 12 planned treatment sessions, the participants in the exercise group attended a mean of 8.8 sessions (SD = 3.5) compared with 9.6 sessions (SD = 3).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Costantino 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, not specified **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 54 (E1 = 27, E2 = 27) **Chronic LBP duration:** 23.4 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 73 **Sex (female):** 44%
Interventions	**Exercise Group 1 (E1):** Hydrotherapy: warm-up, stretching and strengthening exercises for abdominal, gluteal, back, and hip muscles; type = strengthening & stretching; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Stretching and strengthening exercises for abdominal, gluteal, back, and hip muscles; type = strengthening & stretching; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = psychological therapy & back school
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Costantino 2014 *(Continued)*

Random sequence generation (selection bias)	Low risk	The participants were randomly allocated using computer randomisation software (RANDI2 software version 0.6.1).
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Two before treatment started; none during
Participants analysed in group allocated (attrition bias)	Low risk	All analyses were performed on the basis of intention-to-treat.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences in age, sex, weight, body mass index and symptoms duration
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Both rehabilitation programmes had no side effects or complications and we also had high participant compliance.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Critchley 2007

Study characteristics	
Methods	**Study design:** RCT (ISRCTN56323917) **Setting:** England, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 212 (E1 = 71, E2 = 72, E3 = 69) **Chronic LBP duration:** 92 weeks (long) **Neurological/radicular symptoms:** Some participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Critchley 2007 *(Continued)*

	Mean age (years): 44 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Trunk muscle retraining, stretches, and general spinal mobility for home practice (as part of physiotherapy); type = mixed; duration = not specified weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Exercise Group 2 (E2):** Transversus abdominis and lumbar multifidus muscle training, spinal stability exercises; type = core strengthening; duration = 26 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 3 (E3):** General strengthening, stretching, and light aerobic exercises progressed according to pacing principles; type = mixed; duration = not specified weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); work (days not working due to back pain in the previous 6 months); HRQoL (EuroQol 5D) **Follow-up time periods available for syntheses:** 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Arthritis Research Campaign **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Before the trial started, the randomisation protocol was computer-generated and held by a trials unit independent of and distant from the trial setting.
Allocation concealment (selection bias)	Low risk	After clinical assessment, participants were assigned to their intervention by clinic staff telephoning the trials unit.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Masking of participants or clinicians was neither possible nor desirable.
Blinding of care provider (performance bias)	High risk	Masking of participants or clinicians was neither possible nor desirable.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Author contact: unmasking appears to have occurred at six months.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Randomised 212 participants between March 2002 and September 2003; reassessed 169 participants at six months, 154 at 12 months and 160 at the final time point.
Participants analysed in group allocated (attrition bias)	Low risk	Clinical outcomes were analysed on both intention-to-treat and complete case basis according to a previously prepared data analysis plan.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The three treatment groups had similar characteristics and baseline values of outcome measures.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Critchley 2007 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	Author contact: only 4% of exercise group had manual treatment.
Compliance acceptable in all groups (performance bias)	Low risk	Retention tended to be lowest in pain management (47 of 69, 68%) and highest in individual physiotherapy (59 of 71, 83%).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Cruz-Diaz 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Spain, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 112 (E1 = 57, C1 = 55) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 71 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Physical therapy and 2, 1-hour pilates sessions weekly for 6 weeks including: strengthening exercises involving the use of implements such as fitballs, magic rings and TheraBands; flexibility and joint mobility exercises; breathing exercises; and motor control and posture correction tasks; type = Pilates; duration = 6 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = electrotherapy & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated table of random numbers
Allocation concealment (selection bias)	Low risk	Treatment allocations were concealed using sealed, opaque and consecutively numbered envelopes kept in a locked location; opened in sequence by an independent administrator.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their group assignment due to the nature of the interventions; 2. Patients encouraged to make no changes to current exercise regimen; no information given on deviations

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cruz-Diaz 2016 *(Continued)*

Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from the intended intervention, as the only contact with control patients was in physiotherapy sessions, which were structured.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were patients themselves, who could not be blinded to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Because there was a significant difference between Pilates and control group interventions, there was likely a bias in outcome assessments due to lack of blinding; study results supported this.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 8% (9/112) overall.
Participants analysed in group allocated (attrition bias)	Unclear risk	1. Stated that intention-to-treat was done, however, excluded those who were non-compliant; 2. Dropout rate was sufficiently small that the missing data (and lack of intention-to-treat analysis) was unlikely to have affected overall conclusions.
Selective reporting (reporting bias)	Low risk	1. All planned analyses were executed and reported for all primary and secondary outcomes.
Groups similar at baseline (selection bias)	High risk	At baseline, the control group was significantly older (by 3 years) than the Pilates group; no report of duration of symptoms; baseline pain and function were similar.
Co-interventions avoided or similar (performance bias)	High risk	Patients were allowed to continue their regular extra-curricular exercises without regulation.
Compliance acceptable in all groups (performance bias)	Low risk	Participants were excluded from the study if they missed more than two therapeutic sessions, thus compliance was at least 83% (10/12 sessions).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All patient outcome assessments were identical, regardless of treatment group allocation; 2. Numeric Rating Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Cruz-Diaz 2017

Study characteristics	
Methods	**Study design:** RCT (NCT02371837) **Setting:** Spain, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 102 (E1 = 34, E2 = 34, C1 = 34) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 64%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cruz-Diaz 2017 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Equipment-based pilates; type = Pilates; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Mat-based pilates; type = Pilates; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated table of random numbers
Allocation concealment (selection bias)	Low risk	Allocations were sealed in opaque and consecutively numbered envelopes kept in a locked location; opened in sequence by an independent administrator not involved in the study.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients were not blinded; 2. Pilates protocols were explicit, unlikely to be accessible outside of study.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Pilates interventions were distinct, one with equipment one without, unlikely to have deviations; control group had minimal contact with physiotherapists.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was very small, only 4%.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication, all analyses and outcomes were fully reported; no obvious omissions.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline, except study did not measure duration of symptoms.
Co-interventions avoided or similar (performance bias)	Low risk	No information reported on co-interventions

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cruz-Diaz 2017 *(Continued)*

Compliance acceptable in all groups (performance bias)	High risk	No information reported on compliance, adherence or attendance, despite the study having measured it.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Visual Analogue Scale (for pain) and Roland-Morris Disability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Cruz-Diaz 2018

Study characteristics	
Methods	**Study design:** RCT (NCT02371837) **Setting:** Spain, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 64 (E1 = 32, C1 = 32) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 37 **Sex (female):** 66%
Interventions	**Exercise Group 1 (E1):** Pilates training; type = Pilates; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation performed by an independent researcher in a 1:1 ratio
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed; randomisation was performed by an independent researcher and allocation was placed in sealed, opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to intervention due to the nature of the treatment groups; 2. Lack of patient blinding was unlikely to have led to deviations from intended intervention because patients had little control over intervention; cross-over between groups was unlikely.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Lack of care provider blinding likely did not deviate intended in-

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cruz-Diaz 2018 *(Continued)*

		terventions because providers only interacted with one treatment group; little contact with control outside outcome assessments
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors were patients, who could not be blinded due to the nature of the treatment groups; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 3% (2/64).
Participants analysed in group allocated (attrition bias)	Low risk	1. Not explicitly stated, however, no indication otherwise; numbers matched intention-to-treat analysis.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all outcomes and analyses fully reported; no obvious omissions
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar at baseline on all relevant characteristics, except the exercise group had a slightly higher mean body mass index than control; duration of symptoms was not reported.
Co-interventions avoided or similar (performance bias)	High risk	No information on co-intervention use was reported in this study.
Compliance acceptable in all groups (performance bias)	Low risk	Little explicit reporting of compliance, except for this quote (from discussion): "participants showed high adherence to treatment".
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All patients were assessed identically, regardless of treatment group; 2. Roland-Morris Disability Questionnaire (for disability) and Visual Analogue Scale (for pain) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Cuesta-Vargas 2012

Study characteristics	
Methods	**Study design:** RCT (NCT00888524) **Setting:** Spain, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 58 (E1 = 29, C1 = 29) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 38 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Deep water running; type = aerobic; duration = 16 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Usual care/no treatment (control group: usual general practice)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cuesta-Vargas 2012 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (12-Item Short Form Survey (Spanish)) **Follow-up time periods available for syntheses:** 16 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** National Health Service of Andalusia **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The patients were randomly assigned to one of two groups.
Allocation concealment (selection bias)	Low risk	The assignments, which were generated by a computer, were presented in sealed, sequentially numbered envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences in baseline characteristics.
Co-interventions avoided or similar (performance bias)	Unclear risk	Author contact: did not measure this; it could have affected results.
Compliance acceptable in all groups (performance bias)	Low risk	The dropout rate was acceptable (17.2 %), suggesting that compliance to the study in both groups was adequate.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cuesta-Vargas 2015

Study characteristics	
Methods	**Study design:** RCT **Setting:** Spain, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 36 (E1 = 18, E2 = 18) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Mobility and motor control exercises (10 minutes each), resistance and strengthening exercises (20 minutes), deep water running (20 minutes), 3 times weekly; type = mixed; duration = 52 weeks; dose = high; design = individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Mobility and motor control exercises (10 minutes each), resistance and strengthening exercises (20 minutes), deep water running (20 minutes); type = mixed; duration = 52 weeks; dose = high; design = individualised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (12-Item Short Form Survey; EuroQol 5D) **Follow-up time periods available for syntheses:** 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The patients were randomly assigned (using sealed envelopes).
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Unclear
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Unclear
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Cuesta-Vargas 2015 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	At the baseline, no statistically significant differences were found between the experimental groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Curnow 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Australia, not specified **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 39 (E1 = 12, E2 = 14, E3 = 13) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Four basic Pilates exercises, neutral spine exercise and relaxation posture; type = Pilates; duration = 6 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = advice/education **Exercise Group 2 (E2):** Four basic Pilates exercises and a relaxation posture; type = Pilates; duration = 6 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = advice/education **Exercise Group 3 (E3):** Four basic Pilates exercises; type = Pilates; duration = 6 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Low Back Pain Rating Scale (Manniche)); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed. Sufficient data not available for inclusion in meta-analyses
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Curnow 2009 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Subjects were taught four basic exercises before being randomly allocated to a specific intervention group (A, B or C).
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Unclear risk	Unclear
Groups similar at baseline (selection bias)	Unclear risk	Not described
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	High risk	Compliance was an issue in second eight-week period.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Dalichau 2000

Study characteristics	
Methods	**Study design:** RCT **Setting:** Germany, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 63 (E1 = 21, E2 = 21, C1 = 21)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Dalichau 2000 *(Continued)*

	Chronic LBP duration: Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** Not reported **Sex (female):** 99%
Interventions	**Exercise Group 1 (E1):** Warm-up aerobic exercises, whole body equipment strengthening exercises with lumbar support (60 minutes); type = strengthening & aerobic; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = lumbar support **Exercise Group 2 (E2):** Warm-up aerobic exercises, whole body equipment strengthening exercises (60 minutes); type = strengthening & aerobic; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 35 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Dalichau 2000 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Da Luz 2014

Study characteristics	
Methods	**Study design:** RCT (RBR-7tyg5j) **Setting:** Brazil, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 86 (E1 = 43, E2 = 43) **Chronic LBP duration:** 42 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 41 **Sex (female):** 76%
Interventions	**Exercise Group 1 (E1):** Mat Pilates, performed on the ground using Swiss ball, and elastic bands; type = Pilates; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Equipment-based Pilates, performed on the Cadillac, Reformer, Ladder Barrel, and Step Chair; type = Pilates; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Council of Technological and Scientific Development (CNPq), Brazil (479645/2011-6) **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	A simple randomisation schedule was performed on Microsoft Excel for Windows.
Allocation concealment (selection bias)	Low risk	Using sealed, opaque, and sequentially numbered envelopes

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Da Luz 2014 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	It was not possible to blind the participant and the physical therapist due to the interventions.
Blinding of care provider (performance bias)	High risk	It was not possible to blind the participant and the physical therapist due to the interventions.
Blinding of outcome assessment (detection bias) All outcomes	High risk	It was not possible to blind the participant and the physical therapist due to the interventions.
Incomplete outcome data (attrition bias) All outcomes	Low risk	In Figure
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Not described
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	In the 90%+ for both groups
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Devasahayam 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Singapore, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 28 (E1 = 14, E2 = 14) **Chronic LBP duration:** 57 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 54 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Functional back exercises, practising functional tasks, and home exercises; type = other (not specified); duration = 4 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Devasahayam 2014 *(Continued)*

	Exercise Group 2 (E2): Generic mat exercise: stretches (cat/camel, quads, hamstrings, hip rotators), core stability exercises (bridging, dead bugs, etc.) and home exercises; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The participants were randomised using a sealed envelope.
Allocation concealment (selection bias)	Low risk	Sealed envelope
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Support for judgement was not available.
Blinding of care provider (performance bias)	Unclear risk	The physiotherapists who instructed the exercise classes were blinded to the data collected from their groups.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Thirteen of 28 subjects did not complete the follow-up measures, but all completed treatment.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Chronicity seemed very different across groups.
Co-interventions avoided or similar (performance bias)	Low risk	Participants in both groups were advised not to seek any other treatments in order to standardise the treatments received.
Compliance acceptable in all groups (performance bias)	Unclear risk	The physiotherapist attending the experimental group reported greater compliance among participants.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Devasahayam 2014 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Deyo 1990

Study characteristics

Methods	**Study design:** RCT **Setting:** USA, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 125 (E1 = 63, C1 = 62) **Chronic LBP duration:** 260 weeks (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 51 **Sex (female):** 59%
Interventions	**Exercise Group 1 (E1):** Twelve sequential relaxation and stretching exercises to improve flexibility, home exercises with repeated instruction; type = stretching; duration = 4 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = advice/education & electrotherapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Sickness Impact Profile (Physical)); HRQoL (Sickness Impact Profile); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall improvement rating (6-point)) **Follow-up time periods available for syntheses:** 4 weeks (short); 17 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Robert Wood Johnson Foundation; Multipurpose Arthritis Center Grant; National Institutes of Health; Northwest Health Services Research and Development Field Program, Seattle Veterans Affairs Medical Centre **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Deyo 1990 *(Continued)*
All outcomes

Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Diaz-Arribas 2009

Study characteristics	
Methods	**Study design:** RCT (FIS-PI051650) **Setting:** Spain, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 126 (E1 = 63, C1 = 63) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 39 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Muscle balancing and postural exercises for lumbar spine and pelvis (Godelieve Denys-Struyf Method); type = other (co-ordination); duration = 7 weeks; dose = low; design = standardised; delivery = group; additional intervention = not specified **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Spanish National Institute of Health (NIH) **Other:** SDs imputed

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Diaz-Arribas 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A simple randomisation procedure was used to assign patients to each treatment group.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The physiotherapists who treated and assessed the patients were not the same: the latter were unaware of the treatment received by each participant.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Described in Figure 1
Participants analysed in group allocated (attrition bias)	High risk	Analyses only included patients who completed treatment and gave follow-up data.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	A baseline descriptive analysis was performed on the basic demographic and clinical features of each group.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Djavid 2007

Study characteristics

Methods	**Study design:** RCT **Setting:** Iran, healthcare **Exercise groups:** 2

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Djavid 2007 *(Continued)*

	Comparison groups: 1
Participants	**Number of participants:** 53 (E1 = 19, E2 = 18, C1 = 16) **Chronic LBP duration:** 29 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 38 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Strengthening, stretching, mobilising, co-ordination, and stabilising exercises for abdominal, back, pelvic and lower limb muscles; type = mixed; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = electrotherapy **Exercise Group 2 (E2):** Strengthening, stretching, mobilising, co-ordination, and stabilising exercises for abdominal, back, pelvic and lower limb muscles; type = mixed; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = placebo **Comparison Group 1 (C1):** Other conservative treatment (electrotherapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Allocation of participants was concealed.
Allocation concealment (selection bias)	Low risk	Allocation of participants was concealed.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Patients received laser therapy or placebo laser therapy on Saturday and Wednesday for 12 sessions.
Blinding of care provider (performance bias)	Low risk	Patients received laser therapy or placebo laser therapy on Saturday and Wednesday for 12 sessions.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All outcomes were measured on admission to the trial, at week six (after the last session of intervention) and at week 12 by physicians blinded to group.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Sixty-one patients were randomised into one of the three groups (Figure I). Eight participants withdrew from the trial during the intervention or follow-up period.
Participants analysed in group allocated (attrition bias)	Low risk	Analyses were performed on an intention-to-treat basis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There was no statistically significant difference between the three groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Djavid 2007 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Donzelli 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 43 (E1 = 21, E2 = 22) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Pilates CovaTech exercises: sitting, antalgic, stretching, mobilisation and proprioceptivity improvement exercises; type = Pilates & stretching; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Extension and strengthening exercises of paravertebral muscles and lower limbs, mobilising exercises for the spine; type = core strengthening & flexibility/mobilising; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall benefit from treatment (3-point)) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	After undergoing the pretreatment examination, the patients were divided into two groups.
Allocation concealment (selection bias)	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Donzelli 2006 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	The patients did not know whether they were in the experimental treatment group or the control group.
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The post-treatment evaluations were performed by a blinded assessor and patients were also blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Three dropouts, for health or personal reasons.
Participants analysed in group allocated (attrition bias)	High risk	Forty patients used in analyses
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Groups could be defined as homogeneous.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	High risk	In total, 45% of the Back School group and 28% of the Pilates group had done their exercises.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Same for both groups

Dougherty 2014

Study characteristics	
Methods	**Study design:** RCT (ISRCTN30511490) **Setting:** USA, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 88 (E1 = 28, C1 = 60) **Chronic LBP duration:** 181.8 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 56 **Sex (female):** 34%
Interventions	**Exercise Group 1 (E1):** Directional preference exercises, lumbar stabilisation, general flexibility and specific training exercises; type = mixed; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Dougherty 2014 *(Continued)*

	Comparison Group 1 (C1): Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 5 weeks (short); 12 weeks (short); 24 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Department of Health and Human Services, Health Resources and Service Administration Chiropractic Demonstration grant (Award: R18HP07641-03-03, Grant: R18HP07641) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation to treatments was through a random number producing algorithm.
Allocation concealment (selection bias)	Low risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Attempt to control for expectation of treatment
Blinding of care provider (performance bias)	Low risk	The screening clinician, statistician, and the treating clinician were all blinded to the status of the modified clinical prediction rules.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Attempt to control for expectation of treatment
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropout reasons given and are acceptable
Participants analysed in group allocated (attrition bias)	Low risk	With multiple imputation, using SPSS missing values module
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Anything that differed was included as a covariate in the analysis.
Co-interventions avoided or similar (performance bias)	Low risk	Participants agreed not to undergo any new or different treatment during the intervention and follow-up, but they were allowed to continue medications.
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: compliance similar in both groups
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Dufour 2010

Study characteristics	
Methods	**Study design:** RCT (NCT00256373) **Setting:** Denmark, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 272 (E1 = 129, E2 = 143) **Chronic LBP duration:** 1.2 years (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 41 **Sex (female):** 44%
Interventions	**Exercise Group 1 (E1):** Warm-up (aerobic training), back, gluteus, and abdominal strengthening using machines and circuit training, sports; type = mixed; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy & other & physiotherapy **Exercise Group 2 (E2):** Specific muscle training exercises to strengthen and shorten the muscles in back and gluteus region (developed by Oefeldt); type = core strengthening; duration = 12 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); work (ability to work); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global perceived outcome (3-point)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Apotekerfonden af 1999, Sygekassernes Helsefond; The Danish National Board of Health **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Patients were allocated by a separate secretary.
Allocation concealment (selection bias)	Low risk	Allocated by a separate secretary
Blinding of participants and personnel (performance bias) All outcomes	High risk	The study was designed as a prospective, single-blinded clinical trial.
Blinding of care provider (performance bias)	Low risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All physical examinations at trial visits were performed by one physician who was blinded to the treatment group.
Incomplete outcome data (attrition bias)	Low risk	Eleven patients in each group: A (9%)/B (8%) dropped out during the treatment period.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Dufour 2010 *(Continued)*

All outcomes

Participants analysed in group allocated (attrition bias)	Low risk	Whether analysed on an intention-to-treat basis or based on actual data, the outcome obtained using comparative statistics was similar.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	All patients were encouraged to increase physical activity during everyday life.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Elnaggar 1991

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 56 (E1 = 28, E2 = 28) **Chronic LBP duration:** 322 weeks (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 39 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Extension exercises ('prone, raising trunk' – described as back muscle strengthening); type = core strengthening; duration = 2 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Flexion exercises (described as 'mobilising'); type = McKenzie; duration = 2 weeks; dose = low; design = partially individualised; delivery = independent with follow-up; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score) **Follow-up time periods available for syntheses:** 2 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Bias	Authors' judgement	Support for judgement

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Elnaggar 1991 *(Continued)*

Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Engbert 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** Germany, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 23 (E1 = 10, E2 = 13) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Engbert 2011 *(Continued)*

	Mean age (years): 51 **Sex (female):** 52%
Interventions	**Exercise Group 1 (E1):** Therapeutic climbing on gym climbing wall with a focus on strengthening, stretching, mobilisation, co-ordination and stabilisation of abdominal, back, pelvic, and leg muscles; type = mixed; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Strengthening, stretching, mobilisation, co-ordination, and stabilisation for abdominal, back, pelvic, and lower limb muscles; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (36-Item Short Form Survey); function (Hannover Functional Ability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were randomly assigned to the therapeutic climbing or the standard exercise training by means of a computer-generated randomisation schedule.
Allocation concealment (selection bias)	Unclear risk	Numbered envelopes were prepared by a physical therapist not involved in the study and assigned to patients in a sequential order.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Patients did not know which intervention was being evaluated.
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Patients who did not attend more than 30% of treatment sessions were excluded from subsequent data analysis. Overall, five patients were excluded.
Participants analysed in group allocated (attrition bias)	High risk	Data analysis was based on 23 patients, 10 from the therapeutic climbing group and 13 from the standard exercise group.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences in sex and age were noted between treatment groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Engbert 2011 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Participants of both groups were free to do unspecific exercise, such as walking, in their free time.
Compliance acceptable in all groups (performance bias)	Low risk	Patients who did not attend more than 30% of treatment sessions were excluded from subsequent data analysis.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Farajzadeh 2017

Study characteristics	
Methods	**Study design:** RCT (1N2015051022202IRCT) **Setting:** Iran, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 30 (E1 = 15, E2 = 15) **Chronic LBP duration:** 15.465 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 22 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** McGill Stabilizer exercises (curl up, bird dog, side bridge); type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = none **Exercise Group 2 (E2):** Conventional physiotherapy exercises (knee to chest, prone lying with a pilot, one leg sliding, cycling in supine, bridging); type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Quebec Back Pain Disability Scale) **Follow-up time periods available for syntheses:** 10 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Research Council of Tabriz University of Medical Sciences **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Patients were "randomly assigned".
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Farajzadeh 2017 *(Continued)*

Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Protocols were explicit; both were relatively similar, unlikely to be able to deviate within the study protocol.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective and could be altered by knowledge of the intervention; 3. Likely that lack of patient blinding could have caused bias in outcome assessment because the experimental treatment could be perceived as better than the control.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No description of dropout rate; 2. Analyses did not compensate for missing data and no sensitivity analysis was done; 3. Increased pain and disability could cause missingness; 4. No information on dropout rate in either group; 5. No information on reasoning for any dropouts
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication analyses and outcomes were fully reported; no standard outcome/analysis left out
Groups similar at baseline (selection bias)	Low risk	At baseline, there were no clinically significant differences in age, duration of symptoms, pain and disability; sex distribution not reported.
Co-interventions avoided or similar (performance bias)	Low risk	No mention of co-interventions in this study
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance, adherence or attendance in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Quebec Low Back Pain Disability Questionnaire (for function) Visual Analogue Scale (for pain) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Feldwieser 2018

Study characteristics	
Methods	**Study design:** RCT **Setting:** Germany, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 63 (E1 = 22, E2 = 20, C1 = 21) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 74 **Sex (female):** 71%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Feldwieser 2018 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Bridging exercises with biofeedback: participants performed a series of floor exercises aimed at strengthening their core while laying flat on the ground using Wii Fit balance boards; type = core strengthening; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Bridging exercises only: participants performed a series of floor exercises aimed at strengthening their core while laying flat on the ground; type = core strengthening; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Seventh Framework Programme of the European Commission (Agreement No. 611218) **Other:** Information modified for author contact; sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was performed with drawing pre-prepared opaque envelopes containing the group assignment after inclusion and measuring all baseline data.
Allocation concealment (selection bias)	Low risk	Randomisation was performed with drawing pre-prepared opaque envelopes containing the group assignment after inclusion and measuring all baseline data.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their intervention due to the nature of the treatments; 2. Unlikely lack of patient blinding caused deviations; actual exercises in each group were identical and with same therapists; unlikely for participants (65+) to seek biofeedback intervention outside of study
Blinding of care provider (performance bias)	Low risk	1. Care providers (therapists) could not be blinded to the intervention as they were responsible for delivering the interventions; 2. Nature of the intervention made this unlikely, either used the biofeedback or did not; no transfer of opinion because patients had low user acceptance anyway
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for function outcome were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Exercise versus no treatment was likely to be altered by knowledge of assignment; also possible biofeedback group was frustrated and this influenced outcome assessment
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 27% (17/63), which is not an acceptable rate; 2. Reasons for withdrawal from study for most participants were conflicting patient schedules due to a delay in the original study timeline; no data on how dropouts compared to the rest
Participants analysed in group allocated (attrition bias)	Low risk	1. Frequent connectivity issues led to the biofeedback group performing exercises similar to the standard group; were still analysed in biofeedback group

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Feldwieser 2018 *(Continued)*

Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication fully reported and analysed; did not report traditional pain scales, but no obvious omissions from the analysis
Groups similar at baseline (selection bias)	Low risk	At baseline, participants across all treatment groups were similar on age, body mass index, function, sleep quality and fall risk; there were no data reported on sex or duration of symptoms.
Co-interventions avoided or similar (performance bias)	Low risk	No mention of whether or not the study limited or measured co-interventions
Compliance acceptable in all groups (performance bias)	High risk	Participant attendance was apparently recorded, but exclusion was only reported when participants missed > = 6 sessions (out of 24 total); 24 patients did not complete due to equipment issues.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All participants were assessed identically regardless of treatment group assignment; 2. Oswestry Disability Index and Roland-Morris Disability Questionnaire (both for function) are both well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Ferreira 2007

Study characteristics	
Methods	**Study design:** RCT (ACTRN012605000053628) **Setting:** Australia, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 240 (E1 = 80, E2 = 80, C1 = 80) **Chronic LBP duration:** 48 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 54 **Sex (female):** 69%
Interventions	**Exercise Group 1 (E1):** Strengthening and stretching exercises for main muscle groups, aerobic exercises; type = mixed; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy & relaxation **Exercise Group 2 (E2):** Motor control exercises: contracting transversus abdominal and multifidus muscles in isolation using ultrasound biofeedback, home exercises; type = core strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Arthritis Foundation of New South Wales; Motor Accidents Authority of New South Wales; University of Sydney

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ferreira 2007 *(Continued)*

Other: None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was by a random sequence of randomly permuted blocks of sizes six, nine and 15.
Allocation concealment (selection bias)	Low risk	It was concealed from patients and the other investigators using consecutively numbered, sealed, opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Participants reported their outcomes to a trial physical therapist who was blinded to allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Of 240 participants, 93% were followed up at eight weeks and 88% were followed up at six and 12 months.
Participants analysed in group allocated (attrition bias)	High risk	Analysis was by intention-to-treat in the sense that data were analysed for all randomised subjects for whom follow-up data were available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The groups were similar for most baseline characteristics.
Co-interventions avoided or similar (performance bias)	Low risk	Participants in all groups were asked not to seek other treatments and, where possible, not to change current medications for the eight-week trial period.
Compliance acceptable in all groups (performance bias)	Low risk	There was a high degree of adherence to all three interventions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Franca 2012

Study characteristics

Methods	**Study design:** RCT (NCT01124201) **Setting:** Brazil, healthcare

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Franca 2012 *(Continued)*

	Exercise groups: 2 **Comparison groups:** 0
Participants	**Number of participants:** 30 (E1 = 15, E2 = 15) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Segmental stabilisation exercises: transversus abdominal and multifidus muscles; type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Stretching of the erector spinae, hamstrings, and calf muscles; type = stretching; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** State of São Paulo Research Foundation **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	They were randomised (random number in opaque envelopes) in one of two treatment groups.
Allocation concealment (selection bias)	Low risk	They were randomised (random number in opaque envelopes) in one of two treatment groups.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	High risk	Sessions were supervised by the investigator.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Participants were assessed at baseline and at the end of study treatment by an investigator (physical therapist) blinded to the randomisation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Unclear risk	Sex was not reported.
Groups similar at baseline (selection bias)	Low risk	No significant differences were seen for age, weight, and height.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Franca 2012 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	Participants were instructed not to participate in any other physical programme during the study as well as not to exercise while at home.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Fransoo 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** Belgium, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 58 (E1 = 20, E2 = 19, C1 = 19) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 44 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Progressive strengthening and stabilising muscle exercises with loading, stretching, Back School-type exercises; type = mixed; duration = 7 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Exercise Group 2 (E2):** Progressive strengthening and stabilising muscle exercises without load and stretching; type = core strengthening & stretching; duration = 5 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (manual therapy, education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 7 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias)	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Fransoo 2006 *(Continued)*
All outcomes

Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Unclear risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Frost 1995

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 71 (E1 = 36, C1 = 35) **Chronic LBP duration:** 90 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 52%
Interventions	**Exercise Group 1 (E1):** Warm-up, stretching, progressive exercises, and light aerobic exercise; type = mixed; duration = 4 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = back school **Comparison Group 1 (C1):** Other conservative treatment (back school)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Frost 1995 *(Continued)*

	Follow-up time periods available for syntheses: 7 weeks (short); 26 weeks (moderate); 104 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Back Pain Association; Oxfordshire locally organised research scheme **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not availablelable
Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Frost 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 286 (E1 = 144, C1 = 142) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 41 **Sex (female):** 53%
Interventions	**Exercise Group 1 (E1):** Stretching, strengthening, and mobility exercises; type = mixed; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & other & physiotherapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall perceived benefit from treatment)) **Follow-up time periods available for syntheses:** 8 weeks (short); 24 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Arthritis Research Campaign **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Frost 2004 *(Continued)*

Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Garcia 2013

Study characteristics	
Methods	**Study design:** RCT (ACTRN12610000435088) **Setting:** Brazil, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 148 (E1 = 74, E2 = 74) **Chronic LBP duration:** 22.5 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 54 **Sex (female):** 73%
Interventions	**Exercise Group 1 (E1):** Progressive McKenzie exercises after individual assessment, home exercises; type = McKenzie; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Breathing exercises, kinaesthetic training, stretching of lower back, quadriceps, and hamstring muscles, abdominal strengthening exercises, home exercises; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = back school
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (World Health Organization Quality of Life–BREF [WHOQOL-BREF]) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Fundacao de Amparo a Pesquisa do Estado de Sao Paulo **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Garcia 2013 *(Continued)*

Allocation concealment (selection bias)	Low risk	The allocation was concealed by using consecutively numbered, sealed, opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Given the nature of the interventions, it was not possible for the therapist or the patients to be blinded.
Blinding of care provider (performance bias)	High risk	Given the nature of the interventions, it was not possible for the therapist or the patients to be blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Given the nature of the interventions, it was not possible for the therapist or the patients to be blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Refer to flow chart
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The baseline characteristics of both groups were similar.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Back School group attended a mean of 3.64 sessions (SD 0.08); this was 3.72 sessions (SD 0.06) for participants allocated to the McKenzie group.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Garcia 2017

Study characteristics	
Methods	**Study design:** RCT (NCT02123394) **Setting:** Brazil, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 148 (E1 = 74, C1 = 73) **Chronic LBP duration:** 41.68 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 56 **Sex (female):** 76%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Garcia 2017 *(Continued)*

Interventions	**Exercise Group 1 (E1):** McKenzie method of mechanical diagnosis and therapy, repeated exercises with patient specific direction (3 possible treatment groups: derangement, dysfunction, postural); type = McKenzie; duration = 5 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education **Comparison Group 1 (C1):** Placebo: disconnected ultrasound and diathermy
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 13 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** The care provider who treated patients in the MDT group has completed first level McKenzie training, however has no involvement with the McKenzie Institute. This trial did not receive either funding from McKenzie Institute or any assistance in writing/analysing the results of this trial. Authors do not have any involvement with the McKenzie Institute. MJH, LOPC and ANG receive funding from International Mechanical Diagnosis and Therapy Research Foundation for the following studies on MDT: (1) Hancock MJ, Maher CG, Mota da Silva T, Clare H, Steffens D (2016). Secondary prevention of a recurrence of low back pain. (2) Hancock MJ, Garcia AN, Costa LdCM, Costa LOP (2014). Identifying patients with back pain who respond best to MDT. MJH was keynote speaker at the 2017 McKenzie conference and his travel costs were paid. **Funding source:** São Paulo Research Foundation; Coordination for the Improvement of Higher Education Personnel **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"A simple randomisation schedule was computer generated".
Allocation concealment (selection bias)	Low risk	"The allocation was concealed by using consecutive numbered, sealed and opaque envelopes".
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. No patients were not truly blinded to the treatment they received; 2. Unlikely that patients deviated as study attempted to keep them unaware of the intervention received by other group; also hard to access the McKenzie Method of Mechanical Diagnosis and Therapy outside of study.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Interventions were conducted by different physiotherapists, treatments were audited by authors, minimal decision-making by therapist
Blinding of outcome assessment (detection bias) All outcomes	Low risk	1. Outcomes in question were assessed by the patient who are not blinded to the intervention; 2. Attempted to keep patients unaware of intervention received by other group
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. One of 147 was effectively zero at the 3-month mark.
Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was used in this study.
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported according to registered protocol (NCT02123394).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Garcia 2017 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Co-intervention use was tracked and compared between groups; there were no significant differences between groups on co-intervention use.
Compliance acceptable in all groups (performance bias)	Low risk	Out of 10 sessions that could be completed, the patients allocated to the McKenzie Method of Mechanical Diagnosis and Therapy group attended a mean of 9.01 sessions compared with a mean of 9.23 sessions for patients allocated to the placebo group.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Numeric Rating Scale (for pain) and Roland-Morris Disability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	There were no other sources of bias noticed; generally well-reported.

Gatti 2011

Study characteristics

Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 79 (E1 = 34, E2 = 45) **Chronic LBP duration:** 71.4 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 58 **Sex (female):** 65%
Interventions	**Exercise Group 1 (E1):** Trunk balance exercises and standard trunk flexibility exercises; type = mixed; duration = 5 weeks; dose = low; design = standardised; delivery = group; additional intervention = not specified **Exercise Group 2 (E2):** Strengthening exercises and standard trunk flexibility exercises; type = mixed; duration = 5 weeks; dose = low; design = standardised; delivery = group; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (12-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (patient-report, improvement or worsening in specific positions)) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Using a computer-generated list of random numbers created and managed by a physiotherapist who was blinded

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gatti 2011 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Five/45 in the control group dropped out because of the intervention.
Participants analysed in group allocated (attrition bias)	Low risk	Data were analysed using an intention-to-treat approach.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The two groups were homogeneous at the beginning of the rehabilitation period.
Co-interventions avoided or similar (performance bias)	Unclear risk	No suggestion to the participants on drug management was given.
Compliance acceptable in all groups (performance bias)	Low risk	All participants who completed the study were fully compliant with the intervention programme.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Geisser 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 4 **Comparison groups:** 0
Participants	**Number of participants:** 100 (E1 = 26, E2 = 25, E3 = 24, E4 = 25) **Chronic LBP duration:** 63 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 40 **Sex (female):** 66%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Geisser 2005 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Sahrmann & Bookhour exercises (stretches, strengthening, self-corrections) to address specific dysfunction (i.e. innominate self-correction, unilateral prone press-ups); type = Pilates & stretching; duration = 5 weeks; dose = low; design = individualised; delivery = individual; additional intervention = manual therapy **Exercise Group 2 (E2):** Sahrmann & Bookhour exercises (stretches, strengthening, self-corrections) to address specific dysfunction (i.e. innominate self-correction, unilateral prone press-ups); type = Pilates & stretching; duration = 5 weeks; dose = low; design = individualised; delivery = individual; additional intervention = placebo **Exercise Group 3 (E3):** General stretches (hamstring, knee to chest, quads, and prone on elbows), aerobic exercise (walking at fast pace for 20 minutes); type = stretching & aerobic; duration = 5 weeks; dose = low; design = individualised; delivery = individual; additional intervention = manual therapy **Exercise Group 4 (E4):** General stretches (hamstring, knee to chest, quads, and prone on elbows), aerobic exercise (walking at fast pace for 20 minutes); type = stretching & aerobic; duration = 5 weeks; dose = low; design = individualised; delivery = individual; additional intervention = placebo
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Quebec Back Pain Disability Scale) **Follow-up time periods available for syntheses:** 5 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Center for Medical Rehabilitation Research (R03-HD35893); National Institute of Child and Human Development; National Institutes of Health **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Author contact: 100 allocated to each grp, assigned to group based on order of entry
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The treating therapist was not blind to the treatment group of the patient, but attempted to keep patients blind to their group assignment.
Blinding of care provider (performance bias)	High risk	The treating therapist was not blind to the treatment group of the patient, but attempted to keep patients blind to their group assignment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The self-report measures were re-administered following the last visit with the therapist by the principal investigator, who was blinded to the treatment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Out of the 100 patients recruited, 72 patients completed the study.
Participants analysed in group allocated (attrition bias)	High risk	Only patients who completed treatment were used.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Geisser 2005 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Chi-square tests and analyses of variance (ANOVA) were used to compare the groups in terms of dropouts, age, gender, compensation status, surgical history.
Co-interventions avoided or similar (performance bias)	Low risk	Patients were allowed to continue their use of pain medications, but were asked to not change their usage during the course of the study.
Compliance acceptable in all groups (performance bias)	High risk	Examination of compliance with exercise revealed that six patients who completed the study were unable to reproduce one or more of their exercises.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Author contact: most were measured at five weeks.

Ghaderi 2016

Study characteristics	
Methods	**Study design:** RCT (IRCT2014081218760N1) **Setting:** Iran, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 60 (E1 = 30, E2 = 30) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 53 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Pelvic floor muscle exercises plus usual physiotherapy care with transcutaneous electrical nerve stimulation, hot pack and therapeutic ultrasound; type = core strengthening; duration = 12 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = electrotherapy & heat/ice **Exercise Group 2 (E2):** Regular physiotherapy exercises (focussing on abdominal and paravertebral muscles); type = core strengthening; duration = 12 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = electrotherapy & heat/ice
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Physical Medicine and Rehabilitation Research Center, Tabriz University of Medical Sciences **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were "randomly assigned".
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ghaderi 2016 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviation from intended interventions because patients had no control over the delivery of intervention
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were all defined with specific protocols
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective and responses could be altered by awareness of intervention; 3. Likely that lack of patient blinding could have caused bias in outcome assessment because the experimental group could be perceived as better than the control
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No description of dropout rate 2. Analysis dropped any missing values and there was no sensitivity analysis; 3. Increased pain and disability could cause missing data; 4. No information on dropout rate in either group; 5. No information on reasoning for any dropouts
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all analyses and outcomes were fully reported.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on age, sex (all female patients), body size, function and pain at baseline; duration of symptoms not reported
Co-interventions avoided or similar (performance bias)	High risk	No information on co-interventions presented in this study
Compliance acceptable in all groups (performance bias)	Low risk	No information on compliance, adherence or attendance in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Gladwell 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 34 (E1 = 20, C1 = 14)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gladwell 2006 *(Continued)*

	Chronic LBP duration: 10 years (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 41 **Sex (female):** 79%
Interventions	**Exercise Group 1 (E1):** Ten Pilates exercises using neutral spine and pelvis, recruitment of core muscles; type = Pilates; duration = 6 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: continued with normal activity)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (12-Item Short Form Survey) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Author contact: random number allocated to physical therapy corresponding with allocation
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Pre and post the six weeks intervention, a questionnaire-based assessment and a functional assessment were performed.
Incomplete outcome data (attrition bias) All outcomes	High risk	Twenty-five participants were allocated to the Pilates group and 24 to the control group. Fifteen participants did not complete the trial.
Participants analysed in group allocated (attrition bias)	High risk	Thirty-four participants completed all aspects of the trial with 20 in Pilates group and 14 in the control group.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	High risk	No significant differences were found between the control group and Pilates group in baseline data.
Co-interventions avoided or similar (performance bias)	Low risk	Pilates was applied as an additional intervention to the current drug treatment, including analgesics, with both groups encouraged to make no changes.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gladwell 2006 *(Continued)*

Compliance acceptable in all groups (performance bias)	Low risk	The attendance of classes by the Pilates group was excellent with an overall attendance of 16/20 participants attending all sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Goldby 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, mixed **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 302 (E1 = 84, C1 = 40, C2 = 89) **Chronic LBP duration:** 11.72 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 30 **Sex (female):** 78%
Interventions	**Exercise Group 1 (E1):** Functionally progressive exercises for selective retraining of transversus abdominis, multifidus, pelvic floor, and diaphragm muscles with biofeedback; type = core strengthening; duration = 10 weeks; dose = low; design = individualised; delivery = group; additional intervention = electrotherapy & back school **Comparison Group 1 (C1):** Other conservative treatment (manual therapy) **Comparison Group 2 (C2):** Other conservative treatment (back school)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index); HRQoL (Nottingham Health profile) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Professional Organizational Funds **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Patients were randomly allocated to one of the groups using a stratification procedure.
Allocation concealment (selection bias)	Low risk	Author contact: researcher and managing physio knew treatment group but not research assistant nor treating physio.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Goldby 2006 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Author contact: Research assistant was blinded and researcher too; care was taken to document this process.
Incomplete outcome data (attrition bias) All outcomes	High risk	Of subjects, 10% were lost to follow-up between the three and 12-month stage, and 50% between the 12 and 24-month stage (Figure 2).
Participants analysed in group allocated (attrition bias)	High risk	The study population was defined as subjects remaining from baseline data collection up until and including the 12-month stage.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The study population was defined as subjects remaining from baseline data collection up until and including the 12-month stage.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: all patients had access to National Health Service and public physio.
Compliance acceptable in all groups (performance bias)	High risk	Patient attendance differed among the three groups.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Goode 2018

Study characteristics	
Methods	**Study design:** RCT (NCT02327325) **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 60 (E1 = 20, E2 = 20, C1 = 20) **Chronic LBP duration:** 5 years (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 70 **Sex (female):** 7%
Interventions	**Exercise Group 1 (E1):** Physical activity (core strengthening, stretching and aerobics), instructions for at-home exercise programme; type = mixed; duration = 12 weeks; dose = low; design = partially individualised; delivery = independent; additional intervention = psychological therapy **Exercise Group 2 (E2):** Physical activity: instructions for at-home exercise programme (written instructions with pictures, exercise video, physiotherapist instructions). Exercises included core strengthening, stretching and aerobics; type = mixed; duration = 12 weeks; dose = low; design = partially individualised; delivery = independent; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Goode 2018 *(Continued)*

Notes	**Conflicts of interest:** S. Taylor reported employment/money received from the Department of Veterans Affairs, Health Services Research & Development. **Funding source:** Department of Veterans Affairs Rehabilitation and Research Development (1I21RX001569- 01A1); Center of Innovation for Health Services Research in Primary Care (CIN 13-410), Durham Veteran's Affairs Health Care System **Other:** SDs imputed; sufficient data not available for inclusion in meta-analyses.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were randomised by use of a computer-generated randomisation table programmed into the study tracking database; participants were randomised with equal allocation to 1 of 3 study arms.
Allocation concealment (selection bias)	High risk	No information reported about whether or not treatment allocation was concealed until delivery to participant
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the treatments;2. Unlikely that lack of patient blinding caused significant deviation from intended intervention, as they all required instruction from a trained facilitator (except waiting-list group)
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from intended intervention; the only significant interaction participants had beyond baseline was in the CBT group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessor for outcome of interest (disability) were participants, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. There were clear degrees of intervention complexity and rank of "trendiness"; likely to alter assessment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 10 of 60 withdrew, 17%.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared safe to assume that intention-to-treat analysis was conducted
Selective reporting (reporting bias)	Low risk	1. No published protocol, but all analyses were performed and reported as described in methods.
Groups similar at baseline (selection bias)	Low risk	Most characteristics were similar at baseline among treatment groups; marriage status, education; every important characteristic was measured and reported.
Co-interventions avoided or similar (performance bias)	High risk	No information was reported about co-interventions during the treatment period; somewhat unlikely that waiting-list patients did not use any co-interventions
Compliance acceptable in all groups (performance bias)	High risk	No compliance information was reported; patients followed up by telephone weekly, but no report on reported physical activity completion, unlikely that compliance was very high (unsupervised, home-based exercise)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Goode 2018 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Within each outcome, all participants were measured with the same tools, thresholds and at the same time point; 2. Roland-Morris Disability Questionnaire is a well-validated tool for measuring disability in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Groessl 2017

Study characteristics	
Methods	**Study design:** RCT (NCT02524158) **Setting:** USA, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 152 (E1 = 76, C1 = 76) **Chronic LBP duration:** 15 years (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 53 **Sex (female):** 26%
Interventions	**Exercise Group 1 (E1):** Yoga; type = yoga; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = advice/education & other **Comparison Group 1 (C1):** Usual care/no treatment (delayed treatment and usual care)
Outcomes	**Core outcomes reported:** Pain (Brief Pain Inventory); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Veteran Affairs Rehabilitation Research and Development (Grant #RX000474) **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Binary non-stratified randomisation sequence was computer-generated, at a 1:1 allocation ratio in blocks of 10 participants to facilitate balanced group assignment.
Allocation concealment (selection bias)	Low risk	Study co-ordinator used a secure, web-based data management system to allocate treatment group.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. 3 of 76 admitted to not wanting to do yoga, compliance was frequently followed, minimal deviations from control.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Only yoga instructors were part of the trial; external care providers did not seem to be linked to the study.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcomes in question were assessed by the patient who was not blinded to the intervention; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in a no

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Groessl 2017 *(Continued)*

		exercise versus exercise study are likely to be altered by knowledge of assignment due to low expectations of no treatment; done in group setting; could be mixed between intervention.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 27 discontinued the yoga intervention; 120 of 152 were followed at 12 weeks; 21% lost; 2. Sensitivity analysis for missing data showed no indication of bias, although reporting on this was brief; used last-data-point-forward imputation which does not properly correct
Participants analysed in group allocated (attrition bias)	Low risk	1. "An 'intent-to-treat' approach was followed for all study outcomes".
Selective reporting (reporting bias)	Low risk	1. Primary outcomes were analysed and reported in line with registered protocol (NCT02524158); secondary outcomes not reported here: pain interference, fatigue, quality of life, etc.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Use of opioids, other medical pain treatments and self-help pain treatments were similar between groups during the study.
Compliance acceptable in all groups (performance bias)	High risk	Median number of instructor-led sessions attended by subjects randomised to yoga was 14 of 24 classes; compliance in control group was 96% (did not do yoga within study period).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Roland-Morris Disability Questionnaire (for function), and Brief Pain Inventory (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Gunay 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 63 (E1 = 31, E2 = 32) **Chronic LBP duration:** 20.4 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 84%
Interventions	**Exercise Group 1 (E1):** Muscle endurance training with 4 levels of stability exercises, 300 repetitions held for 10 seconds with regular rest; type = core strengthening; duration = 6 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = advice/education & heat/ice **Exercise Group 2 (E2):** Stretching and strengthening of abdominal, back and gluteal muscles (described as classical strengthening exercises); type = strengthening & stretching; duration = 6 weeks;

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gunay 2014 *(Continued)*

	dose = high; design = standardised; delivery = not specified; additional intervention = advice/education & heat/ice
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Author contact: block randomisation by computer-generated random number list
Allocation concealment (selection bias)	Low risk	Cards in unmarked envelopes
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The patients were blind to the intervention.
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Author contact: 'assessor blinded'; self-report with author report that patients were blinded to 'good' intervention
Incomplete outcome data (attrition bias) All outcomes	Low risk	Fourteen (of 77) patients withdrew because they were not able to complete the treatment programme.
Participants analysed in group allocated (attrition bias)	High risk	Only the 63 who completed were shown.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There was no significant difference between the two groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Author contact: medications were allowed, advised no other treatment or exercises
Compliance acceptable in all groups (performance bias)	Unclear risk	Author contact: not assessed formally (but observed compliance)
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gur 2003

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 75 (E1 = 25, E2 = 25, C1 = 25) **Chronic LBP duration:** 60 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 71%
Interventions	**Exercise Group 1 (E1):** Stretching and strengthening exercises of lumbar and extremity muscle groups; type = mixed; duration = 4 weeks; dose = low; design = partially individualised; delivery = independent; additional intervention = none **Exercise Group 2 (E2):** Stretching and strengthening exercises of lumbar and extremity muscle groups; type = strengthening & stretching; duration = 4 weeks; dose = low; design = partially individualised; delivery = independent; additional intervention = electrotherapy **Comparison Group 1 (C1):** Other conservative treatment (electrotherapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gur 2003 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Hagen 2010

Study characteristics	
Methods	**Study design:** RCT **Setting:** Norway, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 246 (E1 = 124, C1 = 122) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 41 **Sex (female):** 51%
Interventions	**Exercise Group 1 (E1):** Warm-up, circuit training (10 stations of strengthening and balance exercises), stretching (after a 2 week spine clinic intervention); type = mixed; duration = 8 weeks; dose = high; design = partially individualised; delivery = independent; additional intervention = advice/education & relaxation **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Low Back Pain Rating Scale (Manniche)); function (Roland-Morris Disability Questionnaire); work (sick leave (total length of leave, frequency of sick leave periods)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Norwegian Foundation for Health and Rehabilitation (Grant No. Nkr 840 000) **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hagen 2010 *(Continued)*

Random sequence generation (selection bias)	Low risk	The concealed randomisation was made according to a list prepared in advance at the University of Bergen (Norwegian Back Pain Network, Research Unit).
Allocation concealment (selection bias)	Low risk	A secretary at the Research Unit was responsible for the randomisation procedures and was not involved in the treatment of the patients.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	At the first visit at the spine clinic, before randomisation, all patients were examined by a physiotherapist using six different tests.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Eight participants dropped out.
Participants analysed in group allocated (attrition bias)	Low risk	Those with zero attendances were still considered as participants in the exercise group, according to the intention-to-treat principle.
Selective reporting (reporting bias)	Unclear risk	Not described
Groups similar at baseline (selection bias)	Unclear risk	There were few baseline characteristics available to be able to assess group differences.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Range of attendances in the physical exercise programme for the intervention group was 0-24; median 15.0.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Hall 2011

Study characteristics	
Methods	**Study design:** RCT (ACTRN12608000270314) **Setting:** Australia, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 160 (E1 = 80, C1 = 80) **Chronic LBP duration:** Not specified (not specified)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hall 2011 *(Continued)*

	Neurological/radicular symptoms: Some participants **Mean age (years):** 44 **Sex (female):** 74%
Interventions	**Exercise Group 1 (E1):** Tai Chi with warm-up and cool-down; type = other (Tai Chi); duration = 10 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 10 weeks (short)
Notes	**Conflicts of interest:** Dr. Lam has received royalties from the instructional video Tai Chi for Back Pain. **Funding source:** Arthritis Foundation of Australia; Arthritis Care of the United Kingdom **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number function in Microsoft Excel
Allocation concealment (selection bias)	Low risk	Sealed opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Envelopes were opened and patients were assigned Tai Chi or waiting-list control, therefore unblinded.
Blinding of care provider (performance bias)	High risk	Treatment provider was with the patient when envelope was opened, assigning groups.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The groups were not statistically significantly different at baseline with regard to age, sex, self-reported chronic pain grade.
Co-interventions avoided or similar (performance bias)	Unclear risk	All participants were instructed to continue with their normal health care (general practioner visits and recommendations) + fitness.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hall 2011 *(Continued)*

Compliance acceptable in all groups (performance bias)	High risk	In the intervention group, 28.8% adhered to the intervention; 57.5% of the total treatment group attended 50% or more Tai Chi sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Halliday 2016

Study characteristics	
Methods	**Study design:** RCT (ACTRN12611000971932) **Setting:** Australia, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 70 (E1 = 35, E2 = 35) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 49 **Sex (female):** 80%
Interventions	**Exercise Group 1 (E1):** Motor control exercises: independent contraction of the deep stabilisation muscles using pelvic floor contraction; type = core strengthening; duration = 8 weeks; dose = high; design = standardised; delivery = independent with follow-up; additional intervention = none **Exercise Group 2 (E2):** McKenzie exercises guided by therapist; type = McKenzie; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Patient Specific Functional Scale); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Dr. Clare is currently a member of The McKenzie Institute International Board of Trustees and is currently the International Director of Education for The McKenzie Institute International. **Funding source:** International MDT Research Foundation; Disability Services, University of Sydney **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated numbers by a researcher not involved in the data analyses
Allocation concealment (selection bias)	Low risk	Sequentially numbered, opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	1. Patients received either McKenzie exercises or not; unclear if patients thought all treatment's were plausible (possible)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Halliday 2016 *(Continued)*

Blinding of care provider (performance bias)	High risk	1. Care provider was aware of what intervention he/she was providing.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Patient report
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Three individuals from the intervention group and five individuals from the motor control group were lost to follow-up.
Participants analysed in group allocated (attrition bias)	Unclear risk	1. Stated intention-to-treat, but analysed complete case with no imputation; partial; patients with follow-up kept in randomised groups; low risk of bias with consensus
Selective reporting (reporting bias)	Low risk	1. See Table 2
Groups similar at baseline (selection bias)	Unclear risk	Twenty per cent difference in off work at baseline and longer median duration in motor control group
Co-interventions avoided or similar (performance bias)	Low risk	Other co-interventions not reported
Compliance acceptable in all groups (performance bias)	Low risk	Similar number of sessions attended
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Baseline and eight-week follow-up

Hansen 1993

Study characteristics	
Methods	**Study design:** RCT **Setting:** Denmark, mixed **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 180 (E1 = 60, C1 = 61, C2 = 59) **Chronic LBP duration:** 28.6 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 41 **Sex (female):** 32%
Interventions	**Exercise Group 1 (E1):** Intensive dynamic back-muscle training: extension strengthening; type = strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy) **Comparison Group 2 (C2):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall treatment effect))

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hansen 1993 *(Continued)*

	Follow-up time periods available for syntheses: 4 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** The Danish Rheumatism Association; The Health Insurance Foundation; The Rockwood Foundation **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Harris 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Norway, occupational **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 214 (E1 = 60, C1 = 99, C2 = 55) **Chronic LBP duration:** 11.48 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Group Physical Exercise and Brief Cognitive Intervention: exercises were adapted by physiotherapist, strength and endurance training, relaxation; type = functional restoration; duration = 12 weeks; dose = high; design = individualised; delivery = group; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (psychological therapy) **Comparison Group 2 (C2):** Other conservative treatment (psychological therapy)
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index); work (% still on sick leave) **Follow-up time periods available for syntheses:** 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Research Council of Norway (175466/V50); Norwegian Extra Foundation for Health and Rehabilitation (EXTRA funds) **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was adequate, performed according to a computer-generated randomisation list, generated by the trial statistician; stratified by clinic and gender.
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed using a central telephone randomisation system.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their assignment due to the nature of the treatments; 2. Unlikely that lack of patient blinding altered the intended interventions, as intervention delivery was controlled by care providers
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Different care providers specialised in different treatments provided the care.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors were patients, who could not be blinded to treatment assignment due to nature of interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Because all of the treatment groups were quite distinct and there were clearly "better" groups, it is quite likely that lack of patient blinding led to biased outcome assessments.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 68 of 215 had no follow-up information for the outcomes of interest, 32% lost; 2. Mixed ANOVA dropped observations with missing data; 3. Missingness

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Harris 2017 *(Continued)*

		could be caused by increased disability from low back pain; 4. 25%, 25.5%, 39%
Participants analysed in group allocated (attrition bias)	Low risk	1. Intent-to-treat analysis was included.
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: all planned analyses were executed and reported for all primary and secondary outcomes.
Groups similar at baseline (selection bias)	Low risk	All relevant characteristics were measured at baseline and were similar across treatment groups.
Co-interventions avoided or similar (performance bias)	High risk	No mention of co-interventions in the study or in its protocols
Compliance acceptable in all groups (performance bias)	High risk	No mention of compliance or adherence to the interventions
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All patients were assessed for outcomes identically, regardless of treatment group; 2. Oswestry Disability Index (for function) is a well-validated tool in the low back pain context.
Other bias	Low risk	To increase statistical power, all participants allocated to Brief Cognitive Intervention regardless of centre were compared to the two experimental groups; broke requirements for RCT; sensitivity analysis done but vague

Harts 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** Netherlands, occupational **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 65 (E1 = 23, E2 = 21, C1 = 21) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 42 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Progressive resistance exercises for the isolated lumbar extensor muscle groups (load set at 50% of maximal isometric extension strength of participant at baseline); type = strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Nonprogressive low-intensity resistance exercise programme (load set at a maximum of 20% of the maximal isometric extension strength of participant at baseline); type = strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Harts 2008 *(Continued)*

	Follow-up time periods available for syntheses: 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Concealed randomisation was performed by means of a computer-generated table of random numbers with a block size of six.
Allocation concealment (selection bias)	Low risk	Randomised, three-arm trial with concealed allocation, assessor blinding
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Measurement of outcomes was carried out by the principal investigator or research assistants who were not aware of group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Six participants withdrew during the intervention period.
Participants analysed in group allocated (attrition bias)	Unclear risk	All analyses were carried out according to the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were only minor differences in demographic characteristics.
Co-interventions avoided or similar (performance bias)	Low risk	No co-interventions were reported during both the intervention and the waiting-list period.
Compliance acceptable in all groups (performance bias)	Low risk	A total of 31 participants received the high-intensity strengthening programme.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hartvigsen 2010

Study characteristics	
Methods	**Study design:** RCT (NCT00209820) **Setting:** Denmark, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 136 (E1 = 45, E2 = 46, C1 = 45) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 47 **Sex (female):** 72%
Interventions	**Exercise Group 1 (E1):** Supervised outdoor sessions of Nordic walking, individual pace allowed to vary; type = aerobic; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** One hour Nordic walking instruction and advice to Nordic walk at home; type = aerobic; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Low Back Pain Rating Scale (Manniche)); function (Low Back Pain Rating Scale (function)); work (time off work) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomisation was carried out by a project secretary after collection of the baseline data. Participants drew a sealed opaque envelope containing info.
Allocation concealment (selection bias)	Low risk	Randomisation was carried out by a project secretary after collection of the baseline data.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	During the eight intervention weeks, five, four, and one participants dropped out of the groups and did not contribute with follow-up data at any point.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hartvigsen 2010 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	Data were analysed on an intention-to-treat basis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The randomisation resulted in three groups comparable in all baseline variables.
Co-interventions avoided or similar (performance bias)	Low risk	Also use of over-the-counter pain medication or use of concurrent treatment during the one-year follow-up period were not statistically significant.
Compliance acceptable in all groups (performance bias)	Low risk	Altogether 25 (50.4%) of participants in the supervised nordic walking group and 29 (65.2%) in the unsupervised nordic walking group contributed data for seven or more days.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Hasanpour-Dehkordi 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Iran, not specified **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 36 (E1 = 12, E2 = 12, C1 = 12) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Pilates, with a stretching and walking cool-down portion; type = Pilates; duration = 6 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** McKenzie programme, performed extension and flexion exercises in supine and sitting positions; type = McKenzie; duration = 6 weeks; dose = high; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score Pain Questionnaire) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hasanpour-Dehkordi 2017 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Random number was used to randomly enrol, but no mention of randomisation to group made other than the vague "randomly assigned"
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Patients in the control group were likely to seek out an exercise as they received no treatment in the study; they would be able to access at least one of the interventions (Pilates); 3. No information on deviations from protocol; 4. Control group seeking exercise interventions could improve their low back pain and disability.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were all quite distinct
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	High risk	1. No description of dropout rate 2. ANOVA does not handle missing data; no sensitivity analyses were conducted; 3. Missingness could be caused by increased disability from low back pain; 4. No information on dropout rate in either group; 5. No information on reasoning for any dropouts
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. Mentioned study protocol but could not be found: within this publication, no pre-planned outcomes were indicated in methods, but standard outcomes reported analyses in methods were fully reported.
Groups similar at baseline (selection bias)	High risk	At baseline, groups had significantly different pain and general health; no information on duration of symptoms (age was balanced, all patients were male)
Co-interventions avoided or similar (performance bias)	Low risk	Study excluded patients who were undergoing other therapies during the study period (co-interventions).
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance, adherence or attendance in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. McGill Pain Score Questionnaire (for pain) is a well-validated tool in the low back pain context.
Other bias	Low risk	No other sources of bias noticed; generally very poorly reported

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Helmhout 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** Netherlands, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 81 (E1 = 41, E2 = 40) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 41 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Progressive resistance exercises for the isolated lumbar extensor muscle groups (load set at 50% of maximal isometric extension strength of participant at baseline); type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Nonprogressive low-intensity resistance exercise programme (load set at a maximum of 20% of the maximal isometric extension strength of participant at baseline); type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was done by means of a computer-generated table.
Allocation concealment (selection bias)	Low risk	The study is an observer-blinded trial.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Same physio delivered both programmes.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The training sessions and data collection (both from questionnaires and from strength measurements) were conducted by different people.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Six at post-treatment, five more at six months, eight more at nine months = 62/81 total completed

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Helmhout 2004 *(Continued)*

Participants analysed in group allocated (attrition bias)	High risk	Analyses were done using all randomised participants who provided any post-baseline data.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Table 1 shows demographic properties of the population as well as baseline values for the outcome measures.
Co-interventions avoided or similar (performance bias)	Low risk	Participants were asked to abandon other treatments In both groups; participants reported no co-interventions during the 12-week treatment period.
Compliance acceptable in all groups (performance bias)	High risk	The high-intensity training group showed a higher treatment compliance than the low-intensity training group: 29 participants (71%) in the high-intensity training group, versus 19 participants (48%) in the low-intensity training group
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Hemmilä 1997

Study characteristics	
Methods	**Study design:** RCT **Setting:** Finland, mixed **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 114 (E1 = 35, C1 = 34, C2 = 45) **Chronic LBP duration:** 390 weeks (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 43%
Interventions	**Exercise Group 1 (E1):** Bending, rotation exercises; "auto-stretching when appropriate"; type = core strengthening & stretching; duration = 6 weeks; dose = low; design = standardised; delivery = independent; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy) **Comparison Group 2 (C2):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Finnish Slot Machine Association **Other:** SDs imputed
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hemmilä 1997 *(Continued)*

Random sequence generation (selection bias)	High risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Henchoz 2010

Study characteristics	
Methods	**Study design:** RCT **Setting:** Switzerland, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 103 (E1 = 55, E2 = 48) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Henchoz 2010 *(Continued)*

	Mean age (years): 40 **Sex (female):** 39%
Interventions	**Exercise Group 1 (E1):** Post-functional restoration programme: warm-up (15 minutes), core strengthening exercises, exercise for trunk flexor and extensor, upper and lower limb muscle using equipment, functional tasks, stretching (15 minutes); type = mixed; duration = 15 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Post-functional restoration programme: written description of the exercises they performed during functional restoration, recommended to continue at home; type = mixed; duration = 15 weeks; dose = high; design = individualised; delivery = independent with follow-up; additional intervention = psychological therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 15 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation occurred on completion of functional multidisciplinary rehabilitation by means of sequentially numbered envelopes that were prepared in advance by a person who was not involved.
Allocation concealment (selection bias)	Low risk	Randomisation occurred on completion of functional multidisciplinary rehabilitation by means of sequentially numbered envelopes that were prepared in advance by a person who was not involved in the study.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Evaluators were not blinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	Thirteen subjects (23%) out of the 56 patients enrolled in the exercise programme group did not complete the exercise programme.
Participants analysed in group allocated (attrition bias)	Low risk	According to the intention-to-treat principle, analyses were done using all randomised participants who provided any follow-up data.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant difference was found between groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Henchoz 2010 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	No patient had to be excluded because of missing more than six exercise sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Henry 2014

Study characteristics	
Methods	**Study design:** RCT (NCT01362049) **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 101 (E1 = 25, E2 = 76) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 42 **Sex (female):** 51%
Interventions	**Exercise Group 1 (E1):** Spinal stability exercise: motor control of the deep trunk muscles; strengthening of the flexor, extensor, and oblique trunk muscles; "standard" physiotherapy exercises; home exercise log; type = core strengthening; duration = 6 weeks; dose = high; design = standardised; delivery = individual; additional intervention = advice/education & other & physiotherapy **Exercise Group 2 (E2):** Movement System Impairment: specific trunk movements and postures focussed on pain-free and functional activity Oswestry Disability Indexification, "standard" physiotherapy exercises, home exercise log; type = other (not specified); duration = 6 weeks; dose = high; design = standardised; delivery = individual; additional intervention = advice/education & other & physiotherapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 7 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** National Institutes of Health (NIH/NCMRR/R01HD040909 **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The statistician used computer-generated randomisation with centralised allocation concealment to randomise subjects.
Allocation concealment (selection bias)	Low risk	The statistician used computer-generated randomisation with centralised allocation concealment to randomise subjects.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Henry 2014 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Physical therapy clinicians who provided the treatment and the subjects were masked to strata assignment but not to treatment.
Blinding of care provider (performance bias)	High risk	Physical therapy clinicians who provided the treatment and the subjects were masked to strata assignment but not to treatment.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Physical therapy clinicians who provided the treatment and the subjects were masked to strata assignment but not to treatment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	Low risk	Author contact: the primary analysis was an intention-to-treat analysis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no differences in baseline characteristics between the two groups.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: changed to yes
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: changed to yes
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Highland 2018

Study characteristics	
Methods	**Study design:** RCT (NCT02132910) **Setting:** USA, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 68 (E1 = 34, C1 = 34) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 44 **Sex (female):** 63%
Interventions	**Exercise Group 1 (E1):** Yoga with a special focus on major muscles affected by chronic LBP (back and core); type = yoga; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = other

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Highland 2018 *(Continued)*

	Comparison Group 1 (C1): Usual care/no treatment (usual care based on participants provider: medication, physiotherapy, chiropractic care, massage, supplements, etc.)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (Patient-Reported Outcomes Measurement Information System-29 Symptom Burden) **Follow-up time periods available for syntheses:** 13 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** U.S. Department of the Army, Telemedicine and Advanced Technology Research Center; U.S. Army Medical and Materiel Command; Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. under Cooperative Agreement (W81XWH-11-2-0201) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A computerised random number generator produced the randomisation table.
Allocation concealment (selection bias)	Low risk	Participants opened pre-sealed envelopes labelled with the sequential enrolment number containing group assignment to conceal treatment allocation.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from the intended interventions because patients had little control over intervention delivery
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely for regular care providers to recommend or conduct yoga type treatment in a clinical setting; participants were randomised to yoga instructors.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Control was care-as-usual whereas the experimental group received a somewhat "trendy" intervention, which increased risk of bias; follow-up was 4 weeks after intervention, 3 months after baseline, potentially lowering risk of bias.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Follow-up data was available for all but 6 of 68 participants at the 3-month follow-up, 9%.
Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported according to registered protocol (NCT02132910), though a few more outcomes were in the article than the protocol (Physical Functioning and Symptom Burden).
Groups similar at baseline (selection bias)	High risk	Several characteristics were collected; only statistical test results were reported; data were not presented for interpretation.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Highland 2018 *(Continued)*

Co-interventions avoided or similar (performance bias)	High risk	Control group (treatment-as-usual) purposely allowed patients to use any medications, physiotherapy or other therapy that their care provider felt was necessary.
Compliance acceptable in all groups (performance bias)	Low risk	Compliance in treatment group was 82% (334 estimated attended sessions out of 408 maximum possible for 34 yoga group members); assumed 100% compliance for treatment-as-usual
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Roland-Morris Disability Questionnaire (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Hildebrandt 2000

Study characteristics	
Methods	**Study design:** RCT **Setting:** Netherlands, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 222 (E1 = 112, C1 = 110) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 42 **Sex (female):** 49%
Interventions	**Exercise Group 1 (E1):** Postural exercises (Cesar therapy); type = other (co-ordination); duration = 12 weeks; dose = low; design = not specified; delivery = individual; additional intervention = not specified **Comparison Group 1 (C1):** Usual care/no treatment (usual general practitioner care)
Outcomes	**Core outcomes reported:** Global Perceived Health or Recovery (Global Perceived Health or Recovery (global improvement (recovered)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias)	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hildebrandt 2000 *(Continued)*
All outcomes

Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Unclear risk	Support for judgement is not available

Iversen 2018

Study characteristics	
Methods	**Study design:** RCT (NCT02420236) **Setting:** Norway, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 99 (E1 = 50, E2 = 49) **Chronic LBP duration:** 1 year (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Elastic resistance band training; type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = independent with follow-up; additional intervention = advice/education & psychological therapy **Exercise Group 2 (E2):** General physical activity; type = mixed; duration = 12 weeks; dose = low; design = partially individualised; delivery = independent with follow-up; additional intervention = advice/education & psychological therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Iversen 2018 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index); work (Work Ability Index 0-10); HRQoL (EuroQol 5D); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Rating of Change Scale)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Kommunal Landspensjonskasse, The Norwegian Research Council **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was adequate, performed at 1:1, using block randomisation with unknown block sizes varying between 10 and 20, by a third party.
Allocation concealment (selection bias)	Low risk	Treatment allocation concealment was adequate as randomisation was performed by a third party.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their intervention due to the nature of the treatments; 2. Unlikely that lack of patient blinding led to deviations, as most of the control over the intervention was held by training therapists; both groups received exercise interventions.
Blinding of care provider (performance bias)	Low risk	1. Care providers (physiotherapist, physicians, social workers, psychologists) could not be blinded to the intervention as they were responsible for delivering the exercise interventions; 2. Unlikely that lack of care provider blinding led to deviations from intended interventions, as interventions were standardised and there was not a strong indication as to which treatment was "better".
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. For the outcomes of pain and function, outcome assessors were the patients themselves, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Unlikely that the assessment of outcomes was altered by lack of patient blinding because the treatment groups were quite similar, with no clear "better" group.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 24 + 22 followed up at 12 weeks; 46 of 99 dropped out before the 12-week follow-up; 46% dropout rate; 2. Dropouts likely did not bias the type of patient who was used in the final analysis of the study; "There was no significant difference between participants that completed and those who dropped out".
Participants analysed in group allocated (attrition bias)	Low risk	1. "Primary and secondary outcomes were analysed in accordance with the intention-to-treat principle".
Selective reporting (reporting bias)	Low risk	1. All outcomes and analyses were performed as outlined in previously published study protocol; 6 and 12-month outcome data forthcoming
Groups similar at baseline (selection bias)	Low risk	No significant baseline differences were observed between participants in the trial and the reference group, except for a higher proportion of people being sick listed in the trial.
Co-interventions avoided or similar (performance bias)	Low risk	Only co-intervention that was directly assessed was use of analgesics; there was comparable use of analgesics among all treatment groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Iversen 2018 *(Continued)*

Compliance acceptable in all groups (performance bias)	High risk	Only report of compliance was that 14/24 participants completed > = 60% of the elastic band exercises; no difference in conclusion from either intent-to-treat or per-protocol analysis
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All participants were assessed for outcomes identically, regardless of treatment group; 2. Oswestry Disability Index (for function) and Numerical Rating Scale (for pain) are both well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Jans 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** Netherlands, not specified **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 201 (E1 = 96, C1 = 105) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 34 **Sex (female):** 67%
Interventions	**Exercise Group 1 (E1):** Postural exercises (Cesar therapy) (after a 12-week exercise intervention); type = other (co-ordination); duration = not specified weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Graded Chronic Pain Scale (von Korff)); function (Quebec Back Pain Disability Scale); Global Perceived Health or Recovery (Global Perceived Health or Recovery (self-report recovery; general health (categorical)) **Follow-up time periods available for syntheses:** 26 weeks (moderate); 78 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jans 2006 *(Continued)*

Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Jaromi 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** Hungary, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 111 (E1 = 56, C1 = 55) **Chronic LBP duration:** 29 weeks (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 32 **Sex (female):** 17%
Interventions	**Exercise Group 1 (E1):** Ergonomic training exercises, strengthening and stretching exercises, home exercise; type = strengthening & stretching; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education & back school **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate); 52 weeks (long)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jaromi 2012 *(Continued)*

Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	The participant's names were listed in alphabetic order and each second participant was selected for intervention group.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Some dropouts
Participants analysed in group allocated (attrition bias)	High risk	Only the 111 were analysed.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	During the intervention study, participants were not given any other further therapies based on the medical recommendations.
Compliance acceptable in all groups (performance bias)	Unclear risk	Only mentioned how many days per week on average people exercised
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Measurements were taken at baseline, after the interventions and at the six-months and one-year follow-up visits.

Jaromi 2018

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jaromi 2018 *(Continued)*

Methods	**Study design:** RCT **Setting:** Hungary, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 146 (E1 = 73, C1 = 73) **Chronic LBP duration:** 26.6 weeks (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 42 **Sex (female):** 93%
Interventions	**Exercise Group 1 (E1):** Spine Care for Nurses (Back School): different types of exercises (core strengthening, stretching, functional restoration, aerobic) practised for 20 minutes per day, 5 days per week at home; type = mixed; duration = 12 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Subjects were randomly assigned to either the experimental or control groups by drawing lots using the group numbers.
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	High risk	1. Difficult to blind participants to allocation, given the distinct nature of experimental and control groups; 2. All participants were nurses at a few different hospitals, but could have been co-workers; they likely discussed their experiences during the trial, sharing information from both treatments; 3. If participants did indeed discuss their trial experiences during the treatment period, then the control group likely stood to gain more than the experimental group by deviating from protocol; 4. If control group participants deviated from the protocol, then likely their outcomes would have been improved compared to a participant who adhered to the control group protocol
Blinding of care provider (performance bias)	Low risk	1. Care providers (physiotherapist) could not be blinded to allocation because they were responsible for delivering the interventions; 2. Lack of care provider could have led to deviation from protocol, but it appeared that the control group only had brief interaction with the provider, so unlikely to have received substantial deviated treatment.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For our primary outcome (pain), the outcome assessor was the participant, who was not blinded to allocation. For any other outcomes measured, the research staff assessing the outcome was blinded; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Reasonably likely that knowledge of allocation led to altered judgement by participants assessing the pain outcome, due to the distinct nature of the interventions

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jaromi 2018 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 137 of 146 were followed up.
Participants analysed in group allocated (attrition bias)	Low risk	1. Patients were analysed according to treatment allocation (see flow diagram).
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: within this context all outcomes and analyses were fully reported; did not measure function via questionnaire but had various examinations of lifting techniques
Groups similar at baseline (selection bias)	Low risk	No clinically significant differences in age, sex, body mass index, years of work, duration of pain, pain intensity or any other outcome measures
Co-interventions avoided or similar (performance bias)	High risk	Co-interventions not reported; possible that participants could have been using medication or over-the-counter therapies (heat pack, etc.) However, excluded participants with physiotherapy in past 3 months
Compliance acceptable in all groups (performance bias)	High risk	No reporting of attendance to sessions; weekly educational sessions were supervised, but at home (or at work) intervention practice not reported
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Within each outcome, all participants were measured with the same tools and at the same time point; 2. Visual Analogue Scale is a well-validated tool for measuring pain in the context of low back pain.
Other bias	Low risk	No other sources of bias noticed

Javadian 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** Iran, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 30 (E1 = 15, E2 = 15) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Warm-up (cycling, stretching, 15 minutes); training exercises including knee to chest, bridging, supine cycling, heel slides, abdominal crunches; type = mixed; duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Bracing and hollowing exercises in supine, bridging, kneeling, sitting and standing, progressing to on wobble board and Swiss ball; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Tehran University of Medical Sciences

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Javadian 2012 *(Continued)*

Other: SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were divided randomly between the control and experimental groups.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unclear
Participants analysed in group allocated (attrition bias)	Unclear risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Not described
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Eight weeks (post-treatment), three months

Jensen 2012

Study characteristics

Methods	**Study design:** RCT (NCT00454792) **Setting:** Denmark, healthcare

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jensen 2012 *(Continued)*

	Exercise groups: 1 **Comparison groups:** 1
Participants	**Number of participants:** 100 (E1 = 51, C1 = 49) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 46 **Sex (female):** 68%
Interventions	**Exercise Group 1 (E1):** Exercises for stabilising muscles in the low back and abdomen, dynamic exercises, exercises for postural instability, light physical fitness training; type = core strengthening & aerobic; duration = 10 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (advised to avoid hard physical activity and to rest lying down 2 x 1 hour per day and instructed to wear a lumbar belt)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); work (sick leave (by text)); HRQoL (EuroQol 5D); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global assessment (7-point)) **Follow-up time periods available for syntheses:** 10 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** VELUX Foundation **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	On receiving the completed baseline questionnaire, the project secretary allocated each patient into one of the two intervention groups.
Allocation concealment (selection bias)	Low risk	The researchers were masked to group assignment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The researchers analysing the data were blinded to patient identification as the study participants were given a random identification number.
Incomplete outcome data (attrition bias) All outcomes	Low risk	At the end of the intervention period, 78 patients had completed the full treatment programme (dropout rate 22%).
Participants analysed in group allocated (attrition bias)	Low risk	All data were analysed using the intention-to-treat (ITT) principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jensen 2012 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the two groups regarding any of the baseline variables, including patient expectations of treatment effectiveness.
Co-interventions avoided or similar (performance bias)	Low risk	interventions equal between groups
Compliance acceptable in all groups (performance bias)	Low risk	On average, patients participated in 84% of the sessions in the rest group or 91% of sessions if only those who completed the intervention were considered.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Johanssen 1995

Study characteristics	
Methods	**Study design:** RCT **Setting:** Denmark, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 27 (E1 = 14, E2 = 13) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 38 **Sex (female):** 48%
Interventions	**Exercise Group 1 (E1):** Aerobics exercises emphasising co-ordination, balance, stability, stretching; type = mixed; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Aerobics, dynamic exercises emphasising muscle endurance, abdominal, shoulder, hip muscles stretching; type = mixed; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Activities of Daily Living Scale); work (sick leave) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Johanssen 1995 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Jousset 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** France, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 85 (E1 = 43, E2 = 42) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 40 **Sex (female):** 33%
Interventions	**Exercise Group 1 (E1):** Functional restoration; type = functional restoration; duration = 5 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = psychological therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jousset 2004 *(Continued)*

	Exercise Group 2 (E2): Active individual therapy (strengthening, stretching, aerobics recommended); type = mixed; duration = 5 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education & psychological therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Quebec Back Pain Disability Scale); work (sick leave days at 6 months) **Follow-up time periods available for syntheses:** 24 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Union Regionale des Caisses d'Assurance Maladie des Pays de Loire **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	High risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jousset 2004 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Kaapa 2006

Study characteristics

Methods	**Study design:** RCT **Setting:** Finland, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 125 (E1 = 64, E2 = 61) **Chronic LBP duration:** 15 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 46 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** General exercises including stationary bikes, low impact aerobics, walking; muscle strengthening exercises using equipment, spinal stabilising exercises, functional exercises for posture and control, home exercises; type = mixed; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = psychological therapy & back school **Exercise Group 2 (E2):** Light exercise, stretching, spine mobilisation, core stability, home exercises; type = mixed; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = electrotherapy & manual therapy
Outcomes	**Core outcomes reported:** Pain (Low Back Pain Rating Scale (Manniche)); function (Oswestry Disability Index); work (subjective working capacity (rated on a scale of 0-10, sick leave due to back pain (classified scale: 0 days, 1-30 days, over 30 days during past 12 months)); Global Perceived Health or Recovery (Global Perceived Health or Recovery (general well-being after treatment)) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Foundation funds (source not indicated) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation list was generated by an independent biostatistician using a table of random numbers, and results were kept in sealed envelope.
Allocation concealment (selection bias)	Low risk	The physiotherapist randomised each patient into one of the two groups by opening an opaque sealed envelope.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kaapa 2006 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	In total, 96% (six months follow-up), 89% (12 months follow-up), and 79% (24 months follow-up) of the included patients provided follow-up information.
Participants analysed in group allocated (attrition bias)	Low risk	All patients were included in the analysis on the basis of their intervention allocation.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	At baseline, patients were comparable in each treatment arm (Table 2).
Co-interventions avoided or similar (performance bias)	Unclear risk	Significant use of passive interventions (30-40 minutes out of 60 minutes) in individual physiotherapy arm and multidisciplinary management in multidisciplinary rehabilitation arm.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kader 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** United Kingdom, not specified **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 63 (E1 = 20, E2 = 22, E3 = 21) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Mulligan and McKenzie mobilisation techniques, deep core strengthening, swimming and walking; type = mixed; duration = 10 weeks; dose = high; design = not specified; delivery = individual; additional intervention = advice/education & other & physiotherapy **Exercise Group 2 (E2):** Gym ball exercise including intense use of a gym ball in the clinic and at home; type = core strengthening; duration = 10 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = advice/education **Exercise Group 3 (E3):** Home exercise at least 3 times weekly (walking and cycling); type = aerobic; duration = 10 weeks; dose = high; design = standardised; delivery = independent; additional intervention = anti-inflammatory/analgesics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kader 2012 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (McGill Pain Score); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 10 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The randomisation process was not stratified.
Allocation concealment (selection bias)	Unclear risk	Randomised using sealed envelopes to one of three treatment groups
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Participants completed questionnaires.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	Unclear risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Stated that "The three groups were similar at trial entry" but no statistics were presented
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kankaanpaa 1999

Study characteristics	
Methods	**Study design:** RCT **Setting:** Finland, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 54 (E1 = 30, C1 = 24) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 35%
Interventions	**Exercise Group 1 (E1):** Strengthening (machines) with stretching, co-ordination, David Beck Clinic programme (functional restoration); type = mixed; duration = 12 weeks; dose = high; design = individualised; delivery = group; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Pain Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Device status category: 2 **Funding source:** Ministry of Education; Academy of Finland; Finnish Work Environmental Fund; The Finnish Medical Society Duodecim; Yrjo Jahnsson; Eemil Aaltonen; Instrumentarium Science Foundations; Kuopio University EVO Fund (496115) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kankaanpaa 1999 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Keane 2017

Study characteristics

Methods	**Study design:** RCT **Setting:** England, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 29 (E1 = 10, E2 = 10, C1 = 9) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 47 **Sex (female):** 83%
Interventions	**Exercise Group 1 (E1):** Aquastretch: assisted stretching of the whole body in water; type = stretching; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Land based stretching; type = stretching; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Modified Oswestry Disability Index Oswestry Low Back Pain (MOLBPQ)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Subjects were assigned randomly; sequence generation not described.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Keane 2017 *(Continued)*

Allocation concealment (selection bias)	Low risk	Sealed envelopes were used.
Blinding of participants and personnel (performance bias) All outcomes	High risk	1. Not mentioned, but due to nature of intervention patients were likely not blinded; 2. Control group maintained pre-trial regimen; may not be consistent controls; two exercise groups were different enough that expectations may have differed; 3. Control group maintained pre-trial regimen (i.e. exercise) but did not receive anything from this study; this may bias them as control; 4. Treatments in the control group may have altered outcomes.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that control group had no contact with care providers and other interventions were explicit (land versus pool)
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcomes in question were pain and function which were assessed by the patients who were not blinded; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. One group was control, while the other two received different interventions.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 69% of people recruited took part, but one dropped out after the study started.
Participants analysed in group allocated (attrition bias)	Low risk	1. Not explicitly mentioned, however, it seemed like intention-to-treat was done; no dropouts and no deviations mentioned
Selective reporting (reporting bias)	Low risk	1. No obvious omissions despite lack of linked protocol
Groups similar at baseline (selection bias)	Low risk	No significant differences in baseline age; there were other important variables such as body mass index and pain duration that were not reported on.
Co-interventions avoided or similar (performance bias)	Low risk	Co-interventions were avoided for exercise groups; no new interventions were recommended for the controls.
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance with assigned intervention; unlikely to be perfect compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All subjects submitted visual analogue scale measures by email once a week; all Modified Oswestry Low Back Pain Questionnaire questionnaires were completed at the first meeting and again at weeks 6 and 12 2. Visual Analogue Scale and Modified Low Back Pain Disability Questionnaire are valid measures.
Other bias	Low risk	No other obvious sources of bias

Kell 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Canada, general population **Exercise groups:** 2 **Comparison groups:** 1

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kell 2009 *(Continued)*

Participants	**Number of participants:** 27 (E1 = 9, E2 = 9, C1 = 9) **Chronic LBP duration:** 27.6 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 37 **Sex (female):** 40%
Interventions	**Exercise Group 1 (E1):** Resistance training for upper and lower body using free weights, machines, and body weight; type = strengthening; duration = 16 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Aerobic training including any form of aerobic exercise which was of interest to the participant (i.e. elliptical trainer, treadmill walking or jogging); type = aerobic; duration = 16 weeks; dose = high; design = individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no description)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 16 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Saskatchewan Health Research Foundation (New Investigator Grant); University of Alberta, Augustana Campus (travel grant) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The 27 subjects were randomly assigned to one of three groups.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	High risk	Initially, there were 33 subjects in the study, but six subjects dropped out, leaving 27 who completed the 16-week study.
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences were apparent among the groups at baseline, week eight, or week 16 for body composition (Table 1).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kell 2009 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	The present study sought to eliminate this problem by isolating two forms of exercise rehabilitation, resistance training and aerobic training, and stopping all other exercise modes.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kendall 1968

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 42 (E1 = 14, E2 = 14, E3 = 14) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Mobilising, strengthening, posture; type = mixed; duration = 12 weeks; dose = low; design = standardised; delivery = independent; additional intervention = advice/education **Exercise Group 2 (E2):** Isometric flexion exercises: strengthening abdominal and trunk muscles (2 exercises); type = core strengthening; duration = 12 weeks; dose = low; design = standardised; delivery = independent; additional intervention = advice/education **Exercise Group 3 (E3):** Strengthening extension muscles, posture, lifting; type = strengthening & flexibility/mobilising; duration = 12 weeks; dose = low; design = standardised; delivery = independent; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Global Perceived Health or Recovery (Global Perceived Health or Recovery (number improved, symptom-free)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kendall 1968 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	High risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Unclear risk	Support for judgement is not available

Kendall 2015

Study characteristics	
Methods	**Study design:** RCT (NCT01567566) **Setting:** Canada, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 80 (E1 = 40, E2 = 40) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 37 **Sex (female):** 52%
Interventions	**Exercise Group 1 (E1):** Lumbopelvic motor control (core and pelvic strengthening exercises) with biofeedback, open and closed kinetic chain hip strengthening exercises; home exercises; type = core

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kendall 2015 *(Continued)*

	strengthening; duration = 6 weeks; dose = high; design = partially individualised; delivery = not specified; additional intervention = advice/education **Exercise Group 2 (E2):** Lumbopelvic motor control (core and pelvic strengthening exercises) with biofeedback, home exercises, home exercise log; type = core strengthening; duration = 6 weeks; dose = high; design = partially individualised; delivery = not specified; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Workers Compensation Board Alberta's Research Grant **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Simple randomisation using a computer-generated table of random numbers
Allocation concealment (selection bias)	Low risk	Opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Not clear - trial reported "single blinded"
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Four/40 and 5/40 participants were lost to follow-up, with good reasons given.
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The randomisation was successful in creating equivalent groups at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Participants were asked to refrain from seeking any other types of rehabilitation treatments, yoga, or Pilates during the trial.
Compliance acceptable in all groups (performance bias)	Low risk	Adherence to the programme was similar in both groups and was considered excellent.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kendall 2015 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kim 2015a

Study characteristics

Methods	**Study design:** RCT **Setting:** South Korea, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 74 (E1 = 37, C1 = 37) **Chronic LBP duration:** 9.65 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 29 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** CORE exercise programme; type = core strengthening; duration = 8 weeks; dose = high; design = standardised; delivery = independent with follow-up; additional intervention = electrotherapy & heat/ice **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 8 weeks (short); 16 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	All participants were randomly assigned to the CORE group or the control group using random allocation software.
Allocation concealment (selection bias)	Low risk	Randomised by an independent examiner who was not involved in participant recruitment
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	1. Impossible for patients to be blinded to exercises versus control protocols; 2. Control versus exercise; 3. Control versus exercise; 4. Control group was provided some form of care (transcutaneous electrical nerve stimulation/hot pack)
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Likely different providers giving CORE protocol and transcutaneous electrical nerve stimulation
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome was pain, which was assessed by the patients who were not blinded; 2. Pain questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Exercise versus control
Incomplete outcome data (attrition bias)	Low risk	1. 21 of 74 dropped out, 28%; 2. Dropouts were excluded from analysis; 3. Increased pain and function could cause missingness; 4. 10 and 11 in each group

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kim 2015a *(Continued)*

All outcomes

Participants analysed in group allocated (attrition bias)	High risk	1. Non-compliant were not analysed, and those who got additional therapies were excluded; 2. 15 excluded due to taking medication; likely that these patients had different outcomes
Selective reporting (reporting bias)	Low risk	1. Study protocol mentioned in methods
Groups similar at baseline (selection bias)	Low risk	Age, weight, height, duration, and sex were all similar.
Co-interventions avoided or similar (performance bias)	Low risk	Patients who received additional therapies were excluded.
Compliance acceptable in all groups (performance bias)	Low risk	In CORE group, the eight-week programme compliance rate was 85.5%, though no compliance reported in control group
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Same blinded examiner helped to deliver outcome measures in order to reduce measurement error; 2. Visual Analogue Scale valid and reliable
Other bias	Low risk	Appeared free from other sources of bias

Kim 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** South Korea, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 30 (E1 = 15, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 40 **Sex (female):** 43%
Interventions	**Exercise Group 1 (E1):** Proprioceptive neuromuscular facilitation: abdominal muscle strength training, including warm-up and cool-down stretching; type = core strengthening; duration = 6 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kim 2017 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Participants were "randomized".
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Control group still getting some treatment despite not getting exercise treatment
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviation from intended intervention because interventions were highly prescribed with no room for deviation
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in a exercise versus no treatment study likely to be altered by knowledge of intervention assignment
Incomplete outcome data (attrition bias) All outcomes	High risk	1. No description of dropout rate 2. Paired t-test did not compensate for missing data, no sensitivity analysis; 3. Increased pain and disability could cause missing values; 4. No information on dropout rate in either group; 5. No information on reasoning for any dropouts
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all analyses and outcomes were fully reported.
Groups similar at baseline (selection bias)	Low risk	Groups were similar on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	High risk	No information on co-interventions in this study
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance, adherence or attendance in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Kim 2018a

Study characteristics	
Methods	**Study design:** RCT

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kim 2018a *(Continued)*

	Setting: South Korea, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 77 (E1 = 38, E2 = 39) **Chronic LBP duration:** 9.65 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 43 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Sling exercises with elastic bands, focussing on engaging core muscles, with warm-up (stretching); type = core strengthening; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Traditional trunk stabilisation exercises focussing on core muscles Includes stretching warm-up as well; type = core strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was performed "using a stratified randomization procedure with a permuted block size of 4 using a computer that balanced ages (< 40 or ≥ 40 years) and sexes (male or female)".
Allocation concealment (selection bias)	Low risk	"Randomization codes were kept in sealed envelopes with consecutive numbering" to keep treatment allocation concealed.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from the intended interventions because neither treatment group was better
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from the intended interventions because the two treatment groups were distinct, yet neither was perceived as better
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Although the sling exercise seemed relatively novel, both groups were receiving similar active treatments; unlikely to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 5%.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kim 2018a *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all outcomes and analyses were fully reported.
Groups similar at baseline (selection bias)	High risk	Groups were balanced on all relevant characteristics at baseline, except one group was almost significantly older than the other.
Co-interventions avoided or similar (performance bias)	Low risk	"Participants were not allowed to receive other treatment for back pain during the intervention period".
Compliance acceptable in all groups (performance bias)	High risk	No information about compliance, adherence or attendance in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function), and Numeric Rating Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Kladny 2003

Study characteristics	
Methods	**Study design:** RCT **Setting:** Germany, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 99 (E1 = 50, E2 = 49) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 40 **Sex (female):** 34%
Interventions	**Exercise Group 1 (E1):** Isometric strengthening transversus abdominal and multifidus muscles; type = mixed; duration = not specified weeks; dose = high; design = individualised; delivery = individual; additional intervention = electrotherapy & manual therapy & heat/ice **Exercise Group 2 (E2):** Strengthening of trunk muscles, stretching, McKenzie; type = strengthening & stretching; duration = not specified weeks; dose = high; design = individualised; delivery = individual; additional intervention = electrotherapy & manual therapy & heat/ice
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kladny 2003 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Kofotolis 2006

Study characteristics	
Methods	**Study design:** RCT **Setting:** Greece, not specified **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 86 (E1 = 28, E2 = 28, C1 = 30)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kofotolis 2006 *(Continued)*

	Chronic LBP duration: Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Alternating trunk flexion-extension isometric contractions against resistance for 10 seconds, with no motion intended, 3 sets of 15 repetitions; type = core strengthening; duration = 4 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Alternating concentric and eccentric contractions (5 seconds) for truck flexion and flexion-extension exercises, 3 sets of 15 repetitions; type = core strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Borg Rate of Perceived Pain Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short); 8 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The randomisation method was not adequately specified.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Of the 108 subjects, a total of 86 subjects completed all training and testing measurements (Table 1).
Participants analysed in group allocated (attrition bias)	High risk	ITT not used
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The groups were homogeneous and showed nonsignificant differences in basic characteristics (Table 1).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kofotolis 2006 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	None of the subjects received additional physical therapy interventions during the study period.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kofotolis 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Greece, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 120 (E1 = 40, E2 = 40, C1 = 40) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 41 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Pilates: focusses on trunk stability and improving pelvis and lumbar spine alignment; type = Pilates; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** General trunk strength training: abdominal and back muscle strengthening and stretching; type = stretching & core strengthening; duration = 8 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no description)
Outcomes	**Core outcomes reported:** Pain (36-Item Short Form Survey); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation method employed a series of randomly-generated numbers.
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed by keep allocation in a sealed envelope until the patient opened it.
Blinding of participants and personnel (performance bias)	Low risk	1. Due to the nature of the intervention patients could not be blinded to the intervention; 2. Unlikely that lack of patient blinding caused deviations from in-

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kofotolis 2016 *(Continued)*

All outcomes		tended interventions because patients had little control over intervention delivery
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Same care provider conducted both active treatments; protocol was explicit; no decision-making by personnel; opinion may have transferred to patients but no clear better intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 16% (19/120).
Participants analysed in group allocated (attrition bias)	Low risk	1. It is likely that patients were analysed in their assigned groups; no mention of analysis used, but no mention of changing interventions; modified to not use missing data
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: all planned analyses were executed and reported for all primary and secondary outcomes.
Groups similar at baseline (selection bias)	Low risk	All treatment groups were similar on all relevant characteristics at baseline, except duration of symptoms was not measured.
Co-interventions avoided or similar (performance bias)	Low risk	"None of the participants received additional physical therapy interventions during the study period"; no description of analgesic use (or other types of co-interventions)
Compliance acceptable in all groups (performance bias)	Low risk	No information presented on compliance, adherence or attendance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Roland-Morris Disability Questionnaire (for function), and 36-Item Short Form Survey (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Koldas 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 55 (E1 = 19, E2 = 18, E3 = 18) **Chronic LBP duration:** 51 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 40

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Koldas 2008 *(Continued)*

	Sex (female): 78%
Interventions	**Exercise Group 1 (E1):** Aerobic exercise performed on the treadmill and flexion, extension, and stretching (double knee to chest, and alternate arm and leg lifts) 15-20 repetitions once daily; type = mixed; duration = 6 weeks; dose = high; design = partially individualised; delivery = not specified; additional intervention = advice/education **Exercise Group 2 (E2):** Flexion, extension, and stretching (double knee to chest, and alternate arm and leg lifts) 15-20 repetitions once daily; type = mixed; duration = 6 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education **Exercise Group 3 (E3):** Flexion, extension, and stretching (double knee to chest, and alternate arm and leg lifts) 15-20 repetitions once daily; type = mixed; duration = 6 weeks; dose = low; design = standardised; delivery = independent; additional intervention = advice/education & electrotherapy & heat/ice
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The patients were randomly assigned to one of the three groups using a sequence of random numbers before baseline assessments were performed.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	One patient from Group 1 could not complete the treatment programme due to vertigo and one patient from Group 2 was dropped out because of surgery.
Participants analysed in group allocated (attrition bias)	High risk	Completers analysed only
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences were observed between the groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Koldas 2008 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	One patient from Group 2, and two patients from Group 3 were dropped out because of the poor compliance to the treatment.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kool 2007

Study characteristics	
Methods	**Study design:** RCT **Setting:** Switzerland, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 174 (E1 = 87, E2 = 87) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 21%
Interventions	**Exercise Group 1 (E1):** Work simulation, strength and endurance training, aerobic exercises (walking and water aerobics), sports, and home exercises; type = functional restoration; duration = 3 weeks; dose = high; design = partially individualised; delivery = independent; additional intervention = advice/education & psychological therapy & anti-inflammatory/analgesics **Exercise Group 2 (E2):** Mobilisation, stretching, strength training, low intensity movement in pool and progressive muscle relaxation; type = mixed; duration = 3 weeks; dose = high; design = partially individualised; delivery = independent; additional intervention = advice/education & psychological therapy & electrotherapy & manual therapy & anti-inflammatory/analgesics
Outcomes	**Core outcomes reported:** Work (number of calendar work days in the follow-up year) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Swiss Federal Office of Health (grant no. 00.00437) **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomisation was concealed and assessment of the primary outcome, work days, was blinded.
Allocation concealment (selection bias)	Unclear risk	Randomisation was concealed and assessment of the primary outcome, work days, was blinded.
Blinding of participants and personnel (performance bias)	High risk	Patients could not be blinded to treatment, but they received no detailed information about the difference between the two treatments.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kool 2007 *(Continued)*
All outcomes

Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assessment of the primary outcome, work days, was blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Two dropouts, seven no responses at follow-up; detailed in Figure 1
Participants analysed in group allocated (attrition bias)	High risk	Analysis was based on the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the groups for most variables.
Co-interventions avoided or similar (performance bias)	Low risk	Assessed health care use in the follow-up period by means of a questionnaire sent to the health insurance provider with return rate of 78%
Compliance acceptable in all groups (performance bias)	Low risk	All patients attended at least 90% of the scheduled treatments, and treatment duration was comparable (Table 2).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Koumantakis 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** United Kingdom, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 55 (E1 = 29, E2 = 26) **Chronic LBP duration:** 12 weeks (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 37 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Warm-up (stretching and stationary biking for 10-15 minutes), general exercises activating extensor (paraspinals) and flexor (abdominals) muscle groups, stabilisation exercises (isometric and low-loading) increasing in contraction and duration; type = mixed; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Warm-up (stretching and stationary bike 10-15 minutes), general exercises activating extensor (paraspinals) and flexor (abdominals) muscle groups; type = mixed; duration = 8 weeks;

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Koumantakis 2005 *(Continued)*

	dose = low; design = partially individualised; delivery = group; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short); 20 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Greek State Scholarship Foundation (1KY), Athens, Greece (grant T104830098); Hospital Saving Association, London, United Kingdom **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	This procedure was undertaken by an independent trial manager.
Allocation concealment (selection bias)	Low risk	Randomisation codes were kept in sealed envelopes with consecutive numbering.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Patients were not aware of the theoretical bases of each of the exercise regimens.
Blinding of care provider (performance bias)	High risk	The clinical physical therapist (FR) who administered the exercise programmes could not be masked to group allocation.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The research physical therapist who was in charge of the study and performed the outcome assessments of subjects was unaware of the group allocation throughout the study.
Incomplete outcome data (attrition bias) All outcomes	Low risk	From the 55 randomly assigned subjects, 10 dropped out of the programme (n = 5 per group), most of them due to time constraints.
Participants analysed in group allocated (attrition bias)	Low risk	All analyses were performed primarily according to the "intention-to-treat" (ITT) principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Only the Visual Analogue Scale baseline data were different between groups (Table 2): all other variables were considered sufficiently similar.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: did not measure, although it is possible.
Compliance acceptable in all groups (performance bias)	Low risk	Adherence data for clinic-based exercise were normally distributed.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Koumantakis 2005 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Krein 2013

Study characteristics

Methods	**Study design:** RCT (NCT00694018) **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 229 (E1 = 111, E2 = 118) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 52 **Sex (female):** 13%
Interventions	**Exercise Group 1 (E1):** Daily walking with a pedometer (as part of an internet-mediated walking programme) and internet support group; type = aerobic; duration = 52 weeks; dose = high; design = partially individualised; delivery = independent with follow-up; additional intervention = advice/education & psychological therapy **Exercise Group 2 (E2):** Daily walking with a pedometer with email reminders; type = aerobic; duration = 52 weeks; dose = high; design = partially individualised; delivery = independent with follow-up; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 24 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Michigan Diabetes Research and Training Center (P60 DK020572); Center for Health Communications Research (P50CA101451); Michigan Institute for Clinical and Health Research (NIH #UL1RR024986) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number generator
Allocation concealment (selection bias)	Low risk	Participants were emailed to inform them of their group assignment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants were emailed to inform them of their group assignment (Internet support or monthly upload).
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias)	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Krein 2013 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from 92% of those in the intervention group and 89% receiving usual care at 12 months
Participants analysed in group allocated (attrition bias)	Low risk	Author contact: Author reported yes, but figures and tables contradict this.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	None of the observed differences in baseline characteristics were statistically significant.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: not limited, but no differences across groups
Compliance acceptable in all groups (performance bias)	Unclear risk	Eighty per cent compliance for uploading data, but participants logged in to the computer only 38% of the time.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kumar 2009a

Study characteristics	
Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 102 (E1 = 51, C1 = 51) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 34 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Dynamic muscular stabilisation: segmental stability training delivered in 4 stages: 1) abdominal hollowing; 2) stability exercises static load; 3) progressed to controlled movement; 4) high speed movement; type = core strengthening; duration = 5 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 25 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Indian Council of Medical Research, New Delhi **Other:** None

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kumar 2009a *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The subjects were randomly assigned equally into two groups by a lottery.
Allocation concealment (selection bias)	Unclear risk	Not clear whether the papers were opaque or able to be read before picking
Blinding of participants and personnel (performance bias) All outcomes	High risk	After group allocations, respective subjects were treated either with conventional treatment or dynamic muscular stabilization techniques in a single-blind manner.
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Severity or level of pain, back pressure changes and abdominal pressure changes were assessed by same tester and same physiotherapist supervising the test procedure at baseline.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The baseline characteristics across groups were found to be the same.
Co-interventions avoided or similar (performance bias)	Low risk	The subjects were not allowed to receive any other treatment, including painkillers.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kumar 2009b

Study characteristics

Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 1

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kumar 2009b *(Continued)*

	Comparison groups: 1
Participants	**Number of participants:** 30 (E1 = 15, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 24 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Segmental stability training delivered in 4 stages: 1) abdominal hollowing; 2) stability exercises static load; 3) progressed to controlled movement; 4) high speed movement; type = core strengthening; duration = 5 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Waddell Disability Index) **Follow-up time periods available for syntheses:** 5 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Indian Council of Medical Research, New Delhi **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The subjects were randomised by lottery method.
Allocation concealment (selection bias)	Unclear risk	No indication of whether envelopes/papers were opaque
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	All tests were performed for dependent variables (walking, stand-ups, climbing, and pain) by the same tester and same physiotherapist.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	On comparing, average baseline characteristics of two groups, they did not differ significantly (P > 0.05).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kumar 2009b *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	The subjects did not allow any other treatment and medication during the treatment.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kumar 2010

Study characteristics	
Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 141 (E1 = 72, C1 = 69) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 35 **Sex (female):** 35%
Interventions	**Exercise Group 1 (E1):** Segmental stability training delivered in 4 stages: 1) abdominal hollowing; 2) stability exercises static load; 3) progressed to controlled movement; 4) high speed movement; type = core strengthening; duration = 5 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Waddell Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Indian Council of Medical Research, New Delhi **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The subjects were randomised equally in two groups by lottery method.
Allocation concealment (selection bias)	Unclear risk	No indication of whether patients could see what was on the paper
Blinding of participants and personnel (performance bias) All outcomes	High risk	After group allocations, respective subjects were treated either with conventional techniques or dynamic muscular stabilization techniques in a single-blind manner.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kumar 2010 *(Continued)*

Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcome variables such as pain severity, physical strength, functional ability
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Participants analysed in group allocated (attrition bias)	Unclear risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Demographic characteristics (Table 2) of two treatment groups were assessed at baseline and found not to be significant.
Co-interventions avoided or similar (performance bias)	Low risk	The subjects were not allowed to get any other treatment options including the painkillers.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Kuukkanen 2000

Study characteristics	
Methods	**Study design:** RCT **Setting:** Finland, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 57 (E1 = 29, C1 = 28) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 51%
Interventions	**Exercise Group 1 (E1):** Strengthening, endurance, balance and co-ordination; type = mixed; duration = 12 weeks; dose = high; design = standardised; delivery = independent; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Borg Rate of Perceived Pain Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kuukkanen 2000 *(Continued)*

Notes	**Conflicts of interest:** Not reported **Funding source:** University of Jyväskylä, the Juho Vainio Foundation; TULES Graduate School; Academy of Finland **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Unclear risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lee 2016a

Study characteristics	
Methods	**Study design:** RCT **Setting:** South Korea, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 36 (E1 = 15, E2 = 15, C1 = 6) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 44 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Strength exercises performed on a mat and with an exercise ball (e.g. bridge, plank, squat, push-ups, back extension), walking exercises were performed with a step box, warm-up and cool-down stretching step exercises; type = mixed; duration = 12 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = none **Exercise Group 2 (E2):** Strength training exercises including bridge, plank, squat, push-ups and back extension, warm-up and cool-down stretching: core focussed; type = stretching & core strengthening; duration = 12 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no description)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Sangmyung University **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Patients were randomly divided; no further information given.
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to intervention because of the distinct nature of each treatment group; 2. Unlikely that patients' lack of blinding led to deviations from intended interventions because patients have little control over how the intervention is delivered
Blinding of care provider (performance bias)	Low risk	1. Participants did not necessarily have a regular care provider; exercise instructors were not blinded to intervention, as they were delivering distinct exercise protocols; 2. Exercise instructors were delivering a group (probably) exercise that was standard, so unlikely that they deviated from exercise protocol for any participants
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Assessors of outcomes of interest (pain and disability) were participants themselves, who were not blinded to allocation because of how different each treatment group was; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Due to the differences between the control group and the active exercise groups, there is sig-

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lee 2016a *(Continued)*

		nificant potential for patients' judgement to be biased when self-evaluating pain and disability outcomes.
Incomplete outcome data (attrition bias) All outcomes	High risk	1. No description of dropout rate; 2. No evidence of non-biased analysis; 3. Increased pain and decreased function could cause missing outcome; 4. No information; 5. No information
Participants analysed in group allocated (attrition bias)	Low risk	1. Reasonable to assume that participants were analysed according to their original group allocation
Selective reporting (reporting bias)	Low risk	1. No linked protocol found but within this paper all outcomes and analyses fully reported; no obvious omissions
Groups similar at baseline (selection bias)	High risk	All groups were similar at baseline on age; and body mass index, pain duration and sex were not reported, and nearly all outcome measures (including pain and disability) were very different among groups.
Co-interventions avoided or similar (performance bias)	High risk	No reporting of co-interventions; participants were all sedentary, and so were not likely participating in any physical activity outside of the study.
Compliance acceptable in all groups (performance bias)	High risk	No information on compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Assessment protocol was identical between groups; assuming survey delivery was the same for the control group; 2. Visual Analogue Scale (for pain) and Roland-Morris Disability Questionnaire (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Under reported, no other apparent sources of bias

Lewis 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 80 (E1 = 40, E2 = 40) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 46 **Sex (female):** 65%
Interventions	**Exercise Group 1 (E1):** Treadmill, exercise bicycle, sit to stand repetitions, spinal stabilisation exercises (supine, prone, 4-point kneeling), upper and lower extremity exercises; type = mixed; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & manual therapy **Exercise Group 2 (E2):** Spinal stabilisation exercises; type = core strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education & manual therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Quebec Back Pain Disability Scale)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lewis 2005 *(Continued)*

	Follow-up time periods available for syntheses: 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Using random number tables, subjects were assigned to Group 1 (exercise class) or Group 2 (individual treatment).
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	The nature of the interventions precluded any blinding of physiotherapists or participants to the assigned treatments.
Blinding of care provider (performance bias)	High risk	The nature of the interventions precluded any blinding of physiotherapists or participants to the assigned treatments.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Two assessors blinded to group allocation administered questionnaires and conducted physical assessments, before group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Out of 80, 17 were lost and are detailed in Figure 1.
Participants analysed in group allocated (attrition bias)	Low risk	In this study, the data from each participant was analysed in accordance with the intention-to-treat approach.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	High risk	After randomisation, there was no statistical difference between the groups.
Co-interventions avoided or similar (performance bias)	Low risk	In the exercise group, three subjects required additional conservative treatment at the end of their treatment programme, five at six months and six at 12 months.
Compliance acceptable in all groups (performance bias)	Low risk	There was a steady decline in the reported compliance of the home exercises over the one-year follow-up period.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Li 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** China, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 20 (E1 = 10, C1 = 10) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Transversus abdominis muscle exercises; type = core strengthening; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment ("traditional treatment")
Outcomes	**Core outcomes reported:** Function (Aberdeen Back Pain Scale) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Assumed from data given
Participants analysed in group allocated (attrition bias)	Low risk	Assumed all patients completed the study
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Li 2008 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Lie 1999

Study characteristics

Methods	**Study design:** RCT **Setting:** Norway, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 27 (E1 = 12, E2 = 15) **Chronic LBP duration:** 390 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 67%
Interventions	**Exercise Group 1 (E1):** Mobilising and stretching, walking in 'flexible way'; type = mixed; duration = 9 weeks; dose = low; design = not specified; delivery = group; additional intervention = not specified **Exercise Group 2 (E2):** Stabilising exercises and walking without instructions; type = core strengthening & aerobic; duration = 9 weeks; dose = low; design = not specified; delivery = independent; additional intervention = not specified
Outcomes	**Core outcomes reported:** Function (Clinical Outcome Score) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias)	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lie 1999 *(Continued)*

All outcomes		
Blinding of care provider (performance bias)	Unclear risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Limke 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 100 (E1 = 49, E2 = 51) **Chronic LBP duration:** 34 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 47 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Step aerobics (10 minutes), stretching (20 minutes), strengthening (1 hour), endurance exercises, 2 sets of resistance exercises, home exercise; type = mixed; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy & heat/ice **Exercise Group 2 (E2):** Step aerobics (10 minutes), stretching (20 minutes), strengthening (1 hour), endurance exercises, 1 set of resistance exercises, home exercise; type = mixed; duration = 6 weeks; dose

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Limke 2008 *(Continued)*

	= low; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy & heat/ice
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomisation was done in blocks of 10, with shuffled envelopes containing group assignment.
Allocation concealment (selection bias)	Unclear risk	Randomisation was done in blocks of 10, with shuffled envelopes containing group assignment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Same treatment but different numbers of exercises
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropout rate was 18% in one group and 20% in the other.
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	All demographic characteristics were similar.
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Little 2008

Study characteristics	
Methods	**Study design:** RCT (NCT0028108728) **Setting:** England, healthcare **Exercise groups:** 4 **Comparison groups:** 4
Participants	**Number of participants:** 579 (E1 = 72, E2 = 72, E3 = 71, E4 = 71, C1 = 72, C2 = 75, C3 = 73, C4 = 73) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Doctor prescription for aerobic exercises (predominantly walking); type = aerobic; duration = 12 weeks; dose = high; design = not specified; delivery = independent; additional intervention = psychological therapy **Exercise Group 2 (E2):** Doctor prescription for aerobic exercises (predominantly walking); type = aerobic; duration = not specified weeks; dose = high; design = not specified; delivery = independent; additional intervention = psychological therapy & manual therapy **Exercise Group 3 (E3):** Doctor prescription for aerobic exercises (predominantly walking); type = aerobic; duration = not specified weeks; dose = high; design = not specified; delivery = independent; additional intervention = advice/education & psychological therapy **Exercise Group 4 (E4):** Doctor prescription for aerobic exercises (predominantly walking); type = aerobic; duration = not specified weeks; dose = high; design = not specified; delivery = independent; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Usual care/no treatment ("normal care") **Comparison Group 2 (C2):** Other conservative treatment (manual therapy) **Comparison Group 3 (C3):** Other conservative treatment (physical therapy) **Comparison Group 4 (C4):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Number of Pain Days); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall improvement (health transitions)) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Medical Research Council (grant number G0001104) **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	At the baseline appointment, after informed written consent had been obtained, participants were randomised to one of eight groups by the practice nurse.
Allocation concealment (selection bias)	Unclear risk	At the baseline appointment, after informed written consent had been obtained, participants were randomised to one of eight groups by the practice nurse.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Little 2008 *(Continued)*

Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	A total of 579 people were randomised and completed the baseline questionnaires.
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics were similar for all variables (Table 2) except there were fewer women in the Alexander technique groups.
Co-interventions avoided or similar (performance bias)	Low risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Good adherence was defined by the trial management group as attending five out of six massage sessions, five out of six lessons in the other group.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	One session a week for six weeks. Two lessons a week for two weeks then one lesson a week for two weeks. Twenty two lessons over five months

Little 2014

Study characteristics	
Methods	**Study design:** RCT (ISRCTN51496752) **Setting:** England, healthcare **Exercise groups:** 2 **Comparison groups:** 2
Participants	**Number of participants:** 68 (E1 = 17, E2 = 18, C1 = 16, C2 = 17) **Chronic LBP duration:** 321.5 days (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 51 **Sex (female):** 62%
Interventions	**Exercise Group 1 (E1):** Motor relearning; stretching, aerobic exercises, exercise targeting motor control of deep abdominal and lumbar paraspinal muscle, home exercise; type = mixed; duration = 4 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = psychological therapy **Exercise Group 2 (E2):** Motor relearning, stretching, aerobic exercises, exercise targeting motor control of deep abdominal and lumbar paraspinal muscle, home exercise; type = mixed; duration = 4 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = psychological therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Little 2014 *(Continued)*

	Comparison Group 1 (C1): Usual care/no treatment (control group: continued with normal activity and treatment)) **Comparison Group 2 (C2):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Graded Chronic Pain Scale (Von Korff)); function (Roland-Morris Disability Questionnaire); HRQoL (EuroQol 5D); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall improvement (health transition)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Paul Little is editor-in-chief of the Programme Grants for Applied Research journal. **Funding source:** Search Results Web results Efficacy and Mechanism Evaluation programme, Medical Research Council, National Institute for Health Research **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers, stratified by history of previous pain
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Eighty-three per cent follow-up at three months, 81% at six months (reasons given in Table 6)
Participants analysed in group allocated (attrition bias)	Low risk	No imputation (described on page 12)
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Even with the relatively small numbers, the groups were reasonably well balanced at baseline.
Co-interventions avoided or similar (performance bias)	Unclear risk	Control group participants were allowed to be allocated physio by their general practitioner.
Compliance acceptable in all groups (performance bias)	Low risk	Described in Table 5

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Little 2014 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Lomond 2014

Study characteristics

Methods	**Study design:** RCT (NCT01611792) **Setting:** USA, not specified **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 38 (E1 = 20, E2 = 18) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 41 **Sex (female):** 53%
Interventions	**Exercise Group 1 (E1):** Strength and conditioning physical therapy including general trunk strengthening and endurance exercises; type = strengthening; duration = 10 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = not specified **Exercise Group 2 (E2):** Stabilisation physical therapy including specific exercises to improve control of deep trunk muscles; type = core strengthening; duration = 10 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 11 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Institutes of Health NIH2R01HD040909 (PI: Henry SM) **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised based on a covariate adaptive randomisation scheme
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Treatment assignments were transmitted to the study treating physiotherapist.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lomond 2014 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	High risk	Dropout rate was high (6/29 in stabilisation group and 7/29 in strengthening group completed).
Participants analysed in group allocated (attrition bias)	Unclear risk	Paper reported yes, but Figure seemed to be missing some subjects.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Long 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** Canada, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 229 (E1 = 80, E2 = 69, E3 = 80) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 47%
Interventions	**Exercise Group 1 (E1):** Matched direction: unidirectional end-range lumbar exercises that matched their directional preference identified during assessment, recommendation to exercise every 2 hours, home exercises and compliance log; type = Mckenzie; duration = 2 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = advice/education **Exercise Group 2 (E2):** Opposite direction: unidirectional end-range lumbar exercises in opposite direction from their directional preference identified during assessment, recommendation to exercise every 2 hours, home exercises and compliance log; type = strengthening; duration = 2 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = advice/education & psychological therapy **Exercise Group 3 (E3):** Multidirectional, midrange lumbar exercises, and stretches for the hip and thigh muscles, recommendation to exercise every 2 hours, home exercises and compliance log; type = mixed; duration = 2 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = advice/education & psychological therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Long 2004 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); work (ability to return to work and leisure activities); Global Perceived Health or Recovery (Global Perceived Health or Recovery (response to treatment)) **Follow-up time periods available for syntheses:** 2 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Foundation funds (source not indicated) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	After giving informed consent, each member of these three directional preference subsets was randomised to one of three treatments.
Allocation concealment (selection bias)	Low risk	After giving informed consent, each member of these three directional preference subsets was randomised to one of three treatments.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Same treatment but matched for directional preference
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome measures were administered at baseline and two weeks by nonmedical reception staff blinded to the study design.
Incomplete outcome data (attrition bias) All outcomes	High risk	Twenty-nine subjects (12.6%) failed to return for treatment appointments and did not provide data at two weeks.
Participants analysed in group allocated (attrition bias)	Low risk	Subjects unable/unwilling to continue for the full two-week protocol were included in the analysis (intention-to-treat).
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics of the 312 recruited subjects are listed in Table 3.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: two-week intervention where patients asked not to seek other care
Compliance acceptable in all groups (performance bias)	Low risk	Sixty-eight per cent of our 201 subjects returned their compliance questionnaires with overall good compliance rating (3-4 sets/day).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Longo 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 14 (E1 = 7, C1 = 7) **Chronic LBP duration:** 14.58 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 52 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Posture called "standing posture with flexion of the trunk" in addition to physiotherapy: position maintained from 30 to 60 seconds and was repeated 5 times, with a one-minute break between repetitions; type = core strengthening & stretching; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomisation not described beyond the fact that patients were "randomly divided" into treatment groups
Allocation concealment (selection bias)	High risk	Not described beyond the fact that patients were "randomly divided" into treatment groups
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Unlikely that the patients could be blinded as to whether or not they received an extra exercise at the end of their physiotherapy routines; 2. Unlikely that a lack of patient blinding led to deviations from the intended intervention because the difference between treatment groups was only a single exercise, and it required physiotherapy guidance
Blinding of care provider (performance bias)	Unclear risk	1. Care providers (physiotherapist) could not be blinded to intervention because they delivered it, and had to know whether or not to deliver the exercise of interest; 2. Very possible that lack of physiotherapy blinding led to deviations from intervention because the article explicitly stated that physiotherapists could use their discretion and did not use a strict protocol; 3. Likely that any deviations from intended intervention due to lack of physiotherapist blinding were balanced between groups because both groups received the same unstructured physio treatment
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Since the outcome assessors for pain and disability were the patients, it is unlikely they were blinded to the intervention allocation; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Longo 2016 *(Continued)*

		of intervention; 3. One group was getting "conventional" care while the other was getting something supplemental; could alter assessment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No dropouts from baseline to time point 1; there were dropouts at time point 2, but these outcome data were incomplete and not extracted for systematic review.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to their assigned treatment group, according to intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: no obvious omissions in outcomes or analyses
Groups similar at baseline (selection bias)	Low risk	No statistically significant between-group differences in baseline age, sex, body mass index, symptom duration, pain or disability; nearly clinically significant difference in body mass index and disability: worse in control
Co-interventions avoided or similar (performance bias)	Low risk	Care providers were not given any strict protocols (especially around co-interventions) to follow in their "usual care", but they could have applied this equally to both treatment groups.
Compliance acceptable in all groups (performance bias)	High risk	Compliance (attendance) was not reported in the study, and so not able to determine exact compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were carried out with the same tools and at the same time for both treatment groups, up until T1 assessments; only experimental followed to T2; 2. Numeric Rating Scale (for pain) and Roland-Morris Disability Questionnaire (for disability) are validated tools in low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Lopes 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Portugal, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 46 (E1 = 23, C1 = 23) **Chronic LBP duration:** 29 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 22 **Sex (female):** 59%
Interventions	**Exercise Group 1 (E1):** Pilates: exercises focussed on deep stabilisers and hip extensors; type = Pilates; duration = .14 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: rest 20 minutes resting in a sitting position)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lopes 2017 *(Continued)*

	Follow-up time periods available for syntheses: 0 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** Information modified for author contact; sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Block randomisation, 1:1
Allocation concealment (selection bias)	Low risk	Allocation concealment was achieved by using numbered sheets inside sealed, opaque envelopes picked up by the participants before baseline data collection.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Intervention only consisted of one session with the assessment occurring right after; it was not possible to deviate from intervention.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from the intended intervention as interventions were very structured and distinct; also, entire trial took place in one day
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in a exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No dropouts in the study because it took place in a single session
Participants analysed in group allocated (attrition bias)	Low risk	1. Not explicitly stated, however, one can assume that in a single intervention trial that intention-to-treat analysis was inevitable.
Selective reporting (reporting bias)	Low risk	1. No protocol or statistical analysis plan found: all planned analyses were executed and reported for all primary and secondary outcomes.
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were balanced on all relevant characteristics at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	No opportunity for co-interventions because the trial took place in a single session
Compliance acceptable in all groups (performance bias)	Low risk	Compliance was 100% because the trial took place in a single session.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lopes 2017 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Macedo 2012

Study characteristics	
Methods	**Study design:** RCT (ACTRN12607000432415) **Setting:** Australia, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 172 (E1 = 86, E2 = 86) **Chronic LBP duration:** 100 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 49 **Sex (female):** 59%
Interventions	**Exercise Group 1 (E1):** Motor control exercises: contracting transversus abdominals and multifidus muscles in isolation using biofeedback, training co-ordination of all trunk muscles during functional tasks, home exercises; type = core strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Graded activity: increasing activity tolerance by performing individualised and submaximal exercises and home exercise; type = other (graded activity); duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & psychological therapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (recovery)) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Australia's National Health; Medical Research Council **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The randomisation sequence was computer-generated by an investigator not involved in recruitment or treatment allocation.
Allocation concealment (selection bias)	Low risk	Allocation was concealed in sequentially numbered, sealed, opaque envelopes by an investigator not involved in the study.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Macedo 2012 *(Continued)*

Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All self-reported measures were collected by an investigator blinded to treatment allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Nine participants withdrew from the study: two from the graded activity group (one for not improving and one unknown) and seven from the motor control exercise group.
Participants analysed in group allocated (attrition bias)	Low risk	Eight-six participants in each group analysed
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics, including demographics and baseline scores for both treatment groups, were reported in Table 3.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: measured and no difference between groups
Compliance acceptable in all groups (performance bias)	Low risk	Adherence to treatment in the initial eight-week period was excellent, with both groups attending a mean of 10.3 (SD 3.6) of the planned 12 sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Same time points for all

Machado 2007

Study characteristics	
Methods	**Study design:** RCT **Setting:** Brazil, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 33 (E1 = 17, C1 = 16) **Chronic LBP duration:** 68.5 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 43 **Sex (female):** 70%
Interventions	**Exercise Group 1 (E1):** Twenty minutes walking, general stretching and strengthening in bridge position; type = mixed; duration = 9 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (psychological therapy)
Outcomes	**Core outcomes reported:** Pain (Low Back Pain Rating Scale (Manniche)); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 9 weeks (short); 26 weeks (moderate)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Machado 2007 *(Continued)*

Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A randomisation sequence was generated using a random numbers table.
Allocation concealment (selection bias)	Low risk	Concealment of allocation was ensured by the use of sequentially numbered sealed opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The patients were assessed by blinded examiners at baseline, nine weeks, and six months.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Twelve patients were not able to return to the triage centre for the nine-week assessment, and six patients could not be contacted for the phone interview.
Participants analysed in group allocated (attrition bias)	Unclear risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Table 1 summarised the sociodemographic and clinical characteristics of the participants in the exercise and psychotherapy groups at baseline.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	High risk	The adherence to treatment was similar in both groups.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Magalhaes 2017

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Magalhaes 2017 *(Continued)*

Methods	**Study design:** RCT (NCT01719276) **Setting:** Brazil, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 66 (E1 = 33, E2 = 33) **Chronic LBP duration:** 67 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 47 **Sex (female):** 74%
Interventions	**Exercise Group 1 (E1):** Graded activity: treadmill for 20 minutes at 70-80% Max heart rate, strengthening of the legs and back; type = mixed; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Physiotherapy: stretches and strengthening of arms, legs, and trunk; type = stretching & core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); work (return to work); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 13 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A randomisation code was created using Microsoft Excel for Windows software.
Allocation concealment (selection bias)	Low risk	Randomisation was performed by a researcher who was not involved in the recruitment process; allocation schedule was concealed by using consecutively numbered, sealed, opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from the intended interventions because the two treatment groups were so similar
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from the intended interventions because the two treatment groups were so similar
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Neither treatment was obviously better than the other; no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 6 patients out of 66 were lost to follow-up at the 3-month follow-up, 9.1%.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Magalhaes 2017 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	1. Analyses "were performed on an intention-to-treat basis".
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported according to registered protocol (NCT01719276).
Groups similar at baseline (selection bias)	Low risk	Several characteristics were collected and were similar; slightly large difference in duration of symptoms was likely accountable to chance; there was no indication of problematic randomisation.
Co-interventions avoided or similar (performance bias)	Low risk	"Participants were instructed not to participate in any other intervention during the treatment period. There was no interference in the use of medication".
Compliance acceptable in all groups (performance bias)	Low risk	"Patients in the physiotherapy exercise attended 91.4% of the sessions, while those in the graded activity group attended 91.9%".
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Roland-Morris Disability Questionnaire (for function), and Numeric Rating Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Magnussen 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** Norway, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 152 (E1 = 79, C1 = 73) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 43 **Sex (female):** 46%
Interventions	**Exercise Group 1 (E1):** Exercise in heated water (stretching and mobilising), group exercise in gym (walking, mobilising, stretching) and individual exercise programme (e.g. cycling, walking); type = mixed; duration = 1 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy & relaxation **Comparison Group 1 (C1):** Usual care/no treatment (control group: referred back to the general practitioner)
Outcomes	**Core outcomes reported:** Pain (Activity Discomfort Scale) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Norwegian Research Council; Norwegian Ministry of Health and Social Affairs, Research Unit of the Norwegian Back Pain Network **Other:** Sufficient data not available for inclusion in meta-analyses

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Magnussen 2005 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation was concealed and done in blocks of 10 by an assistant with no prior knowledge of the patients drawing ID-numbers and group allocation.
Allocation concealment (selection bias)	Low risk	The randomisation was concealed and done in blocks of 10 by an assistant with no prior knowledge of the patients drawing ID-numbers and group allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	After randomisation, letters were sent to the participants informing them of which group they were allocated to.
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	All self-report
Incomplete outcome data (attrition bias) All outcomes	Low risk	During the study period, 16.5 per cent in the treatment group and 19.2 per cent in the control group resigned from the study.
Participants analysed in group allocated (attrition bias)	Unclear risk	Numbers seemed to be without withdrawals.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the groups on demographic factors or sick leave at baseline (see Table 2).
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Manniche 1988

Study characteristics

Methods	**Study design:** RCT **Setting:** Denmark, healthcare

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Manniche 1988 *(Continued)*

	Exercise groups: 3 **Comparison groups:** 0
Participants	**Number of participants:** 105 (E1 = 33, E2 = 36, E3 = 36) **Chronic LBP duration:** 780 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 45 **Sex (female):** 39%
Interventions	**Exercise Group 1 (E1):** Intensive strengthening: trunk, back strengthening routine; type = core strengthening; duration = 12 weeks; dose = high; design = standardised; delivery = individual; additional intervention = not specified **Exercise Group 2 (E2):** Back strengthening routine, similar to intensive less dose; type = core strengthening; duration = 12 weeks; dose = low; design = standardised; delivery = individual; additional intervention = not specified **Exercise Group 3 (E3):** Isometric exercises for lumbar spine; type = strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Activities of Daily Living Scale) **Follow-up time periods available for syntheses:** 12 weeks (short); 39 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Danish Research Council and the Danish Health Foundation, Sygekassemcs Helsefond **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Manniche 1988 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Mannion 1999

Study characteristics	
Methods	**Study design:** RCT **Setting:** Switzerland, general population **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 148 (E1 = 50, E2 = 49, E3 = 49) **Chronic LBP duration:** 566 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 45 **Sex (female):** 57%
Interventions	**Exercise Group 1 (E1):** Aerobics, stretching; type = aerobic & stretching; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Controlled progressive exercises with machines, David Beck Clinic programme (functional restoration); type = functional restoration; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 3 (E3):** Exercises using TheraBand's and general strength training and home exercise (as part of physiotherapy); type = strengthening; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Device status category: 1 **Funding source:** Swiss National Science Foundation (Grant No. 32-50979.97); Schulthess Klinik Research Fund; DBC International **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mannion 1999 *(Continued)*

Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Support for judgement is not available
Blinding of care provider (performance bias)	Low risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Marshall 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** New Zealand, mixed **Exercise groups:** 4 **Comparison groups:** 0
Participants	**Number of participants:** 50 (E1 = 12, E2 = 12, E3 = 13, E4 = 13) **Chronic LBP duration:** 3.85 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 37 **Sex (female):** 49%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Marshall 2008 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Swiss ball exercises (after 4-week non-manipulation treatment) 3 phases: 0-4 weeks (isometric), 4-8 weeks (controlled concentric/eccentric), 8-12 weeks (dynamic); type = core strengthening; duration = 16 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Swiss ball exercises (after 4-week manipulation treatment) 3 phases: 0-4 weeks (isometric), 4-8 weeks (controlled concentric/eccentric), 8-12 weeks (dynamic); type = core strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = manual therapy **Exercise Group 3 (E3):** Home exercises (after 4-week non-manipulation treatment): demonstration and info sheet providing pictures of exercises (lunges, push up, bridge pose, plank); type = core strengthening & core strengthening; duration = 16 weeks; dose = low; design = partially individualised; delivery = independent; additional intervention = advice/education **Exercise Group 4 (E4):** Home exercises (after 4-week manipulation treatment): demonstration and info sheet providing pictures of exercises (lunges, push up, bridge pose, plank); type = core strengthening & core strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = independent; additional intervention = advice/education & manual therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (12-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 40 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	These individuals who had not already undergone self-selected treatment were randomly assigned to either manipulation or nonmanipulative treatment.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The provision of exercise advice to this group ensured that the blinding to the Swiss ball exercise group was successfully achieved.
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Six people dropped out; three through non-attendance, two gave no reasons, one sought other treatment.
Participants analysed in group allocated (attrition bias)	High risk	Fifty subjects analysed only
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Marshall 2008 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	See Table 1 - all non significant
Co-interventions avoided or similar (performance bias)	High risk	There was no restriction placed on what forms of exercise were prescribed by the treating specialists who assisted this study.
Compliance acceptable in all groups (performance bias)	Unclear risk	Attendance at supervised training and programme: compliance within the Swiss ball training groups was comparable.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Marshall 2013

Study characteristics	
Methods	**Study design:** RCT (ACTRN12611000229976) **Setting:** Australia, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 64 (E1 = 32, E2 = 32) **Chronic LBP duration:** 10.3 years (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 36 **Sex (female):** 62%
Interventions	**Exercise Group 1 (E1):** Pilates; type = Pilates; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = not specified **Exercise Group 2 (E2):** Stationary cycling; type = aerobic; duration = 8 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Participants were randomly assigned in blocks of eight with an equal number of participants assigned to each group.
Allocation concealment (selection bias)	Low risk	The allocation sequence was concealed from researchers involved in enrolling and assessing participants.
Blinding of participants and personnel (performance bias)	Low risk	To reduce expectation bias, participants were blinded to the use of different modalities in the trial.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Marshall 2013 *(Continued)*
All outcomes

Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	To reduce expectation bias, participants were blinded to the use of different modalities in the trial.
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1
Participants analysed in group allocated (attrition bias)	Low risk	Data were analysed using SPSS version 20 (IBM, New York, NY) with "intention-to-treat" principles.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No between-group differences were observed at baseline for demographics and self-report scores (Table 2).
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: changed to yes, though at six months follow-up some participants from each group reported regular training using the other exercise.
Compliance acceptable in all groups (performance bias)	Low risk	Specific trunk exercise group attendance was 21.8 ± 1.9 out of the 24 sessions and stationary cycling group attendance was 19.0 ± 4.2 sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Masharawi 2013

Study characteristics	
Methods	**Study design:** RCT **Setting:** Israel, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 40 (E1 = 20, C1 = 20) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 53 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Core stability exercises in non-weight-bearing positions including supine/crook lying, side-lying (left and right) and prone; type = core strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Masharawi 2013 *(Continued)*

	Follow-up time periods available for syntheses: 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	An independent local staff member generated the allocation sequence with sealed envelopes.
Allocation concealment (selection bias)	Low risk	The method used to generate the random allocation sequence was sealed envelope.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	A total of 20 treatment group females and 20 control females completed the study with 100% attendance.
Participants analysed in group allocated (attrition bias)	Low risk	A total of 20 treatment group females and 20 control females completed the study with 100% attendance.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the group's mean baseline (t0) in all measured parameters.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: changed response to yes
Compliance acceptable in all groups (performance bias)	Low risk	One-hundred per cent attendance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Both groups were assessed at baseline and four weeks and only the exercise group was assessed at eight weeks.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Masse-Alarie 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Canada, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 24 (E1 = 12, E2 = 12) **Chronic LBP duration:** 45.05 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 36%
Interventions	**Exercise Group 1 (E1):** Isometric activation of deep core muscles; type = core strengthening; duration = 3 weeks; dose = low; design = partially individualised; delivery = independent with follow-up; additional intervention = advice/education **Exercise Group 2 (E2):** Global exercise targeting the paravertebral muscles; type = core strengthening; duration = 3 weeks; dose = low; design = standardised; delivery = independent with follow-up; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8.7 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Canadian Foundation for Innovation (CS equipment); Fonds de Recherche du Québec – Santé (HMA and LDB PhD studentships); Canadian Institutes of Health Research (HMA studentship) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A member of the laboratory not involved in the study prepared a random list where each new participant was allocated to an exercise group.
Allocation concealment (selection bias)	Low risk	Third party randomisation
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patient could not be blinded due to nature of intervention; 2. Both interventions required supervision and training; one intervention was not clearly superior to the other.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. One intervention was not clearly superior to the other; protocol was explicit.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 2. Pain and function questionnaires are subjective; 3. Both groups were getting exercise interventions.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 2 of 24 dropped out; equal between groups.
Participants analysed in group allocated (attrition bias)	Low risk	1. Not stated but numbers matched

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Masse-Alarie 2016 *(Continued)*

Selective reporting (reporting bias)	Low risk	1. Protocol referenced in publication
Groups similar at baseline (selection bias)	Low risk	Participants similar on gender, age, height, weight, body mass index, sedentary lifestyle, pain duration, and side of pain
Co-interventions avoided or similar (performance bias)	Low risk	No information on additional therapies
Compliance acceptable in all groups (performance bias)	High risk	Exercise log collected but not reported
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcomes assessed similarly; 2. Visual Analogue Scale & Oswestry Disability Index valid.
Other bias	Low risk	Appeared free from other sources of bias

Maul 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** Switzerland, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 148 (E1 = 74, C1 = 74) **Chronic LBP duration:** 12 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 39 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Static and dynamic strengthening exercises (1st phase, 1-4 weeks), Norsk machines for strength and endurance (2nd phase, 5-8 weeks), sequence training (3rd phase, 9-12 weeks); type = strengthening; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = back school **Comparison Group 1 (C1):** Other conservative treatment (back school)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Swiss National Science Foundation (Project NFP 26, No. 4026-27064); BBW (Project ''SOS-LBD'', No. 97.0046) **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Maul 2005 *(Continued)*

Random sequence generation (selection bias)	Low risk	Prior to a first contact they were randomly assigned to the exercise or comparison group by means of a computer randomisation program.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Prior to the start of the intervention study all participants were thoroughly examined and completed several functional tests and questionnaires.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	The total number of subjects who finally participated in the study was 183.
Participants analysed in group allocated (attrition bias)	High risk	For analysis, only subjects that had completed the intervention programme were considered.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics did not differ for low back pain symptoms and demographic variables.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Mazloum 2018

Study characteristics	
Methods	**Study design:** RCT **Setting:** Iran, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 60 (E1 = 20, E2 = 20, C1 = 20) **Chronic LBP duration:** 32 months (moderate) **Neurological/radicular symptoms:** No participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mazloum 2018 *(Continued)*

	Mean age (years): 40 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Selective Pilates moves: shoulder bridge, side kick, one leg stretch, hundred, roll up, swan dive, swimming, one leg circle, double arm stretch, spine twist; type = Pilates; duration = 6 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Extension-based exercises: deep breathing in prone, passive trunk extension on elbows in prone, passive trunk extension on hands in prone, passive trunk extension in standing, knee to chest in crook lying, trunk flexion in sitting on a chair; type = stretching; duration = 6 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 10 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were "randomly allocated".
Allocation concealment (selection bias)	High risk	Insufficient information to determine whether treatment allocation was concealed
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from intended interventions; experimental group was structured and controlled by care providers, and control group was purposely broad
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from intended interventions; experimental group was structured, and control group was purposely broad
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Likely that lack of patient blinding introduced bias to outcome assessments, as the experimental groups were clearly "better" than the control; study results for all outcomes supported this.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 22% (13/60); 2. No comparison of dropout subjects to non-dropouts to determine if missing data introduced selection bias; 3. Missingness could be caused by increased disability from low back pain; 4. Rate of missing subjects in each group was very similar.
Participants analysed in group allocated (attrition bias)	Low risk	1. All patients were analysed according to the allocation to which they were randomised.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this article everything was fully reported.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mazloum 2018 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics, except there was no report of patient sex.
Co-interventions avoided or similar (performance bias)	Low risk	Avoided co-interventions by excluding patients receiving physical therapy or other treatment interventions in the past six months; no mention of medication use
Compliance acceptable in all groups (performance bias)	High risk	No information about compliance/adherence/attendance of interventions in this study
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

McCaskey 2018

Study characteristics

Methods	**Study design:** RCT (NCT02304120) **Setting:** Switzerland, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 22 (E1 = 11, E2 = 11) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 55 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Sensorimotor training: balance training involving standing on a labile, swaying platform; type = core strengthening; duration = 4.5 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = physiotherapy **Exercise Group 2 (E2):** Low intensity cardio on a treadmill, elliptical, or stationary bike for 15 minutes at a comfortable pace; type = aerobic; duration = 4.5 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = physiotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8.5 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Cantonal Department of Health and Social Services, Canton of Argovia; Reha Rheinfelden, Switzerland **Other:** Information modified for author contact, SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomisation schedule using permuted blocks of random sizes; block sizes were not disclosed, to ensure concealment.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

McCaskey 2018 *(Continued)*

Allocation concealment (selection bias)	Low risk	An independent third party created the randomisation list, which was stored away from the study team to ensure concealment of treatment allocation.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Concomitant care was not restricted but sham exercise was also exercise based so the distinction between comparator and treatment group was less obvious; no changes in lifestyle were reported.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Protocol was explicit; no decision-making by study personnel; opinion may have transferred to patients but study personnel were not responsible for other care.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. No obvious reason for patient to alter response; potentially due to the novel nature of the experimental group, but control group was receiving exercise which they presumably thought was effective.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 4.5% (1/22).
Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was performed; those who dropped out before the end had their missing data imputed based on the carry-forward method.
Selective reporting (reporting bias)	Low risk	1. T-test analysis plan in protocol was not executed due to non-normal distribution of data and heteroscedasticity of variance. Instead, a more robust method was used.
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics, except duration of symptoms was not measured.
Co-interventions avoided or similar (performance bias)	High risk	No mention of co-interventions in the study report
Compliance acceptable in all groups (performance bias)	Low risk	81.1% of people attended all sessions and 4 of 22 attended 8 of 9; therefore 100% of people attended at least 8 sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

McDonough 2013

Study characteristics	
Methods	**Study design:** RCT (ISRCTN67030896) **Setting:** Northern Ireland, healthcare **Exercise groups:** 1

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

McDonough 2013 *(Continued)*

	Comparison groups: 1
Participants	**Number of participants:** 56 (E1 = 39, C1 = 17) **Chronic LBP duration:** 10.7 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 49 **Sex (female):** 55%
Interventions	**Exercise Group 1 (E1):** Graded pedometer-driven walking programme focussed on behaviour change; type = aerobic; duration = 8 weeks; dose = high; design = partially individualised; delivery = independent with follow-up; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Quebec Back Pain Disability Scale); HRQoL (EuroQol 5D) **Follow-up time periods available for syntheses:** 9 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Physiotherapy Research Foundation (PRF/08/1); Chartered Society of Physiotherapy, London; Department of Employment and Learning, Belfast, Northern Ireland **Other:** Information modified for author contact, SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Consenting participants were randomised using a computer-generated random allocation sequence.
Allocation concealment (selection bias)	Low risk	Generated the schedule for the random allocation sequence, which was held in a secure cabinet
Blinding of participants and personnel (performance bias) All outcomes	High risk	Because of the nature of the interventions, it was not possible to blind participants or treatment providers.
Blinding of care provider (performance bias)	High risk	Because of the nature of the interventions, it was not possible to blind participants or treatment providers.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Because of the nature of the interventions, it was not possible to blind participants or treatment providers.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Figure 1
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Male/female ratio different and employment; all others similar

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

McDonough 2013 *(Continued)*

Co-interventions avoided or similar (performance bias)	High risk	Author contact: not likely co-interventions during treatment period, but different after nine-week intervention period
Compliance acceptable in all groups (performance bias)	Low risk	Mean percentage adherence with weekly step targets was 70% (95% CI 62% to 77%).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Mehling 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 28 (E1 = 12, E2 = 16) **Chronic LBP duration:** 12 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 49 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Exercises for postural righting, flexibility, and pain relief, stabilisation and strengthening exercises, functional task performance, home exercise daily 20-30 minutes; type = mixed; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Exercise Group 2 (E2):** Daily home exercises (20-30 minutes); type = mixed; duration = 6 weeks; dose = low; design = not specified; delivery = independent; additional intervention = relaxation
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Mount Zion Health Fund; Health Resources & Services Administration Fellowship, US Department of Health and Human Services **Other:** Information modified for author contact; SDs imputed

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Using blocked, stratified randomisation that was performed after completion of all baseline assessments, participants were randomly assigned.
Allocation concealment (selection bias)	Low risk	Group assignments were made using opaque, sequentially-numbered, sealed envelopes that contained the group assignment.
Blinding of participants and personnel (performance bias)	High risk	Participants were not blinded to which intervention they received.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mehling 2005 *(Continued)*
All outcomes

Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Of the 16 subjects undergoing breath therapy, one dropped out after two sessions and was lost to follow-up.
Participants analysed in group allocated (attrition bias)	High risk	Analyses were performed on all available data in an intention-to-treat fashion for 14 subjects in the breath therapy group and 12 subjects in the physical therapy group.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The baseline characteristics of the 28 study subjects were summarised in Table 1.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: measured co-interventions
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: no difference between groups
Timing of outcome assessment similar in all groups (detection bias)	Unclear risk	Support for judgement was not available.

Michaelson 2016

Study characteristics	
Methods	**Study design:** RCT (NCT01061632) **Setting:** Sweden, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 70 (E1 = 35, E2 = 35) **Chronic LBP duration:** 326 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** 56%
Interventions	**Exercise Group 1 (E1):** High load lifting (deadlifts) focussing on strengthening core back muscles; type = core strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Low load lifting: stepped programme aiming to train participants to use their back muscles appropriately during everyday activities; type = core strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Michaelson 2016 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 8.7 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Visare Norr, Sweden; Norrbottens County Council, Sweden **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	For each group, randomisation was performed by applying a computer-generated procedure of n out of N.
Allocation concealment (selection bias)	Low risk	An investigator, who had not met any of the participants, and who was blinded to all patient characteristics, performed a blinded randomisation procedure to provide a concealed allocation.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the interventions; 2. No obvious better intervention; both groups were receiving active care.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Two different physiotherapists delivered the intervention; each was specialised in the intervention; explicit list of interventions; unlikely to have deviated
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for all outcomes were the patients, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Unlikely that the lack of outcome assessor (participant) blinding altered judgement significantly, as there was no clearly "better" intervention; no significant differences between groups
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 3 patients lost to follow-up at earliest time point; one gave the reason of adverse effects, the other 2 had no reason.
Participants analysed in group allocated (attrition bias)	Low risk	1. Participants were analysed according to their allocated treatment group.
Selective reporting (reporting bias)	Low risk	1. Registered protocol NCT01061632: all analyses conducted as described, except secondary outcome "functional capacity" from protocol not reported, seemed to be reported in another article
Groups similar at baseline (selection bias)	Low risk	All relevant baseline characteristics were reported and were similar between treatment groups.
Co-interventions avoided or similar (performance bias)	High risk	Report did not mention any acknowledgement of co-interventions during treatment period.
Compliance acceptable in all groups (performance bias)	High risk	High load lifting participants attended a mean of 11/12 sessions (92%), while the low load motor control group attended a mean of 6.1/12 sessions (51%); compliance was not acceptable in the low load motor control group.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Michaelson 2016 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were similar between treatment groups in terms of timing, tools, etc.; 2. Visual Analogue Scale (for pain) and Roland-Morris Disability Questionnaire (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Michalsen 2016

Study characteristics	
Methods	**Study design:** RCT (DRKS00000373) **Setting:** Germany, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 68 (E1 = 36, C1 = 32) **Chronic LBP duration:** 11.5 years (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 55 **Sex (female):** 76%
Interventions	**Exercise Group 1 (E1):** Home-based exercise programme: booklet containing education on low back pain and exercises that focussed on stretching, strengthening, and joint mobilisation exercises; type = mixed; duration = 8 weeks; dose = low; design = standardised; delivery = independent; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (relaxation)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Else KrönerFresenius-Stiftung, Germany **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was based on the “ranuni” pseudo-random number generator of the SAS/Base statistical software, using a nonstratified block-randomisation with varying block lengths method.
Allocation concealment (selection bias)	Low risk	Patients were randomly allocated to a treatment group by preparing sealed, sequentially numbered opaque envelopes containing the treatment assignments; envelopes were prepared by study biostatistician.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to the intervention due to the nature of the treatment groups; 2. Unlikely that knowledge of the meditation intervention would have significantly altered the way participants in the take-home exercise group performed the intervention
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding led to deviations

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Michalsen 2016 *(Continued)*

		from protocols; meditation was practiced in a standard way, and independent exercise group only saw physician once at the start of study.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for outcomes of interest (pain and disability) were participants, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Possible that either intervention could be considered better with no clear bias in a particular direction
Incomplete outcome data (attrition bias) All outcomes	High risk	1. 16 out of 68 total randomised participants (23.5%) dropped out of the study 2. Dropout rates between treatment groups were not equal, which indicates that there may have been bias in missingness of data between treatment groups; 3. Possible that participants who saw no improvement or worsening of symptoms were more likely to drop out; missingness of outcomes could have been more likely for poorer outcome results; 4. Missingness rates between treatment groups were not equal; 37.5% missing from meditation group, 11% missing from exercise group; 5. Likely that participants who saw no improvement or worsening of symptoms dropped out. Thus, worse outcome results were missing, biasing meditation group to appear better than it truly was.
Participants analysed in group allocated (attrition bias)	Low risk	1. All outcome criteria were analysed by intention-to-treat, including all randomised subjects, irrespective of whether or not they adhered to the protocol, or gave a full set of data.
Selective reporting (reporting bias)	Low risk	1. All analyses conducted and reported as described, except global rating of effectiveness, which was only briefly reported in discussion for one treatment group
Groups similar at baseline (selection bias)	Low risk	All relevant baseline characteristics were measured and similar between treatment groups, except there was a higher proportion of females in the meditation group.
Co-interventions avoided or similar (performance bias)	High risk	Regular analgesic medication should not have been changed in 6 weeks before study start; rescue medication was allowed; use was reduced in 5 subjects of the mediation group and none of exercise group.
Compliance acceptable in all groups (performance bias)	High risk	Study reported "non-compliance" among some individuals, but none of these individuals completed a post-intervention assessment, so we considered them dropouts; no other information reported
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All participants had outcome assessments at the same time points, and for each outcome; were assessed using the same tools; 2. Visual Analogue Scale (for pain) and Roland-Morris Disability Questionnaire (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Miller 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 29 (E1 = 15, E2 = 14)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Miller 2005 *(Continued)*

	Chronic LBP duration: 26 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 49 **Sex (female):** 48%
Interventions	**Exercise Group 1 (E1):** Stabilisation exercises for strengthening the lumbar multifidus and transversus abdominis muscles, home exercises; type = core strengthening; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = not specified **Exercise Group 2 (E2):** McKenzie approach with individual assessments, end-range repeated movements of the spine, home exercises; type = Mckenzie; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education & manual therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Functional Status Questionnaire) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was done using a random number generator to assign each subject a number.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	The examiners were not blinded during data collection.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Of the 30 subjects, 29 completed the study.
Participants analysed in group allocated (attrition bias)	High risk	Twenty-nine of 30 subjects analysed
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The sample consisted of 14 females and 15 males between the ages of 19 and 87 years with a mean age of 47 years.
Co-interventions avoided or similar (performance bias)	Unclear risk	Authors contact: "believed they were compliant but did not assess"

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Miller 2005 *(Continued)*

Compliance acceptable in all groups (performance bias)	Unclear risk	Author contact: tried to monitor compliance but patients not compliant in completing logs
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Miyamoto 2013

Study characteristics	
Methods	**Study design:** RCT (ACTRN12610000523000) **Setting:** Brazil, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 86 (E1 = 43, C1 = 43) **Chronic LBP duration:** 65 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 40 **Sex (female):** 81%
Interventions	**Exercise Group 1 (E1):** Pilates; type = Pilates; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Simple randomisation was conducted using Microsoft Excel for Windows software.
Allocation concealment (selection bias)	Low risk	Allocation was concealed by using consecutively numbered, sealed, opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Due to the nature of the interventions, it was not possible to blind the participants and the therapist involved in the study.
Blinding of care provider (performance bias)	High risk	Due to the nature of the interventions, it was not possible to blind the participants and the therapist involved in the study.
Blinding of outcome assessment (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Miyamoto 2013 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	The analyses followed the intention-to-treat principles.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The baseline data from both groups were similar for most of the characteristics.
Co-interventions avoided or similar (performance bias)	Unclear risk	Participants were allowed to keep taking their medication normally as prescribed by their medical doctor.
Compliance acceptable in all groups (performance bias)	Low risk	There was 90.3% attendance.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Miyamoto 2018

Study characteristics	
Methods	**Study design:** RCT (NCT02241538) **Setting:** Brazil, general population **Exercise groups:** 3 **Comparison groups:** 1
Participants	**Number of participants:** 296 (E1 = 74, E2 = 74, E3 = 74, C1 = 74) **Chronic LBP duration:** 45.75 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 48 **Sex (female):** 76%
Interventions	**Exercise Group 1 (E1):** Pilates: 3 sessions of Pilates per week, including ground exercises (with or without accessories, such as ball, magic circle and toning ball) and apparatus exercises (Barrel, Cadillac, Chair and Reformer—Metalife, Santa Catarina, Brazil) for 6 weeks, warm-up (breathing and mobility exercises) and cool-down (relaxation exercises and massage with a ball); type = Pilates; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Exercise Group 2 (E2):** Pilates: 2 sessions of Pilates per week, ground exercises (with or without accessories, such as ball, magic circle and toning ball) and apparatus exercises (Barrel, Cadillac, Chair and Reformer—Metalife, Santa Catarina, Brazil) for 6 weeks, warm-up and cool-down; type = Pilates; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Exercise Group 3 (E3):** Pilates once per week: ground exercises (with or without accessories, such as ball, magic circle and toning ball) and apparatus exercises (Barrel, Cadillac, Chair and Reformer—Metalife, Santa Catarina, Brazil) for 6 weeks, warm-up and cool-down; type = Pilates; duration = 6 weeks;

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Miyamoto 2018 *(Continued)*

	dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale (−5 to +5)) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** GCM is an instructor of NeoPilates courses. **Funding source:** São Paulo Research Foundation (process: 2013/26321-8 and 2016/07915-2) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Random number table
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed through sealed opaque envelopes sequentially numbered by a researcher, blinded for patient characteristics.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Very difficult to blind the patient to allocation, due to the differences between control and experimental groups; 2. Unlikely that lack of patient blinding led to deviations; intervention required trained personnel; patients explicitly told not to start co-interventions; supposedly monitored but no report
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused meaningful deviations; they were quite clearly prescribed; advice control group versus Pilates; personnel had no contact with the control group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and disability were patients, and so were not likely to have been blinded to allocation, due to the distinct nature of each treatment group; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment; varying levels of intensity likely came with expectations of better care.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Follow-up at 6 weeks for clinical outcomes was 94%.
Participants analysed in group allocated (attrition bias)	Low risk	1. All patients were explicitly reported to have been analysed by the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	1. Protocol NCT02241538: all outcomes analysed and reported as per protocol.
Groups similar at baseline (selection bias)	Low risk	No significant differences between groups on age, sex, body mass index, duration, marital status, education, previous receipt of treatment, pain, disability or other outcomes at baseline

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Miyamoto 2018 *(Continued)*

Co-interventions avoided or similar (performance bias)	High risk	Co-interventions were discouraged in all groups; patients were allowed to use their usual medication, and there were significant differences in use between Pilates and control.
Compliance acceptable in all groups (performance bias)	Low risk	On average 84% of sessions were attended.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were performed with the same tools and at the same time across all treatment groups; 2. Numeric Pain Rating Scale and Roland-Morris Disability Questionnaire (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Mohseni-Bandpei 2011

Study characteristics

Methods	**Study design:** RCT **Setting:** Iran, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 20 (E1 = 10, E2 = 10) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 35 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Back flexor and extensor strengthening and pelvic floor exercises (4 repetitions of 5 second contractions, progressing to 10 repetitions of 10 second contractions) with biofeedback; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = electrotherapy **Exercise Group 2 (E2):** Back flexion and extension strengthening exercises with biofeedback; type = strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** University of Social Welfare and Rehabilitation Sciences; Mazandaran University of Medical Sciences **Other:** SDs imputed. Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The participants who met the inclusion and exclusion criteria were randomly assigned to one of the two groups through a block-style randomisation schedule.
Allocation concealment (selection bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mohseni-Bandpei 2011 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts during treatment; 25% loss at follow-up
Participants analysed in group allocated (attrition bias)	Unclear risk	Three months follow-up was performed on all patients to assess any changes in pain intensity, functional disability, PFM endurance and strength.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	As is shown in Table 1, there was no statistically significant difference between the two groups at baseline.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Same timing

Monticone 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 20 (E1 = 10, E2 = 10) **Chronic LBP duration:** 14.45 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 58 **Sex (female):** 55%
Interventions	**Exercise Group 1 (E1):** Motor training (spinal stabilising exercises), stretching; type = core strengthening; duration = 8 weeks; dose = high; design = individualised; delivery = not specified; additional intervention = advice/education & psychological therapy & manual therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Monticone 2014 *(Continued)*

	Exercise Group 2 (E2): Stretching and muscle strengthening; type = stretching & strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = not specified; additional intervention = advice/education & manual therapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index); work (ability to participate/perform usual activities); HRQoL (36-Item Short Form Survey (Italian)) **Follow-up time periods available for syntheses:** 8 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** EuroSpine Task Force on Research **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Using a list of blinded treatment codes previously generated by a biostatistician
Allocation concealment (selection bias)	Low risk	Automatic assignment system in order to conceal the allocation
Blinding of participants and personnel (performance bias) All outcomes	Low risk	In order to limit expectation bias and reduce problems of cross-over, the patients were blinded to the study hypothesis.
Blinding of care provider (performance bias)	High risk	The physiatrists, the psychologist, the physiotherapists, and the patients could not be blinded.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	In order to limit expectation bias and reduce problems of cross-over, the patients were blinded to the study hypothesis.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All of the participants completed the treatment interventions and all of the assessment tests.
Participants analysed in group allocated (attrition bias)	Low risk	All of the participants completed the treatment interventions and all of the assessment tests.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no differences between the groups in terms of age, body mass index, or the duration of pain before study enrolment.
Co-interventions avoided or similar (performance bias)	Low risk	No major pharmacological agents were allowed, although mild analgesics and non-steroidal anti-inflammatory drugs were permitted.
Compliance acceptable in all groups (performance bias)	Low risk	All of the participants completed the treatment interventions and all of the assessment tests.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Monticone 2014 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Moon 2013

Study characteristics

Methods	**Study design:** RCT **Setting:** South Korea, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 21 (E1 = 11, E2 = 10) **Chronic LBP duration:** 11.8 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 28 **Sex (female):** 33%
Interventions	**Exercise Group 1 (E1):** Sixteen stabilisation exercises to activate the deep lumbar stabilising muscle groups; type = core strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = not specified **Exercise Group 2 (E2):** Fourteen dynamic lumbar strengthening exercises to activate the extensor and flexor muscle groups; type = strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact, SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Using computer-generated random number sequence
Allocation concealment (selection bias)	Unclear risk	Author contact: did not provide clarity
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Moon 2013 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Two patients in the lumbar dynamic strengthening exercise group and one patient in the lumbar stabilisation exercise group dropped out.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences in the general characteristics or the baseline Visual Analogue Scale and Oswestry Disability Index scores.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: advised against other co-interventions, but did not assess
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: modified response
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Moreschi-Guastala 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Brazil, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 43 (E1 = 23, E2 = 20) **Chronic LBP duration:** 82.62 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 51 **Sex (female):** 79%
Interventions	**Exercise Group 1 (E1):** Global posture re-education: patients hold 3 different positions (statically) for 15 minutes each; type = core strengthening; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Isostretching: cycles of holding 9 different postures for 9 breaths each, with 60-second rest between each cycle; type = stretching; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Moreschi-Guastala 2016 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was done through computer-generated numbers and delivered in sealed, opaque envelopes numbered in sequence.
Allocation concealment (selection bias)	Low risk	Sealed, opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Impossible for patients to be blinded to different exercise protocols; 2. Two treatments were similar; patients likely did not have expectations about which treatment would be more effective and thus their expectations of effectiveness would be unbiased.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Same care provider conducted both active treatments; protocol was explicit, no decision-making by personnel; opinion may have transferred to patients but no clear better intervention
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcomes in question were pain and function which were assessed by the patients who were not blinded; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Neither treatment was obviously better than the other; no reason to alter patient response
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was = > 0% in both treatment groups; one dropout was due to a change of address.
Participants analysed in group allocated (attrition bias)	Unclear risk	1. Explicitly stated that intent-to-treat analysis was not performed; It was deemed unnecessary due to the low dropout rate; 2. No mention of patients switching groups; dropout rate was sufficiently small that the missing data (and lack of intention-to-treat analysis) was unlikely to have affected overall conclusions.
Selective reporting (reporting bias)	Low risk	1. All study outcomes were fully analysed according to trial registration (NCT01468298).
Groups similar at baseline (selection bias)	Low risk	A few characteristics were collected and similar; no indication of problematic randomisation
Co-interventions avoided or similar (performance bias)	High risk	Co-interventions not reported; no mention of medications and whether or not participants were allowed to use other forms of treatment during treatment
Compliance acceptable in all groups (performance bias)	High risk	No adherence reporting but limitations listed compliance as a difficulty; all treatment sessions were supervised, and thus compliance could have been reported.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcomes assessments were performed identically for participants in both treatment groups; data collection was performed at the same time for both treatment groups and with the same tools; 2. Visual Analogue Scale and Roland-Morris Disability Questionnaire are very commonly used in low back pain studies and have been validated in this context.
Other bias	Low risk	Under-reported, no other apparent sources of bias

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Morone 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 70 (E1 = 41, C1 = 29) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 60 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Stretching, strengthening, postural exercises, breathing exercises; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = psychological therapy & back school **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention, only medical/pharmacological assistance)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Block randomisation
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Another physician was involved in patients' assessment and was unaware of the treatment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Four patients dropped out at the beginning of the study.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Morone 2011 *(Continued)*

Participants analysed in group allocated (attrition bias)	High risk	Per-protocol analysis
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No statistical differences at baseline
Co-interventions avoided or similar (performance bias)	Low risk	Some analgesics allowed in both groups
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Morone 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 75 (E1 = 25, C1 = 25, C2 = 25) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 55 **Sex (female):** 72%
Interventions	**Exercise Group 1 (E1):** Stretching, strengthening, postural exercises, breathing exercises; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = psychological therapy & back school **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention, only medical/pharmacological assistance) **Comparison Group 2 (C2):** Other conservative treatment (back school)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Morone 2012 *(Continued)*

Random sequence generation (selection bias)	Low risk	Specifically, patients were asked to take a sealed envelope from a box, containing a piece of paper with the assignment which was concealed.
Allocation concealment (selection bias)	Low risk	Patients were asked to take a sealed envelope from a box, containing a piece of paper with the assignment, which was concealed.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Each patient was assessed before and at the end of treatment and also at the 12- and 24-week follow-up by the same physician, who was unaware of the allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All scores of the 25 participants in each group were recorded and analysed statistically.
Participants analysed in group allocated (attrition bias)	Low risk	An intention-to-treat analysis was performed.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	At baseline, no significant differences were observed between groups.
Co-interventions avoided or similar (performance bias)	Low risk	The same drugs were permitted in the three groups.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Moseley 2002

Study characteristics	
Methods	**Study design:** RCT **Setting:** Australia, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 57 (E1 = 29, C1 = 28) **Chronic LBP duration:** 8.8 weeks (moderate) **Neurological/radicular symptoms:** Not specified

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Moseley 2002 *(Continued)*

	Mean age (years): 41 **Sex (female):** 59%
Interventions	**Exercise Group 1 (E1):** Specific trunk muscle training: individualised with home exercise programme (Richardson & Jull); type = core strengthening; duration = 4 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Comparison Group 1 (C1):** Usual care/no treatment (usual general practitioner care; no physiotherapy)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 4 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Flip of a coin by independent person.
Allocation concealment (selection bias)	Low risk	Independent person assigned to group.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	High risk	Sixty-seven per cent completed one year follow-up.
Participants analysed in group allocated (attrition bias)	High risk	Reported intention to treat, but not clear that the 29/28 subjects were all included in the analysis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	Unclear if they sought other physiotherapy treatment; did not measure co-interventions in the intervention group.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Moseley 2002 *(Continued)*

Compliance acceptable in all groups (performance bias)	Low risk	Compliance not assessed in the home programme; not mentioned regarding other component.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Moussouli 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** Greece, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 39 (E1 = 13, E2 = 13, C1 = 13) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 57 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Isometric stabilisation exercises: isometric contractions without motion of core muscle groups, warm-up (5-10 minutes) and cool-down (5-7 minutes); type = core strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none **Exercise Group 2 (E2):** Isotonic stabilisation exercises: dynamic contractions of core muscle groups, warm-up (5-10 minutes) and cool-down (5-7 minutes); type = core strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (36-Item Short Form Survey); function (36-Item Short Form Survey); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 4 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Author contact: only patients in experimental group randomised, control group not randomised
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Moussouli 2014 *(Continued)*

Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	High risk	Only outcomes with significant differences (pain and vitality) were reported.
Groups similar at baseline (selection bias)	High risk	The control group displayed a greater mean age of approximately 62 years compared with the two exercise groups with an average age of 53 years.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	No missing responses were observed.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Muharram 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** Thailand, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 164 (E1 = 82, C1 = 82) **Chronic LBP duration:** 3 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 43 **Sex (female):** 44%
Interventions	**Exercise Group 1 (E1):** Chen style shadowboxing exercise (24 steps, involving multisegmental and multidirectional movement patterns); type = aerobic & strengthening; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no exercise, but given suggestions for healthy living)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Muharram 2011 *(Continued)*

Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised using random number generator
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	During the study period, three control group participants and four group participants dropped out of the study.
Participants analysed in group allocated (attrition bias)	Unclear risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No statistically significant differences were found between the experimental and control groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Although participants in the control group were encouraged to engage in healthy living activities, they were not directed in any specific activities.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Murtezani 2011

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Murtezani 2011 *(Continued)*

Methods	**Study design:** RCT **Setting:** Kosovo, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 101 (E1 = 50, C1 = 51) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** 49%
Interventions	**Exercise Group 1 (E1):** Aerobic exercise: warm-up (10-minutes), biking, treadmill walking, or stair climbing, stretching (5 minutes); type = aerobic & stretching; duration = 12 weeks; dose = high; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Patients who had consented to participate were randomly allocated in accordance with recognised procedures, by computer-generated random allocation.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Fourteen/115 dropped out; reasons were given and none were linked to adverse events.
Participants analysed in group allocated (attrition bias)	High risk	Baseline data and post-treatment data were only given for the completers.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Murtezani 2011 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the two groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Patients' adherence to the treatment was satisfactory.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Nabavi 2018

Study characteristics

Methods	**Study design:** RCT **Setting:** Iran, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 41 (E1 = 20, E2 = 21) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 37 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Stabilisation exercises plus routine physiotherapy consisting of electrotherapy and warm-up exercises which were all stretches; type = core strengthening; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = electrotherapy **Exercise Group 2 (E2):** "Routine" exercises: images portray 16 different poses; type = core strengthening; duration = 4 weeks; dose = not reported; design = standardised; delivery = individual; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** University of Social Welfare and Rehabilitation Sciences **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Block-style randomisation used
Allocation concealment (selection bias)	High risk	No information reported on concealment of treatment allocation

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nabavi 2018 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Treatments were so similar; doubtful that someone would have known if he/she was in the treatment or comparison groups
Blinding of care provider (performance bias)	High risk	1. Physical therapist delivered treatments to both groups, so he/she would have known which group someone was in; 2. Possible that the lack of physiotherapist blinding could have led patients with severe symptoms in the routine treatment group to receive more care at the physiotherapist's discretion; 3. Unlikely that deviations from intended intervention due to lack of physiotherapist blinding were balanced between groups, as routine care group could be perceived to receive less/inadequate care and need more; 4. If the lack of physiotherapist blinding caused patients in routine care to receive better care than expected, then we would expect smaller differences in outcomes between interventions; we did observe this.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain intensity were patients themselves, who could not be blinded due to the nature of the interventions; 2. If patients perceived one intervention as better than the other, then their lack of blinding may have led patients in the "worse" intervention to feel that they had "worse" outcomes; 3. If lack of patient blinding caused bias toward "worse" outcome assessment, we would expect a greater difference in outcomes between intervention groups; we observed no significant group differences.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was zero, as all patients who were randomised completed the trial.
Participants analysed in group allocated (attrition bias)	Low risk	1. No dropouts from baseline, so they were just analysed in the groups they were treated in.
Selective reporting (reporting bias)	Low risk	1. No linked protocol but within the article there were no obvious omissions. Pain reported and biomechanical function reported
Groups similar at baseline (selection bias)	Low risk	No significant differences between the intervention groups on age, pain intensity or body mass index at baseline; neither sex nor duration of symptoms were reported.
Co-interventions avoided or similar (performance bias)	High risk	No report on whether co-interventions were controlled in this study
Compliance acceptable in all groups (performance bias)	High risk	Compliance not assessed
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcomes assessed at same time of day, blinded assessors used across all groups; 2. Visual Analogue Scale for pain (our outcome of interest here)
Other bias	Low risk	Appeared free from other sources of bias

Nagrale 2012

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nagrale 2012 *(Continued)*

Methods	**Study design:** RCT **Setting:** India, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 60 (E1 = 30, E2 = 30) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 38 **Sex (female):** 65%
Interventions	**Exercise Group 1 (E1):** Warm-up on stationary bike (5 minutes), lumbar spine mobilisation to the hypomobile segments (10 minutes), lumbar stabilisation exercises (2 series of 10 repetitions of wall squats, bridges, tilts, and quadruped lifts); type = core strengthening & stretching; duration = 3 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = electrotherapy **Exercise Group 2 (E2):** Warm-up on stationary bike (5 minutes), lumbar spine mobilisation to the hypomobile segments (10 minutes), lumbar stabilisation exercises (2 series of 10 repetitions of wall squats, bridges, tilts, and quadruped lifts), slump-stretch exercises; type = mixed; duration = 3 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = electrotherapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A computer-generated randomised block of numbers obtained before the study was used to determine group assignment
Allocation concealment (selection bias)	Low risk	After the baseline examination was completed, the examining therapist left the room and a second, blinded therapist entered.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Subjects completed several self-report outcome measures.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All 60 subjects completed the study and were included in the analysis.
Participants analysed in group allocated (attrition bias)	Low risk	All 60 subjects completed the study and were included in the analysis.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nagrale 2012 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The baseline characteristics were analysed and found to be similar between groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Nassif 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** France, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 75 (E1 = 37, C1 = 38) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 38%
Interventions	**Exercise Group 1 (E1):** Joint flexion and extension, stretching, stability, co-ordination, and muscle strengthening exercises; type = mixed; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Usual care/no treatment (control group: continued with normal activity and treatment)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); work (subscale of the Dallas Pain Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	All participants signed a written informed consent before the intervention and were assigned by simple randomisation.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nassif 2011 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Because of the complexity of the workplace setting, there could be no patient blinding.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	The compliance of the subjects at each stage of the study are presented in the CONSORT flowchart diagram in Figure 1.
Participants analysed in group allocated (attrition bias)	High risk	Excluded subjects with a large amount of missing data
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Not described
Co-interventions avoided or similar (performance bias)	High risk	Our control group had a significant improvement at six months for Numeric Rating Scale, impact of pain on work and recreational activity, finger to floor distance, and finger to tip of toes distance, even though the individuals did not go through the intervention.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Natour 2015

Study characteristics	
Methods	**Study design:** RCT **Setting:** Brazil, not specified **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 60 (E1 = 30, C1 = 30) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 48 **Sex (female):** 78%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Natour 2015 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Pilates; type = Pilates; duration = 13 weeks; dose = high; design = standardised; delivery = group; additional intervention = anti-inflammatory/analgesics **Comparison Group 1 (C1):** Other conservative treatment (anti-inflammatory/analgesics)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Fundacao Amparo a Pesquisa do Estado de Sao Paulo (2007/53423-5) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Electronically generated randomisation table
Allocation concealment (selection bias)	Low risk	Sealed opaque envelopes were used to ensure the confidentiality of the assignment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Three dropouts (one in control, other two in experimental)
Participants analysed in group allocated (attrition bias)	Low risk	Values were carried forward.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No statistically significant differences
Co-interventions avoided or similar (performance bias)	Low risk	NSAIDS were used comparably in both groups.
Compliance acceptable in all groups (performance bias)	Low risk	In total, 96% of patients attended all sessions while 4% missed one or two sessions.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Natour 2015 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Nemcic 2013

Study characteristics

Methods	**Study design:** RCT **Setting:** Croatia, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 72 (E1 = 36, E2 = 36) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 46 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Water-based exercise: warm-up, range of motion, strengthening, and stretching exercises, water walking; type = mixed; duration = 3 weeks; dose = low; design = standardised; delivery = group; additional intervention = electrotherapy & manual therapy **Exercise Group 2 (E2):** Land-based exercise: warm-up, strengthening and stretching exercises; type = mixed; duration = 3 weeks; dose = low; design = standardised; delivery = group; additional intervention = electrotherapy & manual therapy
Outcomes	**Core outcomes reported:** Function (Pain Disability Index) **Follow-up time periods available for syntheses:** 3 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised into two groups; equally divided according to gender
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The Physical Disability Index is an observer-administered, performance-based instrument so outcome assessors could be blinded.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nemcic 2013 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Assumed no dropouts
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no statistically significant differences in baseline sociodemographic and anthropometric characteristics between the two groups.
Co-interventions avoided or similar (performance bias)	Low risk	No painkillers (paracetamol, tramadol, nonsteroidal anti-inflammatory drugs) were allowed to be changed during the study.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Nguyen 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** France, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 88 (E1 = 45, C1 = 43) **Chronic LBP duration:** 4.9 months (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 47 **Sex (female):** 59%
Interventions	**Exercise Group 1 (E1):** Daily intensive inpatient spa therapy: Spa therapy (pump-jet shower, whirlpool, underwater massage, balneotherapy stretching), exercise therapy (back and abdominal muscle strengthening); type = core strengthening; duration = 1 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education & manual therapy & relaxation **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Quebec Back Pain Disability Scale); work (return to work at 1 year post-randomisation); HRQoL (12-Item Short Form Survey) **Follow-up time periods available for syntheses:** 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Association Française pour la Recherche Termale (AFRETH 2010 programme) **Other:** Information modified for author contact; sufficient data not available for inclusion in meta-analyses

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nguyen 2017 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomisation list with permuted, variable-size blocks; allocation ratio of assignments was 1:1.
Allocation concealment (selection bias)	Low risk	Randomisation and allocation concealment performed by the investigator who included the patient and used a secured dedicated software (CleanWeb).
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Due to the nature of the intervention, patients were not blinded; 2. Difficult for patients in the control group to access the experimental intervention
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Care of the control group was at the discretion of their physician who, it seemed, was external to the study.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for outcomes of interest (pain and disability) were participants, thus were not blinded due to nature of treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Control group participants (receiving less care) were more likely to experience "resentful demoralization", but this was addressed by not informing them of the content of the experimental intervention.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate for experimental group was 27.3% and control group was 23.3%; both over the 20% threshold of acceptability; 2. Imputed missing data; produced a mean area under the curve; could not be assumed to have corrected for bias; 3. Increased low back pain or decreased function could cause missing outcome; 4. Dropout rates and reasons were reasonably balanced between treatment groups.
Participants analysed in group allocated (attrition bias)	Low risk	1. Primary efficacy analysis was conducted as intent-to-treat; all randomised patients were analysed for the primary outcome in their arm of randomisation.
Selective reporting (reporting bias)	High risk	1. Linked protocol cited regression models and Mann-Whitney U test; paper omitted regression models; missing data imputed in final report, protocol stated method of maximum bias; 2. None of the outcomes that were described in protocol were missing from the final report; no reporting bias; 3. It is possible that the method of dealing with missing data was changed in the final report because it gave better results; imputed results did not differ from non-imputed, so no major issue
Groups similar at baseline (selection bias)	Low risk	All characteristics were similar between the treatment groups at baseline; all relevant characteristics were reported.
Co-interventions avoided or similar (performance bias)	Low risk	Pharmacological and non-pharmacological co-interventions were allowed in both groups and were recorded in the electronic case report form.
Compliance acceptable in all groups (performance bias)	Low risk	84.5% of all expected sessions were attended; of 44 experimental group participants, 36 attended all 5 sessions, 1 attended only 4 sessions, and 1 attended only 2 sessions; 6 did not attend any.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Within each outcome, all participants were assessed with the same tool and at the same time point; 2. Numeric Rating Scale (for pain) and Quebec Back

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nguyen 2017 *(Continued)*

		Pain Disability Scale (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Niemisto 2003

Study characteristics

Methods	**Study design:** RCT **Setting:** Finland, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 204 (E1 = 102, C1 = 102) **Chronic LBP duration:** 312 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 37 **Sex (female):** 54%
Interventions	**Exercise Group 1 (E1):** Stabilising exercises aiming to correct lumbo-pelvic rhythm; type = core strengthening; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); work (sick leave days); HRQoL (Health-Related Quality of Life 15D) **Follow-up time periods available for syntheses:** 20 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None declared **Funding source:** Social Insurance Institute of Finland; Finska Lakarsa¨llskapet **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	A total of 102 patients were assigned randomly to the combination group and 102 to the consultation group.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Niemisto 2003 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Of the subjects, 94% in the combination group and 92% in the consultation group visited the physician for the second examination.
Participants analysed in group allocated (attrition bias)	Unclear risk	Continuous outcomes were analysed according to the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The combination and consultation groups were comparable in sociodemographic characteristics, including age, gender, level of education, and smoking habits.
Co-interventions avoided or similar (performance bias)	Low risk	The proportion of patients using analgesics for their back pain decreased from 32% to 15% and from 36% to 15% in these same two groups.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Noori 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** Iran, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 30 (E1 = 17, C1 = 13) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** 53%
Interventions	**Exercise Group 1 (E1):** William's flexor, McKenzie's extensor, stability, or a combination of these exercises; type = core strengthening; duration = 6 weeks; dose = high; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Noori 2011 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Just the term 'random' had been used in the text.
Allocation concealment (selection bias)	Unclear risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	There was no dropout.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Oh 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** South Korea, general population **Exercise groups:** 3 **Comparison groups:** 1
Participants	**Number of participants:** 37 (E1 = 9, E2 = 9, E3 = 10, C1 = 9)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Oh 2014 *(Continued)*

	Chronic LBP duration: 6.5 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 21 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Horse simulator riding (30 min); type = core strengthening; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Horse simulator riding (20 minutes); type = core strengthening; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 3 (E3):** Horse simulator riding (10 minutes); type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no description)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Hanseo University, Republic of Korea **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Subjects were randomly divided into four groups.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Not described
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Four/41 patients were excluded; everyone else completed.
Participants analysed in group allocated (attrition bias)	Low risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No among-group differences were observed, indicating homogeneity in anthropometric characteristics.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Oh 2014 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	During the study, all participants agreed not to change their daily activity patterns or dietary habits.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Ojoawo 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Nigeria, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 63 (E1 = 21, E2 = 21, E3 = 21) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 52 **Sex (female):** 65%
Interventions	**Exercise Group 1 (E1):** Stabilisation in prone and supine: exercises focussed on back and core muscles, performed in both prone and supine positions; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = manual therapy & other **Exercise Group 2 (E2):** Stabilisation in prone: exercises focussed on back and core muscles, performed in prone position; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = manual therapy & other **Exercise Group 3 (E3):** Stabilisation in supine: exercises focussed on back and core muscles, performed in supine position; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = manual therapy & other
Outcomes	**Core outcomes reported:** Pain (Verbal Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Fishbowl technique; participants drawing group from bowl
Allocation concealment (selection bias)	Low risk	Participants chose their own lot from the bowl.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ojoawo 2017 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from intended interventions; all of the treatments were very similar.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. All exercises were explicitly different; unlikely to be able to change exercise programme.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were the patients themselves, who could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Neither treatment was obviously better than the other; no reason to alter patient response
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate was 11% (7/63).
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that patients were analysed according to the treatment group to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication, all outcomes and analyses were fully reported; no obvious omissions
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics, except sex was not reported.
Co-interventions avoided or similar (performance bias)	Low risk	No mention of co-interventions in this study
Compliance acceptable in all groups (performance bias)	Low risk	Presentation was confusing; seemed that 5 were not completely compliant but no reporting over attendance/compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Oswestry Disability Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Okafor 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** Nigeria, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 30 (E1 = 15, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 55

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Okafor 2012 *(Continued)*

	Sex (female): 67%
Interventions	**Exercise Group 1 (E1):** Aerobic dance; type = aerobic; duration = 6 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = advice/education & manual therapy & heat/ice & anti-inflammatory/analgesics **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (Nottingham Health Profile Questionnaire) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were randomised into two groups.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumednot possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Participants analysed in group allocated (attrition bias)	Unclear risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Only age mentioned
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Okafor 2012 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Olivier 2008

Study characteristics

Methods	**Study design:** RCT **Setting:** France, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 90 (E1 = 30, E2 = 30) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Intensive functional rehabilitation programme: flexibility, endurance, and strengthening exercises for all major muscle groups, daily isokinetic training of the trunk muscles; type = mixed; duration = 4 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = relaxation **Exercise Group 2 (E2):** Intensive functional rehabilitation programme: flexibility, endurance, and strengthening exercises for all major muscle groups; type = mixed; duration = 4 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = relaxation
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Dallas Pain Questionnaire) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Support for judgement was not available.
Allocation concealment (selection bias)	Unclear risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Olivier 2008 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.
Other bias	Unclear risk	Not available: non-English publication

Paolucci 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 30 (E1 = 15, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Stretching of the trunk muscles, erector spine and abdominal reinforcement, postural exercises; type = stretching & strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (back school)
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Paolucci 2012 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Thirty individuals, referred to an academic hospital, were enrolled in this study and randomised into two groups.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Not described
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Paolucci 2017

Study characteristics

Methods	**Study design:** RCT (NCT02231554) **Setting:** Italy, healthcare **Exercise groups:** 2

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Paolucci 2017 *(Continued)*

	Comparison groups: 0
Participants	**Number of participants:** 53 (E1 = 26, E2 = 27) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 61 **Sex (female):** 82%
Interventions	**Exercise Group 1 (E1):** Feldenkrais method: sessions aim to increase self-awareness of movement and pain, and to improve trunk mobility by teaching participants how to perform functional movements without pain; type = functional restoration; duration = 5 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Back School: sessions run by physicians taught general anatomical information about the spine, muscle stretching, core reinforcement, postural and diaphragmatic breathing exercises; type = mixed; duration = 5 weeks; dose = low; design = standardised; delivery = group; additional intervention = back school
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Waddell Disability Index Score); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 13 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were randomised to treatment groups at a 1:1 ratio according to a computer-generated randomisation list (using www.random.org).
Allocation concealment (selection bias)	Low risk	Patient's allocation was printed and hidden under a patch by an independent researcher; patch was only removed by another independent researcher with patient; clinical assessor was kept blinded.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the treatments; 2. Both groups receiving novel care; unlikely to be able to access protocols outside of study context
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely here that providers could significantly alter their delivery; most likely situation is for provider in control group to put in extra effort, but here, control intervention was standardised
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors of the outcomes of interest (pain and disability) were the participants themselves, who could not be blind to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Neither treatment is obviously better than the other; no reason to alter patient response.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Two subjects of 53 dropped out during the training and were not considered in the study; one subject was not assessed at T2; thus, we used the last-observation-carried-forward method.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Paolucci 2017 *(Continued)*

Participants analysed in group allocated (attrition bias)	High risk	1. Excluded patients who were not compliant; 2. By excluding non-compliant patients, analysis was biased to make intervention appear more effective than it otherwise would have been.
Selective reporting (reporting bias)	Low risk	1. No published protocol, but all analysis were performed and reported as described in methods.
Groups similar at baseline (selection bias)	Low risk	Participants in the two treatment groups were similar at baseline on age, sex, body mass index, pain and disability; did not measure duration of pain symptoms
Co-interventions avoided or similar (performance bias)	Low risk	All patients were instructed not to take any new medication of nonsteroidal anti-inflammatory drugs during the study protocol and not to undergo other rehabilitation approaches (those who did so dropped out of the study).
Compliance acceptable in all groups (performance bias)	Low risk	Excluded patients who attended fewer than 9 (out of 10) lessons; compliance was at least 90%.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Within each outcome, participants in each treatment group were measured using the same tool and at the same time point; 2. Visual Analogue Scale (for pain) and Waddell score (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Patti 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Italy, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 38 (E1 = 19, C1 = 19) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 41 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Pilates: floor-based Pilates exercise intervention (2 levels of difficulty, basic and intermediate); type = Pilates; duration = 14 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 14 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Patti 2016 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random number sequence
Allocation concealment (selection bias)	Low risk	Allocation sequence was computer-generated, with group allocation conducted by a research assistant who did not participate in any component of the study.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the interventions; 2. Unlikely that lack of patient blinding led to deviations from intended intervention; Pilates group required instructor; control group was allowed any usual activities.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded due to the nature of the interventions; 2. Unlikely that lack of care provider blinding led to deviations from intended interventions because they only interacted with Pilates group; control group were not followed until outcome assessments.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For the outcome of interest (disability) the outcome assessor was the participant, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No dropouts in this study
Participants analysed in group allocated (attrition bias)	Low risk	1. Participants were analysed according to their treatment allocation.
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: only looked at Oswestry but sourced it as comprising of pain and function, so no obvious omissions
Groups similar at baseline (selection bias)	High risk	At baseline, both groups were similar in age and disability; however, there was no report of sex distribution or duration of symptoms.
Co-interventions avoided or similar (performance bias)	High risk	Participants in exercise group did not use nonsteroidal anti-inflammatory drugs, but control group could if they wished (as well as other social activities and "usual care").
Compliance acceptable in all groups (performance bias)	Low risk	Flow diagram stated all patients completed protocol.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome measurement took place at the same time points for all participants, and using the same tools (for each outcome); 2. Oswestry Disability Index (for disability) is a well-validated tool in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Paungmali 2017

Study characteristics	
Methods	**Study design:** RCT (cross-over) **Setting:** Thailand, healthcare **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 75 (E1 = 25, C1 = 25, C2 = 25) **Chronic LBP duration:** 40.36 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 33 **Sex (female):** 72%
Interventions	**Exercise Group 1 (E1):** Lumbo-pelvic core stabilisation exercises: exercises performed in supine position, focussing on exercising the abdominal and trunk muscles facilitated by the Pilates power gym transformer machine; type = core strengthening; duration = 0 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control: participants lie in supine position, with legs passively angled the same way as the active intervention group) **Comparison Group 2 (C2):** Placebo: automated passive cycling (moves participants legs passively in the same manner as the active intervention, without engaging core muscles)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 0 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** TRF & MUA **Other:** Information modified for author contact; sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Patients were randomised to order of exercises; no description on the randomisation process
Allocation concealment (selection bias)	High risk	No description of randomisation process to judge treatment allocation
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to their intervention due to the nature of the interventions; 2. Lack of patient blinding likely did not lead to deviations from intended intervention, as all patients were aware that they would have the chance to participate in all three study interventions.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded due to the nature of the interventions; 2. Unlikely that lack of care provider blinding led to deviations from intended intervention, as all patients would have the chance to participate in all three study interventions.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For the outcome of interest (pain), the outcome assessors were the participants themselves, who could not be blinded due to the nature of the interventions; 2. Pain and function are subjective and are susceptible to awareness of the intervention; 3. Outcomes in a control versus exercise study likely to be altered by awareness of intervention.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Paungmali 2017 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No description of dropout rate; 2. Analyses would exclude missing; 3. Increased pain and decreased function could cause missing outcome; 4. No information on missing data; 5. No information
Participants analysed in group allocated (attrition bias)	Low risk	1. Participants were analysed according to each intervention they received (complicated by cross-over design, but no cause for concern over risk of bias).
Selective reporting (reporting bias)	Low risk	1. All analyses were conducted and reported as described in methods.
Groups similar at baseline (selection bias)	Low risk	Same 25 participants were in all three intervention "groups" due to cross-over study design; baseline characteristics of all three "groups" were identical.
Co-interventions avoided or similar (performance bias)	Low risk	Subjects were advised to refrain from medications, alcohol, and heavy physical activities for at least 8 hours prior to the test.
Compliance acceptable in all groups (performance bias)	Low risk	No information on compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. For all participants, outcome assessments were performed at the same time points after each intervention, and using the same measurement tools (within each outcome); 2. Visual Analogue Scale (for pain) is a well-validated tool in the context of low back pain.
Other bias	Low risk	Risk of insufficient wash-out period between interventions; partially addressed by randomising the order in which patients participated in the three interventions

Petersen 2002

Study characteristics	
Methods	**Study design:** RCT **Setting:** Denmark, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 260 (E1 = 128, E2 = 132) **Chronic LBP duration:** 54 weeks (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 35 **Sex (female):** 46%
Interventions	**Exercise Group 1 (E1):** Strengthening training: stationary bike; intensive dynamic back strengthening in flexion and extension; type = core strengthening & aerobic; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** McKenzie therapy; type = Mckenzie; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Low Back Pain Rating Scale (Manniche)); work (self-reported sick-leave); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global change in back-related quality of life (5-point)) **Follow-up time periods available for syntheses:** 8 weeks (short); 35 weeks (moderate)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Petersen 2002 *(Continued)*

Notes	**Conflicts of interest:** None to declare **Funding source:** Foundation funds (source not indicated) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement was not available.
Allocation concealment (selection bias)	Low risk	Approximately equal number of subjects in each group preferring the opposite treatment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Author contact: changed from not satisfied - an equal number of subjects preferred other treatment
Blinding of care provider (performance bias)	Low risk	No, but they thought their treatment was the superior.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Self-report
Incomplete outcome data (attrition bias) All outcomes	High risk	Sixty-six per cent follow-up
Participants analysed in group allocated (attrition bias)	High risk	Called it intention to treat, but did not include all patients randomised; third reviewer used for consensus
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	General practitioner comparable; 'patients were encouraged not to seek any other'.
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: changed from unclear - stated they assessed compliance at > 70%; all compliant
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Petersen 2011

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Petersen 2011 *(Continued)*

Methods	**Study design:** RCT **Setting:** Denmark, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 350 (E1 = 175, E2 = 175) **Chronic LBP duration:** 94.5 weeks (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 38 **Sex (female):** 56%
Interventions	**Exercise Group 1 (E1):** McKenzie treatment after individual assessment; type = McKenzie; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** General mobilising exercises, lumbar flexion/extension movement and stretching without directional preference; type = flexibility/mobilising & stretching; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Roland-Morris Disability Questionnaire); work (return to work); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 12 weeks (short); 20 weeks (moderate); 64 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Grant, Foundation, and Professional Organizational funds (sources not indicated) **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was carried out by a computer-generated list of random numbers in blocks of 10 using sealed envelopes.
Allocation concealment (selection bias)	Low risk	A secretary who was not involved in the study prepared opaque, sequentially numbered envelopes indicating one of the two treatments.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Blinding of the practitioners in one group to the treatment given in the other group was not possible.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Follow-up assessment was carried out by a secretary blinded to treatment allocation at the end of treatment, after two months, and one-year post-treatment.
Incomplete outcome data (attrition bias) All outcomes	High risk	A statistically significant larger proportion of patients withdrew or was excluded during treatment in the manipulation group.
Participants analysed in group allocated (attrition bias)	Low risk	Intention-to-treat analysis was performed on all participants in the study.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Petersen 2011 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics were similar for the two treatment groups except for the fact that significantly more patients were on sick leave.
Co-interventions avoided or similar (performance bias)	Low risk	Patients were encouraged not to seek any other kind of treatment for the two months period of self-administered exercises.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Phattharasupharerk 2019a

Study characteristics	
Methods	**Study design:** RCT **Setting:** Thailand, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 72 (E1 = 36, C1 = 36) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 35 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Qigong: Wu Chi meditation and static exercises performed standing in place, and involve various repeated movements with the arms and/or hands; type = other (Qigong); duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = relaxation & other **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund) (Grant Number 59, 2560) **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	A computer program randomised participants equally into two groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Phattharasupharerk 2019a *(Continued)*

Allocation concealment (selection bias)	Low risk	Treatment codes were placed sequentially in numbered, sealed opaque envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the intervention (active qigong versus waiting list); 2. Lack of patient blinding unlikely to lead to deviations from protocol, as they required professional instruction, and control group was promised future intervention (on waiting list).
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from the intended interventions because each provider only interacted with one treatment group after allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For outcomes of interest (pain and disability), the outcome assessors were the participants, who could not be blinded due to the nature of the interventions; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Likely that outcome assessment was altered by lack of blinding because participants on waiting list knew they were receiving no treatment; reflected in outcome results.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. In total, 7 out of 72 (9.7%) participants dropped out of the study.
Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	1. No protocol found. All planned analyses were executed and reported for all primary and secondary outcomes within the context of this publication; no obvious omissions.
Groups similar at baseline (selection bias)	Low risk	All characteristics were similar between groups at baseline; only important characteristic that wasn't measured was the duration of symptoms.
Co-interventions avoided or similar (performance bias)	Low risk	Patients were excluded if they took any medication during the study period; participants were asked to avoid other forms of treatment during the study period.
Compliance acceptable in all groups (performance bias)	Low risk	Two participants in the Qigong group missed one session out of six sessions; otherwise, participants were 100% compliant.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. For each outcome, all participants were assessed at the same time point and using the same tools; 2. Visual Analogue Scale (for pain) and Roland-Morris Disability Questionnaire (for disability) are well-validated scales in the context of low back pain.
Other bias	Low risk	Missing data could also be managed by using the last-observation-carried-forward method. Since dropouts were balanced in both groups, unlikely to cause bias to results

Puntumetakul 2013

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Puntumetakul 2013 *(Continued)*

Methods	**Study design:** RCT **Setting:** Thailand, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 42 (E1 = 21, E2 = 21) **Chronic LBP duration:** 45.8 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 81%
Interventions	**Exercise Group 1 (E1):** Core stabilisation exercises: isolation and co-activation of trunk muscles (weeks 1-2), trunk activation with extremity movement (weeks 3-7), trunk co-activation with functional tasks (weeks 8-10); type = core strengthening; duration = 10 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = heat/ice **Exercise Group 2 (E2):** Conventional trunk stretching exercises: alternating single knee to chest and lateral trunk bending in standing; type = stretching; duration = 10 weeks; dose = high; design = standardised; delivery = group; additional intervention = heat/ice
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 10 weeks (short); 22 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Graduate School (GS 54112101) and Back, Neck and Other Joint Pain Research Group, Khon Kaen University, Thailand **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were randomly assigned to one of the two groups by a block randomisation with block sizes of two, four and six.
Allocation concealment (selection bias)	Low risk	Randomisation results were concealed in sealed and opaque envelopes with consequential numbering.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Exercises were demonstrated and supervised by a research assistant blinded to the outcome assessment.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Two participants were lost to follow-up from each group at 10 weeks of intervention.
Participants analysed in group allocated (attrition bias)	Low risk	All analyses were performed on the basis of intention-to-treat, with the last-observation-carried-forward.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Puntumetakul 2013 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	All characteristics and baseline outcomes were equally balanced between groups.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: reported compliance monitored and no co-interventions
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: similar in each group
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Both groups were assessed at the same times.

Puppin 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** Brazil, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 55 (E1 = 30, C1 = 25) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 38 **Sex (female):** 55%
Interventions	**Exercise Group 1 (E1):** Stretching of various muscle groups, in 6 sequences, followed by 30 seconds rest periods between stretches; type = stretching; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 16 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Randomised in order of arriving at clinic
Allocation concealment (selection bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Puppin 2011 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Seven/62 dropped out - mainly because of return to work after sick leave.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant differences between baseline characteristics
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Quinn 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** United Kingdom, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 29 (E1 = 15, C1 = 14) **Chronic LBP duration:** 4 years (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 43 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** One hour guided Pilates classes, self-guided practice 15 minutes per day, 5 days a week; type = Pilates; duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = none

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Quinn 2011 *(Continued)*

	Comparison Group 1 (C1): Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Used sequentially numbered opaque sealed envelopes
Allocation concealment (selection bias)	Low risk	Opaque sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Patients were aware if they were in Pilates classes or not.
Blinding of care provider (performance bias)	High risk	Principal investigator conducted the Pilates classes for the intervention group.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Baseline and final outcome measures were recorded separately by a blinded physiotherapist.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Only 66% completed the intervention.
Participants analysed in group allocated (attrition bias)	Low risk	Last known values were carried forward to substitute missing values for subjects who failed to attend final assessment.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Compared in Table 2 and tested for significant difference
Co-interventions avoided or similar (performance bias)	Low risk	Stated: control group received no further intervention.
Compliance acceptable in all groups (performance bias)	High risk	Treatment group 10/15 (66%) attended 6/8 sessions; 'High level of compliance' reported in 7/10 who completed
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rantonen 2012

Study characteristics	
Methods	**Study design:** RCT (NCT00908102) **Setting:** Finland, occupational **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 126 (E1 = 43, E2 = 43, C1 = 40) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 32%
Interventions	**Exercise Group 1 (E1):** Light mobilisation and exercises, progressive exercises, home exercises; type = mixed; duration = 12 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy & relaxation **Exercise Group 2 (E2):** Progressive exercise using equipment to target trunk muscles, stretching, home exercises; type = mixed; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); work (sickness absence days); HRQoL (HRQoL 15D instrument) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** During physical medicine unit (PMU) interventions, MH was the head physician of the PMU. ST is the Medical Director of DBC International. **Funding source:** Centenary Foundation of Kymi Corporation; Yrjo Jahnsson Foundation; Juho Vainio Foundation; Finnish Cultural Foundation **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	An independent biostatistician prepared the randomisation scheme using a computer-generated randomisation table.
Allocation concealment (selection bias)	Low risk	Based on the randomisation scheme, a research assistant prepared sealed envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Due to the nature of the interventions, the participants and occupational health professionals were not blinded for the group assignment after randomisation.
Blinding of care provider (performance bias)	High risk	Due to the nature of the interventions, the participants and OH professionals were not blinded for the group assignment after randomisation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Author contact: Doctors likely not blinded, patients not blinded
Incomplete outcome data (attrition bias)	Unclear risk	Author contact: dropout differed.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rantonen 2012 *(Continued)*
All outcomes

Participants analysed in group allocated (attrition bias)	Low risk	All statistical analyses were performed at employee level, according to the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Randomisation was successful and the treatment arms were comparable regarding the relevant demographic factors.
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Unclear risk	Author contact: compliance was dropout rate; did not add clarity for this item
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Rasmussen-Barr 2003

Study characteristics	
Methods	**Study design:** RCT **Setting:** Sweden, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 47 (E1 = 24, C1 = 23) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 38 **Sex (female):** 74%
Interventions	**Exercise Group 1 (E1):** Stabilising exercises: activation and control deep abdominal, multifidus muscles (Richardson & Jull); type = core strengthening; duration = 6 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); Global Perceived Health or Recovery (Global Perceived Health or Recovery (general health Visual Analogue Scale)) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Anne-Marie and Ragnar Hemborg Foundation **Other:** None
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rasmussen-Barr 2003 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Cards in a box
Allocation concealment (selection bias)	Unclear risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	High risk	Sixteen lost to follow-up - no info on how different; six at first follow-up
Participants analysed in group allocated (attrition bias)	High risk	Sample sizes in Table 3 missing values
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	High risk	No, more in manual treatment group reported treatment during follow-up period.
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Rasmussen-Barr 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Sweden, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 71 (E1 = 36, E2 = 35)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rasmussen-Barr 2009 *(Continued)*

	Chronic LBP duration: 10 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 38 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Graded stabilising exercises: low-load endurance exercises (Richardson 1999) and home training, incorporating activation of stabilising muscles into daily life with biofeedback; type = core strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Walking at-home programme: daily 30 minutes walk at fastest pace without pain, general home exercises with walking diary; type = aerobic; duration = 8 weeks; dose = high; design = individualised; delivery = independent with follow-up; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Capio Research Foundation; Ann-Marie and Ragnar Hemborg Foundation **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A fixed allocation randomisation procedure guaranteed equal numbers of patients of each sex in each group.
Allocation concealment (selection bias)	Low risk	The assignments were presented in sealed, sequentially numbered envelopes, and the assignment list was maintained by the clinicians secretarial staff.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1 for dropouts
Participants analysed in group allocated (attrition bias)	Low risk	An intention-to-treat procedure was followed (last-observation-carried-forward).
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Clinical and demographic characteristics were similar between the two groups (Table 1).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rasmussen-Barr 2009 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Twenty-two per cent in the exercise group and 46% in the reference group reported a recurrent need for new treatment periods at the 12-month follow-up.
Compliance acceptable in all groups (performance bias)	Low risk	In the exercise group, there was 96% attendance at the physical therapy sessions, and in the reference group, 71% adherence with the daily walks.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Rhee 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** South Korea, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 42 (E1 = 21, C1 = 21) **Chronic LBP duration:** 11 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 52 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Spinal stabilisation exercises, core muscle strengthening through training the isometric holding function of spinal muscles; type = core strengthening; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Agenda Project, Korea Research Council of Fundamental Science & Technology (P-09-JC-LU63-C01); Korea University; Basic Science Research Program through the National Research Foundation of Korea, Ministry of Education, Science and Technology (2010-0003015) **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	A randomisation list was provided, with patients having an equal chance of being allocated to the intervention or control group.
Allocation concealment (selection bias)	Low risk	The co-ordinator ensured anonymity of allocation with respect to randomisation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rhee 2012 *(Continued)*

Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Assumed no dropouts but not specified
Participants analysed in group allocated (attrition bias)	Unclear risk	Assumed no dropouts but not specified
Selective reporting (reporting bias)	Unclear risk	Pain and Oswestry Disability Index at baseline looked to be different between the two groups.
Groups similar at baseline (selection bias)	Unclear risk	Inspected descriptive statistics for sample characteristics and scatter plots of the data to ensure that no outliers existed in the data set
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Patients kept an exercise log, and phone calls were made to ensure compliance with the exercise protocol.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Four weeks

Risch 1993

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 54 (E1 = 31, C1 = 23) **Chronic LBP duration:** 364 weeks (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 37%
Interventions	**Exercise Group 1 (E1):** Dynamic extension strengthening programme (with machines); type = strengthening; duration = 10 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (West Haven-Yale Multidimensional Pain Inventory); function (Sickness Impact Profile) **Follow-up time periods available for syntheses:** 10 weeks (short)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Risch 1993 *(Continued)*

Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Support for judgement is not available
Allocation concealment (selection bias)	High risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Rittweger 2002

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rittweger 2002 *(Continued)*

Methods	**Study design:** RCT **Setting:** Germany, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 60 (E1 = 30, E2 = 30) **Chronic LBP duration:** 678 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 52 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Lumbar extension, repetitive contraction cycles, constant speed, load gradually increased, resistance exercise of the abdominal and thigh muscles; type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = none **Exercise Group 2 (E2):** Specific exercise: platform that oscillates around a resting axis between the subjects feet during exercise units, the subject performed slow movements of the hip and waist, with bending in the sagittal and frontal planes and rotation in the horizontal plane; type = other (co-ordination); duration = 12 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Pain Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short); 24 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Author contact: changed response
Allocation concealment (selection bias)	Low risk	Author contact: changed response
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	25/30 completed the study in both groups.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rittweger 2002 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Ritvanen 2007

Study characteristics

Methods	**Study design:** RCT **Setting:** Finland, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 61 (E1 = 28, C1 = 33) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 41 **Sex (female):** 44%
Interventions	**Exercise Group 1 (E1):** Stretching, trunk stabilisation exercise, general exercises; type = mixed; duration = 8 weeks; dose = low; design = standardised; delivery = individual; additional intervention = manual therapy **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Finland's Slot Machine Association **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The patients were randomised by a closed envelope system. The closed envelopes were set in two boxes (for men and women separately).
Allocation concealment (selection bias)	Low risk	The patients were randomised by a closed envelope system. The closed envelopes were set in two boxes (for men and women separately).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ritvanen 2007 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Those who administered the interventions were not blinded to study data, but the assessor who evaluated the outcomes was blinded to group assignment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome measurements were repeated one month after the last treatment by independent assessors who were unaware of the patients' treatment histories.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Three dropouts in one group; job busy, ill child, long way to go for treatment
Participants analysed in group allocated (attrition bias)	High risk	Twenty-eight analysed/35, 33 analysed/35
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The randomisation was successful because there was no statistically significant difference between the groups in age and anthropometric characteristics.
Co-interventions avoided or similar (performance bias)	Low risk	The patients did not get any other treatment apart from painkillers.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Roche-Leboucher 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** France, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 132 (E1 = 68, E2 = 64) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 40 **Sex (female):** 35%
Interventions	**Exercise Group 1 (E1):** Isotonic muscle strengthening, endurance training (stepping, jogging, and cycling exercises), supervised weightlifting exercises, work simulation workshop; type = functional

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Roche-Leboucher 2011 *(Continued)*

	restoration; duration = 5 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy & relaxation **Exercise Group 2 (E2):** Flexibility exercises and pain management (first 2 weeks), followed by strengthening exercises and functional training with home exercises (stretching, jogging, and swimming); type = mixed; duration = 5 weeks; dose = high; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); work (reduction in number of sick-leave days) **Follow-up time periods available for syntheses:** 5 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Institutional funds (source not indicated) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Patients were randomised by an independent methodologist to one of the two rehabilitation programmes, according to an eight-element permutation table.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Patients were evaluated at baseline and after treatment by a physiotherapist who was not blinded but had not been involved in the treatment.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Nineteen (14.4%) patients were missing at one-year follow-up.
Participants analysed in group allocated (attrition bias)	High risk	The main limitation of this study was the lost to follow-up rate.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There was no significant difference between the two groups.
Co-interventions avoided or similar (performance bias)	Low risk	No other co-interventions or treatments were allowed, except their medication prescribed at baseline.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Roche-Leboucher 2011 *(Continued)*

Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Ryan 2010

Study characteristics

Methods	**Study design:** RCT **Setting:** United Kingdom, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 38 (E1 = 20, C1 = 18) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 66%
Interventions	**Exercise Group 1 (E1):** Circuit-based, graded, aerobic exercise with some core stability exercises (Moffett 2000); type = mixed; duration = 6 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate)
Notes	**Conflicts of interest:** One of the authors is a co-inventor of the activPAL physical activity monitor and a director of PAL technologies Ltd. **Funding source:** School of Health and Social Care of Glasgow Caledonian University **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were recruited from five different physiotherapy departments and randomised using a random number generator.
Allocation concealment (selection bias)	Low risk	Blinding allocation was concealed using sealed envelopes and allocation patients were not allocated until after the education had been received.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants were not blinded.
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome assessing therapist was blinded.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ryan 2010 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Seven participants dropped out.
Participants analysed in group allocated (attrition bias)	Low risk	Participants remained in their group throughout.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Duration of pain varied and was used as a covariate.
Co-interventions avoided or similar (performance bias)	Low risk	While participants were requested not to see co-interventions beyond their general practitioner during the course of the study, one participant from the education only group reported that she received treatment.
Compliance acceptable in all groups (performance bias)	High risk	Three of the participants attended no exercise classes and only six attended all six classes.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Rydeard 2006a

Study characteristics	
Methods	**Study design:** RCT **Setting:** Hong Kong, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 39 (E1 = 21, C1 = 18) **Chronic LBP duration:** 7 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 35 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Pilates (on floor mat and Pilates Reformer) to specifically activate the gluteus maximus and home practice; type = Pilates; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: consultation with healthcare professionals as necessary, continue with current physical activity)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rydeard 2006a *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was administered by independent office staff. Subjects randomly pulled a card from a box of concealed premarked cards to obtain assignment.
Allocation concealment (selection bias)	Low risk	Randomisation was administered by independent office staff with subjects randomly pulling a card from a box of concealed premarked cards to obtain assignment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Data collection monitored both pain intensity and functional status and included two self-report questionnaires administered by the research assistant.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All subjects in the main study completed the four-week treatment intervention according to the study protocol.
Participants analysed in group allocated (attrition bias)	Low risk	The first intention-to-treat analysis, using 'last observation carried forward, revealed significant improvements in Roland-Morris Disability Questionnaire scores over the 12-month period.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Analysis indicated no significant difference between the groups regarding baseline characteristics (Table 1).
Co-interventions avoided or similar (performance bias)	Low risk	Most of the subjects had seen more than one medical specialist over the years and were continuing to seek treatment.
Compliance acceptable in all groups (performance bias)	Low risk	Both groups completed the study and compliance was high, with 100% attendance at scheduled clinic appointments.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Salamat 2017

Study characteristics	
Methods	**Study design:** RCT (IRCT2013121615822N1) **Setting:** Iran, healthcare **Exercise groups:** 2

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Salamat 2017 *(Continued)*

	Comparison groups: 0
Participants	**Number of participants:** 32 (E1 = 16, E2 = 16) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 36 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Stabilisation exercises: co-ordinated training and independent activity of deep trunk muscles including transversus abdominis and multifidus in pain-free positions and movements, also done at home; type = core strengthening; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Movement control exercises: dissociation of lumbo-pelvic movement from thoracic movement in order to reduce excessive lumbar extension, training to perform everyday and work tasks without abnormal bracing or breathing, daily exercises at home; type = functional restoration; duration = 4 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Rehabilitation School, Tehran University of Medical Sciences **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomisation method was not described enough to make an assessment of adequacy; only description was that randomisation was achieved "via a random number sequence".
Allocation concealment (selection bias)	High risk	Not enough information on the randomisation and treatment allocation methods to assess
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients were not blind to intervention because the interventions were too different for patients not to know the difference, however the team "Attempt[ed] to keep subjects blind to their group assignment"; 2. Both groups getting sufficient treatment, unlikely to be dissatisfied; difficult to access outside of study context
Blinding of care provider (performance bias)	Low risk	1. "The treating therapist was not blind to the treatment group of the subjects" (treating therapist = care providers); 2. Explicit set of exercises in protocol; no decision by care provider; since there wasn't a favoured intervention prior, unlikely that providers tried to "balance" care
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors were the patients for outcomes pain and disability, they were not blinded; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Because there was no clear "better" intervention between the two, it is unlikely that the lack of patient blinding led to a serious biasing of judgement for outcome assessment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Dropout rate (25%) was borderline unacceptable for a short-term follow-up (four weeks); 2. Evidence that the analysis was not biased by missing data: the dropout rate was identical for both treatment groups, and dropout reasons were unrelated to low back pain.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Salamat 2017 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analysis was included; all patients were analysed according to the treatment to which they were randomised.
Selective reporting (reporting bias)	Low risk	1. No linked protocol found but within this paper all outcomes and analyses fully reported; no obvious omissions
Groups similar at baseline (selection bias)	Low risk	Groups were similar at baseline on age, body mass index, pain and disability, but did not report sex or duration of back pain symptoms at baseline.
Co-interventions avoided or similar (performance bias)	High risk	No information on how the study dealt with co-interventions (or if they existed) was reported.
Compliance acceptable in all groups (performance bias)	High risk	Compliance was not reported, so it was difficult to assess its acceptability; intervention was supervised, which tends to increase compliance.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcome assessments were similarly timed and measured for all treatment groups; 2. Numeric Rating Scale (for pain) and Oswestry Disability Index (for disability) are well-validated tools in the low back pain field.
Other bias	Low risk	Appeared free from other sources of bias

Salavati 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Iran, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 40 (E1 = 20, C1 = 20) **Chronic LBP duration:** 40.42 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 31 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** Stabilisation exercises plus physiotherapy, including some stretching; type = core strengthening; duration = 4 weeks; dose = not reported; design = partially individualised; delivery = individual; additional intervention = electrotherapy & physiotherapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Salavati 2016 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Subjects were "randomly assigned".
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Both groups getting in depth care
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. One physiotherapist treated both groups; protocol was explicit, no decision-making by personnel; opinion may have transferred to patients but no clear deviations.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcomes in question were assessed by the patients who were not blinded to the intervention; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Although control group was receiving no thorough active component, likely that to the participants it was classified as activity, no reason to alter response
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Technically, the dropout rate was zero because the two patients who did drop out were replaced.
Participants analysed in group allocated (attrition bias)	Unclear risk	1. Subjects were replaced if dropped out; 2. Dropout rate was sufficiently small that the missing data (and lack of intention-to-treat analysis) was unlikely to have affected overall conclusions.
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all outcomes and analyses fully reported; no obvious omissions
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics.
Co-interventions avoided or similar (performance bias)	High risk	No report or description of dealing with co-interventions in this study
Compliance acceptable in all groups (performance bias)	Low risk	Seemed like protocol had to completed in full for inclusion in the analysis; assuming compliance was 100%
Timing of outcome assessment similar in all groups (detection bias)	High risk	1. There were two patients that dropped out part-way through the study, and replaced; while these patients had the same length of study period, their start and end were staggered from the rest; 2. Oswestry Disability Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	High risk	Five subjects were substituted for new ones due to inability to complete the treatment; these patients had an Oswestry and pain higher than average; no mention on how randomisation was preserved

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saner 2016

Study characteristics	
Methods	**Study design:** RCT (ISRCTN80064281) **Setting:** Switzerland, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 106 (E1 = 52, E2 = 54) **Chronic LBP duration:** 10 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 42 **Sex (female):** 38%
Interventions	**Exercise Group 1 (E1):** Movement control exercise, learning to control aberrant movements causing pain; type = core strengthening; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = physiotherapy **Exercise Group 2 (E2):** General exercise focussing on strength and endurance of muscles of the lumbar/pelvic region and legs; type = core strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = physiotherapy
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Swiss National Science Foundation (SNSF) (Project no. 127240) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation schedule used to assign participants to either treatment group was generated electronically using a block size of four.
Allocation concealment (selection bias)	Low risk	Allocation done by someone who was not otherwise involved in the study
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Due the nature of the interventions, patients were not blinded; 2. Interventions required training and supervision; both groups received active treatment.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Both groups received active treatment; protocol was explicit.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Function was self-reported so the assessor could not be blinded; 2. Function was self-reported so the assessor could not be blinded; 3. Both groups received active credible treatment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 16 of 106 missing
Participants analysed in group allocated (attrition bias)	Low risk	1. All participants analysed

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saner 2016 *(Continued)*

Selective reporting (reporting bias)	High risk	1. Protocol published elsewhere; chronic pain grade scale planned but not reported; 2. No reason not to report pain scale, presented as baseline characteristic; 3. Analytic method simple
Groups similar at baseline (selection bias)	Low risk	Similar on age, gender, height, weight, sport participation, work status, and pain duration
Co-interventions avoided or similar (performance bias)	Low risk	Co-intervention monitored and minimal
Compliance acceptable in all groups (performance bias)	High risk	Compliance not reported
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcomes measured after treatment, though no specifics on timing or method included; functional scale consistent across participants; 2. Roland-Morris Disability Questionnaire valid and reliable
Other bias	Low risk	Appeared free from other sources of bias

Saper 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 30 (E1 = 15, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 44 **Sex (female):** 83%
Interventions	**Exercise Group 1 (E1):** Yoga postures and breathing techniques and home practice; type = yoga; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education & relaxation **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Pain Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global improvement (7-point)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Center for Complementary and Alternative Medicine, National Institutes of Health **Other:** Information modified for author contact
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saper 2009 *(Continued)*

Random sequence generation (selection bias)	Low risk	Computer-generated permuted block randomisation
Allocation concealment (selection bias)	Low risk	Treatment assignments were placed in opaque, sequentially numbered envelopes prepared by a biostatistician who had no contact with participants.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants and yoga teachers could not be blinded to treatment allocation.
Blinding of care provider (performance bias)	High risk	Participants and yoga teachers could not be blinded to treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	All study participants met in person with unblinded research staff members to complete paper questionnaires at baseline, six and 12 weeks.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Participant retention was 97% at 12 weeks and 77% at 26 weeks.
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No statistically significant differences between groups at baseline were observed.
Co-interventions avoided or similar (performance bias)	Low risk	Both groups were discouraged from starting any new back pain treatments during the study.
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: they say they assessed compliance, not an issue
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Saper 2013

Study characteristics	
Methods	**Study design:** RCT (NCT01761617) **Setting:** USA, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 95 (E1 = 46, E2 = 49) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saper 2013 *(Continued)*

	Mean age (years): 48 **Sex (female):** 75%
Interventions	**Exercise Group 1 (E1):** Two yoga classes per week and home practice; type = yoga; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = relaxation **Exercise Group 2 (E2):** One yoga class per week and home practice; type = yoga; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = relaxation
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (overall improvement (7-point)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Center for Complementary and Alternative Medicine (Grant 1R01AT005956), National Institutes of Health **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The 1:1 randomisation schedule was created in StudyTrax using a permuted block design with randomly determined block sizes (4, 8, 12).
Allocation concealment (selection bias)	High risk	Due to the nature of the interventions, participants and study staff who scheduled classes could not be masked to treatment allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Due to the nature of the interventions, participants and study staff who scheduled classes could not be masked to treatment allocation.
Blinding of care provider (performance bias)	High risk	Due to the nature of the interventions, participants and study staff who scheduled classes could not be masked to treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	See flow chart
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	A couple of characteristics were different but they were used as confounders.
Co-interventions avoided or similar (performance bias)	Low risk	Reported and seem to be similar across the groups

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saper 2013 *(Continued)*

Compliance acceptable in all groups (performance bias)	High risk	Thirty-two (65%) and 20 (44%) participants assigned to once-weekly and twice-weekly classes, respectively, achieved treatment adherence.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Saper 2017

Study characteristics	
Methods	**Study design:** RCT (NCT01343927) **Setting:** USA, mixed **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 320 (E1 = 127, E2 = 129, C1 = 64) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 46 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Yoga: relaxation and meditation exercises, yoga breathing, and yoga philosophy (30 minutes of daily home practice, facilitated by a video, a manual, and take-home yoga supplies); type = yoga; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Physiotherapy: treatment-based classification, graded exercise, and screening for fear-avoidance beliefs. Exercise involved supervised aerobic exercises For patients with high fear-avoidance beliefs, a therapist provided a copy of a Back Book and counselled participants to reduce their fear. Participants received written instructions and supplies for home practice and logged the number of exercises completed daily; type = core strengthening & aerobic; duration = 12 weeks; dose = low; design = individualised; delivery = individual; additional intervention = psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numerical Rating Scale); function (Roland-Morris Disability Questionnaire); work (work productivity); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Rating of Change Scale)) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Dr. Saper reported grants from the National Center for Complementary and Integrative Health of the National Institutes of Health during the conduct of the study. Dr. Sherman reported grants from National Institutes of Health during the conduct of the study. Dr. Herman reported grants from the National Center for Complementary and Integrative Health during the conduct of the study. Dr. Weinberg reported grants from the National Institutes of Health during the conduct of the study. Authors not named here have disclosed no conflicts of interest. **Funding source:** National Center for Complementary and Integrative Health, National Institutes of Health (5R01-AT005956) **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saper 2017 *(Continued)*

Random sequence generation (selection bias)	Low risk	Computer-generated randomisation sequence using permuted block randomisation with varying block sizes and a 2:2:1 ratio of yoga, physiotherapy, and education
Allocation concealment (selection bias)	Low risk	Randomisation all done by computer program; nothing calculated by hand; study staff did find out before participants (allocation was not revealed to participant and study staff at the same time).
Blinding of participants and personnel (performance bias) All outcomes	High risk	1. Patients could not be blinded to intervention due to nature of treatments; 2. Yoga protocol was defined, it was not unlike the sequence one would receive at an external class; accessible outside of study; 3. No information on any deviations; 4. Seeking the treatment could reduce low back pain.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of provider blinding led to deviations from intended intervention; different providers delivered each treatment, i.e. yoga instructors only saw yoga participants.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For outcomes of interest (pain and disability), outcome assessors were patients, who could not be blinded to intervention; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Particularly for control group participants, lack of blinding likely led to them feeling underserved by their treatment allocation, and were biased in a perception of lack of improvement in outcomes.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. At 12 weeks follow-up was 88% versus 98% versus 95%.
Participants analysed in group allocated (attrition bias)	Low risk	1. Intention-to-treat analyses of the full study population performed
Selective reporting (reporting bias)	Low risk	1. Published protocol (Saper 2014) and registered (NCT01343927): all analysis conducted and reported in these documents; work productivity not reported here but noted it would be in separate publication
Groups similar at baseline (selection bias)	Low risk	All group characteristics were similar at baseline; duration of symptoms was not measured or reported.
Co-interventions avoided or similar (performance bias)	High risk	Medication use during treatment period was reported; education group had slightly higher use of all medications, though it differed by subtype of medication; no other co-interventions discussed
Compliance acceptable in all groups (performance bias)	High risk	Compliance for yoga group: median 7 out of 12 classes attended (58%); compliance for physiotherapy group: median 7 out of 15 appointments (47%); compliance for education: 44% read at least 75% of education book.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcomes were assessed at the same time point; Within each outcome, all participants in all groups were measured using the same tools, etc.; 2. Roland-Morris Disability Questionnaire (for disability) and Numeric Rating Scale (for pain) are well-validated tools in the context of low back pain.
Other bias	Low risk	Powered for non-inferiority; had to adjust for baseline Roland-Morris Disability Questionnaire

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Schenkman 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, mixed **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 61 (E1 = 20, C1 = 20, C2 = not reported) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Functional training (core stability) during performance of actual daily activities (e.g., vacuuming, sweeping, lifting household items), flexibility and core strengthening exercises; type = core strengthening & flexibility/mobilising; duration = not specified weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (education) **Comparison Group 2 (C2):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Foundation for Physical Therapy **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Group assignment was made after the baseline testing and was determined by computer-generated allocation.
Allocation concealment (selection bias)	Low risk	Group assignment was made after the baseline testing and was determined by computer-generated allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Because this was an exercise study, it was not possible to blind subjects or physical therapists to the group assignments.
Blinding of care provider (performance bias)	High risk	Because this was an exercise study, it was not possible to blind subjects or physical therapists to the group assignments.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Assessors of outcome measures were blinded to the group assignment and conducted measures at a separate location from the interventions.
Incomplete outcome data (attrition bias) All outcomes	High risk	Six-seven per cent of the participants provided at least some data on completion of the first two months of the study, and 44% provided data for the entire 12 months.
Participants analysed in group allocated (attrition bias)	High risk	Not included

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Schenkman 2009 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant between-group differences were found at baseline for gender, age, height, weight, pain or physical functional capacity.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Incorporated strategies to enhance adherence and refresher sessions at six and 12 months
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Schinhan 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Austria, not specified **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 30 (E1 = 15, C1 = 15) **Chronic LBP duration:** 67.365 months (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 28 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Climbing: participants climbed 5 bouldering routes in an indoor climbing centre with instruction on proper body position and movement; type = other (rock climbing); duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: instructed not to change lifestyle, allowed to take paracetamol 500 mg 4 times daily if needed and not allowed to go rock climbing)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 14 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated randomisation system, provided by the Institute of Medical Statistics
Allocation concealment (selection bias)	Low risk	Randomisation system was provided by third party (Institute of Medical Statistics).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Schinhan 2016 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded due to the nature of the interventions; 2. Unlikely that lack of participant blinding caused deviations from protocol; participants in climbing group needed instructor, and in control group were allowed to continue their routines without climbing
Blinding of care provider (performance bias)	Low risk	1. Only care providers involved were the climbing instructors, who could not be blinded due to their involvement in delivering the intervention; 2. Lack of care provider blinding was unlikely to lead to deviations from the intended intervention because the only care provider involved was the climbing instructor.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. For the outcomes of interest (pain and disability), participants were the outcomes assessors; they could not be blinded due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No dropouts in this study
Participants analysed in group allocated (attrition bias)	Low risk	1. All participants were analysed according to their allocated treatment group.
Selective reporting (reporting bias)	Low risk	1. No linked protocol found: within this paper no obvious omissions of outcomes or analyses
Groups similar at baseline (selection bias)	Low risk	All relevant characteristics were measured and were similar across groups at baseline.
Co-interventions avoided or similar (performance bias)	Low risk	Control group allowed to take paracetamol if needed, asked to note this; they were asked to not go rock climbing; all patients allowed to continue any of their routine physical activity.
Compliance acceptable in all groups (performance bias)	Low risk	Compliance appeared to have been 100% (only applicable for rock climbing group).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. All outcomes were measured at the same time point for all participants; within each outcome, all participants were measured using the same tool; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Segal-Snir 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** Israel, healthcare **Exercise groups:** 1 **Comparison groups:** 1

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Segal-Snir 2016 *(Continued)*

Participants	**Number of participants:** 45 (E1 = 25, C1 = 20) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 56 **Sex (female):** 100%
Interventions	**Exercise Group 1 (E1):** Rotation exercises were conducted in 5 different weight bearing and non-weight bearing positions including standing, sitting, supine/crook lying, side-lying (left and right) and prone; type = stretching; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); work (number of hours worked during study period) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Article stated that the process was random, but did not explicitly report the method used; allocations were delivered in sealed envelopes.
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Impossible to blind patient to the allocation, since the interventions were so different (exercise versus guidance only); 2. Patient blinding likely did not cause deviations from intervention protocol because exercise group received control treatment plus exercises, and control group could not do exercises without physiotherapist.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Therapists had no contact with control; lack of blinding for care providers unlikely to cause deviations because control patients were wait-listed
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcomes in question were pain and function which were assessed by the non-blinded patients; 2. Treatment groups were very different, so patients may have had different expectations of treatment group effectiveness when assessing self-reported outcomes; 3. Outcomes in an exercise versus no treatment study likely to be altered by knowledge of intervention assignment
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. Unacceptable dropout rate (more than 20%); treatment group: 20%, control group: 25%; 2. Study dropped missing data since it occurred before intervention; 3. All control group dropouts were logistic in nature, while 4 of intervention group dropouts were medical including severe low back pain attacks; 4. 5 from each group
Participants analysed in group allocated (attrition bias)	Low risk	1. Paper cited that intention-to-treat was not done, however it seemed it was done just with the exclusion of missing data. Included patients were analysed in arm they were randomised to.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Segal-Snir 2016 *(Continued)*

Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication all outcomes and analyses fully reported; no obvious omissions
Groups similar at baseline (selection bias)	Low risk	No statistically significant differences between treatment groups on baseline characteristics except lumbar extension
Co-interventions avoided or similar (performance bias)	Low risk	Patients asked to record co-interventions for treatment period; five intervention and six control patients used medications; two intervention and six control did physical activity; three each worked.
Compliance acceptable in all groups (performance bias)	Low risk	100% attendance reported for both treatment groups
Timing of outcome assessment similar in all groups (detection bias)	High risk	1. All outcomes assessed the same way for all participants except for experimental group outcome measures were compared among time points with ANOVA, and control group used paired t-test; 2. Used well known scales, no reason to think outcome measures were invalid or unsuitable
Other bias	Low risk	Inadequate follow-up time but no other apparent sources of bias

Sertpoyraz 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 40 (E1 = 20, E2 = 20) **Chronic LBP duration:** 42.5 months (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 39 **Sex (female):** 77%
Interventions	**Exercise Group 1 (E1):** Warm-up (10-minute walk) and trunk flexion and extension exercise using Cybexercise Norm Dynamometer machines; type = strengthening & aerobic; duration = 3 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Warm-up (10-minute walk) and passive lumbar flexion and extension, spine stretching and mobilisation exercises; type = mixed; duration = 3 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short); 7 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sertpoyraz 2009 *(Continued)*

Random sequence generation (selection bias)	Low risk	The randomisation list was generated by a blinded researcher (the fourth author, experienced in biostatistics) using a table of random numbers.
Allocation concealment (selection bias)	Low risk	The randomisation results were kept in sealed envelopes, one for each patient.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The isokinetic and standard exercise groups showed no difference in demographic and clinical characteristics (Table 1).
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Exercise programmes were carried out under supervision, which resulted in higher patient compliance, attendance at exercise programmes and patient satisfaction.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Shankar 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 60 (E1 = 30, C1 = 30) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shankar 2011 *(Continued)*

	Mean age (years): 35 **Sex (female):** 66%
Interventions	**Exercise Group 1 (E1):** Strengthening exercises and lumbar extension training to improve strength of low back, partial sit-ups and pelvic tilt; type = core strengthening; duration = 3 weeks; dose = high; design = not specified; delivery = not specified; additional intervention = anti-inflammatory/analgesics **Comparison Group 1 (C1):** Other conservative treatment (electrotherapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 3 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	To remove bias, with the help of a computer-generated randomisation list, the patients were divided in two study groups of 30 patients each.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There was no significant difference in mean age and sex ratio between the various groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shankar 2011 *(Continued)*

Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Shaughnessy 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** Ireland, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 41 (E1 = 20, C1 = 21) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 66%
Interventions	**Exercise Group 1 (E1):** Core stability muscles training to activate transversus abdominus and multifidus muscles using biofeedback and home exercises; type = core strengthening; duration = 10 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 10 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Subjects were randomly assigned to either a treatment or control group.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias)	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shaughnessy 2004 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	Initially, 23 subjects (seven male, 16 female) were assigned to the treatment group and 22 subjects (eight male, 14 female) to the control group.
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between the groups mean baseline Oswestry Disability Questionnaire, Roland Morris Disability Questionnaire, or 36-Item Short Form Health Survey scores (P > 0.05).
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Shen 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** China, healthcare **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 45 (E1 = 15, E2 = 15, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 46 **Sex (female):** 0%
Interventions	**Exercise Group 1 (E1):** McKenzie therapy; type = McKenzie; duration = 4 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education & electrotherapy **Exercise Group 2 (E2):** McKenzie therapy; type = McKenzie; duration = 4 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education & electrotherapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Japanese Orthopedic Association Score) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shen 2009 *(Continued)*

Other: SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Sealed, opaque envelopes
Allocation concealment (selection bias)	Low risk	Sealed, opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Sherman 2005

Study characteristics

Methods	**Study design:** RCT (NCT00056212) **Setting:** USA, mixed

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sherman 2005 *(Continued)*

	Exercise groups: 2 **Comparison groups:** 1
Participants	**Number of participants:** 101 (E1 = 36, E2 = 35, C1 = 30) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 44 **Sex (female):** 66%
Interventions	**Exercise Group 1 (E1):** Viniyoga: emphasis on safety with focus on relaxation, strength-building, flexibility, and strengthening hip muscles, asymmetric poses; type = yoga; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = relaxation **Exercise Group 2 (E2):** Series of 7 aerobic exercises and 10 strengthening exercises emphasizing leg, hip, abdominal, and back muscles, stretching; type = mixed; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education & relaxation **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (bothersomeness); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** National Center for Complementary and Alternative Medicine (grant R21AT 001215); National Institute for Arthritis and Musculoskeletal and Skin Diseases (grant P60AR48093) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomly generated treatment assignments for each class series by using a computer program with block sizes of six or nine
Allocation concealment (selection bias)	Low risk	A researcher who was not involved in patient recruitment or randomisation placed the assignments in opaque, sequentially numbered envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Interviewers who were masked to the treatment assignments conducted telephone interviews at baseline and at six, 12, and 26 weeks after randomisation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Follow-up rates remained high, even at 26 weeks (95 of 101 participants completed telephone interviews).
Participants analysed in group allocated (attrition bias)	Low risk	An intention-to-treat approach was used.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sherman 2005 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	No significant differences were found between groups except for pain travelling below knee.
Co-interventions avoided or similar (performance bias)	Low risk	During the 12-week intervention, 11% of participants in the yoga group reported making visits to healthcare providers.
Compliance acceptable in all groups (performance bias)	Low risk	Class attendance was similar in the yoga (median classes attended, 9) and exercise (median classes attended, 8) groups (Figure 1).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Sherman 2011

Study characteristics

Methods	**Study design:** RCT (NCT00447668) **Setting:** USA, mixed **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 228 (E1 = 92, E2 = 91, C1 = 45) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 48 **Sex (female):** 64%
Interventions	**Exercise Group 1 (E1):** Viniyoga: 17 relatively simple postures with variations and adaptations, classes include breathing exercises, 5-11 postures, and guided deep relaxation; type = yoga; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Stretching classes consisting of 15 exercise designed to stretch major muscle groups (52 minutes of stretching) and 4 strengthening exercises; type = stretching & strengthening; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (bothersomeness); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global rating of improvement)) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Center for Complementary and Alternative Medicine (Cooperative Agreement Number U01 AT003208 M) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	After completing the baseline interview at Group Health facilities, participants within each recruitment cohort were randomised by a research assistant.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sherman 2011 *(Continued)*

Allocation concealment (selection bias)	Low risk	After completing the baseline interview at Group Health facilities, participants within each recruitment cohort were randomised by a research assistant.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Telephone interviews were conducted by masked interviewers at baseline and at six, 12, and 26 weeks after randomisation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Overall follow-up rates were 90% or 91% at all time points.
Participants analysed in group allocated (attrition bias)	Low risk	All analyses were conducted assuming intention-to-treat principles using SAS statistical software (version 9 2; SAS Institute Inc, Cary, North Carolina).
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics were well balanced across groups, except the yoga group had greater back dysfunction (Table 1).
Co-interventions avoided or similar (performance bias)	Low risk	Compared with baseline, roughly a quarter to a third fewer participants in the yoga and stretching groups reported using any medications for back pain.
Compliance acceptable in all groups (performance bias)	Low risk	Participants randomised to yoga were more likely than those assigned to stretching to attend at least one class (95% vs 82%, respectively) (Figure 1).
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Shirado 2010

Study characteristics	
Methods	**Study design:** RCT **Setting:** Japan, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 201 (E1 = 103, C1 = 98) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** 55%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shirado 2010 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Trunk strengthening (trunk flexor and extension exercises), stretching (abdominal and back muscles, iliopsoas, gluteals and hamstrings), physician visits to ensure compliance; type = strengthening & stretching; duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Comparison Group 1 (C1):** Other conservative treatment (anti-inflammatory/analgesics)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The allocation sequence was produced by a computer-generated, permuted-block randomisation with a block size of four.
Allocation concealment (selection bias)	Unclear risk	The office managed the patient registration including confirmation of the eligibility criteria, randomisation to an allocated treatment, and data management.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The outcome measures were conducted in a blinded fashion by a person who was not involved in this trial.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Among all eligible participants, two in exercise and six in Non-steroidal anti-inflammatory drugs dropped out during the follow-up period.
Participants analysed in group allocated (attrition bias)	Unclear risk	Analysis was by intention-to-treat.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no statistically significant differences in age, gender, body height, body weight, body-mass index, Visual Analogue Scale, Roland-Morris Disability Questionnaire, Japan Low back pain Evaluation Questionnaire, and finger-floor distance.
Co-interventions avoided or similar (performance bias)	Low risk	A run-in period of two weeks was designed to ensure the stability of patient low back pain before starting the intervention.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shirado 2010 *(Continued)*

Compliance acceptable in all groups (performance bias)	Low risk	The follow-up system for participants was also distinctive compared with other studies with a treatment diary used to record daily exercise and medication.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Shnayderman 2013

Study characteristics	
Methods	**Study design:** RCT **Setting:** Israel, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 52 (E1 = 26, E2 = 26) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 45 **Sex (female):** 79%
Interventions	**Exercise Group 1 (E1):** Active movements and strengthening exercises for the trunk and upper and lower limbs; type = strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = not specified **Exercise Group 2 (E2):** Walking (treadmill); type = aerobic; duration = 6 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = not specified
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Block randomisation (four cells in each block) stratified by age groups (18–44 years, 45–65 years) were set up.
Allocation concealment (selection bias)	Low risk	Assignment was concealed in sealed envelopes and was revealed by an independent researcher not involved in the assessments or intervention.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shnayderman 2013 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Stated in paper that assessors were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Six patients from the study group and three from the control group withdrew from the study (Figure 1).
Participants analysed in group allocated (attrition bias)	Low risk	All analyses were based on an intention-to-treat basis.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no statistically significant baseline differences in demographic or clinical parameters between the study groups.
Co-interventions avoided or similar (performance bias)	Unclear risk	Author contact: not measured
Compliance acceptable in all groups (performance bias)	Low risk	No difference in average number of visits between groups
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Sjogren 2006a

Study characteristics	
Methods	**Study design:** Cluster-RCT (cross-over) **Setting:** Finland, occupational **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 36 (E1 = 21, C1 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** 80%
Interventions	**Exercise Group 1 (E1):** Light resistance training: 6 dynamic symmetrical movements, extension and flexion of upper extremities, trunk and knees; type = strengthening; duration = 15 weeks; dose = high; design = partially individualised; delivery = independent with follow-up; additional intervention = advice/education **Comparison Group 1 (C1):** Usual care/no treatment (control group: no intervention)
Outcomes	**Core outcomes reported:** Function (Support Needs Questionnaire) **Follow-up time periods available for syntheses:** 15 weeks (moderate)
Notes	**Conflicts of interest:** Not reported

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sjogren 2006a *(Continued)*

Funding source: Chydenius Institute; University of Jyvaskyla; Palokka health centre; Finnish Work Environment Fund; Juho Vainio Foundation
Other: Sufficient data not available for inclusion in meta-analyses

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	After the baseline measurements the blinded measurers allocated the workers in each department into the two treatment sequence groups.
Allocation concealment (selection bias)	Low risk	The sequence was concealed from the participants for as long as possible.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	In the Low Back Symptoms Group, three subjects were lost to follow-up during the physical exercise intervention.
Participants analysed in group allocated (attrition bias)	Low risk	In the Low Back Symptoms Group, three subjects were lost to follow-up during the physical exercise intervention.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Unclear risk	Not described
Co-interventions avoided or similar (performance bias)	Low risk	Excluding the light resistance training, the participants were asked to keep the level of intensity and amount of their physical activity unchanged.
Compliance acceptable in all groups (performance bias)	Low risk	Training adherence (69%) was a percentage ratio of self-reported training time and guided target training time.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Smeets 2008

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Smeets 2008 *(Continued)*

Methods	**Study design:** RCT (ISRCTN22714229) **Setting:** Netherlands, healthcare **Exercise groups:** 3 **Comparison groups:** 1
Participants	**Number of participants:** 223 (E1 = 61, E2 = 53, E3 = 58, C1 = 51) **Chronic LBP duration:** 60 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Aerobic training (30 minutes on bike) and strength training (75 minutes) for back and leg muscles and graded activity; type = other (graded activity) & mixed; duration = 10 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy **Exercise Group 2 (E2):** Aerobic training (30 minutes on bike) and strength training (75 minutes) for back and leg muscles; type = aerobic & strengthening; duration = 10 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Exercise Group 3 (E3):** Graded activity: individualised assessment and treatment goal, selected activities done toward final treatment goals; type = other (graded activity with problem-solving training (gap)); duration = 10 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); Global Perceived Health or Recovery (Global Perceived Health or Recovery (perceived improvement (7-point)) **Follow-up time periods available for syntheses:** 10 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Zorgonderzoek Nederland/Medische Wetenschappen (Grant No. 014-32-007) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Each cluster of four consecutive patients was assigned to one of the three active treatments or a waiting list, using block randomisation.
Allocation concealment (selection bias)	Low risk	Opaque, sequentially numbered, sealed envelopes were prepared for each rehabilitation centre before enrolment started.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Patients and therapists were not blinded to treatment allocation.
Blinding of care provider (performance bias)	High risk	Patients and therapists were not blinded to treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The research assistants collecting data were blinded to treatment allocation.
Incomplete outcome data (attrition bias)	Low risk	Follow-up rates for the questionnaires remained high, even at 12 months (156 of 172 patients).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Smeets 2008 *(Continued)*

All outcomes

Participants analysed in group allocated (attrition bias)	Low risk	All statistical analyses were performed according to the intention-to-treat principle.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline status of patients, including distribution of baseline Roland Morris Disability Questionnaire score and the level of treatment expectancy and credibility, did not differ significantly.
Co-interventions avoided or similar (performance bias)	Low risk	Patients were allowed to continue medication prescribed at baseline, but other co-interventions were discouraged.
Compliance acceptable in all groups (performance bias)	Low risk	Compliance was well measured. There was at least 72% compliance for each group.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Smith 2011

Study characteristics	
Methods	**Study design:** RCT **Setting:** England, healthcare **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 46 (E1 = 16, E2 = 17, E3 = 13) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 43 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Lumbar extension training: one set of approximately 8-12 repetitions through the participant's full range of motion with pelvic stabilisation; type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = none **Exercise Group 2 (E2):** Lumbar extension training: 1 set of approximately 8-12 repetitions through the participant's full range of motion without pelvic stabilisation; type = strengthening; duration = 12 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = not specified **Exercise Group 3 (E3):** McKenzie protocol, muscle imbalance protocol, home exercises; type = mixed; duration = 12 weeks; dose = low; design = standardised; delivery = individual; additional intervention = advice/education & manual therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Smith 2011 *(Continued)*

Other: Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants were randomly allocated to one of three groups.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Author contact: outcome assessor was not blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Four dropouts from 46
Participants analysed in group allocated (attrition bias)	High risk	Not described
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Author contact: no significant differences
Co-interventions avoided or similar (performance bias)	Unclear risk	Author contact: people continued physical therapy.
Compliance acceptable in all groups (performance bias)	Low risk	Author contact: participants only missed one or two sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Sorensen 2010

Study characteristics

Methods	**Study design:** RCT (NCT00410319) **Setting:** Denmark, healthcare

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sorensen 2010 *(Continued)*

	Exercise groups: 1 **Comparison groups:** 1
Participants	**Number of participants:** 207 (E1 = 102, C1 = 105) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 39 **Sex (female):** 52%
Interventions	**Exercise Group 1 (E1):** Symptom-based physical training including directional preferences and posture exercises, back stabilisation and strengthening exercises and home exercise; type = core strengthening; duration = not specified weeks; dose = low; design = individualised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Low Back Pain Rating Scale); work (work ability, an 11-item scale on the patients work situation); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global quality of life due to treatment (5-point)) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Institut für Makroökonomie und Konjunkturforschung Foundation; Health Insurance Foundation (Sygekassernes Helsefond); Tryg Foundationen; Funen County Research Foundation; Danish Rheumatism Association **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A secretary managed the randomisation, using unmarked sealed envelopes, containing a note on which was randomly written one of the group names.
Allocation concealment (selection bias)	Unclear risk	A secretary managed the randomisation, using unmarked sealed envelopes, containing a note on which was randomly written one of the group names.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The same investigator (PHS) managed the baseline examination and controlled the follow-up forms, blinded to the treatment group.
Incomplete outcome data (attrition bias) All outcomes	Low risk	The losses to follow-up were all due to non-attendance, even after a second written invitation.
Participants analysed in group allocated (attrition bias)	High risk	Not included

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sorensen 2010 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Participants in both groups (n = 105 and 102) were comparable at baseline, as shown in Table 1 and 2.
Co-interventions avoided or similar (performance bias)	Low risk	There was no statistically significant difference (P = 0.65 to 0.87) between the two groups.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	High risk	Support for judgement was not available.

Soukup 1999

Study characteristics

Methods	**Study design:** RCT **Setting:** Norway, mixed **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 120 (E1 = 39, E2 = 43, C1 = 38) **Chronic LBP duration:** 600 weeks (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 40 **Sex (female):** 54%
Interventions	**Exercise Group 1 (E1):** Mensendiek exercise: aerobic, stretching, strength, co-ordination; type = mixed; duration = 13 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Thirty-minute exercise session; type = mixed; duration = 13 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = back school **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 22 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Device status category: 1 **Funding source:** Norwegian Fund for Post Graduate Training in Physiotherapy; The Royal Norwegian Ministry of Health and Social Affairs **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Soukup 1999 *(Continued)*

Allocation concealment (selection bias)	Low risk	Author contact: changed from unclear
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Unclear risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Soundararajan 2016

Study characteristics	
Methods	**Study design:** RCT **Setting:** India, occupational **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 30 (E1 = 15, E2 = 15) **Chronic LBP duration:** 3.77 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 27 **Sex (female):** 40%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Soundararajan 2016 *(Continued)*

Interventions	**Exercise Group 1 (E1):** Multifidus retraining programme: deep spine stabiliser; type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none **Exercise Group 2 (E2):** Traditional back exercises to strengthen back and core; type = core strengthening; duration = 6 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Purposive random sampling with 15 subjects (9 males and 6 females) in each group
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Both groups were receiving exercise interventions, unlikely to deviate from intervention
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of care provider blinding caused deviations from intended interventions
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Outcome assessors for pain and function were patients themselves, who could not be blinded to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Both groups were receiving exercise programmes; unlikely for patient response to be significantly altered
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. It was highly likely that the dropout rate in this case was zero because the trial was conducted in a specific workplace, encouraging participation and follow-up.
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. Study was analysed and reported as described in the methods section.
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics.
Co-interventions avoided or similar (performance bias)	High risk	No information about co-interventions and whether or not they were comparable

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Soundararajan 2016 *(Continued)*

Compliance acceptable in all groups (performance bias)	High risk	No explicit reference in the study to compliance/adherence/attendance, but because this study was run within a workplace, it was likely quite high
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Sousa 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** Brazil, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 60 (E1 = 30, C1 = 30) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 46 **Sex (female):** 72%
Interventions	**Exercise Group 1 (E1):** Global muscular relaxation, exercises to strengthen the abdominal muscles with and without biofeedback; type = strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = group; additional intervention = psychological therapy & anti-inflammatory/analgesics **Comparison Group 1 (C1):** Other conservative treatment (anti-inflammatory/analgesics)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Patients were randomised in either the control group or the treatment groups.
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sousa 2009 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	The instruments used for assessment were a Visual Analogue Scale, Schober Index, Roland Morris Disability Questionnaire, Becks Depression Scale.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Seven patients were excluded as a result of being absent during the two-month period, and one from the treatment group due to reported symptoms.
Participants analysed in group allocated (attrition bias)	High risk	Not included
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The groups were homogenous in relation to demographic variables (Table 1).
Co-interventions avoided or similar (performance bias)	Unclear risk	Both groups were instructed to use paracetamol, 500 mg, every six hours, if necessary, when feeling pain.
Compliance acceptable in all groups (performance bias)	Low risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Steele 2013

Study characteristics	
Methods	**Study design:** RCT **Setting:** United Kingdom, mixed **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 24 (E1 = 10, E2 = 7, C1 = 7) **Chronic LBP duration:** 12.9 years (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 44 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Lumbar extension resistance exercise on MEDX training system using full range of motion; type = strengthening; duration = 12 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = not specified **Exercise Group 2 (E2):** Lumbar extension resistance exercise on MEDX training system using 50% of range of motion; type = strengthening; duration = 12 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = not specified **Comparison Group 1 (C1):** Usual care/no treatment (control group: continued with normal activity and treatment)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Steele 2013 *(Continued)*

Notes	**Conflicts of interest:** Expert testimony, grants, payment for lecture **Funding source:** No funding received **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Thirty-one participants were randomised using an online randomisation program (Research Randomizer vs 3.0).
Allocation concealment (selection bias)	Unclear risk	Author contact: participants recruited until three available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Flow chart
Participants analysed in group allocated (attrition bias)	High risk	Not included
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Author contact: stated no significant difference at baseline
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: no new interventions; one participant confirmed physio; excluded from analysis
Compliance acceptable in all groups (performance bias)	Unclear risk	Author contact: only recorded for lumbar extension group. Attendance between training groups for lumbar extension training sessions also did not significantly differ.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Stephan 2011

Study characteristics

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Stephan 2011 *(Continued)*

Methods	**Study design:** RCT **Setting:** Germany, general population **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 74 (E1 = 58, C1 = 16) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 44 **Sex (female):** 55%
Interventions	**Exercise Group 1 (E1):** Progressive hypertrophy-oriented strength training on machines with variable resistance, lumbar extensor training with the pelvis stabilised for all muscle groups; type = strengthening; duration = 24 weeks; dose = high; design = partially individualised; delivery = independent with follow-up; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (MOS Pain Severity); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 26 weeks (moderate)
Notes	**Conflicts of interest:** Anika Stephan and Dr. Sven Goebel work for the Research and Development Department at Kieser Training AG. Prof. Dr. Dietmar Schmidtbleicher has been remunerated for his consultancy by Kieser Training AG. **Funding source:** Not reported **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Twenty-two of the 80 participants in the training group dropped out.
Participants analysed in group allocated (attrition bias)	High risk	Fifty-six/80 analysed only
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Stephan 2011 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	Table 1 described the sample and the characteristics of the sample group at the start of the intervention (no significant differences).
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Participants dropped out if they didn't have good compliance; others not discussed
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Sung 2013

Study characteristics	
Methods	**Study design:** RCT **Setting:** South Korea, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 46 (E1 = 25, E2 = 21) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 50 **Sex (female):** 48%
Interventions	**Exercise Group 1 (E1):** Five core stabilisation exercises; type = core strengthening; duration = 4 weeks; dose = low; design = partially individualised; delivery = independent with follow-up; additional intervention = none **Exercise Group 2 (E2):** Five spinal flexibility exercises; type = flexibility/mobilising; duration = 4 weeks; dose = low; design = partially individualised; delivery = independent with follow-up; additional intervention = none
Outcomes	**Core outcomes reported:** Function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Cleveland State University; Korea University and the Basic Science Research Program, National Research Foundation of Korea, Ministry of Education, Science and Technology (2010-0003015) **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The randomisation procedure was conducted by a computer program.
Allocation concealment (selection bias)	Unclear risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sung 2013 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Figure 3
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Overall, there were no differences between groups in age, height, body weight or the months since pain onset.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Adherence was mentioned as being measured but not reported on.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Tekur 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** India, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 80 (E1 = 40, E2 = 40) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 49 **Sex (female):** 45%
Interventions	**Exercise Group 1 (E1):** Walking 30 minutes, exercises throughout the day, breathing practice; type = aerobic; duration = 1 weeks; dose = high; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Tekur 2012 *(Continued)*

	Exercise Group 2 (E2): Special yoga (asana) techniques for back pain progressing slowly over 3 days and Pranayama (breathing techniques); type = yoga; duration = 1 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education & psychological therapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale) **Follow-up time periods available for syntheses:** 1 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Swami Vivekananda Yoga Research Foundation **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation: used two sets of 40 numbers spanning integers 1-80 created by a random number table
Allocation concealment (selection bias)	Low risk	Numbered containers were used to conceal the random allocation before implementation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	In intervention studies of this kind, subjects clearly identify their own treatment: double blinding was not possible.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Statistician who developed randomisation sequence, clinical psychologist, and researcher who carried out assessments were all blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts
Participants analysed in group allocated (attrition bias)	Low risk	No dropouts
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The two groups were similar with respect to sociodemographic and medical characteristics (Table 3).
Co-interventions avoided or similar (performance bias)	Low risk	Residential treatment
Compliance acceptable in all groups (performance bias)	Low risk	Residential treatment
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Teut 2016

Study characteristics	
Methods	**Study design:** RCT (NCT01303588) **Setting:** Germany, general population **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 176 (E1 = 61, E2 = 58, C1 = 57) **Chronic LBP duration:** 18.79 years (long) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 73 **Sex (female):** 89%
Interventions	**Exercise Group 1 (E1):** Yoga; type = yoga; duration = 12 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Qigong; type = other (Qigong); duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = manual therapy **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Hannover Functional Ability Questionnaire); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 13 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** J.K. is a Viniyoga teacher trained by the Berliner Yoga Zentrum. **Funding source:** Karl and Veronica Carstens Foundation **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation carried out with the 'ranuni' function of SAS software with a 1:1:1 ratio, stratified by participant housing (living in retirement home versus alone) and blocked with a fixed, unknown length.
Allocation concealment (selection bias)	Low risk	Randomisation list was held in a secure database, inaccessible to anyone involved in randomisation or treatment; allocation could not be changed or deleted, ensuring allocation concealment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that the lack of patient blinding caused deviations from the intended interventions, as patients had very little control over delivery of interventions
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of provider blinding caused significant deviations; each treatment group highly structured with distinct nature, yoga and qigong delivered by separate trained personnel
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were patients themselves, who could not be blinded to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Likely that lack of patient blinding caused bias in outcome assessments, as the Qigong and yoga groups were clearly "better" than the control; few clear results in support

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Teut 2016 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	1. At the 3-month follow-up, 12 participants were lost to follow-up of 176.
Participants analysed in group allocated (attrition bias)	Low risk	1. "Outcomes were analyzed for the full analysis set, on the basis of the intention-to-treat principle".
Selective reporting (reporting bias)	High risk	1. Study was fully analysed and reported according to the registered trial protocol (NCT01303588), except that there was no Tinetti test results at 6 months nor adverse events reported; 2. Particularly for adverse events, it is suspicious that they were not reported and it was reasonably likely that they were not reported due to their unflattering results 3. There was no apparent reporting bias for the analytic method used in this study.
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics.
Co-interventions avoided or similar (performance bias)	Low risk	Excluded patients using central nervous system pain agents (e.g. opioids) or with preplanned start of physiotherapy within the study duration; there was no difference in pain medication use during study period.
Compliance acceptable in all groups (performance bias)	Low risk	Participated in 75% or more of classes = 74.1% (yoga), 72.7% (qigong) of patients; participated in 50 to 75% of classes = 12.9% (yoga), 18.2% (qigong) of patients
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Functional Rating Index (for pain) and the Hannover Functional Ability Questionnaire (for function) are well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Tilbrook 2011

Study characteristics	
Methods	**Study design:** RCT (ISRCTN81079604) **Setting:** United Kingdom, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 313 (E1 = 156, C1 = 157) **Chronic LBP duration:** 27 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 46 **Sex (female):** 71%
Interventions	**Exercise Group 1 (E1):** Yoga programme adapted for low back pain; type = yoga; duration = 12 weeks; dose = low; design = standardised; delivery = group; additional intervention = relaxation **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (EuroQol 5D) **Follow-up time periods available for syntheses:** 12 weeks (short); 26 weeks (moderate); 52 weeks (long)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Tilbrook 2011 *(Continued)*

Notes	**Conflicts of interest:** Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNumM10-2577. **Funding source:** Arthritis Research United Kingdom **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation sequence was computer-generated by an independent data manager and was stratified by participants prespecified availability.
Allocation concealment (selection bias)	Low risk	Eligible participant details were entered into a randomisation database by the trial co-ordinators and secretary, who were blinded to the allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Questionnaires were posted with a prepaid envelope. Nonresponders were initially followed up with postal reminders and then with a telephone call.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Discussed in Figure and seemed similar across the two groups
Participants analysed in group allocated (attrition bias)	High risk	Author contact: author referred to consort diagram; no confirmation of ITT
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Author contact: still unclear, did not look at co-interventions
Compliance acceptable in all groups (performance bias)	High risk	Sixty per cent of participants offered yoga adhered to the programme: of the remaining participants, 26% attended at least one class but did not fully adhere.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Torstensen 1998

Study characteristics	
Methods	**Study design:** RCT **Setting:** Norway, occupational **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 208 (E1 = 71, E2 = 70, C1 = 67) **Chronic LBP duration:** 327 weeks (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** 50%
Interventions	**Exercise Group 1 (E1):** Medical Exercise Therapy: mobilising and strengthening with special equipment, aerobic exercise warm-up; type = mixed; duration = 12 weeks; dose = high; design = individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Ordinary activity level: walking exercise programme; type = aerobic; duration = 12 weeks; dose = low; design = standardised; delivery = independent; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); work (return to work at 1 year) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** Device status category: 1 **Funding source:** Ministry of Health and Social Affairs, Norwegian national budget (chapter no. 0720. 63/97, project no. 103 10); programme trygd og rehabilitering (May 1993-June 1997); Foundation for Education and Research in Physiotherapy, Norway (July 1997- December 1997) **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation sequence was computer-generated by an independent data manager and was stratified by participants prespecified availability.
Allocation concealment (selection bias)	Low risk	Eligible participant details were entered into a randomisation database by the trial co-ordinators and secretary, who were blinded to the allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Questionnaires were posted with a prepaid envelope. Nonresponders were initially followed up with postal reminders and then with a telephone call.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Discussed in Figure and seemed similar across the two groups

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Torstensen 1998 *(Continued)*

Participants analysed in group allocated (attrition bias)	Low risk	Author contact: author referred to consort diagram; no confirmation of ITT
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: still unclear, did not look at co-interventions
Compliance acceptable in all groups (performance bias)	Low risk	Sixty per cent of participants offered yoga adhered to the programme: of the remaining participants, 26% attended at least one class but did not fully adhere.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Tritilanunt 2001

Study characteristics	
Methods	**Study design:** RCT **Setting:** Thailand, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 72 (E1 = 36, E2 = 36) **Chronic LBP duration:** 150 weeks (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 41 **Sex (female):** 78%
Interventions	**Exercise Group 1 (E1):** Aerobic exercises; type = aerobic; duration = 12 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = advice/education **Exercise Group 2 (E2):** Flexion and postural exercises; type = flexibility/mobilising; duration = 12 weeks; dose = high; design = standardised; delivery = not specified; additional intervention = advice/education & psychological therapy
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Tritilanunt 2001 *(Continued)*

Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Unclear risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Unclear risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Unclear risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Turner 1990

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, mixed **Exercise groups:** 2 **Comparison groups:** 2
Participants	**Number of participants:** 96 (E1 = 24, E2 = 24, C1 = 23, C2 = 25) **Chronic LBP duration:** 670 weeks (long) **Neurological/radicular symptoms:** No participants

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Turner 1990 *(Continued)*

	Mean age (years): 44 **Sex (female):** 48%
Interventions	**Exercise Group 1 (E1):** Progressive aerobic fitness (fast walking to slow jogging), warm-up, cool-down stretching, with spousal support; type = aerobic & stretching; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Progressive aerobic fitness (fast walking to slow jogging), warm-up, cool-down stretching; type = aerobic & stretching; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = psychological therapy **Comparison Group 1 (C1):** Usual care/no treatment (control group: continued with normal activity and treatment) **Comparison Group 2 (C2):** Other conservative treatment (psychological therapy)
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score); function (Sickness Impact Profile) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Institute of Neurological and Communicative Disorders and Stroke (Grants 2 RO1 NS19619 and PO1 NS 16329) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	High risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Turner 1990 *(Continued)*

Co-interventions avoided or similar (performance bias)	High risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	High risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Ulger 2017

Study characteristics	
Methods	**Study design:** RCT (NCT02696057) **Setting:** Turkey, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 144 (E1 = 72, C1 = 72) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** 59%
Interventions	**Exercise Group 1 (E1):** Spinal stabilisation exercises: various exercises focussed on contracting the transverse abdominus and multifidus muscles; type = core strengthening; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (manual therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index); HRQoL (36-Item Short Form Survey) **Follow-up time periods available for syntheses:** 6 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was performed by drawing lots among the patients who had applied for the treatment and met the participation criteria; method also applied block randomisation.
Allocation concealment (selection bias)	High risk	No information on treatment allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients were not blinded because of the differences in interventions; only statistician and outcome assessor (for non-patient-reported outcomes) were blinded; 2. Unlikely that patients could cause deviations from treatment protocols, as they had to be applied by trained physiotherapists

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ulger 2017 *(Continued)*

Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation because they were delivering interventions; 2. Different therapists for each treatment
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	1. Assessors of outcomes of interest (pain and disability) were patients, and thus were not blinded to allocation due to the distinct nature of the treatment groups; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. The two treatment groups had reasonably similarly intense intervention, and thus there was not a strong indication to a non-expert as to which one would be more effective.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. 31 "excluded" ...failed to "make time" and "show up", 20%; 2. Patients were excluded from the analysis; 3. Increased low back pain and decreased function can cause missing outcome; 4. 15 versus 16
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to their allocated treatment group (intention-to-treat)
Selective reporting (reporting bias)	Low risk	1. Protocol (NCT02696057): all analysis reported as described in protocol and article methods section
Groups similar at baseline (selection bias)	High risk	Age, sex body mass index, Oswestry and most Quality of Life subscales were not significantly different at baseline between treatment groups, however, visual analogue scale was notably different at baseline between groups.
Co-interventions avoided or similar (performance bias)	Low risk	Patients with ongoing pharmaceutical medication were excluded from the study (avoiding co-interventions); history of physiotherapy was recorded but not reported.
Compliance acceptable in all groups (performance bias)	Low risk	No reporting of compliance
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Within each outcome, all participants were measured with the same tool and at the same time point; 2. Visual Analogue Scale (for pain) and Oswestry Disability Index (for disability) are well-validated tools in the context of low back pain.
Other bias	Low risk	Appeared free from other sources of bias

Underwood 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** United Kingdom, healthcare **Exercise groups:** 2 **Comparison groups:** 2
Participants	**Number of participants:** 758 (E1 = 185, E2 = 173, C1 = 190, C2 = 210) **Chronic LBP duration:** Not specified (moderate) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 43 **Sex (female):** 56%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Underwood 2004 *(Continued)*

Interventions	**Exercise Group 1 (E1):** "Back to fitness": initial individual assessment then group classes incorporating cognitive behavioural principles and a mix of exercises; type = mixed; duration = 12 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy **Exercise Group 2 (E2):** "Back to fitness": Initial individual assessment then group classes incorporating cognitive behavioural principles and a mix of exercises; type = mixed; duration = 12 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education & psychological therapy & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (education) **Comparison Group 2 (C2):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Graded Chronic Pain Scale (Von Korff)); Function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey; EuroQol 5D) **Follow-up time periods available for syntheses:** 4 weeks (short); 12 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** LL, JM, MU, MV, and KW have received salaries from the MRC. MU has received fees for speaking from Menarini Pharmaceuticals, the manufacturers of dexketoprofen and ketoprofen, and Pfizer, the manufacturers of celecoxib and valdecoxib. **Funding source:** Medical Research Council (research costs); National Health Service in England, Northern Ireland, Scotland, and Wales (excess treatment and service support costs) **Other:** Information modified for author contact

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	After consenting and participants had completed baseline assessments, nurses contacted the remote randomisation service.
Allocation concealment (selection bias)	Low risk	Author contact: Changed to yes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Patients not blinded
Blinding of care provider (performance bias)	High risk	Care providers not blinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	At three months, 1029 (77%) returned questionnaires; at 12 months, 995 (75%) returned questionnaires.
Participants analysed in group allocated (attrition bias)	Low risk	Used two-sided significance tests to analyse the primary outcome, Roland disability questionnaire score, after three or 12 months by intention-to-treat
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The mean (SD) age of participants at randomisation was 43 (11) years; 56% were female, and 9% were not working because of poor health.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Underwood 2004 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	They agreed to avoid physical treatments, other than trial treatments, for three months.
Compliance acceptable in all groups (performance bias)	High risk	Author contact: 47% did not attend Ax + one session.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Unsgaard-Tondel 2010

Study characteristics

Methods	**Study design:** RCT (NCT00201513) **Setting:** Norway, mixed **Exercise groups:** 3 **Comparison groups:** 0
Participants	**Number of participants:** 109 (E1 = 36, E2 = 36, E3 = 37) **Chronic LBP duration:** 6 years (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 40 **Sex (female):** 71%
Interventions	**Exercise Group 1 (E1):** Motor control exercises: isolating and activating transverse abdominus, abdominal muscles, pelvic floor, multifidus using biofeedback, home abdominal exercises; type = core strengthening; duration = 8 weeks; dose = low; design = individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Sling exercises: back exercises in slings to stabilise spine through a range of leg and arm positions and movements; type = core strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = individual; additional intervention = advice/education **Exercise Group 3 (E3):** General exercises: trunk extension, flexion, rotation with resistance and stretching of trunk and extremity muscles with resistance equipment, home flexibility exercises; type = stretching & strengthening; duration = 8 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = advice/education
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 8 weeks (short); 52 weeks (long)
Notes	**Conflicts of interest:** Not reported **Funding source:** Norwegian Fund for Post-Graduate Training in Physiotherapy **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation was administered by an independent study secretary via telephone.
Allocation concealment (selection bias)	Low risk	The secretary consecutively reported group allocation for included participants from a list of random numbers between 0 and 1.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Unsgaard-Tondel 2010 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Although the baseline assessment was performed blinded, the physical therapist conducting the post-intervention evaluation was not blinded to treatment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Twelve of 80 participants who were recruited by announcement at the local hospital and 10 of 29 participants who were recruited from primary care dropped out.
Participants analysed in group allocated (attrition bias)	Low risk	Performed an intention-to-treat analysis, and used mixed linear models to estimate mean scores, to estimate baseline-adjusted between group differences
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Not described
Co-interventions avoided or similar (performance bias)	Low risk	Participants were not allowed to receive other treatment for low back pain during the intervention period.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Valenza 2017

Study characteristics	
Methods	**Study design:** RCT **Setting:** Spain, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 54 (E1 = 27, C1 = 27) **Chronic LBP duration:** 14.5 months (moderate) **Neurological/radicular symptoms:** No participants **Mean age (years):** 39 **Sex (female):** 76%
Interventions	**Exercise Group 1 (E1):** Pilates; type = Pilates; duration = 8 weeks; dose = low; design = partially individualised; delivery = not specified; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Valenza 2017 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short)
Notes	**Conflicts of interest:** None to declare **Funding source:** No funding received **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number generator in blocks of eight with no stratification
Allocation concealment (selection bias)	Low risk	Treatment allocation was concealed by using an independent third party to prepare randomised allocations, which were mailed to the study personnel.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	1. Patients could not be blinded to allocation due to the nature of the treatments; 2. Unlikely that lack of patient blinding caused deviations from intended interventions; experimental group was structured and controlled by care providers, and control group was purposely broad.
Blinding of care provider (performance bias)	Low risk	1. Care providers could not be blinded to allocation due to the nature of the treatments; 2. Control group had minimal contact with care providers.
Blinding of outcome assessment (detection bias) All outcomes	High risk	1. Outcome assessors for pain and function were patients themselves, who could not be blinded to allocation due to the nature of the treatments; 2. Pain and functional questionnaires are subjective, and responses could be altered by awareness of intervention; 3. Likely that lack of patient blinding introduced bias to outcome assessments, as the experimental group was clearly "better" than the control; study results for all outcome supported this.
Incomplete outcome data (attrition bias) All outcomes	Low risk	1. No dropouts in this study
Participants analysed in group allocated (attrition bias)	Low risk	1. Appeared that all patients were analysed according to the allocation to which they were randomised
Selective reporting (reporting bias)	Low risk	1. No linked protocol or statistical analysis plan found: within this publication, all analysis and outcomes were fully reported.
Groups similar at baseline (selection bias)	Low risk	Both treatment groups were similar on all relevant baseline characteristics.
Co-interventions avoided or similar (performance bias)	Low risk	"Participants were instructed to follow their normal schedule of medications and physical activity, without starting any new exercise programme or drug treatment throughout the course of the study".
Compliance acceptable in all groups (performance bias)	High risk	No report on attendance, despite the study claiming to have recorded this information
Timing of outcome assessment similar in all groups (detection bias)	Low risk	1. Outcome assessments were identical for all patients, regardless of treatment group; 2. Roland-Morris Disability Questionnaire and Oswestry Disability

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Valenza 2017 *(Continued)*

		Index (for function), and Visual Analogue Scale (for pain) are all well-validated tools in the low back pain context.
Other bias	Low risk	Appeared free from other sources of bias

Van der Roer 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** Netherlands, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 114 (E1 = 60, E2 = not reported) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** Not reported **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Progressive programme including aerobic, back abdominal and buttock exercises; type = mixed; duration = 6 weeks; dose = high; design = partially individualised; delivery = group; additional intervention = psychological therapy & back school **Exercise Group 2 (E2):** Guideline physiotherapy including exercise therapy with a behavioural approach; type = mixed; duration = 6 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire); work (work absenteeism, measured with the Short Form Health and Labour Questionnaire (HLQ)); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Netherlands Organisation for Health Research and Development (grant no: 945-03-023) **Other:** SDs imputed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	For each research centre, a randomisation list was prepared and permuted blocks of four patients were made to ensure equal distribution of patients.
Allocation concealment (selection bias)	Low risk	The principal investigator (NvdR), who was not involved in the recruitment of patients or treatment allocation, prepared opaque, sealed envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Due to the pragmatic design, both patients and physiotherapists could not be blinded for the interventions.
Blinding of care provider (performance bias)	High risk	Due to the pragmatic design, both patients and physiotherapists could not be blinded for the interventions.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Van der Roer 2008 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Patients who were lost to follow-up (n = 12) were significantly younger at baseline than completers (mean age 34.1 vs 42.6). See Figure 1
Participants analysed in group allocated (attrition bias)	Low risk	An intention-to-treat analysis was conducted for each follow-up moment using multilevel modelling.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Baseline characteristics of the patients were largely similar in both groups (Table 1).
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Despite the training sessions the protocol physiotherapists received, authors found that in 18% of the patients the protocol was not adequately followed.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Vibe Fersum 2013

Study characteristics	
Methods	**Study design:** RCT (NCT01129817) **Setting:** Norway, mixed **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 94 (E1 = 43, E2 = 51) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 42 **Sex (female):** 51%
Interventions	**Exercise Group 1 (E1):** Motor control exercise (as part of physiotherapy); type = mixed; duration = 12 weeks; dose = high; design = individualised; delivery = individual; additional intervention = manual therapy **Exercise Group 2 (E2):** Functional movement exercises, functional activities, walking, biking; type = flexibility/mobilising; duration = 12 weeks; dose = low; design = individualised; delivery = not specified; additional intervention = psychological therapy & manual therapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Oswestry Disability Index); work (extracted from the Orebro Screening Questionnaire using a 10-category variable) **Follow-up time periods available for syntheses:** 12 weeks (short); 52 weeks (long)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Vibe Fersum 2013 *(Continued)*

Notes	**Conflicts of interest:** None to declare **Funding source:** Norwegian Fund for Post-Graduate Training in Physiotherapy **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was performed in permuted blocks of 16.
Allocation concealment (selection bias)	Low risk	One-hundred and sixty sealed opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	Low risk	It was clearly stated that there were two active comparable treatment arms and that, based on current knowledge, patients did not know which was superior.
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	There were two active comparable treatment arms and that, based on current knowledge, assessors did not know which was superior.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Figure 1
Participants analysed in group allocated (attrition bias)	High risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Comparable in terms of baseline characteristics, with the exception of small but significant differences in Hopkins Symptoms Checklist and Fear-Avoidance Beliefs Questionnaire work (Table 1).
Co-interventions avoided or similar (performance bias)	Low risk	There was no difference between the two groups in terms of medication intake before or after the treatment.
Compliance acceptable in all groups (performance bias)	Unclear risk	"A lack of compliance of greater than 50% was set as a withdrawal criterion" but unclear what other compliance rates were
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Vincent 2014a

Study characteristics	
Methods	**Study design:** RCT (NCT01250262) **Setting:** USA, mixed **Exercise groups:** 2 **Comparison groups:** 1
Participants	**Number of participants:** 49 (E1 = 17, E2 = 18, C1 = 14) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 68 **Sex (female):** 67%
Interventions	**Exercise Group 1 (E1):** Total body resistance exercise including lumbar extension (with machines); type = strengthening; duration = 17 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = advice/education **Exercise Group 2 (E2):** Isolated lumbar extension resistance exercise (with machines); type = strengthening; duration = 17 weeks; dose = high; design = partially individualised; delivery = individual; additional intervention = advice/education & psychological therapy **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 17 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health (AR057552- 01A1) **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	A computer-generated list was used to randomly assign the group allocation.
Allocation concealment (selection bias)	Low risk	The assignments per participant number were placed in numbered sealed envelopes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Figure 1

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Vincent 2014a *(Continued)*

Participants analysed in group allocated (attrition bias)	High risk	Not according to the flow chart
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no differences in the physiological characteristics among the three study groups.
Co-interventions avoided or similar (performance bias)	Low risk	Author contact: measured diet and other exercises
Compliance acceptable in all groups (performance bias)	Low risk	Adherence to the training programmes in both isolated lumbar extension resistance exercise group and total body resistance exercise group were excellent.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	All the same

Vollenbroek-Hutten 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** Netherlands, healthcare **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 163 (E1 = 79, C1 = 84) **Chronic LBP duration:** 60 months (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 39 **Sex (female):** Not reported
Interventions	**Exercise Group 1 (E1):** Conditional training and sport, swimming; type = mixed; duration = 8 weeks; dose = high; design = standardised; delivery = group; additional intervention = psychological therapy & electrotherapy & manual therapy & back school **Comparison Group 1 (C1):** Usual care/no treatment (control group: continued with normal activity and treatment)
Outcomes	**Core outcomes reported:** Function (Roland-Morris Disability Questionnaire); HRQoL (EuroQol 5D) **Follow-up time periods available for syntheses:** 8 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Vollenbroek-Hutten 2004 *(Continued)*

Random sequence generation (selection bias)	Low risk	After the baseline measurements, patients were put into either the control group or the treatment group using the minimization method.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Patients could not be blinded for the group to which they were randomised, neither could the therapist who conducted the general physical condition measurement.
Blinding of care provider (performance bias)	High risk	Patients could not be blinded for the group to which they were randomised, neither could the therapist who conducted the general physical condition measurement.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The researchers conducted all other measurements, and they were blinded for the group to which the patients were randomised.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Of the 163 patients who were included in the trial, 21 patients were lost during follow-up (13%).
Participants analysed in group allocated (attrition bias)	Low risk	For all analyses, an intention-to-treat analysis, including patients with protocol deviations, was performed and results were considered statistically.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no differences in any of these baseline characteristics between the treatment and control group
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Wajswelner 2012

Study characteristics	
Methods	**Study design:** RCT (ACTRN12609000536268) **Setting:** Australia, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 87 (E1 = 44, E2 = 43) **Chronic LBP duration:** Not specified (moderate)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Wajswelner 2012 *(Continued)*

	Neurological/radicular symptoms: Some participants **Mean age (years):** 49 **Sex (female):** 55%
Interventions	**Exercise Group 1 (E1):** Series of Pilates exercises performed on the reformer and trapeze equipment, directional-specific; type = Pilates; duration = 6 weeks; dose = low; design = individualised; delivery = group; additional intervention = none **Exercise Group 2 (E2):** Exercise including stationary bike, leg stretches, upper body weights, theraband, Swiss ball, and floor exercises: multidirectional; type = mixed; duration = 6 weeks; dose = low; design = standardised; delivery = group; additional intervention = none
Outcomes	**Core outcomes reported:** Pain (Quebec Back Pain Disability Scale); function (Quebec Back Pain Disability Scale); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (Global Perceived Effect scale)) **Follow-up time periods available for syntheses:** 6 weeks (short); 13 weeks (moderate); 104 weeks (long)
Notes	**Conflicts of interest:** Henry Wajswelner works at a physiotherapy and Pilates clinic that uses clinical Pilates exercises to treat patients. He also teaches clinical Pilates to other physiotherapists. **Funding source:** Mr. Craig Phillips of DMA Clinical Pilates Physiotherapy in South Yarra, Melbourne, Victoria, Australia; Mr. Marcus Pain of Back in Motion Physiotherapy in Brunswick, Melbourne, Victoria, Australia **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	After baseline assessment, participants were randomly allocated in permuted blocks of six and eight, stratified by age.
Allocation concealment (selection bias)	Low risk	Allocation was sealed in opaque and consecutively numbered envelopes held centrally.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Several self-report questionnaire measures were taken including those recommended in the European Guidelines on Chronic Low Back Pain.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Eighty-three participants (96%) completed the six-week intervention, 67 (77%) completed the 12-week follow-up, and 60 (69%) completed the 24-week follow-up.
Participants analysed in group allocated (attrition bias)	Low risk	The primary analysis was by intention-to-treat and was performed in a blinded manner using the Statistical Package for the Social Sciences.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Wajswelner 2012 *(Continued)*

Groups similar at baseline (selection bias)	Low risk	The groups were similar at baseline for demographic and clinical characteristics (Table 1).
Co-interventions avoided or similar (performance bias)	Low risk	There were few reported co-interventions.
Compliance acceptable in all groups (performance bias)	Low risk	Thirteen participants (seven from the clinical Pilates group and six from the general exercise group) failed to complete all 12 exercise class sessions.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Weifen 2013

Study characteristics	
Methods	**Study design:** RCT **Setting:** China, healthcare **Exercise groups:** 4 **Comparison groups:** 1
Participants	**Number of participants:** 320 (E1 = 141, E2 = 47, E3 = 47, E4 = 38, C1 = 47) **Chronic LBP duration:** 2.1 years (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 38 **Sex (female):** 41%
Interventions	**Exercise Group 1 (E1):** Tai Chi (Chen style); type = other (Tai Chi, Qigong); duration = 26 weeks; dose = high; design = standardised; delivery = independent with follow-up; additional intervention = electrotherapy & manual therapy **Exercise Group 2 (E2):** Backward walking; type = aerobic; duration = 26 weeks; dose = high; design = standardised; delivery = independent with follow-up; additional intervention = electrotherapy & manual therapy **Exercise Group 3 (E3):** Jogging; type = aerobic; duration = 26 weeks; dose = high; design = standardised; delivery = independent; additional intervention = electrotherapy & manual therapy **Exercise Group 4 (E4):** Swimming; type = aerobic; duration = 26 weeks; dose = high; design = standardised; delivery = independent with follow-up; additional intervention = electrotherapy & manual therapy **Comparison Group 1 (C1):** Other conservative treatment (physical therapy)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale) **Follow-up time periods available for syntheses:** 26 weeks (moderate)
Notes	**Conflicts of interest:** None to declare **Funding source:** Not reported **Other:** SDs imputed

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Using a random number table generated by SPSS 17.0

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Weifen 2013 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Low risk	At the beginning of our trial, participants were told that these exercises were helpful with alleviating back pain.
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	During the trial, seven patients from the control group and nine patients from the Tai Chi group dropped out of the study.
Participants analysed in group allocated (attrition bias)	Unclear risk	It was not clear whether the dropouts were left out of all the analyses or whether the 320 included the dropouts.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	At the beginning of the trial, Tai Chi and control groups showed no significant differences.
Co-interventions avoided or similar (performance bias)	Low risk	Sports training was vetoed but no mention of medication, other therapies, etc.
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Weiner 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, healthcare **Exercise groups:** 2 **Comparison groups:** 2
Participants	**Number of participants:** 200 (E1 = 50, E2 = 50, C1 = 50, C2 = 50) **Chronic LBP duration:** 7.5 years (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 74 **Sex (female):** 55%

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Weiner 2008 *(Continued)*

Interventions	**Exercise Group 1 (E1):** General conditioning (strength and flexibility), aerobic exercise, home exercise (flexibility and walking); type = mixed; duration = 6 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = placebo **Exercise Group 2 (E2):** General conditioning (strength and flexibility), aerobic exercise, home exercise (flexibility and walking); type = mixed; duration = 6 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = electrotherapy **Comparison Group 1 (C1):** Placebo: sham percutaneous electrical nerve stimulation (PENS) treatment **Comparison Group 2 (C2):** Other conservative treatment (electrotherapy)
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score); function (Roland-Morris Disability Questionnaire); HRQoL (36-Item Short Form Survey); Global Perceived Health or Recovery (Global Perceived Health or Recovery (global change in condition (5-point)) **Follow-up time periods available for syntheses:** 6 weeks (short); 26 weeks (moderate)
Notes	**Conflicts of interest:** Dr. Perera received funding from Eli Lily & Co. to do observational research **Funding source:** National Center for Complementary and Alternative Medicine (Grant R01 AT000985); National Institute on Aging, National Institutes of Health; Pittsburgh Claude D. Pepper Older Americans Independence Center (NIA P30 AG-024827) **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were randomised to one of the four groups, using a stratified blocked randomisation scheme and a statistical software for random deviation.
Allocation concealment (selection bias)	High risk	One of the study investigators created and monitored the implementation of the randomisation scheme.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	One of the study investigators created and monitored the implementation of the randomisation scheme.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The research associate who collected the outcome data was masked to group assignment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	The overall dropout rate was 8% (Figure 1).
Participants analysed in group allocated (attrition bias)	Low risk	The analysis was repeated with last-value-carried-forward and multiple imputation approaches to assess sensitivity of the results.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	There were no significant differences between groups.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Weiner 2008 *(Continued)*

Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Low risk	Adherence with general conditioning and aerobic exercise was also comparable between groups.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Williams 2005

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 44 (E1 = 20, C1 = 24) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 48 **Sex (female):** 68%
Interventions	**Exercise Group 1 (E1):** Yoga: 29 different postures using supportive props to enhance alignment, flexibility, mobility and stability in all muscles and joints that affect spinal alignment and posture; type = yoga; duration = 16 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Other conservative treatment (education)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Pain Disability Index) **Follow-up time periods available for syntheses:** 16 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** West Virginia University **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Subjects were randomised to control or yoga groups using a random number generating program from JMP 4.0 statistical software.
Allocation concealment (selection bias)	Low risk	Author contact: changed to yes; they used envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Williams 2005 *(Continued)*

Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Data collectors were blind to the subject's treatment status.
Incomplete outcome data (attrition bias) All outcomes	High risk	Ten subjects were excluded from the analysis in the yoga group.
Participants analysed in group allocated (attrition bias)	High risk	Not included
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	A one-way ANOVA (unpaired t-test) revealed no significant differences in demographics and medical history between the yoga and control groups (P > 0.05).
Co-interventions avoided or similar (performance bias)	Low risk	Changes from the baseline in drug consumption were evaluated at post-intervention and at three-month follow-up.
Compliance acceptable in all groups (performance bias)	Low risk	Of the 20 subjects completing the yoga intervention, an attendance rate of 91.9% was achieved for the 16-week protocol.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Williams 2009

Study characteristics	
Methods	**Study design:** RCT **Setting:** USA, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 90 (E1 = 43, C1 = 47) **Chronic LBP duration:** Not specified (long) **Neurological/radicular symptoms:** No participants **Mean age (years):** 48 **Sex (female):** 77%
Interventions	**Exercise Group 1 (E1):** Iyengar yoga; type = yoga; duration = 24 weeks; dose = high; design = standardised; delivery = group; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (control group: continued with normal activity and treatment)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 24 weeks (moderate); 50 weeks (long)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Williams 2009 *(Continued)*

Notes	**Conflicts of interest:** None to declare **Funding source:** National Institutes of Health's National Center for Complementary and Alternative Medicine grant (no.1 R21 AT001679-01A2) **Other:** Information modified for author contact; SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Eligible participants were given envelopes with randomly generated group assignment and enrolled in one of four cohorts of 20 to 28 participants each.
Allocation concealment (selection bias)	Low risk	Author contact: changed to yes; they used envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Participants were asked to return at 12 (midway), 24 (immediately after), and 48 weeks (six-month follow-up) after the start of the intervention.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Sixteen participants (12 from the yoga group) did not complete the 24-week protocol.
Participants analysed in group allocated (attrition bias)	Low risk	For intention-to-treat analyses, missing baseline data were replaced by group means while missing data at 12 and 24 weeks were replaced.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	More African-Americans were enrolled in the yoga versus control groups.
Co-interventions avoided or similar (performance bias)	Low risk	Agreed to not get chiropractic treatment, massage therapy, Pilates, or acupuncture or to participate in any other yoga programme
Compliance acceptable in all groups (performance bias)	Low risk	On average, yoga completers attended 88.5% of classes and 87.1% completed home practice.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Woods 2008

Study characteristics	
Methods	**Study design:** RCT **Setting:** Canada, mixed **Exercise groups:** 1 **Comparison groups:** 2
Participants	**Number of participants:** 44 (E1 = 13, C1 = 16, C2 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 46 **Sex (female):** 66%
Interventions	**Exercise Group 1 (E1):** Graded activity: the shaping of healthy behaviours through positive reinforcement of predefined activity quotas (Vlaeyen 2002); type = other (graded activity); duration = 4 weeks; dose = low; design = individualised; delivery = individual; additional intervention = none **Comparison Group 1 (C1):** Usual care/no treatment (waiting-list group) **Comparison Group 2 (C2):** Other conservative treatment (psychological therapy)
Outcomes	**Core outcomes reported:** Pain (McGill Pain Score); function (Pain Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short); 8 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Royal Bank of Canada; Canadian Institutes of Health Research Investigator Award **Other:** None

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	All patients were randomly assigned to one of three conditions via the rolling of a six-sided dice (a dice roll of 1 and 4 = graded in vivo exposure, 2 and 5 = graded activity)
Allocation concealment (selection bias)	High risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described
Blinding of care provider (performance bias)	High risk	Not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcome assessors not blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	Randomised (n = 83); graded in vivo exposure (n = 36) completed (n = 15), dropouts (n = 21); graded activity (n = 25) completed (n = 13), dropouts (n = 12); wait list control (n = 22) completed (n = 16)
Participants analysed in group allocated (attrition bias)	Low risk	An intent-to-treat analysis, using Analysis of Covariance was conducted.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Woods 2008 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Independent samples t-tests (for age) and v2 analyses (for sex, education level, and employment status) comparing demographics between conditions showed no differences in characteristics other than sex.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Xueqiang 2012

Study characteristics	
Methods	**Study design:** RCT **Setting:** China, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 60 (E1 = 28, E2 = 32) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 39 **Sex (female):** 42%
Interventions	**Exercise Group 1 (E1):** Muscle strengthening exercises: abdominal and trunk extensor muscles, warm-up (5 minutes), abdominal extensor exercises (15 minutes), trunk extensor exercises (15 minutes), cool down (5 minutes); type = core strengthening; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = not specified **Exercise Group 2 (E2):** Co-ordination and control of, and strength and endurance of the trunk muscles, neural control and neutral spine exercises (bridging with legs lifts, reverse bridge, etc.); type = core strengthening; duration = 12 weeks; dose = high; design = standardised; delivery = group; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 12 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** National Science Foundation for Distinguished Young Scholars (Grant: 81101391); Science and Technology Foundation Program of Shanghai University of Sport (Grant: YJSCX201120); Shanghai Natural Science Foundation of Chine (Grant: 11ZR1434900) **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Xueqiang 2012 *(Continued)*

Random sequence generation (selection bias)	Low risk	All subjects were randomly assigned by using a computer-generated random number sequence to either core stability training group or control group
Allocation concealment (selection bias)	Low risk	With the use of opaque closed envelopes and stratified by centre, included patients were randomised to core stability training group or control group by an independent co-ordinator.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	This study was a two-armed randomised, controlled trial with blinding of patients and assessors with respect to the nature of therapy.
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All assessments were done by three independent, experienced physical therapists, who were not working in the participating rehabilitation centres.
Incomplete outcome data (attrition bias) All outcomes	Low risk	After 12-week programme, there were 55 subjects.
Participants analysed in group allocated (attrition bias)	Unclear risk	Two patients did not attend the final evaluation session.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	The groups were well matched at the baseline assessment, with no differences in key outcome variables.
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Post-treatment (12 weeks)

Yelland 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** Australia, mixed **Exercise groups:** 1 **Comparison groups:** 1
Participants	**Number of participants:** 110 (E1 = 55, C1 = 55) **Chronic LBP duration:** 700 weeks (long)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yelland 2004 *(Continued)*

	Neurological/radicular symptoms: No participants **Mean age (years):** 50 **Sex (female):** 43%
Interventions	**Exercise Group 1 (E1):** Sagittal loading flexibility and mobilising exercises; type = flexibility/mobilising; duration = 24 weeks; dose = low; design = standardised; delivery = independent; additional intervention = electrotherapy & anti-inflammatory/analgesics **Comparison Group 1 (C1):** Usual care/no treatment (control group: continue normal activity and exercises)
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire); HRQoL (12-Item Short Form Survey) **Follow-up time periods available for syntheses:** 10 weeks (short); 24 weeks (moderate); 52 weeks (long)
Notes	**Conflicts of interest:** None to declare **Funding source:** Australian General Practice Evaluation Program; Australian Association of Musculoskeletal Medicine; Musculoskeletal Research Foundation of Australia **Other:** None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available
Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yelland 2004 *(Continued)*

Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

Yeung 2003

Study characteristics	
Methods	**Study design:** RCT **Setting:** Hong Kong, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 52 (E1 = 26, E2 = 26) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** Some participants **Mean age (years):** 53 **Sex (female):** 83%
Interventions	**Exercise Group 1 (E1):** Stretching, mobilising, strengthening exercises; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = group; additional intervention = advice/education **Exercise Group 2 (E2):** Stretching, mobilising, strengthening exercises; type = mixed; duration = 4 weeks; dose = low; design = standardised; delivery = independent; additional intervention = advice/education & electrotherapy
Outcomes	**Core outcomes reported:** Pain (Numeric Rating Scale); function (Aberdeen Back Pain Scale) **Follow-up time periods available for syntheses:** 8 weeks (short); 16 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** Hong Kong Polytechnic University Area of Strategy Development Fund; Tung Wah Board Fund **Other:** Information modified for author contact

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Support for judgement is not available
Allocation concealment (selection bias)	Low risk	Support for judgement is not available
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement is not available

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yeung 2003 *(Continued)*

Blinding of care provider (performance bias)	High risk	Support for judgement is not available
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement is not available
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement is not available
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement is not available
Selective reporting (reporting bias)	Low risk	Support for judgement is not available
Groups similar at baseline (selection bias)	Low risk	Support for judgement is not available
Co-interventions avoided or similar (performance bias)	Low risk	Support for judgement is not available
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement is not available
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement is not available

You 2014

Study characteristics	
Methods	**Study design:** RCT **Setting:** South Korea, general population **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 40 (E1 = 20, E2 = 20) **Chronic LBP duration:** 20.53 months (moderate) **Neurological/radicular symptoms:** Not specified **Mean age (years):** 51 **Sex (female):** 52%
Interventions	**Exercise Group 1 (E1):** Core stabilisation exercises in hook lying position, adding ankle dorsiflexion to the drawing in of the abdominal wall with ultrasound biofeedback; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = advice/education & manual therapy **Exercise Group 2 (E2):** Core stabilisation exercises in hook lying position: add ankle dorsiflexion to the drawing in of the abdominal wall with ultrasound biofeedback; type = core strengthening; duration = 8 weeks; dose = low; design = standardised; delivery = not specified; additional intervention = advice/education & manual therapy

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

You 2014 *(Continued)*

Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Roland-Morris Disability Questionnaire) **Follow-up time periods available for syntheses:** 8 weeks (short); 16 weeks (moderate)
Notes	**Conflicts of interest:** Not reported **Funding source:** No funding received **Other:** SDs imputed

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomisation was done with sealed envelopes.
Allocation concealment (selection bias)	Unclear risk	The investigator prepared group allocation on a sheet of paper and gave it to subjects in a blinded manner.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Assumed not possible
Blinding of care provider (performance bias)	High risk	Assumed not possible
Blinding of outcome assessment (detection bias) All outcomes	High risk	Assumed not possible
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.
Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	No significant difference which indicated that the groups had similar demographic characteristics
Co-interventions avoided or similar (performance bias)	Unclear risk	Not described
Compliance acceptable in all groups (performance bias)	Unclear risk	Not described
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yozbatiran 2004

Study characteristics	
Methods	**Study design:** RCT **Setting:** Turkey, healthcare **Exercise groups:** 2 **Comparison groups:** 0
Participants	**Number of participants:** 30 (E1 = 15, E2 = 15) **Chronic LBP duration:** Not specified (not specified) **Neurological/radicular symptoms:** No participants **Mean age (years):** 39 **Sex (female):** 77%
Interventions	**Exercise Group 1 (E1):** Warm-up, stretching, progressive exercises, light aerobics on land; type = mixed; duration = 4 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = not specified **Exercise Group 2 (E2):** Warm-up, stretching, progressive exercises, light aerobics in water; type = mixed; duration = 4 weeks; dose = low; design = partially individualised; delivery = group; additional intervention = not specified
Outcomes	**Core outcomes reported:** Pain (Visual Analogue Scale); function (Oswestry Disability Index) **Follow-up time periods available for syntheses:** 4 weeks (short)
Notes	**Conflicts of interest:** Not reported **Funding source:** Not reported **Other:** None

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Support for judgement was not available.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Support for judgement was not available.
Blinding of care provider (performance bias)	High risk	Support for judgement was not available.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Support for judgement was not available.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Support for judgement was not available.
Participants analysed in group allocated (attrition bias)	Low risk	Support for judgement was not available.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yozbatiran 2004 *(Continued)*

Selective reporting (reporting bias)	Low risk	Support for judgement was not available.
Groups similar at baseline (selection bias)	Low risk	Support for judgement was not available.
Co-interventions avoided or similar (performance bias)	Unclear risk	Support for judgement was not available.
Compliance acceptable in all groups (performance bias)	Low risk	Support for judgement was not available.
Timing of outcome assessment similar in all groups (detection bias)	Low risk	Support for judgement was not available.

For studies assessed using ROB 2 tool (indicated by numbering in the support for judgement column), the numbering indicates:
Blinding of participants: 1. Were participants blinded to the intervention?; 2. Did lack of patient blinding lead to deviations from the intended intervention due to the experimental context?; 3. Were these deviations balanced between groups (patient)?; 4. Were these deviations likely to have affected the outcome (patient)?
Blinding of care provider: 1. Were care providers blinded to the intervention?; 2. Did lack of care provider blinding lead to deviations from the intended intervention due to the experimental context?; 3. Were these deviations balanced between groups (provider)?; 4. Were these deviations likely to have affected the outcome (provider)?
Blinding of outcome assessment: 1. Were outcome assessors blinded to the intervention?; 2. Could the assessment have been altered by a lack of blinding (i.e. do the measures being assessed require judgement?); 3. Is it likely that the assessment of outcome was altered by a lack of blinding?
Incomplete outcome data: 1. Was the dropout rate described and acceptable? (please provide %, if available); 2. Is there evidence that the analysis was NOT biased by missing data?; 3. Could missingness in outcome depend on its true value?; 4. Do the missingness rates differ between groups?; 5. Is it likely that missingness in outcome depends on its true value?
Participants analysed in group allocated: 1. Was the intention to treat analysis included?; 2. Was it likely to have affected the result?
Selective reporting: 1. Was the study analysed and fully reported according to a pre-specified plan?; 2. Was the outcome reported likely selected on the basis of the results it gave?; 3. Was the analytic method reported likely selected on the basis of the results it gave?
Timing of outcome assessment similar in all groups: 1. Were the outcome assessments similar between groups (timing, tools, scales, and thresholds,etc)?; 2. Were tools/scales used to measure the outcome(s) valid?
*Note that there are inconsistencies with some of our risk of bias judgements and reasons provided across studies. These inconsistencies have resulted from differing consensus decisions for sets of reviewers over time (assessments were conducted over 15 years) and some consistent differences between studies assessments of studies with the RoB 1 vs RoB 2 tools (i.e. different assessment criteria and introduced with the algorithm used to convert RoB 2 to RoB 1 for presentation here). The different judgements would not affect the certainty of evidence.
ANOVA: analysis of variance; HRQoL: health-related quality of life; ITT: intention-to-treat; LBP: low back pain; RCT: randomized controlled trial; SD: standard deviation

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

Study	Reason for exclusion
Akbari 2008	Article judged to contain plagiarism and included inadequate reporting of CONSORT items
Aleksiev 2014	Not a chronic population
Ali 2013	Article judged to contain plagiarism, has been rated as high risk of bias and included inadequate reporting of CONSORT items

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Anema 2007	Population did not have a majority with chronic low back pain (subacute only)
Bergstrom 2012	Population did not have a majority with chronic low back pain (mixed population, neck pain)
Bhatnagar 2017	Article judged to contain plagiarism and included inadequate reporting of CONSORT items
Browder 2007	Population did not have a majority chronic with low back pain (mixed population, some surgical)
Celestini 2005	Population was not nonspecific chronic low back pain
Chatzitheodorou 2008	No comparison for exercise group (exercises did not differ between groups)
Cho 2014	Article judged to be published in a predatory journal and has been rated as having high risk of bias
Chul 2016	Population was not chronic low back pain (intermittent back pain)
Demirel 2008	No comparison for exercise group (exercises did not differ between groups)
Diaz-Arribas 2015	Population did not have a majority with chronic low back pain
Dillen 2016	No comparison for exercise group
Dundar 2009	Article judged to contain plagiarism and included inadequate reporting of CONSORT items
Ferreira 2016	Population did not have a majority with chronic low back pain (main complaint leg pain)
Ford 2016	Exercise not the main component of intervention
Froholdt 2012	Population was not nonspecific chronic low back pain
Galantino 2004	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Gao 2006	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Haufe 2017	Population did not have chronic low back pain at baseline
Hosseinifar 2013	Article judged to be published in a predatory journal, judged to contain plagiarism and included inadequate reporting of CONSORT items
Jeong 2015	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Jeong 2016	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Ju 2015	Not judged to be exercise treatment (passive mobilisations only)
Ju-Hyun 2017	Population did not have a majority with chronic low back pain (main complaint leg pain)
Kachanathu 2012	Article judged to be published in a predatory journal, was rated as having high risk of bias and included inadequate reporting of CONSORT items
Kaeding 2017	Exercise not judged to be the main component of intervention

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Kamali 2014	No comparison for exercise group (exercises did not differ between groups)
Kamali 2018	Population did not have a majority with chronic low back pain
Kang 2016	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Kim 2013	Article judged to be published in a predatory journal and has been rated as high risk of bias
Kim 2014	Article judged to be published in a predatory journal and has been rated as high risk of bias
Kim 2015b	Article judged to be published in a predatory journal, was rated as high risk of bias and included inadequate reporting of CONSORT items
Ko 2018	Article judged to be published in a predatory journal, judged to contain plagiarism and included inadequate reporting of CONSORT items
Konstantinou 2007	Exercise not judged to be the main component of intervention
Kumar 2011	Population did not have a majority with chronic low back pain (mixed population, duration unclear)
Lee 2014a	Article judged to be published in a predatory journal, was rated as high risk of bias and included inadequate reporting of CONSORT items
Lee 2014b	Article judged to be published in a predatory journal, was rated as high risk of bias and included inadequate reporting of CONSORT items
Lee 2016b	Population with low back pain at baseline was unclear
Lee 2017	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Lee 2018	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Maciaszek 2016	Population did not have chronic low back pain
Majiwala 2017	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Masse-Alarie 2017	No comparison for exercise group (exercises did not differ between groups)
Mirovsky 2006	Article judged to contain plagiarism, has been rated as high risk of bias and included inadequate reporting of CONSORT items
Moon 2017	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Nambi 2014	Article judged to contain plagiarism, has been rated as having high risk of bias and included inadequate reporting of CONSORT items
Nazzal 2013	Article judged to contain plagiarism and included inadequate reporting of CONSORT items
Park 2013	Article judged to be published in a predatory journal, was rated as having high risk of bias and included inadequate reporting of CONSORT items

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Park 2019	No comparison for exercise group (exercises did not differ between groups)
Paungmali 2018	Outcomes of interest not measured
Pengel 2007	Population did not have a majority with chronic low back pain (subacute)
Rasmussen 2015	Population did not have low back pain at baseline
Rhon 2018	Population did not have a majority with nonspecific chronic low back pain
Schaller 2017	Not judged to be exercise treatment (movement coaching)
Schiltenwolf 2006	Population did not have chronic low back pain (subacute)
Sjogren 2006b	No comparison for exercise group
Suni 2006	Population did not have a majority with chronic low back pain
Suni 2017	Population did not have a majority with nonspecific chronic low back pain
Tang 2016	Population did not have nonspecific chronic low back pain
Telles 2016	Population did not have nonspecific chronic low back pain
Tsauo 2009	Article judged to contain plagiarism and included inadequate reporting of CONSORT items
Wand 2004	Population did not have chronic low back pain (acute)
Waqqar 2016	Population did not have chronic low back pain
Whitfill 2010	Population did not have chronic low back pain (acute)
Woo 2016	Article judged to be published in a predatory journal, was rated as having high risk of bias and included inadequate reporting of CONSORT items
Wright 2005	Population did not have chronic low back pain
Wu 2004	Population did not have a majority with chronic low back pain (mixed population)
Yilmaz Yelvar 2017	No comparison for exercise group (exercises did not differ between groups)
Yoo 2012	Article judged to be published in a predatory journal, was rated as having high risk of bias and included inadequate reporting of CONSORT items
Yoon 2013	Population did not have chronic low back pain
Yu 2016	Article judged to be published in a predatory journal and included inadequate reporting of CONSORT items
Zahin 2018	No comparison for exercise group (exercises did not differ between groups)

Detailed information about excluded studies is available on request.
CONSORT: Consolidated Standards of Reporting Trials

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Characteristics of studies awaiting classification *[ordered by study ID]*

Abadi 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ahmadi 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ahmadizadeh 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ak 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Akodu 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Alfuth 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Alikhajeh 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Almhdawi 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Altinbilek 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Amorim 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ansari 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Areeudomwong 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Areeudomwong 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Barbosa 2013

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Barni 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Barradas 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Batool 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Bello 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Brodsky 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Bruehl 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Buttagat 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Calatayud 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Canaway 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Cavalcanti 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Chen 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Chhabra 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Cimarras-Otal 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Cox 2010

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Das 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Da Silva 2014

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Daulat 2014

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Daulat 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

De Giogrio 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Delicia 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Del Pozo 2011

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Demirel 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Dimer 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Dineshkumar 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ehsani 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Fang 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Fatoye 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ferreira 2017

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

França 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Galan-Martin 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gao 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Gardner 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ghasemi 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Gholami 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gohil 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Groessl 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Gupta 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Harper 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Heidari 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Hosseinifar 2017

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ibrahim 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Jahantiqh 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jinnouchi 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Joseph 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kamali 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kang 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Karimzadeh 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kell 2011

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Khandhar 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kim 2018b

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kim 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kim 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kim 2020a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kim 2020b

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Krishna 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kumar 2013

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kuvacic 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Kwon 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Leeuw 2008

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Leonard 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Li 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Li 2017

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lima 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Liu 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Liu 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Madadi-Shad 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mataran-Penarrocha 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Matheve 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Matos 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Mazloum 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Mbada 2014

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Mbada 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Moncelon 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Nambi 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nambi 2020a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Nambi 2021

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Narouei 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Natraj 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Neyaz 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ning 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Noormohammadpour 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

O'Keeffe 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Oh 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Oka 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Oliveira 2011

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Oliveira 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Oratsch 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Owen 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Pacrez-de-la-Cruz 2017

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Park 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Park 2020a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Patil 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Prado 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Rabiei 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rahbar 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Rajpal 2008

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ramos 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Ranaalsadat 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Raoul 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Rodriguez-Romero 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Rungthip 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Rydeard 2006b

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saiklang 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Salas 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Samir 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Sarker 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sawant 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Schmidt 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Schulz 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Sedaghati 2017

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Seo 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Shabnam 2011

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Shamsi 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Shamsi 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shamsi 2020a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Shariat 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Shebib 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Sheeran 2013

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sherman 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Singh 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Sipaviciene 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Sokhanguei 2017

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Song 2008

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Springrova 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Srivastav 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Srivastava 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Suh 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Surkitt 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Tagliaferri 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Takinaci 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Teixeira 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Teychenne 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Thomas 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Tul 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Turner 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Van Baal 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Vasseljen 2010a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Verbrugghe 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Verbrugghe 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Verra 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Vincent 2014b

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Wahyuddin 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Walsh 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Wang 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Wang 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Waseem 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Wattamwar 2012

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Weissenfels 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Yalfani 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Yang 2015

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Yao 2020

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Yuan 2016

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Zadro 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Zakari 2019

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Zarzycka 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Zhang 2019a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Zhongliang 2018

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Zou 2019a

Methods	Data extraction not yet completed
Participants	
Interventions	
Outcomes	
Notes	

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

DATA AND ANALYSES

Comparison 1. PRIMARY ANALYSES

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Pain (/100): Earliest follow-up	99	9041	Mean Difference (IV, Random, 95% CI)	-11.24 [-13.82, -8.66]
1.1.1 Exercise vs. no treatment or usual care	35	2746	Mean Difference (IV, Random, 95% CI)	-15.22 [-18.26, -12.18]
1.1.2 Exercise vs. other conservative treatment	64	6295	Mean Difference (IV, Random, 95% CI)	-9.10 [-12.63, -5.57]
1.2 Pain (/100): Short-term follow-up (6-12 weeks)	73	7121	Mean Difference (IV, Random, 95% CI)	-11.42 [-14.71, -8.14]
1.2.1 Exercise vs. no treatment or usual care	26	2247	Mean Difference (IV, Random, 95% CI)	-16.36 [-20.32, -12.40]
1.2.2 Exercise vs. other conservative treatment	47	4874	Mean Difference (IV, Random, 95% CI)	-8.61 [-13.08, -4.13]
1.3 Pain (/100): Medium-term follow-up (~6 months)	55	5569	Mean Difference (IV, Random, 95% CI)	-8.77 [-12.03, -5.51]
1.3.1 Exercise vs. no treatment or usual care	17	1491	Mean Difference (IV, Random, 95% CI)	-10.57 [-15.03, -6.12]
1.3.2 Exercise vs. other conservative treatment	38	4078	Mean Difference (IV, Random, 95% CI)	-8.11 [-12.31, -3.90]
1.4 Pain (/100): Long-term follow-up (> 12 months)	24	2988	Mean Difference (IV, Random, 95% CI)	-6.22 [-11.78, -0.66]
1.4.1 Exercise vs. no treatment or usual care	5	498	Mean Difference (IV, Random, 95% CI)	-11.80 [-21.98, -1.62]
1.4.2 Exercise vs. other conservative treatment	19	2490	Mean Difference (IV, Random, 95% CI)	-5.22 [-11.27, 0.84]
1.5 Function (/100): Earliest follow-up	88		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.5.1 Exercise vs. no treatment or usual care	38	2942	Mean Difference (IV, Random, 95% CI)	-6.81 [-8.32, -5.31]
1.5.2 Exercise vs. other conservative treatment	52	6004	Mean Difference (IV, Random, 95% CI)	-4.11 [-5.98, -2.23]
1.6 Function (/100): Short-term follow-up (6-12 weeks)	72		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.6.1 Exercise vs. no treatment or usual care	30	2555	Mean Difference (IV, Random, 95% CI)	-7.42 [-9.24, -5.60]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.6.2 Exercise vs. other conservative treatment	44	5541	Mean Difference (IV, Random, 95% CI)	-3.95 [-5.99, -1.91]
1.7 Function (/100): Medium-term follow-up (~6 months)	49		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.7.1 Exercise vs. no treatment or usual care	22	1831	Mean Difference (IV, Random, 95% CI)	-6.55 [-8.29, -4.80]
1.7.2 Exercise vs. other conservative treatment	29	3531	Mean Difference (IV, Random, 95% CI)	-2.22 [-4.53, 0.10]
1.8 Function (/100): Long-term follow-up (> 12 months)	21	3366	Mean Difference (IV, Random, 95% CI)	-0.72 [-2.80, 1.37]
1.8.1 Exercise vs. no treatment or usual care	6	550	Mean Difference (IV, Random, 95% CI)	-4.40 [-7.29, -1.52]
1.8.2 Exercise vs. other conservative treatment	15	2816	Mean Difference (IV, Random, 95% CI)	0.26 [-2.16, 2.68]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

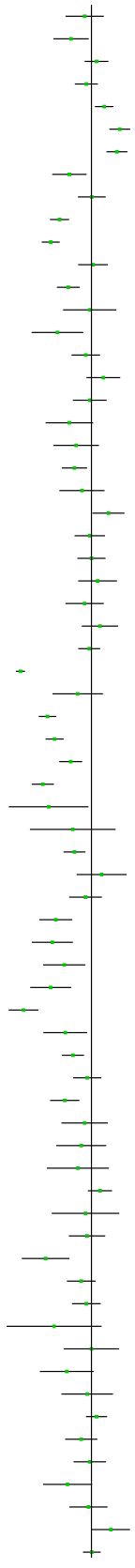
Analysis 1.1. Comparison 1: PRIMARY ANALYSES, Outcome 1: Pain (/100): Earliest follow-up

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.1.1 Exercise vs. no treatment or usual care								
Albaladejo 2010 (1)	60	22.22	100	70	22.22	109	0.9%	-10.00 [-16.03 , -3.97]
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	0.8%	-9.50 [-21.15 , 2.15]
Cabak 2017 (2)	37	17.67	22	56	17.67	46	0.8%	-19.00 [-27.98 , -10.02]
Chen 2014	21.7	14.2	64	34.8	17.7	63	0.9%	-13.10 [-18.69 , -7.51]
Cortell-Tormo 2018	15	15	11	44	14	8	0.7%	-29.00 [-42.14 , -15.86]
Costa 2009b	46	28	77	56	26	77	0.8%	-10.00 [-18.53 , -1.47]
Cruz-Diaz 2017 (3)	21	13.6	34	49.6	13.1	15	0.8%	-28.60 [-36.65 , -20.55]
Cruz-Diaz 2017 (3)	17	14.1	34	49.6	13.1	15	0.8%	-32.60 [-40.75 , -24.45]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	0.8%	-14.90 [-23.47 , -6.33]
Dalichau 2000 (4)	31	17.67	21	55	17.67	11	0.7%	-24.00 [-36.89 , -11.11]
Dalichau 2000 (4)	21	17.67	21	55	17.67	10	0.7%	-34.00 [-47.31 , -20.69]
Garcia 2017	39.5	27.3	73	47	29.7	73	0.8%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	0.9%	-2.00 [-7.75 , 3.75]
Hall 2011 (2)	34	20.31	80	47	20.54	80	0.9%	-13.00 [-19.33 , -6.67]
Highland 2018	27.5	24.3	33	33.5	17.9	29	0.8%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	0.8%	-5.00 [-13.44 , 3.44]
Keane 2017 (3)	26	23.7	10	48.9	33.3	4	0.3%	-22.90 [-58.69 , 12.89]
Keane 2017 (3)	26.5	18.6	10	48.9	33.3	5	0.4%	-22.40 [-53.78 , 8.98]
Kell 2009 (4)	33	5	9	48	7	4	0.8%	-15.00 [-22.60 , -7.40]
Kell 2009 (4)	48	8	9	48	7	5	0.8%	0.00 [-8.06 , 8.06]
Lee 2016a (3)	22	11.3	15	35.8	17.2	3	0.6%	-13.80 [-34.09 , 6.49]
Lee 2016a (3)	33.1	20	15	35.8	17.2	3	0.5%	-2.70 [-24.64 , 19.24]
Masharawi 2013	16.8	8.2	20	38.8	15.4	20	0.8%	-22.00 [-29.65 , -14.35]
Mazloum 2018 (3)	30	9	15	69	16	8	0.8%	-39.00 [-50.99 , -27.01]
Mazloum 2018 (3)	48	11	16	69	16	8	0.7%	-21.00 [-33.33 , -8.67]
Morone 2011	45	23	41	76	21	29	0.8%	-31.00 [-41.39 , -20.61]
Morone 2012 (5)	42.4	21.85	25	72.08	17.19	25	0.8%	-29.68 [-40.58 , -18.78]
Moseley 2002	19	15	29	31	15	28	0.8%	-12.00 [-19.79 , -4.21]
Muharram 2011 (2)	41.2	35.15	78	53.1	37.88	80	0.8%	-11.90 [-23.29 , -0.51]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	0.7%	-16.50 [-28.91 , -4.09]
Oh 2014 (4)	34.4	14.11	9	38	14.11	3	0.6%	-3.60 [-22.04 , 14.84]
Oh 2014 (4)	11.1	14.11	9	38	14.11	3	0.6%	-26.90 [-45.34 , -8.46]
Oh 2014 (4)	14.4	21.23	10	38	14.11	3	0.6%	-23.60 [-44.29 , -2.91]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	0.8%	-39.50 [-49.57 , -29.43]
Puppin 2011	24	27	30	41	27	25	0.7%	-17.00 [-31.33 , -2.67]
Quinn 2011 (2)	30.93	24.06	15	35.2	24.06	14	0.6%	-4.27 [-21.79 , 13.25]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	0.8%	-15.60 [-24.89 , -6.31]
Schinhan 2016	13.3	10.5	15	21	17.3	15	0.8%	-7.70 [-17.94 , 2.54]
Smeets 2008 (3)	42.31	25.56	55	53.35	22.6	17	0.7%	-11.04 [-23.73 , 1.65]
Smeets 2008 (3)	44.63	28.86	52	53.35	22.6	17	0.7%	-8.72 [-22.02 , 4.58]
Smeets 2008 (3)	37.76	24.33	55	53.35	22.6	17	0.7%	-15.59 [-28.11 , -3.07]
Steele 2013 (4)	25	10.97	7	25.91	14.89	3	0.6%	-0.91 [-19.62 , 17.80]
Steele 2013 (4)	16.43	25.76	10	25.91	14.89	4	0.6%	-9.48 [-31.11 , 12.15]
Teut 2016 (4)	39.04	19.76	55	44.05	18.63	26	0.8%	-5.01 [-13.87 , 3.85]
Teut 2016 (4)	37.56	19.94	56	44.05	18.63	27	0.8%	-6.49 [-15.25 , 2.27]
Williams 2009 (2)	24.3	17.9	43	36.9	19.81	47	0.8%	-12.60 [-20.39 , -4.81]
Yelland 2004	42	26.5	55	37	45.4	55	0.7%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**1531**			**1215**	**34.9%**	**-15.22 [-18.26 , -12.18]**
Heterogeneity: Tau² = 75.46; Chi² = 180.77, df = 46 (P < 0.00001); I² = 75%								
Test for overall effect: Z = 9.83 (P < 0.00001)								
1.1.2 Exercise vs. other conservative treatment								
Alayat 2014 (3)	37.1	13	24	56.5	10.4	10	0.8%	-19.40 [-27.68 , -11.12]
Alayat 2014 (3)	26.4	12.5	28	56.5	10.4	10	0.8%	-30.10 [-38.04 , -22.16]
Alexandre 2001	6	27.2	14	37	35.61	14	0.5%	-31.00 [-54.47 , -7.53]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	0.8%	-23.10 [-31.54 , -14.66]
Barberini 2011 (2)	52	17.38	10	81	17.38	12	0.7%	-29.00 [-43.59 , -14.41]
Bellido-Fernandez 2018 (3)	30	18.52	9	30	22.22	5	0.5%	0.00 [-22.93 , 22.93]
Bellido-Fernandez 2018 (3)	30	22.22	9	30	22.22	4	0.5%	0.00 [-26.17 , 26.17]
Bendix 1995 (3)	27	23	40	56	29	18	0.7%	-29.00 [-44.18 , -13.82]
Bendix 1995 (3)	44	29	31	56	29	17	0.6%	-12.00 [-29.15 , 5.15]
Bid 2017	5.62	8.7	61	30.6	14.2	64	0.9%	-24.98 [-29.09 , -20.87]
Bronfort 1996 (4)	35	22	40	39	21	21	0.8%	-4.00 [-15.28 , 7.28]
Bronfort 1996 (4)	27	20	56	39	21	21	0.8%	-12.00 [-22.40 , -1.60]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.1. (Continued)

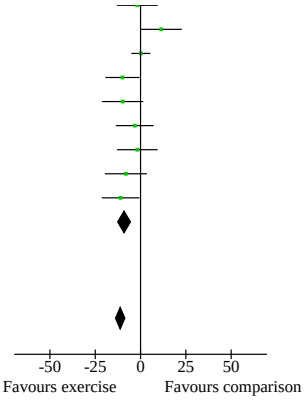
Bronfort 1996 (4)	35	22	40	39	21	21	0.8%	-4.00 [-15.28 , 7.28]
Bronfort 1996 (4)	27	20	56	39	21	21	0.8%	-12.00 [-22.40 , -1.60]
Bronfort 2011 (3)	32	22	91	29	19	49	0.8%	3.00 [-3.98 , 9.98]
Bronfort 2011 (3)	26	21	93	29	19	50	0.9%	-3.00 [-9.78 , 3.78]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	0.9%	7.50 [2.05 , 12.95]
Cecchi 2010 (3)	25	20	68	8.33	11.67	35	0.9%	16.67 [10.54 , 22.80]
Cecchi 2010 (3)	23.33	20	68	8.33	11.67	34	0.9%	15.00 [8.84 , 21.16]
Chan 2017	30	23	49	43	26	43	0.8%	-13.00 [-23.09 , -2.91]
Chown 2008 (5)	73.4	17.2	22	73.2	11.5	33	0.8%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016 (2)	38.1	12.1	53	56.9	16.3	48	0.9%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	0.9%	-24.00 [-29.30 , -18.70]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	0.8%	0.90 [-7.83 , 9.63]
Diaz-Arribas 2009 (2)	23	18.5	63	36.6	20	63	0.9%	-13.60 [-20.33 , -6.87]
Djavid 2007 (3)	43	16	18	44	20	8	0.7%	-1.00 [-16.71 , 14.71]
Djavid 2007 (3)	24	14	19	44	20	8	0.7%	-20.00 [-35.22 , -4.78]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	0.8%	-3.32 [-11.72 , 5.08]
Ferreira 2007 (3)	48	24	73	41	26	38	0.8%	7.00 [-2.93 , 16.93]
Ferreira 2007 (3)	40	25	74	41	26	39	0.8%	-1.00 [-10.95 , 8.95]
Fransoo 2006 (4)	7	17.67	19	20	17.67	9	0.7%	-13.00 [-27.01 , 1.01]
Fransoo 2006 (4)	11	17.67	20	20	17.67	10	0.7%	-9.00 [-22.41 , 4.41]
Frost 1995	12.1	9.9	36	22.1	20.1	35	0.8%	-10.00 [-17.40 , -2.60]
Goldby 2006 (5)	28.81	28.14	78	34.4	36.43	37	0.7%	-5.59 [-18.89 , 7.71]
Gur 2003 (3)	29	13	25	19	14	12	0.8%	10.00 [0.58 , 19.42]
Gur 2003 (3)	18	12	25	19	14	13	0.8%	-1.00 [-9.95 , 7.95]
Hansen 1993 (1)	33.33	21.25	44	33.33	21.25	59	0.8%	0.00 [-8.30 , 8.30]
Hartvigsen 2010 (3)	45.6	19.82	32	42	21.3	21	0.8%	3.60 [-7.81 , 15.01]
Hartvigsen 2010 (3)	37.97	19.38	36	42	21.3	20	0.8%	-4.03 [-15.31 , 7.25]
Hemmilä 1997 (1)	30	22.63	35	25	22.63	34	0.8%	5.00 [-5.68 , 15.68]
Jaromi 2012 (2)	5.75	17.38	56	7.07	17.38	55	0.9%	-1.32 [-7.79 , 5.15]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	0.9%	-41.79 [-44.48 , -39.10]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	0.7%	-8.10 [-22.94 , 6.74]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	0.9%	-25.90 [-31.06 , -20.74]
Kim 2017	24	5.4	15	45.7	9	15	0.9%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	0.9%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	0.9%	-28.60 [-35.11 , -22.09]
Longo 2016	36.2	22	7	61.4	22.7	7	0.5%	-25.20 [-48.62 , -1.78]
Machado 2007	47	40	17	58	28	11	0.5%	-11.00 [-36.21 , 14.21]
Maul 2005	40	20	74	50	20	74	0.9%	-10.00 [-16.44 , -3.56]
McDonough 2013 (2)	45	27.08	39	39	25.24	17	0.7%	6.00 [-8.70 , 20.70]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	0.8%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	0.8%	-21.00 [-30.72 , -11.28]
Miyamoto 2018 (3)	33	23	73	56	26	22	0.8%	-23.00 [-35.08 , -10.92]
Miyamoto 2018 (3)	40	27	67	56	26	23	0.7%	-16.00 [-28.44 , -3.56]
Miyamoto 2018 (3)	32	24	69	56	26	23	0.8%	-24.00 [-36.04 , -11.96]
Murtezani 2011	20	17	50	60	27	51	0.8%	-40.00 [-48.78 , -31.22]
Natour 2015	40.4	24.2	30	55.8	26.8	30	0.7%	-15.40 [-28.32 , -2.48]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	0.9%	-10.90 [-17.43 , -4.37]
Noori 2011	41.5	15.3	17	44	7.5	13	0.8%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	0.8%	-15.70 [-24.32 , -7.08]
Rantonen 2012 (3)	31	20	43	35	28	20	0.7%	-4.00 [-17.65 , 9.65]
Rantonen 2012 (3)	29	27	43	35	28	20	0.7%	-6.00 [-20.69 , 8.69]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	0.6%	-8.00 [-26.27 , 10.27]
Ritvanen 2007 (2)	28	14.11	28	23	14.11	33	0.8%	5.00 [-2.11 , 12.11]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.6%	-3.50 [-23.38 , 16.38]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	0.8%	-2.60 [-13.31 , 8.11]
Saper 2009	44	21	15	71	18	15	0.7%	-27.00 [-41.00 , -13.00]
Saper 2017 (3)	50	21	127	56	22	32	0.8%	-6.00 [-14.45 , 2.45]
Saper 2017 (3)	53	21	129	56	22	32	0.8%	-3.00 [-11.44 , 5.44]
Schenkman 2009 (5)	19	21	17	41	40	9	0.4%	-22.00 [-49.97 , 5.97]
Segal-Snir 2016	70	30	20	70	19	15	0.7%	0.00 [-16.29 , 16.29]
Shen 2009 (4)	11	17.67	15	25.5	17.67	7	0.7%	-14.50 [-30.35 , 1.35]
Shen 2009 (4)	23	17.67	15	25.5	17.67	8	0.7%	-2.50 [-17.66 , 12.66]
Sorensen 2010	49	22	91	46	21	94	0.9%	3.00 [-3.20 , 9.20]
Soukup 1999 (3)	18	15.7	34	24	17	18	0.8%	-6.00 [-15.46 , 3.46]
Soukup 1999 (3)	23	16	38	24	17	17	0.8%	-1.00 [-10.55 , 8.55]
Sousa 2009	33.5	24.8	27	47.6	28	25	0.7%	-14.10 [-28.52 , 0.32]
Torstensen 1998 (4)	37.2	25.3	70	39	28	33	0.8%	-1.80 [-13.04 , 9.44]
Torstensen 1998 (4)	50.4	27.2	71	39	28	34	0.8%	11.40 [0.06 , 22.74]
Ulger 2017	21	16	56	20.8	12	57	0.9%	0.20 [-5.02 , 5.42]



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.1. (Continued)

Torstensen 1998 (4)	50.4	27.2	71	39	28	34	0.8%	11.40 [0.06 , 22.74]
Ulger 2017	21	16	56	20.8	12	57	0.9%	0.20 [-5.02 , 5.42]
Valenza 2017 (2)	39	17.67	27	49	17.67	27	0.8%	-10.00 [-19.43 , -0.57]
Weifen 2013 (4)	22.5	17.67	38	32.4	17.67	12	0.8%	-9.90 [-21.37 , 1.57]
Weifen 2013 (4)	29.2	17.67	141	32.4	17.67	12	0.8%	-3.20 [-13.61 , 7.21]
Weifen 2013 (4)	30.6	17.67	47	32.4	17.67	12	0.8%	-1.80 [-13.00 , 9.40]
Weifen 2013 (4)	24.3	17.67	47	32.4	17.67	11	0.8%	-8.10 [-19.70 , 3.50]
Williams 2005	10	11	20	21	23	24	0.8%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**3758**			**2537**	**65.1%**	**-9.10 [-12.63 , -5.57]**
Heterogeneity: Tau² = 240.03; Chi² = 1169.79, df = 84 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 5.06 (P < 0.00001)								
Total (95% CI)			**5289**			**3752**	**100.0%**	**-11.24 [-13.82 , -8.66]**
Heterogeneity: Tau² = 191.37; Chi² = 1353.87, df = 131 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 8.54 (P < 0.00001)								
Test for subgroup differences: Chi² = 6.64, df = 1 (P = 0.010), I² = 84.9%								



Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Standard deviation(s) required imputation due to missing or implausible value.
(3) Study has multiple exercise groups available: comparison group data divided.
(4) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(5) Study has other comparison groups available for this meta-analysis.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

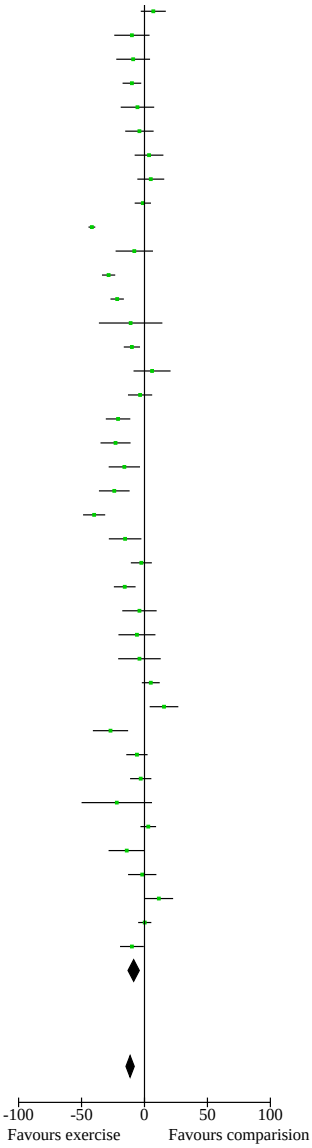
Analysis 1.2. Comparison 1: PRIMARY ANALYSES, Outcome 2: Pain (/100): Short-term follow-up (6-12 weeks)

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.2.1 Exercise vs. no treatment or usual care								
Albaladejo 2010 (1)	60	22.22	100	70	22.22	109	1.1%	-10.00 [-16.03 , -3.97]
Cabak 2017 (2)	37	17.67	22	56	17.67	46	1.1%	-19.00 [-27.98 , -10.02]
Cortell-Tormo 2018	15	15	11	44	14	8	1.0%	-29.00 [-42.14 , -15.86]
Costa 2009b	46	28	77	56	26	77	1.1%	-10.00 [-18.53 , -1.47]
Cruz-Diaz 2017 (3)	21	13.6	34	49.6	13.1	15	1.1%	-28.60 [-36.65 , -20.55]
Cruz-Diaz 2017 (3)	17	14.1	34	49.6	13.1	15	1.1%	-32.60 [-40.75 , -24.45]
Dalichau 2000 (4)	21	17.67	21	55	17.67	10	1.0%	-34.00 [-47.31 , -20.69]
Dalichau 2000 (4)	31	17.67	21	55	17.67	11	1.0%	-24.00 [-36.89 , -11.11]
Garcia 2017	39.5	27.3	73	47	29.7	73	1.1%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	1.2%	-2.00 [-7.75 , 3.75]
Hall 2011 (2)	34	20.31	80	47	20.54	80	1.1%	-13.00 [-19.33 , -6.67]
Highland 2018	27.5	24.3	33	33.5	17.9	29	1.1%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	1.1%	-5.00 [-13.44 , 3.44]
Keane 2017 (3)	26	23.7	10	48.9	33.3	4	0.5%	-22.90 [-58.69 , 12.89]
Keane 2017 (3)	26.5	18.6	10	48.9	33.3	5	0.6%	-22.40 [-53.78 , 8.98]
Lee 2016a (3)	33.1	20	15	35.8	17.2	3	0.8%	-2.70 [-24.64 , 19.24]
Lee 2016a (3)	22	11.3	15	35.8	17.2	3	0.8%	-13.80 [-34.09 , 6.49]
Mazloum 2018 (3)	30	9	15	69	16	8	1.0%	-39.00 [-50.99 , -27.01]
Mazloum 2018 (3)	48	11	16	69	16	8	1.0%	-21.00 [-33.33 , -8.67]
Morone 2011	45	23	41	76	21	29	1.1%	-31.00 [-41.39 , -20.61]
Morone 2012 (5)	42.4	21.85	25	72.08	17.19	25	1.1%	-29.68 [-40.58 , -18.78]
Muharram 2011 (2)	41.2	35.15	78	53.1	37.88	80	1.0%	-11.90 [-23.29 , -0.51]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	1.0%	-16.50 [-28.91 , -4.09]
Oh 2014 (4)	14.4	21.23	9	38	14.11	3	0.8%	-23.60 [-44.75 , -2.45]
Oh 2014 (4)	11.1	14.11	10	38	14.11	3	0.9%	-26.90 [-45.10 , -8.70]
Oh 2014 (4)	34.4	14.11	9	38	14.11	3	0.9%	-3.60 [-22.04 , 14.84]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	1.1%	-39.50 [-49.57 , -29.43]
Puppin 2011	15	16	30	38	24	25	1.1%	-23.00 [-34.01 , -11.99]
Quinn 2011 (2)	30.93	24.06	15	35.2	24.06	14	0.9%	-4.27 [-21.79 , 13.25]
Smeets 2008 (3)	37.76	24.33	52	53.35	22.6	17	1.0%	-15.59 [-28.21 , -2.97]
Smeets 2008 (3)	42.31	25.56	55	53.35	22.6	17	1.0%	-11.04 [-23.73 , 1.65]
Smeets 2008 (3)	44.63	28.86	55	53.35	22.6	17	1.0%	-8.72 [-21.90 , 4.46]
Steele 2013 (4)	16.43	25.76	7	25.91	14.89	3	0.7%	-9.48 [-34.94 , 15.98]
Steele 2013 (4)	25	10.97	10	25.91	14.89	4	0.9%	-0.91 [-17.01 , 15.19]
Teut 2016 (4)	39.04	19.76	55	44.05	18.63	26	1.1%	-5.01 [-13.87 , 3.85]
Teut 2016 (4)	37.56	19.94	56	44.05	18.63	27	1.1%	-6.49 [-15.25 , 2.27]
Yelland 2004	42	26.5	55	37	45.4	55	1.0%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**1276**			**971**	**36.4%**	**-16.36 [-20.32 , -12.40]**
Heterogeneity: Tau² = 105.95; Chi² = 163.43, df = 36 (P < 0.00001); I² = 78%								
Test for overall effect: Z = 8.10 (P < 0.00001)								
1.2.2 Exercise vs. other conservative treatment								
Alayat 2014 (3)	37.1	13	24	56.5	10.4	10	1.1%	-19.40 [-27.68 , -11.12]
Alayat 2014 (3)	26.4	12.5	28	56.5	10.4	10	1.1%	-30.10 [-38.04 , -22.16]
Barberini 2011 (2)	54	17.38	10	64	17.38	12	1.0%	-10.00 [-24.59 , 4.59]
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.2%	-24.98 [-29.09 , -20.87]
Bronfort 1996 (4)	35	22	56	39	21	21	1.1%	-4.00 [-14.67 , 6.67]
Bronfort 1996 (4)	27	20	40	39	21	21	1.1%	-12.00 [-22.91 , -1.09]
Bronfort 2011 (3)	32	22	93	29	19	50	1.1%	3.00 [-3.91 , 9.91]
Bronfort 2011 (3)	26	21	91	29	19	49	1.1%	-3.00 [-9.85 , 3.85]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.2%	7.50 [2.05 , 12.95]
Cecchi 2010 (3)	25	20	68	8.33	11.67	34	1.1%	16.67 [10.51 , 22.83]
Cecchi 2010 (3)	23.33	20	68	8.33	11.67	35	1.1%	15.00 [8.87 , 21.13]
Chan 2017	30	23	49	43	26	43	1.1%	-13.00 [-23.09 , -2.91]
Chown 2008 (5)	73.4	17.2	22	73.2	11.5	33	1.1%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016 (2)	38.1	12.1	53	56.9	16.3	48	1.2%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.2%	-24.00 [-29.30 , -18.70]
Diaz-Arribas 2009 (2)	23	18.5	63	36.6	20	63	1.1%	-13.60 [-20.33 , -6.87]
Djavid 2007 (3)	46	17	18	60	16	8	1.0%	-14.00 [-27.59 , -0.41]
Djavid 2007 (3)	53	21	19	60	16	8	1.0%	-7.00 [-21.56 , 7.56]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.1%	-3.32 [-11.72 , 5.08]
Ferreira 2007 (3)	40	25	74	41	26	39	1.1%	-1.00 [-10.95 , 8.95]
Ferreira 2007 (3)	48	24	73	41	26	38	1.1%	7.00 [-2.93 , 16.93]
Fransoo 2006 (4)	15	17.67	19	25	17.67	9	1.0%	-10.00 [-24.01 , 4.01]
Fransoo 2006 (4)	16	17.67	20	25	17.67	10	1.0%	-9.00 [-22.41 , 4.41]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.2. (Continued)

Ferreira 2007 (3)	46	24	73	41	26	38	1.1%	7.00 [-2.93 , 16.93]
Fransoo 2006 (4)	15	17.67	19	25	17.67	9	1.0%	-10.00 [-24.01 , 4.01]
Fransoo 2006 (4)	16	17.67	20	25	17.67	10	1.0%	-9.00 [-22.41 , 4.41]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.1%	-10.00 [-17.40 , -2.60]
Goldby 2006 (5)	28.81	28.14	78	34.4	36.43	37	1.0%	-5.59 [-18.89 , 7.71]
Hartvigsen 2010 (3)	37.97	19.38	36	42	21.3	20	1.0%	-4.03 [-15.31 , 7.25]
Hartvigsen 2010 (3)	45.6	19.82	32	42	21.3	21	1.0%	3.60 [-7.81 , 15.01]
Hemmilä 1997 (1)	30	22.63	35	25	22.63	34	1.1%	5.00 [-5.68 , 15.68]
Jaromi 2012 (2)	5.75	17.38	56	7.07	17.38	55	1.1%	-1.32 [-7.79 , 5.15]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.2%	-41.79 [-44.48 , -39.10]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	1.0%	-8.10 [-22.94 , 6.74]
Kim 2015a	20.6	8.1	27	49.1	11.1	26	1.2%	-28.50 [-33.75 , -23.25]
Kim 2017	24	5.4	15	45.7	9	15	1.2%	-21.70 [-27.01 , -16.39]
Machado 2007	47	40	17	58	28	11	0.7%	-11.00 [-36.21 , 14.21]
Maul 2005	40	20	74	50	20	74	1.1%	-10.00 [-16.44 , -3.56]
McDonough 2013 (2)	45	27.08	39	39	25.24	17	1.0%	6.00 [-8.70 , 20.70]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.1%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	1.1%	-21.00 [-30.72 , -11.28]
Miyamoto 2018 (3)	33	23	69	56	26	23	1.0%	-23.00 [-34.93 , -11.07]
Miyamoto 2018 (3)	40	27	67	56	26	23	1.0%	-16.00 [-28.44 , -3.56]
Miyamoto 2018 (3)	32	24	73	56	26	22	1.0%	-24.00 [-36.18 , -11.82]
Murtezani 2011	20	17	50	60	27	51	1.1%	-40.00 [-48.78 , -31.22]
Natour 2015	40.4	24.2	30	55.8	26.8	30	1.0%	-15.40 [-28.32 , -2.48]
Noori 2011	41.5	15.3	17	44	7.5	13	1.1%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.1%	-15.70 [-24.32 , -7.08]
Rantonen 2012 (3)	31	20	43	35	28	20	1.0%	-4.00 [-17.65 , 9.65]
Rantonen 2012 (3)	29	27	43	35	28	20	1.0%	-6.00 [-20.69 , 8.69]
Rasmussen-Barr 2003	20	17.65	22	24	33.82	19	0.9%	-4.00 [-20.90 , 12.90]
Ritvanen 2007 (2)	28	14.11	28	23	14.11	33	1.1%	5.00 [-2.11 , 12.11]
Ryan 2010	23.9	23.3	18	8.4	7.5	16	1.0%	15.50 [4.13 , 26.87]
Saper 2009	44	21	15	71	18	15	1.0%	-27.00 [-41.00 , -13.00]
Saper 2017 (3)	50	21	127	56	22	32	1.1%	-6.00 [-14.45 , 2.45]
Saper 2017 (3)	53	21	129	56	22	32	1.1%	-3.00 [-11.44 , 5.44]
Schenkman 2009 (5)	19	21	17	41	40	9	0.6%	-22.00 [-49.97 , 5.97]
Sorensen 2010	49	22	91	46	21	94	1.1%	3.00 [-3.20 , 9.20]
Sousa 2009	33.5	24.8	27	47.6	28	25	1.0%	-14.10 [-28.52 , 0.32]
Torstensen 1998 (4)	37.2	25.3	70	39	28	33	1.0%	-1.80 [-13.04 , 9.44]
Torstensen 1998 (4)	50.4	27.2	71	39	28	34	1.0%	11.40 [0.06 , 22.74]
Ulger 2017	21	16	56	20.8	12	57	1.2%	0.20 [-5.02 , 5.42]
Valenza 2017 (2)	39	17.67	27	49	17.67	27	1.1%	-10.00 [-19.43 , -0.57]
Subtotal (95% CI)			**2882**			**1992**	**63.6%**	**-8.61 [-13.08 , -4.13]**
Heterogeneity: Tau² = 282.67; Chi² = 1085.63, df = 59 (P < 0.00001); I² = 95%								
Test for overall effect: Z = 3.77 (P = 0.0002)								
Total (95% CI)			**4158**			**2963**	**100.0%**	**-11.42 [-14.71 , -8.14]**
Heterogeneity: Tau² = 235.38; Chi² = 1252.53, df = 96 (P < 0.00001); I² = 92%								
Test for overall effect: Z = 6.82 (P < 0.00001)								
Test for subgroup differences: Chi² = 6.47, df = 1 (P = 0.01), I² = 84.5%								



Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Standard deviation(s) required imputation due to missing or implausible value.
(3) Study has multiple exercise groups available: comparison group data divided.
(4) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(5) Study has other comparison groups available for this meta-analysis.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

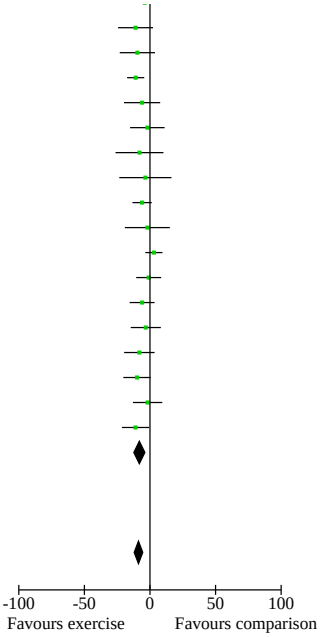
Analysis 1.3. Comparison 1: PRIMARY ANALYSES, Outcome 3: Pain (/100): Medium-term follow-up (~6 months)

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.3.1 Exercise vs. no treatment or usual care								
Albaladejo 2010 (1)	50	14.81	100	80	22.22	109	1.6%	-30.00 [-35.08 , -24.92]
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	1.4%	-9.50 [-21.15 , 2.15]
Chen 2014	21.7	14.2	64	34.8	17.7	63	1.6%	-13.10 [-18.69 , -7.51]
Costa 2009b	50	29	77	56	25	77	1.5%	-6.00 [-14.55 , 2.55]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	1.5%	-14.90 [-23.47 , -6.33]
Dalichau 2000 (2)	52	17.67	21	56	17.67	10	1.3%	-4.00 [-17.31 , 9.31]
Dalichau 2000 (2)	33	17.67	21	56	17.67	11	1.3%	-23.00 [-35.89 , -10.11]
Garcia 2017	44.7	28.4	73	50.3	29	73	1.5%	-5.60 [-14.91 , 3.71]
Highland 2018	27.9	23.4	31	28.6	20.1	28	1.4%	-0.70 [-11.80 , 10.40]
Kell 2009 (2)	33	5	9	48	7	5	1.5%	-15.00 [-21.95 , -8.05]
Kell 2009 (2)	48	8	9	48	7	4	1.5%	0.00 [-8.62 , 8.62]
Morone 2011	44	25	41	65	19	29	1.4%	-21.00 [-31.31 , -10.69]
Morone 2012 (3)	44.4	25.01	25	63.6	20.18	25	1.3%	-19.20 [-31.80 , -6.60]
Nassif 2011	31.5	23	29	35.3	24.7	23	1.3%	-3.80 [-16.91 , 9.31]
Puppin 2011	24	27	30	41	27	25	1.3%	-17.00 [-31.33 , -2.67]
Schinhan 2016	13.3	10.5	15	21	17.3	15	1.4%	-7.70 [-17.94 , 2.54]
Teut 2016 (2)	34.14	20.93	54	41.25	19.22	27	1.5%	-7.11 [-16.26 , 2.04]
Teut 2016 (2)	42.05	20.25	53	41.25	19.22	26	1.5%	0.80 [-8.38 , 9.98]
Williams 2009 (4)	24.3	17.9	43	36.9	19.81	47	1.5%	-12.60 [-20.39 , -4.81]
Yelland 2004	33	26.5	55	32	30.3	55	1.4%	1.00 [-9.64 , 11.64]
Subtotal (95% CI)			**795**			**696**	**28.7%**	**-10.57 [-15.03 , -6.12]**
Heterogeneity: Tau² = 77.48; Chi² = 89.66, df = 19 (P < 0.00001); I² = 79%								
Test for overall effect: Z = 4.65 (P < 0.00001)								
1.3.2 Exercise vs. other conservative treatment								
Alexandre 2001	6	27.2	14	37	35.61	14	0.9%	-31.00 [-54.47 , -7.53]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.5%	-23.10 [-31.54 , -14.66]
Barberini 2011 (4)	52	17.38	10	81	17.38	12	1.2%	-29.00 [-43.59 , -14.41]
Bendix 1995 (5)	27	23	31	56	29	17	1.2%	-29.00 [-44.99 , -13.01]
Bendix 1995 (5)	44	29	40	56	29	18	1.2%	-12.00 [-28.13 , 4.13]
Bronfort 2011 (5)	31	21	89	33	24	46	1.5%	-2.00 [-10.19 , 6.19]
Bronfort 2011 (5)	29	21	90	33	24	47	1.5%	-4.00 [-12.12 , 4.12]
Cambron 2006	23.8	20.65	74	19.2	18.97	90	1.6%	4.60 [-1.52 , 10.72]
Cecchi 2010 (5)	23.33	18.33	68	13.33	11.67	35	1.6%	10.00 [4.18 , 15.82]
Cecchi 2010 (5)	23.33	16.67	68	13.33	11.67	34	1.6%	10.00 [4.42 , 15.58]
Chan 2017	31	24	48	40	25	39	1.4%	-9.00 [-19.38 , 1.38]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.5%	0.90 [-7.83 , 9.63]
Diaz-Arribas 2009 (4)	20.4	20.3	63	56.7	20.9	63	1.5%	-36.30 [-43.49 , -29.11]
Djavid 2007 (5)	43	16	18	44	20	8	1.2%	-1.00 [-16.71 , 14.71]
Djavid 2007 (5)	24	14	19	44	20	8	1.2%	-20.00 [-35.22 , -4.78]
Dougherty 2014	45.4	30.34	89	38.47	26.53	92	1.5%	6.93 [-1.38 , 15.24]
Ferreira 2007 (5)	48	26	68	43	26	36	1.4%	5.00 [-5.50 , 15.50]
Ferreira 2007 (5)	43	26	71	43	26	36	1.4%	0.00 [-10.43 , 10.43]
Fransoo 2006 (2)	7	17.67	20	20	17.67	10	1.3%	-13.00 [-26.41 , 0.41]
Fransoo 2006 (2)	11	17.67	19	20	17.67	9	1.3%	-9.00 [-23.01 , 5.01]
Goldby 2006 (3)	23.16	27.43	73	30.25	31.68	40	1.4%	-7.09 [-18.75 , 4.57]
Hansen 1993 (1)	44.44	21.25	42	44.44	21.25	55	1.5%	0.00 [-8.53 , 8.53]
Hartvigsen 2010 (5)	40.33	24.24	30	39.1	20	17	1.3%	1.23 [-11.64 , 14.10]
Hartvigsen 2010 (5)	36.89	23.6	30	39.1	20	18	1.3%	-2.21 [-14.73 , 10.31]
Hemmilä 1997 (1)	29	22.63	35	25	22.63	34	1.4%	4.00 [-6.68 , 14.68]
Jaromi 2012 (4)	5.11	17.38	56	59.56	17.38	55	1.6%	-54.45 [-60.92 , -47.98]
Kankaanpaa 1999	35.9	27.4	30	46	20.5	24	1.3%	-10.10 [-22.88 , 2.68]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.6%	-25.90 [-31.06 , -20.74]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.6%	-12.20 [-18.94 , -5.46]
Machado 2007	47	38	14	59	30	13	0.8%	-12.00 [-37.73 , 13.73]
Maul 2005	40	20	70	50	20	71	1.6%	-10.00 [-16.60 , -3.40]
McDonough 2013 (4)	38	25.49	39	41	26.3	17	1.2%	-3.00 [-17.84 , 11.84]
Miyamoto 2013	45	22	43	53	23	43	1.5%	-8.00 [-17.51 , 1.51]
Miyamoto 2018 (5)	44	29	67	54	27	20	1.3%	-10.00 [-23.72 , 3.72]
Miyamoto 2018 (5)	50	29	68	54	27	20	1.3%	-4.00 [-17.69 , 9.69]
Miyamoto 2018 (5)	43	26	69	54	27	20	1.3%	-11.00 [-24.33 , 2.33]
Natour 2015	42	27.8	30	51.6	25.3	30	1.3%	-9.60 [-23.05 , 3.85]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.3. (Continued)

Miyamoto 2018 (5)	43	26	69	54	27	20	1.3%	-11.00 [-24.33 , 2.33]
Natour 2015	42	27.8	30	51.6	25.3	30	1.3%	-9.60 [-23.05 , 3.85]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.6%	-10.90 [-17.43 , -4.37]
Rantonen 2012 (5)	29	26	43	35	26	20	1.3%	-6.00 [-19.79 , 7.79]
Rantonen 2012 (5)	33	22	43	35	26	20	1.3%	-2.00 [-15.16 , 11.16]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	1.1%	-8.00 [-26.27 , 10.27]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	1.0%	-3.50 [-23.38 , 16.38]
Saper 2009	39	6	8	45	12	15	1.5%	-6.00 [-13.36 , 1.36]
Schenkman 2009 (3)	26	26	13	28	14	8	1.1%	-2.00 [-19.14 , 15.14]
Sorensen 2010	48	21	89	45	23	87	1.6%	3.00 [-3.51 , 9.51]
Soukup 1999 (5)	23	16	34	24	17	18	1.5%	-1.00 [-10.52 , 8.52]
Soukup 1999 (5)	18	15.7	38	24	17	17	1.5%	-6.00 [-15.50 , 3.50]
Weifen 2013 (2)	29.2	17.67	38	32.4	17.67	12	1.4%	-3.20 [-14.67 , 8.27]
Weifen 2013 (2)	24.3	17.67	47	32.4	17.67	11	1.4%	-8.10 [-19.70 , 3.50]
Weifen 2013 (2)	22.5	17.67	141	32.4	17.67	12	1.4%	-9.90 [-20.31 , 0.51]
Weifen 2013 (2)	30.6	17.67	47	32.4	17.67	12	1.4%	-1.80 [-13.00 , 9.40]
Williams 2005	10	11	20	21	23	24	1.4%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**2445**			**1633**	**71.3%**	**-8.11 [-12.31 , -3.90]**
Heterogeneity: Tau² = 202.63; Chi² = 491.07, df = 51 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 3.78 (P = 0.0002)								
Total (95% CI)			**3240**			**2329**	**100.0%**	**-8.77 [-12.03 , -5.51]**
Heterogeneity: Tau² = 166.01; Chi² = 594.55, df = 71 (P < 0.00001); I² = 88%								
Test for overall effect: Z = 5.27 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.62, df = 1 (P = 0.43), I² = 0%								



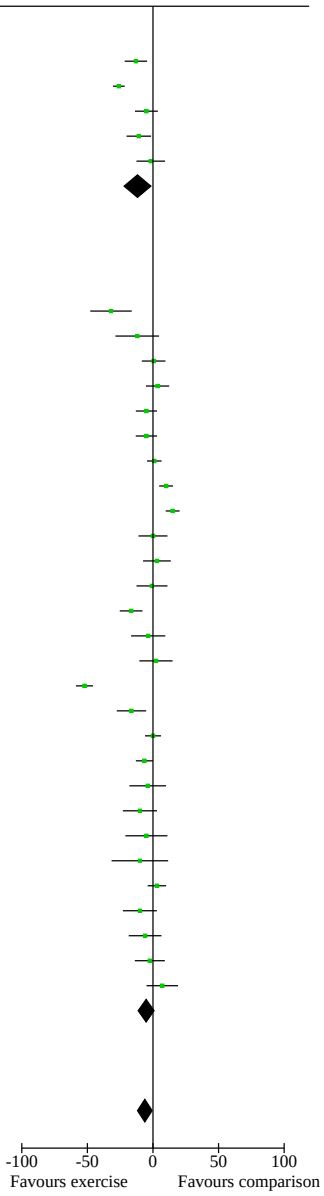
Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(3) Study has other comparison groups available for this meta-analysis.
(4) Standard deviation(s) required imputation due to missing or implausible value.
(5) Study has multiple exercise groups available: comparison group data divided.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.4. Comparison 1: PRIMARY ANALYSES, Outcome 4: Pain (/100): Long-term follow-up (> 12 months)

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.4.1 Exercise vs. no treatment or usual care								
Costa 2009b	50	29	77	63	25	77	3.2%	-13.00 [-21.55 , -4.45]
Cuesta-Vargas 2012	10	8.1	25	36	7.8	24	3.3%	-26.00 [-30.45 , -21.55]
Jensen 2012	43	24	46	48	19	49	3.2%	-5.00 [-13.74 , 3.74]
Williams 2009 (1)	27.7	22.56	43	38.5	22.63	47	3.1%	-10.80 [-20.14 , -1.46]
Yelland 2004	34.1	28.2	55	35.8	30	55	3.0%	-1.70 [-12.58 , 9.18]
Subtotal (95% CI)			**246**			**252**	**15.8%**	**-11.80 [-21.98 , -1.62]**
Heterogeneity: Tau² = 115.90; Chi² = 33.00, df = 4 (P < 0.00001); I² = 88%								
Test for overall effect: Z = 2.27 (P = 0.02)								
1.4.2 Exercise vs. other conservative treatment								
Bendix 1995 (2)	33	34	38	65	25	18	2.7%	-32.00 [-47.82 , -16.18]
Bendix 1995 (2)	53	33	31	65	25	17	2.6%	-12.00 [-28.62 , 4.62]
Bronfort 1996 (3)	29	19	36	28.5	14.11	18	3.1%	0.50 [-8.50 , 9.50]
Bronfort 1996 (3)	32	22	52	28.5	14.11	18	3.1%	3.50 [-5.35 , 12.35]
Bronfort 2011 (2)	28	22	81	33	21	40	3.2%	-5.00 [-13.08 , 3.08]
Bronfort 2011 (2)	28	23	82	33	21	41	3.2%	-5.00 [-13.13 , 3.13]
Cambron 2006	21.6	17.66	78	20.6	19.6	96	3.3%	1.00 [-4.54 , 6.54]
Cecchi 2010 (2)	21.67	15	68	11.67	11.67	35	3.3%	10.00 [4.74 , 15.26]
Cecchi 2010 (2)	26.67	15	68	11.67	11.67	34	3.3%	15.00 [9.70 , 20.30]
Ferreira 2007 (2)	49	29	65	49	26	36	3.0%	0.00 [-11.04 , 11.04]
Ferreira 2007 (2)	52	28	73	49	26	37	3.0%	3.00 [-7.56 , 13.56]
Goldby 2006 (4)	29.23	28.1	71	30	31.68	40	3.0%	-0.77 [-12.56 , 11.02]
Hansen 1993 (5)	27.77	21.25	40	44.44	21.25	55	3.2%	-16.67 [-25.32 , -8.02]
Hartvigsen 2010 (2)	37.96	24	30	41.6	20	16	2.9%	-3.64 [-16.67 , 9.39]
Hartvigsen 2010 (2)	43.89	23.83	31	41.6	20	17	2.9%	2.29 [-10.39 , 14.97]
Jaromi 2012 (1)	8.32	17.38	56	60.51	17.38	55	3.3%	-52.19 [-58.66 , -45.72]
Kankaanpaa 1999	28	18.6	30	44.4	22.5	24	3.0%	-16.40 [-27.60 , -5.20]
Maul 2005	40	10	64	40	20	49	3.3%	0.00 [-6.11 , 6.11]
Niemisto 2003	25.7	23.3	96	32.2	23.3	100	3.3%	-6.50 [-13.03 , 0.03]
Rantonen 2012 (2)	35	27	43	39	26	20	2.8%	-4.00 [-17.96 , 9.96]
Rantonen 2012 (2)	29	21	43	39	26	20	2.9%	-10.00 [-23.01 , 3.01]
Rasmussen-Barr 2003	13	14.71	17	18	29.41	16	2.7%	-5.00 [-21.02 , 11.02]
Schenkman 2009 (4)	27	30	11	37	14	5	2.3%	-10.00 [-31.56 , 11.56]
Sorensen 2010	48	22	78	45	24	86	3.2%	3.00 [-4.04 , 10.04]
Soukup 1999 (2)	22	22	34	32	23	18	2.9%	-10.00 [-22.95 , 2.95]
Soukup 1999 (2)	26	19	38	32	23	17	2.9%	-6.00 [-18.49 , 6.49]
Torstensen 1998 (3)	40.5	24.4	71	42.9	29.5	34	3.0%	-2.40 [-13.83 , 9.03]
Torstensen 1998 (3)	50	28	70	42.9	29.5	33	2.9%	7.10 [-4.91 , 19.11]
Subtotal (95% CI)			**1495**			**995**	**84.2%**	**-5.22 [-11.27 , 0.84]**
Heterogeneity: Tau² = 235.15; Chi² = 336.96, df = 27 (P < 0.00001); I² = 92%								
Test for overall effect: Z = 1.69 (P = 0.09)								
Total (95% CI)			**1741**			**1247**	**100.0%**	**-6.22 [-11.78 , -0.66]**
Heterogeneity: Tau² = 235.74; Chi² = 428.50, df = 32 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 2.19 (P = 0.03)								
Test for subgroup differences: Chi² = 1.19, df = 1 (P = 0.28), I² = 15.8%								



Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value.
(2) Study has multiple exercise groups available: comparison group data divided.
(3) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(4) Study has other comparison groups available for this meta-analysis.
(5) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.5. Comparison 1: PRIMARY ANALYSES, Outcome 5: Function (/100): Earliest follow-up

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.5.1 Exercise vs. no treatment or usual care								
Albaladejo 2010 (1)	31.25	24.69	100	33.33	21.6	109	2.9%	-2.08 [-8.39 , 4.23]
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.1%	-8.50 [-16.99 , -0.01]
Costa 2009b	40	27.08	77	49.58	24.58	77	2.2%	-9.58 [-17.75 , -1.41]
Cruz-Diaz 2017 (2)	19.83	20.42	34	43.37	23.33	15	1.0%	-23.54 [-37.20 , -9.88]
Cruz-Diaz 2017 (2)	26.46	22.08	34	43.37	23.33	15	1.0%	-16.91 [-30.86 , -2.96]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	2.6%	-10.00 [-17.14 , -2.86]
Dalichau 2000 (3)	42	13.92	21	49	13.92	10	1.5%	-7.00 [-17.48 , 3.48]
Dalichau 2000 (3)	40	13.92	21	49	13.92	11	1.6%	-9.00 [-19.15 , 1.15]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	1.9%	-7.83 [-16.79 , 1.13]
Groessl 2017 (4)	30.62	19.06	75	37.54	19.06	75	3.0%	-6.92 [-13.02 , -0.82]
Hall 2011 (4)	29.21	21.49	80	33.75	21.87	80	2.7%	-4.54 [-11.26 , 2.18]
Harts 2008 (2)	25.42	25	20	21.67	16.25	9	0.8%	3.75 [-11.51 , 19.01]
Harts 2008 (2)	14.17	16.67	20	21.67	16.25	10	1.2%	-7.50 [-19.94 , 4.94]
Highland 2018	18.46	19.25	33	29.33	18.96	29	1.8%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	1.8%	0.43 [-8.93 , 9.79]
Keane 2017 (2)	25.8	20.1	10	31.56	18.24	4	0.4%	-5.76 [-27.55 , 16.03]
Keane 2017 (2)	33.2	18.84	10	31.56	18.24	5	0.5%	1.64 [-18.16 , 21.44]
Kell 2009 (3)	35.9	13.27	9	39.1	13.27	4	0.8%	-3.20 [-18.83 , 12.43]
Kell 2009 (3)	24.2	13.27	9	39.1	13.27	5	0.9%	-14.90 [-29.41 , -0.39]
Kofotolis 2006 (3)	20	13.92	28	28	13.92	15	2.0%	-8.00 [-16.73 , 0.73]
Kofotolis 2006 (3)	21.2	13.92	28	28	13.92	15	2.0%	-6.80 [-15.53 , 1.93]
Kofotolis 2016 (2)	33.25	10.17	37	45.25	16.17	14	1.9%	-12.00 [-21.08 , -2.92]
Kofotolis 2016 (2)	17.04	8.08	36	45.25	16.17	14	2.0%	-28.21 [-37.08 , -19.34]
Kuukkanen 2000	10	7.6	27	11	7.8	26	4.2%	-1.00 [-5.15 , 3.15]
Lee 2016a (2)	4.58	3.75	15	7.08	6.67	3	2.3%	-2.50 [-10.28 , 5.28]
Lee 2016a (2)	7.5	8.75	15	7.08	6.67	3	2.0%	0.42 [-8.33 , 9.17]
Little 2014 (4)	27.95	13.92	15	31.08	19.87	13	1.1%	-3.13 [-16.03 , 9.77]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	0.9%	-21.08 [-36.02 , -6.14]
Mazloum 2018 (2)	23.1	7.5	16	26.6	5	8	3.6%	-3.50 [-8.55 , 1.55]
Mazloum 2018 (2)	22.9	3.6	15	26.6	5	8	4.3%	-3.70 [-7.61 , 0.21]
Morone 2011	18	12.9	41	25.8	14.1	29	2.9%	-7.80 [-14.28 , -1.32]
Morone 2012 (5)	13.68	9.39	25	25.28	14.03	25	2.8%	-11.60 [-18.22 , -4.98]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	1.4%	-4.50 [-15.82 , 6.82]
Patti 2016	13	8	19	16.8	15.6	19	2.3%	-3.80 [-11.68 , 4.08]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	2.7%	-8.34 [-15.23 , -1.45]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	2.3%	-7.10 [-14.82 , 0.62]
Quinn 2011 (4)	22.5	19.06	15	31.25	19.06	14	1.0%	-8.75 [-22.63 , 5.13]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	4.2%	-5.00 [-9.08 , -0.92]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	3.6%	-5.06 [-10.19 , 0.07]
Shaughnessy 2004	26	14	20	44	16	21	1.9%	-18.00 [-27.19 , -8.81]
Smeets 2008 (2)	49.6	24.56	55	57.83	19.92	17	1.3%	-8.23 [-19.71 , 3.25]
Smeets 2008 (2)	46.74	23.01	55	57.83	19.92	17	1.4%	-11.09 [-22.34 , 0.16]
Smeets 2008 (2)	47.5	21.88	52	57.83	19.92	17	1.4%	-10.33 [-21.51 , 0.85]
Steele 2013 (3)	14.86	17.95	7	23.2	17.95	3	0.4%	-8.34 [-32.62 , 15.94]
Steele 2013 (3)	17.98	17.95	10	23.2	17.95	4	0.5%	-5.22 [-26.03 , 15.59]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	2.9%	-2.87 [-9.31 , 3.57]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	2.8%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	2.7%	-0.83 [-7.72 , 6.06]
Williams 2009 (4)	17.9	10.49	43	20.8	10.28	47	4.1%	-2.90 [-7.20 , 1.40]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	0.7%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**1664**			**1278**	**100.0%**	**-6.81 [-8.32 , -5.31]**
Heterogeneity: Tau² = 9.64; Chi² = 79.00, df = 49 (P = 0.004); I² = 38%								
Test for overall effect: Z = 8.89 (P < 0.00001)								
1.5.2 Exercise vs. other conservative treatment								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.0%	-12.84 [-15.31 , -10.37]
Bellido-Fernandez 2018 (2)	8	3.7	9	8	6.67	4	1.6%	0.00 [-6.97 , 6.97]
Bellido-Fernandez 2018 (2)	12	5.93	9	8	6.67	5	1.6%	4.00 [-3.01 , 11.01]
Bid 2017	7.29	8.54	64	24.12	12	64	1.9%	-16.83 [-20.44 , -13.22]
Bronfort 1996 (3)	20.9	17	56	20.8	17.3	21	1.4%	0.10 [-8.54 , 8.74]
Bronfort 1996 (3)	15.1	17.4	43	20.8	17.3	21	1.4%	-5.70 [-14.74 , 3.34]
Bronfort 2011 (2)	18.7	19.57	92	21.3	21.74	50	1.6%	-2.60 [-9.83 , 4.63]
Bronfort 2011 (2)	16.96	20	91	21.3	21.74	49	1.6%	-4.34 [-11.68 , 3.00]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.8%	0.83 [-3.79 , 5.45]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

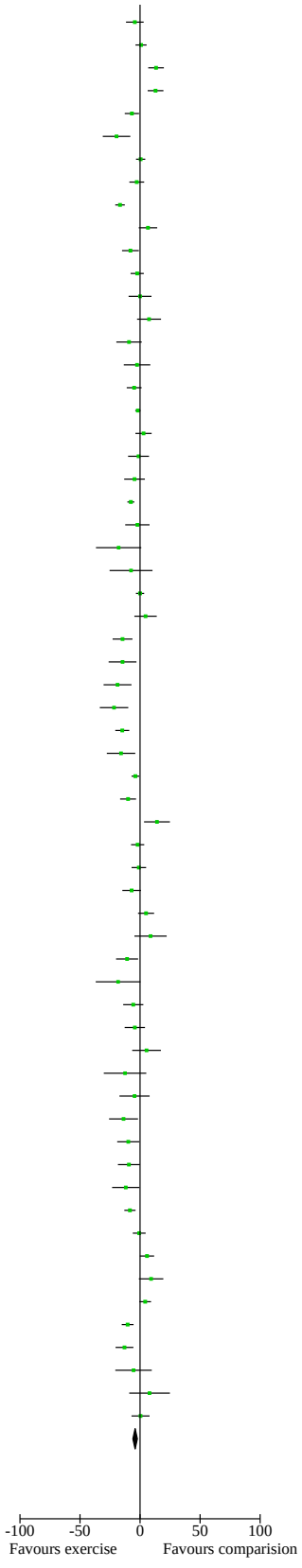
Analysis 1.5. (Continued)

Bronfort 2011 (2)	16.96	20	91	21.3	21.74	49	1.6%	-4.34 [-11.68 , 3.00]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.8%	0.83 [-3.79 , 5.45]
Cecchi 2010 (2)	22.5	19.58	68	9.17	13.75	35	1.6%	13.33 [6.82 , 19.84]
Cecchi 2010 (2)	22.08	19.58	68	9.17	13.75	34	1.6%	12.91 [6.35 , 19.47]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.7%	-6.80 [-12.74 , -0.86]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Chown 2008 (5)	20.8	8.4	35	20.3	8	35	1.9%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016 (4)	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]
Djavid 2007 (2)	48.2	10.4	19	41.6	8.8	8	1.5%	6.60 [-1.08 , 14.28]
Djavid 2007 (2)	33.6	7.4	18	41.6	8.8	8	1.6%	-8.00 [-14.99 , -1.01]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	1.7%	-2.37 [-7.87 , 3.13]
Ferreira 2007 (2)	32.92	23.75	74	32.92	25	39	1.3%	0.00 [-9.53 , 9.53]
Ferreira 2007 (2)	40.42	26.25	73	32.92	25	38	1.3%	7.50 [-2.47 , 17.47]
Fransoo 2006 (3)	4.58	13.92	20	13.75	13.92	10	1.2%	-9.17 [-19.74 , 1.40]
Fransoo 2006 (3)	11.25	13.92	19	13.75	13.92	9	1.2%	-2.50 [-13.54 , 8.54]
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.7%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.0%	-1.80 [-3.96 , 0.36]
Goldby 2006 (5)	31	17.07	78	28.1	17.34	37	1.6%	2.90 [-3.85 , 9.65]
Gur 2003 (2)	26.25	14.58	25	27.5	12.08	13	1.4%	-1.25 [-9.96 , 7.46]
Gur 2003 (2)	22.91	13.33	25	27.5	12.08	12	1.4%	-4.59 [-13.19 , 4.01]
Kim 2017	21.27	3.37	15	28.93	4.61	15	1.9%	-7.66 [-10.55 , -4.77]
Little 2014 (4)	23.16	13.92	14	25.37	13.92	15	1.3%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.7%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	0.7%	-7.50 [-25.37 , 10.37]
Maul 2005	25	8.33	74	25	12.5	74	1.9%	0.00 [-3.42 , 3.42]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]
Miyamoto 2018 (2)	32.5	21.67	69	47.08	25.42	23	1.1%	-14.58 [-26.16 , -3.00]
Miyamoto 2018 (2)	28.33	21.67	67	47.08	25.42	23	1.1%	-18.75 [-30.36 , -7.14]
Miyamoto 2018 (2)	25.42	22.92	73	47.08	25.42	22	1.1%	-21.66 [-33.51 , -9.81]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	1.7%	-14.80 [-20.62 , -8.98]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.1%	-15.83 [-27.67 , -3.99]
Niemisto 2003	14.7	11.6	96	18.6	11	100	1.9%	-3.90 [-7.07 , -0.73]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.6%	-9.97 [-16.60 , -3.34]
Okafor 2012 (6)	25	20.8	15	10.9	4.46	15	1.2%	14.10 [3.33 , 24.87]
Rantonen 2012 (2)	14	11	43	16	10	20	1.7%	-2.00 [-7.48 , 3.48]
Rantonen 2012 (2)	15	14	43	16	10	20	1.7%	-1.00 [-7.06 , 5.06]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	1.5%	-7.00 [-14.73 , 0.73]
Ritvanen 2007 (4)	17	13.27	28	12	13.27	33	1.6%	5.00 [-1.68 , 11.68]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.0%	8.75 [-4.68 , 22.18]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.4%	-10.80 [-19.97 , -1.63]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.7%	-18.27 [-36.90 , 0.36]
Saper 2017 (2)	47.83	21.3	129	53.48	21.74	32	1.4%	-5.65 [-14.03 , 2.73]
Saper 2017 (2)	49.13	22.17	127	53.48	21.74	32	1.4%	-4.35 [-12.81 , 4.11]
Schenkman 2009 (5)	15.81	10.4	17	10.4	16.64	9	1.1%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.7%	-12.50 [-30.19 , 5.19]
Sherman 2005 (2)	22.08	18.75	36	26.67	21.67	15	1.1%	-4.59 [-17.15 , 7.97]
Sherman 2005 (2)	12.92	15	35	26.67	21.67	15	1.1%	-13.75 [-25.79 , -1.71]
Sherman 2011 (3)	18.74	16.28	91	28.52	20.61	22	1.4%	-9.78 [-19.02 , -0.54]
Sherman 2011 (3)	19.26	17.56	92	28.52	20.61	23	1.4%	-9.26 [-18.42 , -0.10]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.2%	-11.88 [-23.25 , -0.51]
Tilbrook 2011 (4)	23.75	20.31	156	32.17	20.31	136	1.8%	-8.42 [-13.09 , -3.75]
Torstensen 1998 (3)	46.2	13.1	71	46.9	13.1	34	1.8%	-0.70 [-6.05 , 4.65]
Torstensen 1998 (3)	52.7	16.6	70	46.9	13.1	33	1.7%	5.80 [-0.12 , 11.72]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.3%	9.20 [-0.98 , 19.38]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	1.8%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	1.8%	-10.33 [-15.25 , -5.41]
Valenza 2017 (4)	16.96	13.92	27	29.92	13.92	27	1.5%	-12.96 [-20.39 , -5.53]
Vincent 2014a (2)	20.83	17.08	18	26.25	17.5	7	0.9%	-5.42 [-20.60 , 9.76]
Vincent 2014a (2)	34.17	22.92	17	26.25	17.5	7	0.8%	7.92 [-9.01 , 24.85]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.5%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**3534**			**2470**	**100.0%**	**-4.11 [-5.98 , -2.23]**

Heterogeneity: Tau² = 44.54; Chi² = 382.92, df = 68 (P < 0.00001); I² = 82%

Test for overall effect: Z = 4.30 (P < 0.0001)

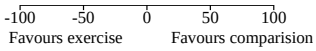
Test for subgroup differences: Chi² = 4.90, df = 1 (P = 0.03), I² = 79.6%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.5. (Continued)

Test for subgroup differences: Chi² = 4.90, df = 1 (P = 0.03), I² = 79.6%



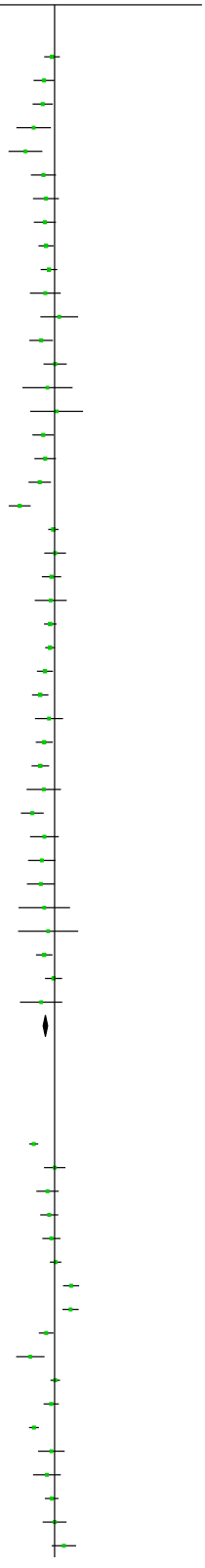
Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Study has multiple exercise groups available: comparison group data divided.
(3) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(4) Standard deviation(s) required imputation due to missing or implausible value.
(5) Study has other comparison groups available for this meta-analysis.
(6) Large baseline differences.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.6. Comparison 1: PRIMARY ANALYSES, Outcome 6: Function (/100): Short-term follow-up (6-12 weeks)

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.6.1 Exercise vs. no treatment or usual care								
Albaladejo 2010 (1)	31.25	24.69	100	33.33	21.6	109	3.6%	-2.08 [-8.39 , 4.23]
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.7%	-8.50 [-16.99 , -0.01]
Costa 2009b	40	27.08	77	49.58	24.58	77	2.8%	-9.58 [-17.75 , -1.41]
Cruz-Diaz 2017 (2)	26.46	22.08	34	43.37	23.33	15	1.4%	-16.91 [-30.86 , -2.96]
Cruz-Diaz 2017 (2)	19.83	20.42	34	43.37	23.33	15	1.4%	-23.54 [-37.20 , -9.88]
Dalichau 2000 (3)	40	13.92	21	49	13.92	11	2.1%	-9.00 [-19.15 , 1.15]
Dalichau 2000 (3)	42	13.92	21	49	13.92	10	2.1%	-7.00 [-17.48 , 3.48]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	2.5%	-7.83 [-16.79 , 1.13]
Groessl 2017 (4)	30.62	19.06	75	37.54	19.06	75	3.7%	-6.92 [-13.02 , -0.82]
Hall 2011 (4)	29.21	21.49	80	33.75	21.87	80	3.4%	-4.54 [-11.26 , 2.18]
Harts 2008 (2)	14.17	16.67	20	21.67	16.25	10	1.6%	-7.50 [-19.94 , 4.94]
Harts 2008 (2)	25.42	25	20	21.67	16.25	9	1.2%	3.75 [-11.51 , 19.01]
Highland 2018	18.46	19.25	33	29.33	18.96	29	2.3%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	2.4%	0.43 [-8.93 , 9.79]
Keane 2017 (2)	25.8	20.1	10	31.56	18.24	5	0.7%	-5.76 [-26.03 , 14.51]
Keane 2017 (2)	33.2	18.84	10	31.56	18.24	4	0.7%	1.64 [-19.71 , 22.99]
Kofotolis 2006 (3)	20.4	13.92	28	29.6	13.92	15	2.6%	-9.20 [-17.93 , -0.47]
Kofotolis 2006 (3)	22	13.92	28	29.6	13.92	15	2.6%	-7.60 [-16.33 , 1.13]
Kofotolis 2016 (2)	33.25	10.17	37	45.25	16.17	14	2.5%	-12.00 [-21.08 , -2.92]
Kofotolis 2016 (2)	17.04	8.08	36	45.25	16.17	14	2.6%	-28.21 [-37.08 , -19.34]
Kuukkanen 2000	10	7.6	27	11	7.8	26	4.8%	-1.00 [-5.15 , 3.15]
Lee 2016a (2)	7.5	8.75	15	7.08	6.67	3	2.6%	0.42 [-8.33 , 9.17]
Lee 2016a (2)	4.58	3.75	15	7.08	6.67	3	3.0%	-2.50 [-10.28 , 5.28]
Little 2014 (4)	27.95	13.92	15	31.08	19.87	13	1.5%	-3.13 [-16.03 , 9.77]
Mazloum 2018 (2)	23.1	7.5	16	26.6	5	8	4.3%	-3.50 [-8.55 , 1.55]
Mazloum 2018 (2)	22.9	3.6	15	26.6	5	8	5.0%	-3.70 [-7.61 , 0.21]
Morone 2011	18	12.9	41	25.8	14.1	29	3.6%	-7.80 [-14.28 , -1.32]
Morone 2012 (5)	13.68	9.39	25	25.28	14.03	25	3.5%	-11.60 [-18.22 , -4.98]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	1.8%	-4.50 [-15.82 , 6.82]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	3.4%	-8.34 [-15.23 , -1.45]
Puppin 2011	13.6	12.3	30	25.1	14.2	25	3.3%	-11.50 [-18.60 , -4.40]
Quinn 2011 (4)	22.5	19.06	15	31.25	19.06	14	1.4%	-8.75 [-22.63 , 5.13]
Shaughnessy 2004	26	14	20	44	16	21	2.4%	-18.00 [-27.19 , -8.81]
Smeets 2008 (2)	49.6	24.56	52	57.83	19.92	17	1.8%	-8.23 [-19.82 , 3.36]
Smeets 2008 (2)	47.5	21.88	55	57.83	19.92	17	1.9%	-10.33 [-21.43 , 0.77]
Smeets 2008 (2)	46.74	23.01	55	57.83	19.92	17	1.9%	-11.09 [-22.34 , 0.16]
Steele 2013 (3)	14.86	17.95	10	23.2	17.95	4	0.7%	-8.34 [-29.15 , 12.47]
Steele 2013 (3)	17.98	17.95	7	23.2	17.95	3	0.5%	-5.22 [-29.50 , 19.06]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	3.5%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	3.4%	-0.83 [-7.72 , 6.06]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	1.0%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**1445**			**1110**	**100.0%**	**-7.42 [-9.24 , -5.60]**
Heterogeneity: Tau² = 13.40; Chi² = 70.46, df = 40 (P = 0.002); I² = 43%								
Test for overall effect: Z = 7.98 (P < 0.00001)								
1.6.2 Exercise vs. other conservative treatment								
Bid 2017	7.29	8.54	64	24.12	12	64	2.2%	-16.83 [-20.44 , -13.22]
Bronfort 1996 (3)	20.9	17	56	20.8	17.3	21	1.7%	0.10 [-8.54 , 8.74]
Bronfort 1996 (3)	15.1	17.4	43	20.8	17.3	21	1.6%	-5.70 [-14.74 , 3.34]
Bronfort 2011 (2)	16.96	20	92	21.3	21.74	50	1.8%	-4.34 [-11.62 , 2.94]
Bronfort 2011 (2)	18.7	19.57	91	21.3	21.74	49	1.8%	-2.60 [-9.90 , 4.70]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	2.1%	0.83 [-3.79 , 5.45]
Cecchi 2010 (2)	22.5	19.58	68	9.17	13.75	35	1.9%	13.33 [6.82 , 19.84]
Cecchi 2010 (2)	22.08	19.58	68	9.17	13.75	34	1.9%	12.91 [6.35 , 19.47]
Chan 2017	17.6	12.6	47	24.4	15.8	43	2.0%	-6.80 [-12.74 , -0.86]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.4%	-19.58 [-31.05 , -8.11]
Chown 2008 (5)	20.8	8.4	35	20.3	8	35	2.2%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016 (4)	16.55	15.78	53	19.29	15.78	48	2.0%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	2.2%	-16.67 [-20.63 , -12.71]
Djavid 2007 (2)	55	13.4	19	57.6	12.8	8	1.4%	-2.60 [-13.32 , 8.12]
Djavid 2007 (2)	51.4	14.8	18	57.6	12.8	8	1.4%	-6.20 [-17.40 , 5.00]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	2.0%	-2.37 [-7.87 , 3.13]
Ferreira 2007 (2)	32.92	23.75	73	32.92	25	38	1.6%	0.00 [-9.64 , 9.64]
Ferreira 2007 (2)	40.42	26.25	74	32.92	25	39	1.5%	7.50 [-2.37 , 17.37]



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

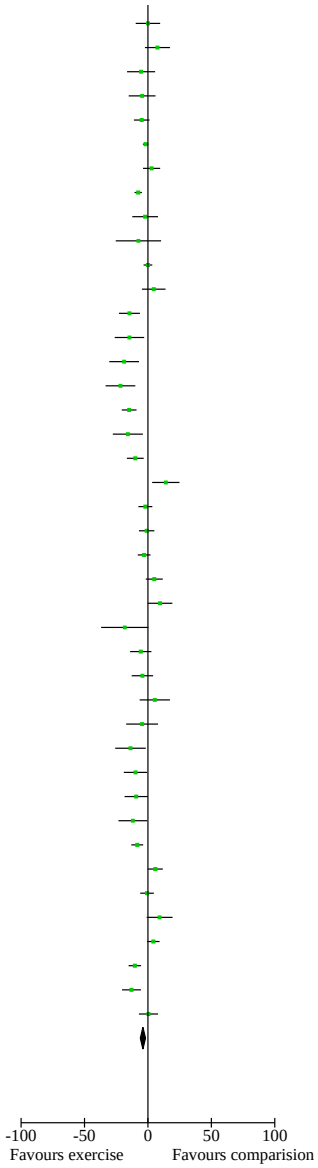
Analysis 1.6. (Continued)

Ferreira 2007 (2)	32.92	23.75	73	32.92	25	38	1.6%	0.00 [-9.64 , 9.64]
Ferreira 2007 (2)	40.42	26.25	74	32.92	25	39	1.5%	7.50 [-2.37 , 17.37]
Fransoo 2006 (3)	12.5	13.92	19	17.91	13.92	9	1.4%	-5.41 [-16.45 , 5.63]
Fransoo 2006 (3)	13.33	13.92	20	17.91	13.92	10	1.5%	-4.58 [-15.15 , 5.99]
Frost 1995	16.3	10.3	29	21.2	14.2	32	2.0%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.3%	-1.80 [-3.96 , 0.36]
Goldby 2006 (5)	31	17.07	78	28.1	17.34	37	1.9%	2.90 [-3.85 , 9.65]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.3%	-7.66 [-10.55 , -4.77]
Little 2014 (4)	23.16	13.92	14	25.37	13.92	15	1.5%	-2.21 [-12.35 , 7.93]
Machado 2007	50	22.92	17	57.5	23.96	11	0.8%	-7.50 [-25.37 , 10.37]
Maul 2005	25	8.33	74	25	12.5	74	2.2%	0.00 [-3.42 , 3.42]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.6%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.7%	-14.58 [-22.85 , -6.31]
Miyamoto 2018 (2)	32.5	21.67	67	47.08	25.42	23	1.3%	-14.58 [-26.19 , -2.97]
Miyamoto 2018 (2)	28.33	21.67	73	47.08	25.42	22	1.3%	-18.75 [-30.48 , -7.02]
Miyamoto 2018 (2)	25.42	22.92	69	47.08	25.42	23	1.3%	-21.66 [-33.37 , -9.95]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	2.0%	-14.80 [-20.62 , -8.98]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.3%	-15.83 [-27.67 , -3.99]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.9%	-9.97 [-16.60 , -3.34]
Okafor 2012 (6)	25	20.8	15	10.9	4.46	15	1.4%	14.10 [3.33 , 24.87]
Rantonen 2012 (2)	14	11	43	16	10	20	2.0%	-2.00 [-7.48 , 3.48]
Rantonen 2012 (2)	15	14	43	16	10	20	2.0%	-1.00 [-7.06 , 5.06]
Rasmussen-Barr 2003	9	10.29	22	12	5.88	19	2.1%	-3.00 [-8.05 , 2.05]
Ritvanen 2007 (4)	17	13.27	28	12	13.27	33	1.9%	5.00 [-1.68 , 11.68]
Ryan 2010	23.33	16.25	18	13.75	12.5	16	1.5%	9.58 [-0.11 , 19.27]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.8%	-18.27 [-36.90 , 0.36]
Saper 2017 (2)	47.83	21.3	129	53.48	21.74	32	1.7%	-5.65 [-14.03 , 2.73]
Saper 2017 (2)	49.13	22.17	127	53.48	21.74	32	1.7%	-4.35 [-12.81 , 4.11]
Schenkman 2009 (5)	15.81	10.4	17	10.4	16.64	9	1.3%	5.41 [-6.53 , 17.35]
Sherman 2005 (2)	22.08	18.75	36	26.67	21.67	15	1.2%	-4.59 [-17.15 , 7.97]
Sherman 2005 (2)	12.92	15	35	26.67	21.67	15	1.3%	-13.75 [-25.79 , -1.71]
Sherman 2011 (3)	18.74	16.28	91	28.52	20.61	22	1.6%	-9.78 [-19.02 , -0.54]
Sherman 2011 (3)	19.26	17.56	92	28.52	20.61	23	1.6%	-9.26 [-18.42 , -0.10]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.4%	-11.88 [-23.25 , -0.51]
Tilbrook 2011 (4)	23.75	20.31	156	32.17	20.31	136	2.1%	-8.42 [-13.09 , -3.75]
Torstensen 1998 (3)	52.7	16.6	70	46.9	13.1	33	2.0%	5.80 [-0.12 , 11.72]
Torstensen 1998 (3)	46.2	13.1	71	46.9	13.1	34	2.0%	-0.70 [-6.05 , 4.65]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.5%	9.20 [-0.98 , 19.38]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	2.1%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	2.1%	-10.33 [-15.25 , -5.41]
Valenza 2017 (4)	16.96	13.92	27	29.92	13.92	27	1.8%	-12.96 [-20.39 , -5.53]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.8%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**3275**			**2266**	**100.0%**	**-3.95 [-5.99 , -1.91]**

Heterogeneity: Tau² = 45.44; Chi² = 318.52, df = 57 (P < 0.00001); I² = 82%

Test for overall effect: Z = 3.80 (P = 0.0001)

Test for subgroup differences: Chi² = 6.18, df = 1 (P = 0.01), I² = 83.8%



Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Study has multiple exercise groups available: comparison group data divided.
(3) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(4) Standard deviation(s) required imputation due to missing or implausible value.
(5) Study has other comparison groups available for this meta-analysis.
(6) Large baseline differences.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.7. Comparison 1: PRIMARY ANALYSES, Outcome 7: Function (/100): Medium-term follow-up (~6 months)

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.7.1 Exercise vs. no treatment or usual care								
Albaladejo 2010 (1)	25	21.6	100	37.5	26.23	109	5.3%	-12.50 [-18.99 , -6.01]
Costa 2009b	42.92	29.17	77	50.83	27.92	77	3.2%	-7.91 [-16.93 , 1.11]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	4.6%	-10.00 [-17.14 , -2.86]
Dalichau 2000 (2)	52	13.92	21	52	13.92	11	2.6%	0.00 [-10.15 , 10.15]
Dalichau 2000 (2)	45	13.92	21	52	13.92	10	2.4%	-7.00 [-17.48 , 3.48]
Garcia 2017	34.71	30.08	73	41.21	30.62	73	2.7%	-6.50 [-16.35 , 3.35]
Groessl 2017 (3)	25.13	19.06	75	39.21	19.06	75	5.8%	-14.08 [-20.18 , -7.98]
Highland 2018	13.54	13.13	31	27.17	22.21	28	2.9%	-13.63 [-23.07 , -4.19]
Kell 2009 (2)	35.9	13.27	9	39.1	13.27	4	1.2%	-3.20 [-18.83 , 12.43]
Kell 2009 (2)	24.2	13.27	9	39.1	13.27	5	1.4%	-14.90 [-29.41 , -0.39]
Kofotolis 2006 (2)	20	13.92	28	28	13.92	15	3.3%	-8.00 [-16.73 , 0.73]
Kofotolis 2006 (2)	21.2	13.92	28	28	13.92	15	3.3%	-6.80 [-15.53 , 1.93]
Kuukkanen 2000	8	6.6	25	12	9	26	8.9%	-4.00 [-8.32 , 0.32]
Little 2014 (3)	23.58	13.92	15	34.04	18.08	13	1.9%	-10.46 [-22.55 , 1.63]
Morone 2011	16.8	14.2	41	26	16.1	29	4.4%	-9.20 [-16.50 , -1.90]
Morone 2012 (4)	14.72	12.59	25	26.08	16.05	25	3.8%	-11.36 [-19.36 , -3.36]
Nassif 2011	41.79	21.33	29	44.17	22.33	23	1.9%	-2.38 [-14.36 , 9.60]
Patti 2016	13	8	19	16.8	15.6	19	3.9%	-3.80 [-11.68 , 4.08]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	4.1%	-7.10 [-14.82 , 0.62]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	7.3%	-5.06 [-10.19 , 0.07]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	5.4%	-2.87 [-9.31 , 3.57]
Vollenbroek-Hutten 2004	41.67	20.83	68	45.83	20.83	72	4.8%	-4.16 [-11.06 , 2.74]
Weiner 2008	33.75	17.92	50	31.25	19.58	50	4.4%	2.50 [-4.86 , 9.86]
Williams 2009 (3)	17.9	10.49	43	20.8	10.28	47	9.0%	-2.90 [-7.20 , 1.40]
Yelland 2004	32.63	33.06	55	41.33	41.33	55	1.4%	-8.70 [-22.69 , 5.29]
Subtotal (95% CI)			**970**			**861**	**100.0%**	**-6.55 [-8.29 , -4.80]**
Heterogeneity: Tau² = 4.04; Chi² = 30.60, df = 24 (P = 0.17); I² = 22%								
Test for overall effect: Z = 7.34 (P < 0.00001)								
1.7.2 Exercise vs. other conservative treatment								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	3.7%	-12.84 [-15.31 , -10.37]
Bronfort 2011 (5)	21.3	22.61	89	21.3	22.61	47	2.7%	0.00 [-7.99 , 7.99]
Bronfort 2011 (5)	18.26	18.26	84	21.3	22.61	46	2.7%	-3.04 [-10.65 , 4.57]
Cambron 2006	14.17	18.4	78	10.83	15.9	91	3.2%	3.34 [-1.89 , 8.57]
Cecchi 2010 (5)	24.17	20.83	68	11.25	14.17	35	2.9%	12.92 [6.10 , 19.74]
Cecchi 2010 (5)	22.5	19.58	68	11.25	14.17	34	2.9%	11.25 [4.59 , 17.91]
Chan 2017	16.5	12.9	47	22.8	18.1	38	2.9%	-6.30 [-13.14 , 0.54]
Djavid 2007 (5)	48.2	10.4	19	41.6	8.8	8	2.7%	6.60 [-1.08 , 14.28]
Djavid 2007 (5)	33.6	7.4	18	41.6	8.8	8	2.9%	-8.00 [-14.99 , -1.01]
Dougherty 2014	26.41	18.87	89	25.6	15.72	92	3.3%	0.81 [-4.26 , 5.88]
Ferreira 2007 (5)	35	26.67	71	32.08	25.83	36	2.2%	2.92 [-7.55 , 13.39]
Ferreira 2007 (5)	42.08	29.17	68	32.08	25.83	36	2.1%	10.00 [-0.92 , 20.92]
Fransoo 2006 (2)	4.58	13.92	20	13.75	13.92	10	2.2%	-9.17 [-19.74 , 1.40]
Fransoo 2006 (2)	11.25	13.92	19	13.75	13.92	9	2.1%	-2.50 [-13.54 , 8.54]
Frost 1995	15.1	8.3	29	23.4	15.2	32	3.1%	-8.30 [-14.37 , -2.23]
Frost 2004	18.23	11.59	144	19.8	10.61	142	3.7%	-1.57 [-4.14 , 1.00]
Goldby 2006 (4)	25.81	17.82	73	23.9	17.75	40	2.9%	1.91 [-4.94 , 8.76]
Little 2014 (3)	21.54	13.92	13	22.12	13.92	15	2.2%	-0.58 [-10.92 , 9.76]
Machado 2007	37.5	27.08	17	50	21.87	13	1.2%	-12.50 [-30.02 , 5.02]
Maul 2005	25	12.5	70	20.83	12.5	71	3.5%	4.17 [0.04 , 8.30]
Miyamoto 2013	18.75	18.75	43	27.92	23.33	43	2.5%	-9.17 [-18.12 , -0.22]
Miyamoto 2018 (5)	26.67	23.33	69	42.5	25.42	20	1.8%	-15.83 [-28.26 , -3.40]
Miyamoto 2018 (5)	32.92	27.08	67	42.5	25.42	20	1.8%	-9.58 [-22.47 , 3.31]
Miyamoto 2018 (5)	36.67	22.92	68	42.5	25.42	20	1.8%	-5.83 [-18.23 , 6.57]
Natour 2015	29.33	22.66	30	28.29	22.25	30	2.0%	1.04 [-10.32 , 12.40]
Niemisto 2003	14.7	11.6	96	18.6	11	100	3.6%	-3.90 [-7.07 , -0.73]
Rantonen 2012 (5)	13	12	43	14	13	20	2.9%	-1.00 [-7.73 , 5.73]
Rantonen 2012 (5)	14	14	43	14	13	20	2.9%	0.00 [-7.07 , 7.07]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	2.7%	-7.00 [-14.73 , 0.73]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.7%	8.75 [-4.68 , 22.18]
Saper 2009	28.71	11.31	8	36.1	12.61	15	2.2%	-7.39 [-17.50 , 2.72]
Schenkman 2009 (4)	14.56	15.39	13	10.4	17.06	8	1.5%	4.16 [-10.32 , 18.64]
Sherman 2005 (5)	13.33	16.25	36	27.92	20	15	2.0%	-14.59 [-26.02 , -3.16]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

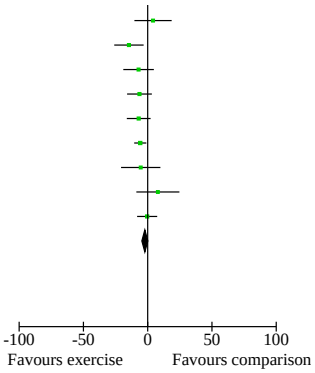
Analysis 1.7. (Continued)

Schenkman 2009 (4)	14.56	15.39	13	10.4	17.06	8	1.5%	4.16 [-10.32 , 18.64]
Sherman 2005 (5)	13.33	16.25	36	27.92	20	15	2.0%	-14.59 [-26.02 , -3.16]
Sherman 2005 (5)	20.83	18.75	35	27.92	20	15	1.9%	-7.09 [-18.97 , 4.79]
Sherman 2011 (2)	18.52	20.31	91	24.91	20.76	22	2.3%	-6.39 [-16.02 , 3.24]
Sherman 2011 (2)	17.91	17.98	92	24.91	20.76	23	2.4%	-7.00 [-16.25 , 2.25]
Tilbrook 2011 (3)	22.58	20.31	156	28.37	20.31	132	3.4%	-5.79 [-10.50 , -1.08]
Vincent 2014a (5)	20.83	17.08	17	26.25	17.5	7	1.4%	-5.42 [-20.72 , 9.88]
Vincent 2014a (5)	34.17	22.92	18	26.25	17.5	7	1.3%	7.92 [-8.82 , 24.66]
Weiner 2008	34.17	22.08	50	34.58	17.5	50	2.7%	-0.41 [-8.22 , 7.40]
Subtotal (95% CI)			**2112**			**1419**	**100.0%**	**-2.22 [-4.53 , 0.10]**

Heterogeneity: Tau² = 35.80; Chi² = 166.97, df = 39 (P < 0.00001); I² = 77%

Test for overall effect: Z = 1.88 (P = 0.06)

Test for subgroup differences: Chi² = 8.56, df = 1 (P = 0.003), I² = 88.3%



Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.

(2) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.

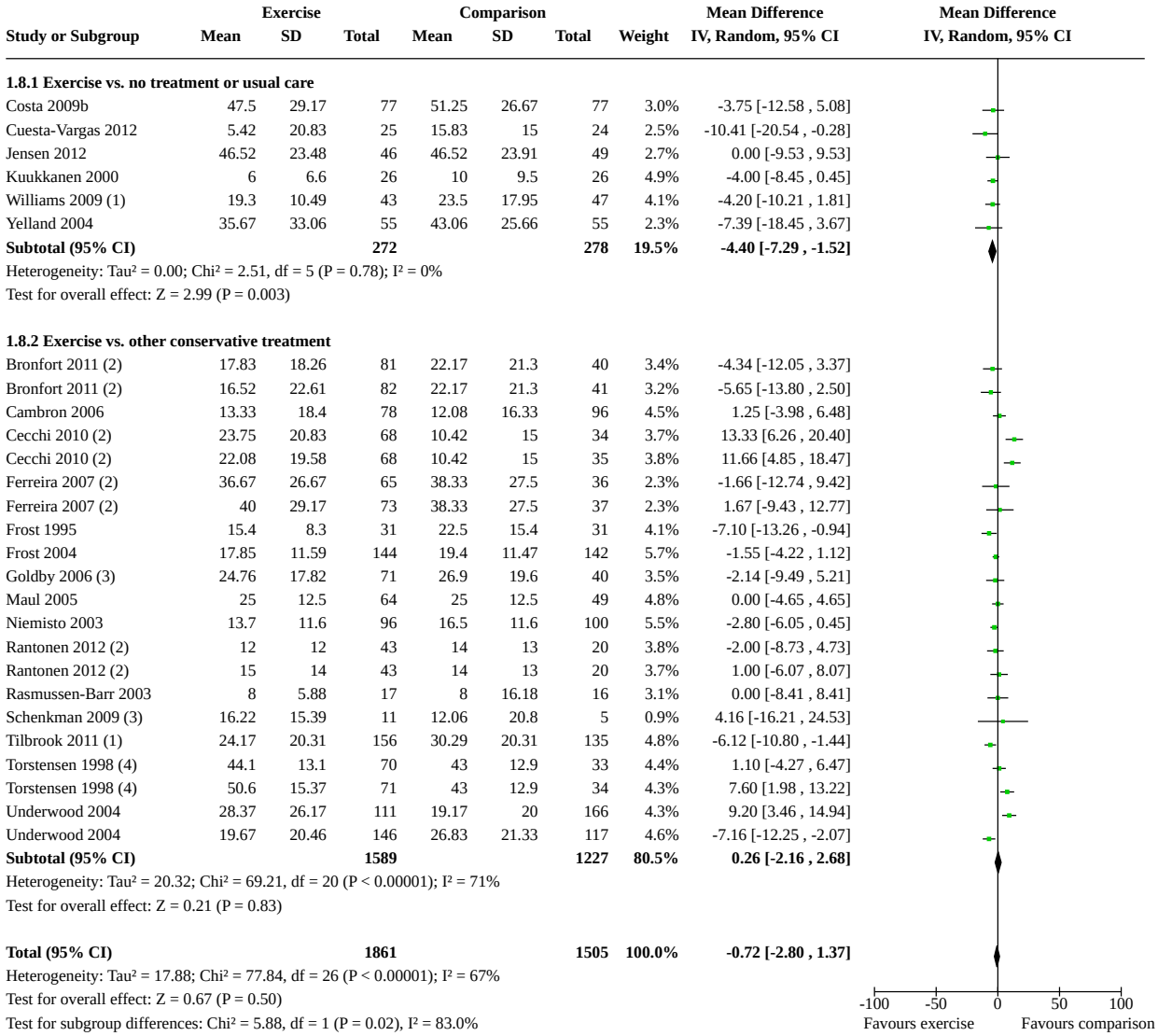
(3) Standard deviation(s) required imputation due to missing or implausible value.

(4) Study has other comparison groups available for this meta-analysis.

(5) Study has multiple exercise groups available: comparison group data divided.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.8. Comparison 1: PRIMARY ANALYSES, Outcome 8: Function (/100): Long-term follow-up (> 12 months)



Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
1.8.1 Exercise vs. no treatment or usual care								
Costa 2009b	47.5	29.17	77	51.25	26.67	77	3.0%	-3.75 [-12.58 , 5.08]
Cuesta-Vargas 2012	5.42	20.83	25	15.83	15	24	2.5%	-10.41 [-20.54 , -0.28]
Jensen 2012	46.52	23.48	46	46.52	23.91	49	2.7%	0.00 [-9.53 , 9.53]
Kuukkanen 2000	6	6.6	26	10	9.5	26	4.9%	-4.00 [-8.45 , 0.45]
Williams 2009 (1)	19.3	10.49	43	23.5	17.95	47	4.1%	-4.20 [-10.21 , 1.81]
Yelland 2004	35.67	33.06	55	43.06	25.66	55	2.3%	-7.39 [-18.45 , 3.67]
Subtotal (95% CI)			**272**			**278**	**19.5%**	**-4.40 [-7.29 , -1.52]**
Heterogeneity: Tau² = 0.00; Chi² = 2.51, df = 5 (P = 0.78); I² = 0%								
Test for overall effect: Z = 2.99 (P = 0.003)								
1.8.2 Exercise vs. other conservative treatment								
Bronfort 2011 (2)	17.83	18.26	81	22.17	21.3	40	3.4%	-4.34 [-12.05 , 3.37]
Bronfort 2011 (2)	16.52	22.61	82	22.17	21.3	41	3.2%	-5.65 [-13.80 , 2.50]
Cambron 2006	13.33	18.4	78	12.08	16.33	96	4.5%	1.25 [-3.98 , 6.48]
Cecchi 2010 (2)	23.75	20.83	68	10.42	15	34	3.7%	13.33 [6.26 , 20.40]
Cecchi 2010 (2)	22.08	19.58	68	10.42	15	35	3.8%	11.66 [4.85 , 18.47]
Ferreira 2007 (2)	36.67	26.67	65	38.33	27.5	36	2.3%	-1.66 [-12.74 , 9.42]
Ferreira 2007 (2)	40	29.17	73	38.33	27.5	37	2.3%	1.67 [-9.43 , 12.77]
Frost 1995	15.4	8.3	31	22.5	15.4	31	4.1%	-7.10 [-13.26 , -0.94]
Frost 2004	17.85	11.59	144	19.4	11.47	142	5.7%	-1.55 [-4.22 , 1.12]
Goldby 2006 (3)	24.76	17.82	71	26.9	19.6	40	3.5%	-2.14 [-9.49 , 5.21]
Maul 2005	25	12.5	64	25	12.5	49	4.8%	0.00 [-4.65 , 4.65]
Niemisto 2003	13.7	11.6	96	16.5	11.6	100	5.5%	-2.80 [-6.05 , 0.45]
Rantonen 2012 (2)	12	12	43	14	13	20	3.8%	-2.00 [-8.73 , 4.73]
Rantonen 2012 (2)	15	14	43	14	13	20	3.7%	1.00 [-6.07 , 8.07]
Rasmussen-Barr 2003	8	5.88	17	8	16.18	16	3.1%	0.00 [-8.41 , 8.41]
Schenkman 2009 (3)	16.22	15.39	11	12.06	20.8	5	0.9%	4.16 [-16.21 , 24.53]
Tilbrook 2011 (1)	24.17	20.31	156	30.29	20.31	135	4.8%	-6.12 [-10.80 , -1.44]
Torstensen 1998 (4)	44.1	13.1	70	43	12.9	33	4.4%	1.10 [-4.27 , 6.47]
Torstensen 1998 (4)	50.6	15.37	71	43	12.9	34	4.3%	7.60 [1.98 , 13.22]
Underwood 2004	28.37	26.17	111	19.17	20	166	4.3%	9.20 [3.46 , 14.94]
Underwood 2004	19.67	20.46	146	26.83	21.33	117	4.6%	-7.16 [-12.25 , -2.07]
Subtotal (95% CI)			**1589**			**1227**	**80.5%**	**0.26 [-2.16 , 2.68]**
Heterogeneity: Tau² = 20.32; Chi² = 69.21, df = 20 (P < 0.00001); I² = 71%								
Test for overall effect: Z = 0.21 (P = 0.83)								
Total (95% CI)			**1861**			**1505**	**100.0%**	**-0.72 [-2.80 , 1.37]**
Heterogeneity: Tau² = 17.88; Chi² = 77.84, df = 26 (P < 0.00001); I² = 67%								
Test for overall effect: Z = 0.67 (P = 0.50)								
Test for subgroup differences: Chi² = 5.88, df = 1 (P = 0.02), I² = 83.0%								

Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value.
(2) Study has multiple exercise groups available: comparison group data divided.
(3) Study has other comparison groups available for this meta-analysis.
(4) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.

Comparison 2. SUBGROUP ANALYSES (Population source)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Pain (/100); Earliest follow-up; Exercise vs. no treatment or usual care	34	2716	Mean Difference (IV, Random, 95% CI)	-15.40 [-18.49, -12.31]
2.1.1 Healthcare populations	18	1454	Mean Difference (IV, Random, 95% CI)	-16.94 [-20.94, -12.95]

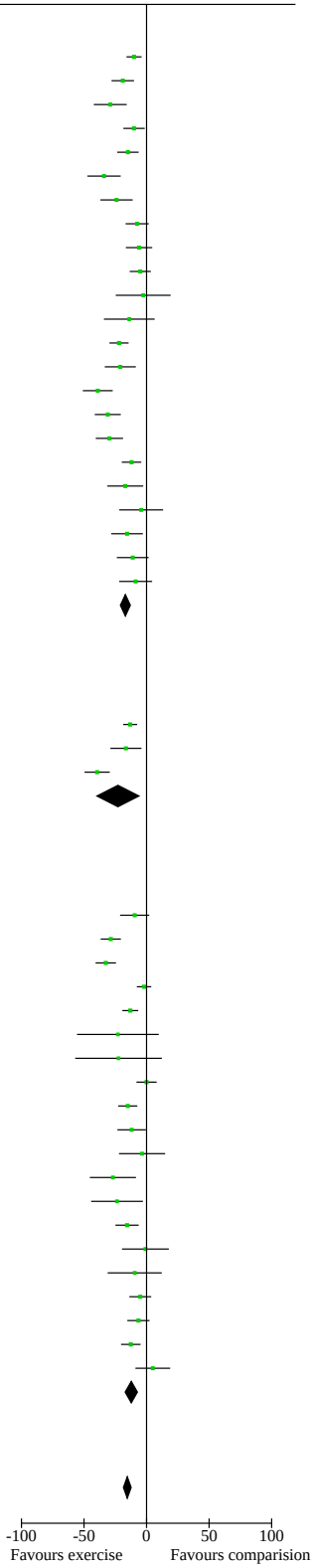
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1.2 Occupational populations	3	252	Mean Difference (IV, Random, 95% CI)	-22.87 [-39.68, -6.05]
2.1.3 General populations or mixed	13	1010	Mean Difference (IV, Random, 95% CI)	-12.20 [-17.10, -7.30]
2.2 Pain (/100); Earliest follow-up; Exercise vs. other conservative treatment	63	6235	Mean Difference (IV, Random, 95% CI)	-9.03 [-12.59, -5.47]
2.2.1 Healthcare populations	33	2602	Mean Difference (IV, Random, 95% CI)	-8.02 [-12.17, -3.87]
2.2.2 Occupational populations	9	935	Mean Difference (IV, Random, 95% CI)	-14.61 [-28.41, -0.81]
2.2.3 General populations or mixed	21	2698	Mean Difference (IV, Random, 95% CI)	-8.35 [-12.76, -3.93]
2.3 Function (/100); Earliest follow-up; Exercise vs. no treatment or usual care	35	2788	Mean Difference (IV, Random, 95% CI)	-7.01 [-8.67, -5.36]
2.3.1 Healthcare populations	22	1900	Mean Difference (IV, Random, 95% CI)	-7.86 [-10.03, -5.70]
2.3.2 Occupational populations	4	237	Mean Difference (IV, Random, 95% CI)	-3.35 [-7.06, 0.35]
2.3.3 General populations or mixed	9	651	Mean Difference (IV, Random, 95% CI)	-5.65 [-8.43, -2.88]
2.4 Function (/100); Earliest follow-up; Exercise vs. other conservative treatment	51	5944	Mean Difference (IV, Random, 95% CI)	-3.97 [-5.85, -2.09]
2.4.1 Healthcare populations	28	2488	Mean Difference (IV, Random, 95% CI)	-3.25 [-6.15, -0.36]
2.4.2 Occupational populations	4	583	Mean Difference (IV, Random, 95% CI)	-2.06 [-6.97, 2.86]
2.4.3 General populations or mixed	19	2873	Mean Difference (IV, Random, 95% CI)	-5.41 [-8.10, -2.71]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.1. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 1: Pain (/100); Earliest follow-up; Exercise vs. no treatment or usual care

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
2.1.1 Healthcare populations								
Albaladejo 2010	60	22.22	100	70	22.22	109	2.9%	-10.00 [-16.03 , -3.97]
Cabak 2017	37	17.67	22	56	17.67	46	2.5%	-19.00 [-27.98 , -10.02]
Cortell-Tormo 2018	15	15	11	44	14	8	2.0%	-29.00 [-42.14 , -15.86]
Costa 2009b	46	28	77	56	26	77	2.6%	-10.00 [-18.53 , -1.47]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	2.6%	-14.90 [-23.47 , -6.33]
Dalichau 2000	21	17.67	21	55	17.67	10	2.0%	-34.00 [-47.31 , -20.69]
Dalichau 2000	31	17.67	21	55	17.67	11	2.1%	-24.00 [-36.89 , -11.11]
Garcia 2017	39.5	27.3	73	47	29.7	73	2.5%	-7.50 [-16.75 , 1.75]
Highland 2018	27.5	24.3	33	33.5	17.9	29	2.4%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	2.6%	-5.00 [-13.44 , 3.44]
Lee 2016a	33.1	20	15	35.8	17.2	3	1.2%	-2.70 [-24.64 , 19.24]
Lee 2016a	22	11.3	15	35.8	17.2	3	1.4%	-13.80 [-34.09 , 6.49]
Masharawi 2013	16.8	8.2	20	38.8	15.4	20	2.7%	-22.00 [-29.65 , -14.35]
Mazloum 2018	48	11	15	69	16	8	2.1%	-21.00 [-33.41 , -8.59]
Mazloum 2018	30	9	16	69	16	8	2.2%	-39.00 [-50.93 , -27.07]
Morone 2011	45	23	41	76	21	29	2.4%	-31.00 [-41.39 , -20.61]
Morone 2012	42.4	21.85	25	72.08	17.19	25	2.3%	-29.68 [-40.58 , -18.78]
Moseley 2002	19	15	29	31	15	28	2.7%	-12.00 [-19.79 , -4.21]
Puppin 2011	24	27	30	41	27	25	1.9%	-17.00 [-31.33 , -2.67]
Quinn 2011	30.93	24.06	15	35.2	24.06	14	1.6%	-4.27 [-21.79 , 13.25]
Smeets 2008	37.76	24.33	52	53.35	22.6	17	2.1%	-15.59 [-28.21 , -2.97]
Smeets 2008	42.31	25.56	55	53.35	22.6	17	2.1%	-11.04 [-23.73 , 1.65]
Smeets 2008	44.63	28.86	55	53.35	22.6	17	2.0%	-8.72 [-21.90 , 4.46]
Subtotal (95% CI)			**808**			**646**	**50.9%**	**-16.94 [-20.94 , -12.95]**
Heterogeneity: Tau² = 60.48; Chi² = 68.98, df = 22 (P < 0.00001); I² = 68%								
Test for overall effect: Z = 8.32 (P < 0.00001)								
2.1.2 Occupational populations								
Chen 2014	21.7	14.2	64	34.8	17.7	63	2.9%	-13.10 [-18.69 , -7.51]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	2.1%	-16.50 [-28.91 , -4.09]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	2.4%	-39.50 [-49.57 , -29.43]
Subtotal (95% CI)			**129**			**123**	**7.5%**	**-22.87 [-39.68 , -6.05]**
Heterogeneity: Tau² = 196.68; Chi² = 20.37, df = 2 (P < 0.0001); I² = 90%								
Test for overall effect: Z = 2.67 (P = 0.008)								
2.1.3 General populations or mixed								
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	2.2%	-9.50 [-21.15 , 2.15]
Cruz-Diaz 2017	21	13.6	34	49.6	13.1	15	2.7%	-28.60 [-36.65 , -20.55]
Cruz-Diaz 2017	17	14.1	34	49.6	13.1	15	2.6%	-32.60 [-40.75 , -24.45]
Gladwell 2006	22	9	20	24	8	14	2.9%	-2.00 [-7.75 , 3.75]
Hall 2011	34	20.31	80	47	20.54	80	2.9%	-13.00 [-19.33 , -6.67]
Keane 2017	26	23.7	10	48.9	33.3	5	0.7%	-22.90 [-55.58 , 9.78]
Keane 2017	26.5	18.6	10	48.9	33.3	4	0.6%	-22.40 [-57.01 , 12.21]
Kell 2009	48	8	9	48	7	5	2.7%	0.00 [-8.06 , 8.06]
Kell 2009	33	5	9	48	7	4	2.7%	-15.00 [-22.60 , -7.40]
Muharram 2011	41.2	35.15	78	53.1	37.88	80	2.3%	-11.90 [-23.29 , -0.51]
Oh 2014	34.4	14.11	9	38	14.11	3	1.5%	-3.60 [-22.04 , 14.84]
Oh 2014	11.1	14.11	9	38	14.11	3	1.5%	-26.90 [-45.34 , -8.46]
Oh 2014	14.4	21.23	10	38	14.11	3	1.3%	-23.60 [-44.29 , -2.91]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	2.5%	-15.60 [-24.89 , -6.31]
Steele 2013	25	10.97	7	25.91	14.89	3	1.5%	-0.91 [-19.62 , 17.80]
Steele 2013	16.43	25.76	10	25.91	14.89	4	1.3%	-9.48 [-31.11 , 12.15]
Teut 2016	39.04	19.76	56	44.05	18.63	27	2.6%	-5.01 [-13.74 , 3.72]
Teut 2016	37.56	19.94	55	44.05	18.63	26	2.6%	-6.49 [-15.38 , 2.40]
Williams 2009	24.3	17.9	43	36.9	19.81	47	2.7%	-12.60 [-20.39 , -4.81]
Yelland 2004	42	26.5	55	37	45.4	55	2.0%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**579**			**431**	**41.6%**	**-12.20 [-17.10 , -7.30]**
Heterogeneity: Tau² = 81.62; Chi² = 78.82, df = 19 (P < 0.00001); I² = 76%								
Test for overall effect: Z = 4.88 (P < 0.00001)								
Total (95% CI)			**1516**			**1200**	**100.0%**	**-15.40 [-18.49 , -12.31]**
Heterogeneity: Tau² = 76.80; Chi² = 179.31, df = 45 (P < 0.00001); I² = 75%								
Test for overall effect: Z = 9.76 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.97, df = 2 (P = 0.23), I² = 32.7%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.1. (Continued)

Test for overall effect: Z = 9.76 (P < 0.00001)
Test for subgroup differences: Chi² = 2.97, df = 2 (P = 0.23), I² = 32.7%

-100 -50 0 50 100
Favours exercise Favours comparision

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

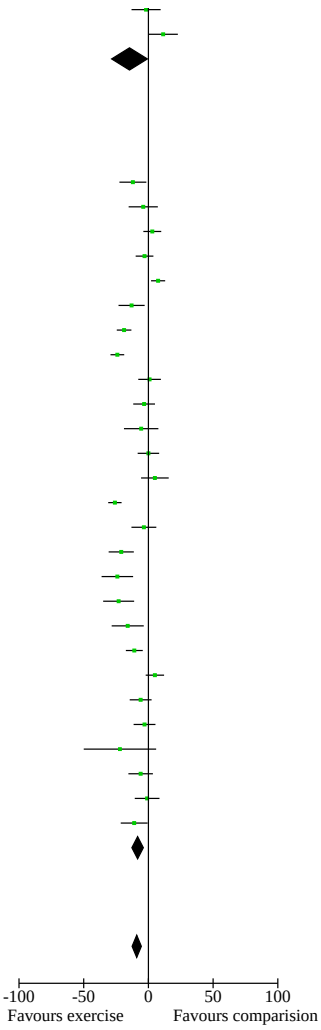
Analysis 2.2. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 2: Pain (/100); Earliest follow-up; Exercise vs. other conservative treatment

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
2.2.1 Healthcare populations								
Alayat 2014	37.1	13	24	56.5	10.4	10	1.3%	-19.40 [-27.68 , -11.12]
Alayat 2014	26.4	12.5	28	56.5	10.4	10	1.3%	-30.10 [-38.04 , -22.16]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.3%	-23.10 [-31.54 , -14.66]
Bellido-Fernandez 2018	30	22.22	9	30	22.22	4	0.8%	0.00 [-26.17 , 26.17]
Bellido-Fernandez 2018	30	18.52	9	30	22.22	5	0.9%	0.00 [-22.93 , 22.93]
Bendix 1995	44	29	31	56	29	17	1.0%	-12.00 [-29.15 , 5.15]
Bendix 1995	27	23	40	56	29	18	1.1%	-29.00 [-44.18 , -13.82]
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.3%	-24.98 [-29.09 , -20.87]
Cecchi 2010	25	20	68	8.33	11.67	34	1.3%	16.67 [10.51 , 22.83]
Cecchi 2010	23.33	20	68	8.33	11.67	35	1.3%	15.00 [8.87 , 21.13]
Chown 2008	73.4	17.2	22	73.2	11.5	33	1.3%	0.20 [-7.99 , 8.39]
Diaz-Arribas 2009	23	18.5	63	36.6	20	63	1.3%	-13.60 [-20.33 , -6.87]
Djavid 2007	43	16	18	44	20	8	1.1%	-1.00 [-16.71 , 14.71]
Djavid 2007	24	14	19	44	20	8	1.1%	-20.00 [-35.22 , -4.78]
Ferreira 2007	48	24	74	41	26	39	1.2%	7.00 [-2.82 , 16.82]
Ferreira 2007	40	25	73	41	26	38	1.2%	-1.00 [-11.06 , 9.06]
Fransoo 2006	11	17.67	19	20	17.67	9	1.1%	-9.00 [-23.01 , 5.01]
Fransoo 2006	7	17.67	20	20	17.67	10	1.1%	-13.00 [-26.41 , 0.41]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.3%	-10.00 [-17.40 , -2.60]
Gur 2003	18	12	25	19	14	12	1.2%	-1.00 [-10.21 , 8.21]
Gur 2003	29	13	25	19	14	13	1.3%	10.00 [0.84 , 19.16]
Hartvigsen 2010	45.6	19.82	36	42	21.3	20	1.2%	3.60 [-7.76 , 14.96]
Hartvigsen 2010	37.97	19.38	32	42	21.3	21	1.2%	-4.03 [-15.35 , 7.29]
Kim 2017	24	5.4	15	45.7	9	15	1.3%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.3%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.3%	-28.60 [-35.11 , -22.09]
Longo 2016	36.2	22	7	61.4	22.7	7	0.9%	-25.20 [-48.62 , -1.78]
Machado 2007	47	40	17	58	28	11	0.8%	-11.00 [-36.21 , 14.21]
McDonough 2013	45	27.08	39	39	25.24	17	1.1%	6.00 [-8.70 , 20.70]
Noori 2011	41.5	15.3	17	44	7.5	13	1.3%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.3%	-15.70 [-24.32 , -7.08]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	1.0%	-8.00 [-26.27 , 10.27]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	1.0%	-3.50 [-23.38 , 16.38]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	1.2%	-2.60 [-13.31 , 8.11]
Saper 2009	44	21	15	71	18	15	1.1%	-27.00 [-41.00 , -13.00]
Segal-Snir 2016	70	30	20	70	19	15	1.1%	0.00 [-16.29 , 16.29]
Shen 2009	11	17.67	15	25.5	17.67	7	1.1%	-14.50 [-30.35 , 1.35]
Shen 2009	23	17.67	15	25.5	17.67	8	1.1%	-2.50 [-17.66 , 12.66]
Sorensen 2010	49	22	91	46	21	94	1.3%	3.00 [-3.20 , 9.20]
Sousa 2009	33.5	24.8	27	47.6	28	25	1.1%	-14.10 [-28.52 , 0.32]
Ulger 2017	21	16	56	20.8	12	57	1.3%	0.20 [-5.02 , 5.42]
Valenza 2017	39	17.67	27	49	17.67	27	1.2%	-10.00 [-19.43 , -0.57]
Weifen 2013	22.5	17.67	47	32.4	17.67	12	1.2%	-9.90 [-21.10 , 1.30]
Weifen 2013	29.2	17.67	47	32.4	17.67	11	1.2%	-3.20 [-14.80 , 8.40]
Weifen 2013	24.3	17.67	38	32.4	17.67	12	1.2%	-8.10 [-19.57 , 3.37]
Weifen 2013	30.6	17.67	141	32.4	17.67	12	1.2%	-1.80 [-12.21 , 8.61]
Subtotal (95% CI)			**1588**			**1014**	**53.8%**	**-8.02 [-12.17 , -3.87]**
Heterogeneity: Tau² = 167.43; Chi² = 391.25, df = 45 (P < 0.00001); I² = 88%								
Test for overall effect: Z = 3.79 (P = 0.0002)								
2.2.2 Occupational populations								
Alexandre 2001	6	27.2	14	37	35.61	14	0.9%	-31.00 [-54.47 , -7.53]
Barberini 2011	52	17.38	10	81	17.38	12	1.1%	-29.00 [-43.59 , -14.41]
Jaromi 2012	5.75	17.38	56	7.07	17.38	55	1.3%	-1.32 [-7.79 , 5.15]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.4%	-41.79 [-44.48 , -39.10]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	1.1%	-8.10 [-22.94 , 6.74]
Maul 2005	40	20	74	50	20	74	1.3%	-10.00 [-16.44 , -3.56]
Murtezani 2011	20	17	50	60	27	51	1.3%	-40.00 [-48.78 , -31.22]
Rantonen 2012	31	20	43	35	28	20	1.1%	-4.00 [-17.65 , 9.65]
Rantonen 2012	29	27	43	35	28	20	1.1%	-6.00 [-20.69 , 8.69]
Torstensen 1998	37.2	25.3	70	39	28	33	1.2%	-1.80 [-13.04 , 9.44]
Torstensen 1998	50.4	27.2	71	39	28	34	1.2%	11.40 [0.06 , 22.74]
Subtotal (95% CI)			**528**			**407**	**12.9%**	**-14.61 [-28.41 , -0.81]**

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.2. (Continued)

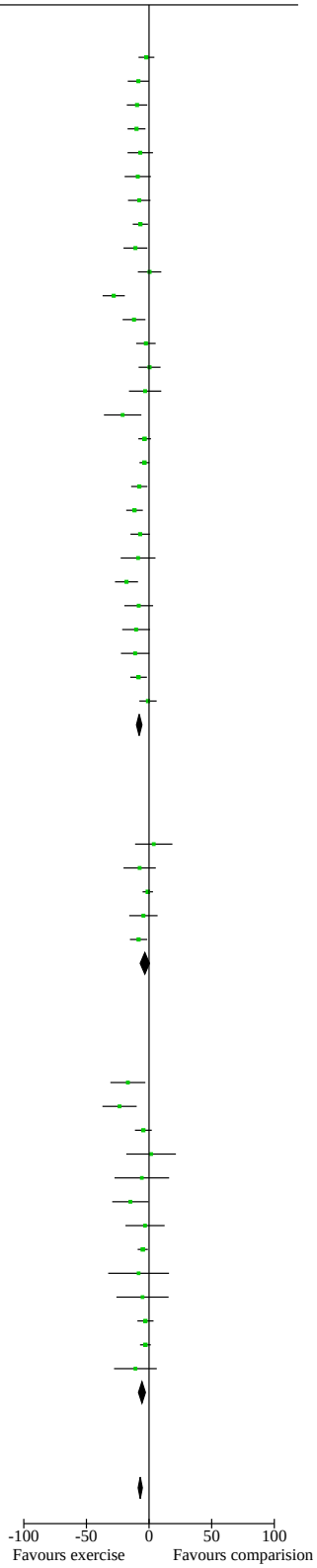
Torstensen 1998	50.4	27.2	71	39	28	34	1.2%	11.40 [0.06 , 22.74]
Subtotal (95% CI)			**528**			**407**	**12.9%**	**-14.61 [-28.41 , -0.81]**
Heterogeneity: Tau² = 504.77; Chi² = 297.84, df = 10 (P < 0.00001); I² = 97%								
Test for overall effect: Z = 2.08 (P = 0.04)								
2.2.3 General populations or mixed								
Bronfort 1996	27	20	56	39	21	21	1.2%	-12.00 [-22.40 , -1.60]
Bronfort 1996	35	22	40	39	21	21	1.2%	-4.00 [-15.28 , 7.28]
Bronfort 2011	32	22	93	29	19	50	1.3%	3.00 [-3.91 , 9.91]
Bronfort 2011	26	21	91	29	19	49	1.3%	-3.00 [-9.85 , 3.85]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.3%	7.50 [2.05 , 12.95]
Chan 2017	30	23	49	43	26	43	1.2%	-13.00 [-23.09 , -2.91]
Cruz-Diaz 2016	38.1	12.1	53	56.9	16.3	48	1.3%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.3%	-24.00 [-29.30 , -18.70]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.3%	0.90 [-7.83 , 9.63]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.3%	-3.32 [-11.72 , 5.08]
Goldby 2006	28.81	28.14	78	34.4	36.43	37	1.1%	-5.59 [-18.89 , 7.71]
Hansen 1993	33.33	21.25	44	33.33	21.25	59	1.3%	0.00 [-8.30 , 8.30]
Hemmilä 1997	30	22.63	35	25	22.63	34	1.2%	5.00 [-5.68 , 15.68]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.3%	-25.90 [-31.06 , -20.74]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.2%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	1.2%	-21.00 [-30.72 , -11.28]
Miyamoto 2018	32	24	73	56	26	22	1.2%	-24.00 [-36.18 , -11.82]
Miyamoto 2018	33	23	67	56	26	23	1.2%	-23.00 [-34.97 , -11.03]
Miyamoto 2018	40	27	69	56	26	23	1.2%	-16.00 [-28.39 , -3.61]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.3%	-10.90 [-17.43 , -4.37]
Ritvanen 2007	28	14.11	28	23	14.11	33	1.3%	5.00 [-2.11 , 12.11]
Saper 2017	50	21	127	56	22	32	1.3%	-6.00 [-14.45 , 2.45]
Saper 2017	53	21	129	56	22	32	1.3%	-3.00 [-11.44 , 5.44]
Schenkman 2009	19	21	17	41	40	9	0.7%	-22.00 [-49.97 , 5.97]
Soukup 1999	18	15.7	34	24	17	18	1.2%	-6.00 [-15.46 , 3.46]
Soukup 1999	23	16	38	24	17	17	1.2%	-1.00 [-10.55 , 8.55]
Williams 2005	10	11	20	21	23	24	1.2%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**1612**			**1086**	**33.3%**	**-8.35 [-12.76 , -3.93]**
Heterogeneity: Tau² = 112.77; Chi² = 194.14, df = 26 (P < 0.00001); I² = 87%								
Test for overall effect: Z = 3.70 (P = 0.0002)								
Total (95% CI)			**3728**			**2507**	**100.0%**	**-9.03 [-12.59 , -5.47]**
Heterogeneity: Tau² = 241.76; Chi² = 1170.88, df = 83 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 4.97 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.81, df = 2 (P = 0.67), I² = 0%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.3. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 3: Function (/100); Earliest follow-up; Exercise vs. no treatment or usual care

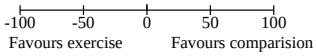
Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
2.3.1 Healthcare populations								
Albaladejo 2010	31.25	24.69	100	33.33	21.6	109	3.2%	-2.08 [-8.39 , 4.23]
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.3%	-8.50 [-16.99 , -0.01]
Costa 2009b	40	27.08	77	49.58	24.58	77	2.4%	-9.58 [-17.75 , -1.41]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	2.8%	-10.00 [-17.14 , -2.86]
Dalichau 2000	42	13.92	21	49	13.92	11	1.8%	-7.00 [-17.15 , 3.15]
Dalichau 2000	40	13.92	21	49	13.92	10	1.8%	-9.00 [-19.48 , 1.48]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	2.2%	-7.83 [-16.79 , 1.13]
Groessl 2017	30.62	19.06	75	37.54	19.06	75	3.3%	-6.92 [-13.02 , -0.82]
Highland 2018	18.46	19.25	33	29.33	18.96	29	2.0%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	2.1%	0.43 [-8.93 , 9.79]
Kofotolis 2016	17.04	8.08	37	45.25	16.17	14	2.2%	-28.21 [-37.07 , -19.35]
Kofotolis 2016	33.25	10.17	36	45.25	16.17	14	2.1%	-12.00 [-21.10 , -2.90]
Lee 2016a	4.58	3.75	15	7.08	6.67	3	2.6%	-2.50 [-10.28 , 5.28]
Lee 2016a	7.5	8.75	15	7.08	6.67	3	2.2%	0.42 [-8.33 , 9.17]
Little 2014	27.95	13.92	15	31.08	19.87	13	1.3%	-3.13 [-16.03 , 9.77]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	1.0%	-21.08 [-36.02 , -6.14]
Mazloum 2018	23.1	7.5	16	26.6	5	8	3.9%	-3.50 [-8.55 , 1.55]
Mazloum 2018	22.9	3.6	15	26.6	5	8	4.5%	-3.70 [-7.61 , 0.21]
Morone 2011	18	12.9	41	25.8	14.1	29	3.1%	-7.80 [-14.28 , -1.32]
Morone 2012	13.68	9.39	25	25.28	14.03	25	3.1%	-11.60 [-18.22 , -4.98]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	2.6%	-7.10 [-14.82 , 0.62]
Quinn 2011	22.5	19.06	15	31.25	19.06	14	1.2%	-8.75 [-22.63 , 5.13]
Shaughnessy 2004	26	14	20	44	16	21	2.1%	-18.00 [-27.19 , -8.81]
Smeets 2008	49.6	24.56	55	57.83	19.92	17	1.5%	-8.23 [-19.71 , 3.25]
Smeets 2008	47.5	21.88	55	57.83	19.92	17	1.6%	-10.33 [-21.43 , 0.77]
Smeets 2008	46.74	23.01	52	57.83	19.92	17	1.6%	-11.09 [-22.44 , 0.26]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	3.1%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	2.9%	-0.83 [-7.72 , 6.06]
Subtotal (95% CI)			**1062**			**838**	**66.6%**	**-7.86 [-10.03 , -5.70]**
Heterogeneity: Tau² = 15.69; Chi² = 53.86, df = 27 (P = 0.002); I² = 50%								
Test for overall effect: Z = 7.11 (P < 0.00001)								
2.3.2 Occupational populations								
Harts 2008	25.42	25	20	21.67	16.25	10	1.0%	3.75 [-11.13 , 18.63]
Harts 2008	14.17	16.67	20	21.67	16.25	9	1.3%	-7.50 [-20.39 , 5.39]
Kuukkanen 2000	10	7.6	27	11	7.8	26	4.4%	-1.00 [-5.15 , 3.15]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	1.6%	-4.50 [-15.82 , 6.82]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	2.9%	-8.34 [-15.23 , -1.45]
Subtotal (95% CI)			**132**			**105**	**11.2%**	**-3.35 [-7.06 , 0.35]**
Heterogeneity: Tau² = 2.39; Chi² = 4.52, df = 4 (P = 0.34); I² = 12%								
Test for overall effect: Z = 1.77 (P = 0.08)								
2.3.3 General populations or mixed								
Cruz-Diaz 2017	26.46	22.08	34	43.37	23.33	15	1.1%	-16.91 [-30.86 , -2.96]
Cruz-Diaz 2017	19.83	20.42	34	43.37	23.33	15	1.2%	-23.54 [-37.20 , -9.88]
Hall 2011	29.21	21.49	80	33.75	21.87	80	3.0%	-4.54 [-11.26 , 2.18]
Keane 2017	33.2	18.84	10	31.56	18.24	5	0.6%	1.64 [-18.16 , 21.44]
Keane 2017	25.8	20.1	10	31.56	18.24	4	0.5%	-5.76 [-27.55 , 16.03]
Kell 2009	24.2	13.27	9	39.1	13.27	5	1.1%	-14.90 [-29.41 , -0.39]
Kell 2009	35.9	13.27	9	39.1	13.27	4	0.9%	-3.20 [-18.83 , 12.43]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	4.4%	-5.00 [-9.08 , -0.92]
Steele 2013	14.86	17.95	7	23.2	17.95	3	0.4%	-8.34 [-32.62 , 15.94]
Steele 2013	17.98	17.95	10	23.2	17.95	4	0.6%	-5.22 [-26.03 , 15.59]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	3.1%	-2.87 [-9.31 , 3.57]
Williams 2009	17.9	10.49	43	20.8	10.28	47	4.3%	-2.90 [-7.20 , 1.40]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	0.8%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**380**			**271**	**22.2%**	**-5.65 [-8.43 , -2.88]**
Heterogeneity: Tau² = 3.47; Chi² = 14.02, df = 12 (P = 0.30); I² = 14%								
Test for overall effect: Z = 4.00 (P < 0.0001)								
Total (95% CI)			**1574**			**1214**	**100.0%**	**-7.01 [-8.67 , -5.36]**
Heterogeneity: Tau² = 11.93; Chi² = 78.42, df = 45 (P = 0.001); I² = 43%								
Test for overall effect: Z = 8.29 (P < 0.00001)								
Test for subgroup differences: Chi² = 4.63, df = 2 (P = 0.10), I² = 56.8%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.3. (Continued)

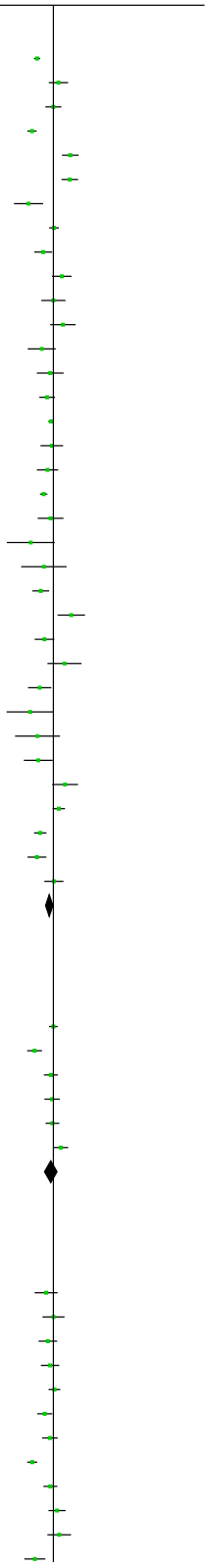
Test for overall effect: Z = 8.29 (P < 0.00001)
Test for subgroup differences: Chi² = 4.63, df = 2 (P = 0.10), I² = 56.8%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.4. Comparison 2: SUBGROUP ANALYSES (Population source), Outcome 4: Function (/100); Earliest follow-up; Exercise vs. other conservative treatment

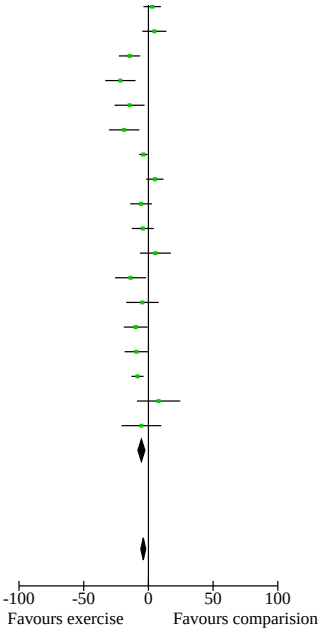
Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
2.4.1 Healthcare populations								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.0%	-12.84 [-15.31 , -10.37]
Bellido-Fernandez 2018	12	5.93	9	8	6.67	4	1.5%	4.00 [-3.60 , 11.60]
Bellido-Fernandez 2018	8	3.7	9	8	6.67	5	1.7%	0.00 [-6.33 , 6.33]
Bid 2017	7.29	8.54	64	24.12	12	64	1.9%	-16.83 [-20.44 , -13.22]
Cecchi 2010	22.5	19.58	68	9.17	13.75	34	1.7%	13.33 [6.77 , 19.89]
Cecchi 2010	22.08	19.58	68	9.17	13.75	35	1.7%	12.91 [6.40 , 19.42]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Chown 2008	20.8	8.4	35	20.3	8	35	1.9%	0.50 [-3.34 , 4.34]
Djavid 2007	33.6	7.4	18	41.6	8.8	8	1.6%	-8.00 [-14.99 , -1.01]
Djavid 2007	48.2	10.4	19	41.6	8.8	8	1.5%	6.60 [-1.08 , 14.28]
Ferreira 2007	32.92	23.75	74	32.92	25	39	1.4%	0.00 [-9.53 , 9.53]
Ferreira 2007	40.42	26.25	73	32.92	25	38	1.3%	7.50 [-2.47 , 17.47]
Fransoo 2006	4.58	13.92	19	13.75	13.92	9	1.2%	-9.17 [-20.21 , 1.87]
Fransoo 2006	11.25	13.92	20	13.75	13.92	10	1.3%	-2.50 [-13.07 , 8.07]
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.7%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.0%	-1.80 [-3.96 , 0.36]
Gur 2003	26.25	14.58	25	27.5	12.08	12	1.4%	-1.25 [-10.16 , 7.66]
Gur 2003	22.91	13.33	25	27.5	12.08	13	1.5%	-4.59 [-12.98 , 3.80]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.0%	-7.66 [-10.55 , -4.77]
Little 2014	23.16	13.92	14	25.37	13.92	15	1.3%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.7%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	0.7%	-7.50 [-25.37 , 10.37]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.6%	-9.97 [-16.60 , -3.34]
Okafor 2012	25	20.8	15	10.9	4.46	15	1.2%	14.10 [3.33 , 24.87]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	1.5%	-7.00 [-14.73 , 0.73]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.0%	8.75 [-4.68 , 22.18]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.4%	-10.80 [-19.97 , -1.63]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.7%	-18.27 [-36.90 , 0.36]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.7%	-12.50 [-30.19 , 5.19]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.2%	-11.88 [-23.25 , -0.51]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.3%	9.20 [-0.98 , 19.38]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	1.8%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	1.8%	-10.33 [-15.25 , -5.41]
Valenza 2017	16.96	13.92	27	29.92	13.92	27	1.6%	-12.96 [-20.39 , -5.53]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.6%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**1355**			**1133**	**50.5%**	**-3.25 [-6.15 , -0.36]**
Heterogeneity: Tau² = 56.38; Chi² = 249.20, df = 34 (P < 0.00001); I² = 86%								
Test for overall effect: Z = 2.20 (P = 0.03)								
2.4.2 Occupational populations								
Maul 2005	25	8.33	74	25	12.5	74	1.9%	0.00 [-3.42 , 3.42]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	1.7%	-14.80 [-20.62 , -8.98]
Rantonen 2012	14	11	43	16	10	20	1.8%	-2.00 [-7.48 , 3.48]
Rantonen 2012	15	14	43	16	10	20	1.7%	-1.00 [-7.06 , 5.06]
Torstensen 1998	46.2	13.1	70	46.9	13.1	33	1.8%	-0.70 [-6.12 , 4.72]
Torstensen 1998	52.7	16.6	71	46.9	13.1	34	1.7%	5.80 [-0.06 , 11.66]
Subtotal (95% CI)			**351**			**232**	**10.6%**	**-2.06 [-6.97 , 2.86]**
Heterogeneity: Tau² = 30.26; Chi² = 26.88, df = 5 (P < 0.0001); I² = 81%								
Test for overall effect: Z = 0.82 (P = 0.41)								
2.4.3 General populations or mixed								
Bronfort 1996	15.1	17.4	43	20.8	17.3	21	1.4%	-5.70 [-14.74 , 3.34]
Bronfort 1996	20.9	17	56	20.8	17.3	21	1.4%	0.10 [-8.54 , 8.74]
Bronfort 2011	16.96	20	91	21.3	21.74	49	1.6%	-4.34 [-11.68 , 3.00]
Bronfort 2011	18.7	19.57	92	21.3	21.74	50	1.6%	-2.60 [-9.83 , 4.63]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.8%	0.83 [-3.79 , 5.45]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.7%	-6.80 [-12.74 , -0.86]
Cruz-Diaz 2016	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	1.8%	-2.37 [-7.87 , 3.13]
Goldby 2006	31	17.07	78	28.1	17.34	37	1.6%	2.90 [-3.85 , 9.65]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.4. (Continued)

Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]
Miyamoto 2018	25.42	22.92	69	47.08	25.42	23	1.1%	-21.66 [-33.37 , -9.95]
Miyamoto 2018	32.5	21.67	67	47.08	25.42	23	1.2%	-14.58 [-26.19 , -2.97]
Miyamoto 2018	28.33	21.67	73	47.08	25.42	22	1.1%	-18.75 [-30.48 , -7.02]
Niemisto 2003	14.7	11.6	96	18.6	11	100	2.0%	-3.90 [-7.07 , -0.73]
Ritvanen 2007	17	13.27	28	12	13.27	33	1.6%	5.00 [-1.68 , 11.68]
Saper 2017	47.83	21.3	127	53.48	21.74	32	1.5%	-5.65 [-14.04 , 2.74]
Saper 2017	49.13	22.17	129	53.48	21.74	32	1.5%	-4.35 [-12.80 , 4.10]
Schenkman 2009	15.81	10.4	17	10.4	16.64	9	1.1%	5.41 [-6.53 , 17.35]
Sherman 2005	12.92	15	36	26.67	21.67	15	1.1%	-13.75 [-25.76 , -1.74]
Sherman 2005	22.08	18.75	35	26.67	21.67	15	1.1%	-4.59 [-17.19 , 8.01]
Sherman 2011	18.74	16.28	91	28.52	20.61	22	1.4%	-9.78 [-19.02 , -0.54]
Sherman 2011	19.26	17.56	92	28.52	20.61	23	1.4%	-9.26 [-18.42 , -0.10]
Tilbrook 2011	23.75	20.31	156	32.17	20.31	136	1.8%	-8.42 [-13.09 , -3.75]
Vincent 2014a	34.17	22.92	18	26.25	17.5	7	0.8%	7.92 [-8.82 , 24.66]
Vincent 2014a	20.83	17.08	17	26.25	17.5	7	0.9%	-5.42 [-20.72 , 9.88]
Subtotal (95% CI)			**1798**			**1075**	**38.9%**	**-5.41 [-8.10 , -2.71]**
Heterogeneity: Tau² = 32.41; Chi² = 92.94, df = 26 (P < 0.00001); I² = 72%								
Test for overall effect: Z = 3.93 (P < 0.0001)								
Total (95% CI)			**3504**			**2440**	**100.0%**	**-3.97 [-5.85 , -2.09]**
Heterogeneity: Tau² = 44.31; Chi² = 379.29, df = 67 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 4.14 (P < 0.0001)								
Test for subgroup differences: Chi² = 1.89, df = 2 (P = 0.39), I² = 0%								



Comparison 3. SUBGROUP ANALYSES (Symptom duration)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Pain (/100): Earliest follow-up; Exercise vs. no treatment or usual care	35	2746	Mean Difference (IV, Random, 95% CI)	-15.20 [-18.24, -12.16]
3.1.1 Moderate duration of symptoms (12 weeks - 3 years)	8	523	Mean Difference (IV, Random, 95% CI)	-16.41 [-22.55, -10.26]
3.1.2 Long duration of symptoms (> 3 years)	13	1318	Mean Difference (IV, Random, 95% CI)	-8.54 [-11.05, -6.02]
3.1.3 Duration of chronic symptoms not specified	14	905	Mean Difference (IV, Random, 95% CI)	-21.55 [-26.57, -16.52]
3.2 Pain (/100): Earliest follow-up; Exercise vs. other conservative treatment	64	6295	Mean Difference (IV, Random, 95% CI)	-9.10 [-12.63, -5.58]
3.2.1 Moderate duration of symptoms (12 weeks - 3 years)	17	1555	Mean Difference (IV, Random, 95% CI)	-10.57 [-18.12, -3.03]
3.2.2 Long duration of symptoms (> 3 years)	17	2094	Mean Difference (IV, Random, 95% CI)	-6.75 [-11.68, -1.82]
3.2.3 Duration of chronic symptoms not specified	30	2646	Mean Difference (IV, Random, 95% CI)	-9.46 [-14.22, -4.70]
3.3 Function (/100): Earliest follow-up; Exercise vs. no treatment or usual care	38	2942	Mean Difference (IV, Random, 95% CI)	-6.82 [-8.32, -5.32]

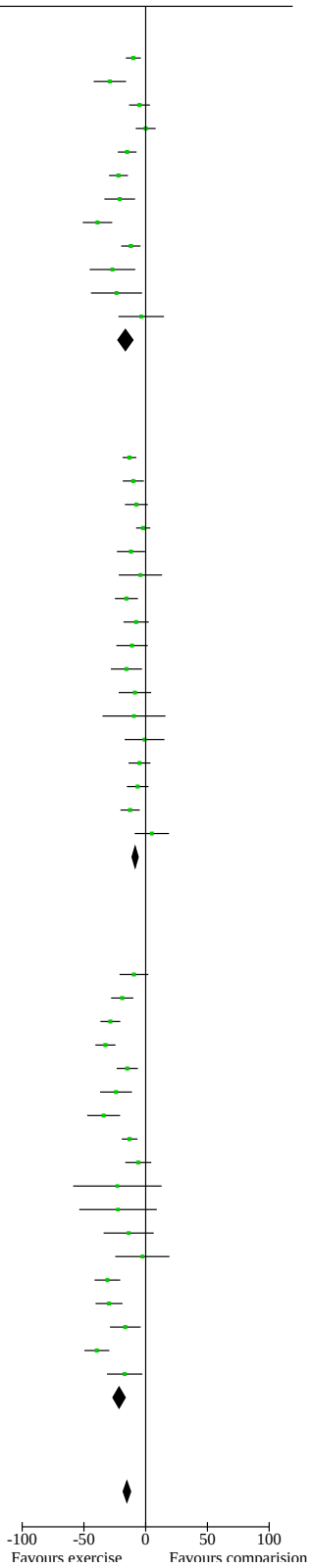
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.3.1 Moderate duration of symptoms (12 weeks - 3 years)	8	543	Mean Difference (IV, Random, 95% CI)	-4.74 [-7.08, -2.41]
3.3.2 Long duration of symptoms (> 3 years)	13	1295	Mean Difference (IV, Random, 95% CI)	-5.58 [-7.41, -3.74]
3.3.3 Duration of chronic symptoms not specified	17	1104	Mean Difference (IV, Random, 95% CI)	-8.50 [-11.44, -5.56]
3.4 Function (/100): Earliest follow-up; Exercise vs. other conservative treatment	52	6004	Mean Difference (IV, Random, 95% CI)	-4.11 [-5.98, -2.24]
3.4.1 Moderate duration of symptoms (12 weeks - 3 years)	15	1736	Mean Difference (IV, Random, 95% CI)	-4.80 [-7.78, -1.81]
3.4.2 Long duration of symptoms (>3 years)	15	2076	Mean Difference (IV, Random, 95% CI)	-4.87 [-8.52, -1.23]
3.4.3 Duration of chronic symptoms not specified	22	2192	Mean Difference (IV, Random, 95% CI)	-3.00 [-6.20, 0.20]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.1. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 1: Pain (/100): Earliest follow-up; Exercise vs. no treatment or usual care

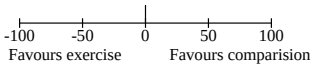
Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
3.1.1 Moderate duration of symptoms (12 weeks - 3 years)								
Albaladejo 2010	60	22.22	100	70	22.22	109	2.8%	-10.00 [-16.03 , -3.97]
Cortell-Tormo 2018	15	15	11	44	14	8	2.0%	-29.00 [-42.14 , -15.86]
Jensen 2012	45	21	42	50	19	45	2.6%	-5.00 [-13.44 , 3.44]
Kell 2009	48	8	9	48	7	5	2.6%	0.00 [-8.06 , 8.06]
Kell 2009	33	5	9	48	7	4	2.6%	-15.00 [-22.60 , -7.40]
Masharawi 2013	16.8	8.2	20	38.8	15.4	20	2.6%	-22.00 [-29.65 , -14.35]
Mazloum 2018	48	11	15	69	16	8	2.1%	-21.00 [-33.41 , -8.59]
Mazloum 2018	30	9	16	69	16	8	2.1%	-39.00 [-50.93 , -27.07]
Moseley 2002	19	15	29	31	15	28	2.6%	-12.00 [-19.79 , -4.21]
Oh 2014	11.1	14.11	9	38	14.11	3	1.5%	-26.90 [-45.34 , -8.46]
Oh 2014	14.4	21.23	10	38	14.11	3	1.3%	-23.60 [-44.29 , -2.91]
Oh 2014	34.4	14.11	9	38	14.11	3	1.5%	-3.60 [-22.04 , 14.84]
Subtotal (95% CI)			**279**			**244**	**26.3%**	**-16.41 [-22.55 , -10.26]**
Heterogeneity: Tau² = 83.48; Chi² = 49.20, df = 11 (P < 0.00001); I² = 78%								
Test for overall effect: Z = 5.23 (P < 0.00001)								
3.1.2 Long duration of symptoms (> 3 years)								
Chen 2014	21.7	14.2	64	34.8	17.7	63	2.9%	-13.10 [-18.69 , -7.51]
Costa 2009b	46	28	77	56	26	77	2.5%	-10.00 [-18.53 , -1.47]
Garcia 2017	39.5	27.3	73	47	29.7	73	2.5%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	2.9%	-2.00 [-7.75 , 3.75]
Muharram 2011	41.2	35.15	78	53.1	37.88	80	2.2%	-11.90 [-23.29 , -0.51]
Quinn 2011	30.93	24.06	15	35.2	24.06	14	1.5%	-4.27 [-21.79 , 13.25]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	2.4%	-15.60 [-24.89 , -6.31]
Schinhan 2016	13.3	10.5	15	21	17.3	15	2.3%	-7.70 [-17.94 , 2.54]
Smeets 2008	42.31	25.56	55	53.35	22.6	17	2.0%	-11.04 [-23.73 , 1.65]
Smeets 2008	37.76	24.33	52	53.35	22.6	17	2.1%	-15.59 [-28.21 , -2.97]
Smeets 2008	44.63	28.86	55	53.35	22.6	17	2.0%	-8.72 [-21.90 , 4.46]
Steele 2013	16.43	25.76	7	25.91	14.89	3	1.0%	-9.48 [-34.94 , 15.98]
Steele 2013	25	10.97	10	25.91	14.89	4	1.7%	-0.91 [-17.01 , 15.19]
Teut 2016	39.04	19.76	55	44.05	18.63	26	2.5%	-5.01 [-13.87 , 3.85]
Teut 2016	37.56	19.94	56	44.05	18.63	27	2.5%	-6.49 [-15.25 , 2.27]
Williams 2009	24.3	17.9	43	36.9	19.81	47	2.6%	-12.60 [-20.39 , -4.81]
Yelland 2004	42	26.5	55	37	45.4	55	1.9%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**751**			**567**	**37.5%**	**-8.54 [-11.05 , -6.02]**
Heterogeneity: Tau² = 3.33; Chi² = 18.22, df = 16 (P = 0.31); I² = 12%								
Test for overall effect: Z = 6.65 (P < 0.00001)								
3.1.3 Duration of chronic symptoms not specified								
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	2.2%	-9.50 [-21.15 , 2.15]
Cabak 2017	37	17.67	22	56	17.67	46	2.5%	-19.00 [-27.98 , -10.02]
Cruz-Diaz 2017	21	13.6	34	49.6	13.1	15	2.6%	-28.60 [-36.65 , -20.55]
Cruz-Diaz 2017	17	14.1	34	49.6	13.1	15	2.6%	-32.60 [-40.75 , -24.45]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	2.5%	-14.90 [-23.47 , -6.33]
Dalichau 2000	31	17.67	21	55	17.67	11	2.0%	-24.00 [-36.89 , -11.11]
Dalichau 2000	21	17.67	21	55	17.67	10	2.0%	-34.00 [-47.31 , -20.69]
Hall 2011	34	20.31	80	47	20.54	80	2.8%	-13.00 [-19.33 , -6.67]
Highland 2018	27.5	24.3	33	33.5	17.9	29	2.3%	-6.00 [-16.54 , 4.54]
Keane 2017	26	23.7	10	48.9	33.3	4	0.6%	-22.90 [-58.69 , 12.89]
Keane 2017	26.5	18.6	10	48.9	33.3	5	0.7%	-22.40 [-53.78 , 8.98]
Lee 2016a	22	11.3	15	35.8	17.2	3	1.3%	-13.80 [-34.09 , 6.49]
Lee 2016a	33.1	20	15	35.8	17.2	3	1.2%	-2.70 [-24.64 , 19.24]
Morone 2011	45	23	41	76	21	29	2.3%	-31.00 [-41.39 , -20.61]
Morone 2012	42.4	21.85	25	72.08	17.19	25	2.3%	-29.68 [-40.58 , -18.78]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	2.1%	-16.50 [-28.91 , -4.09]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	2.4%	-39.50 [-49.57 , -29.43]
Puppin 2011	24	27	30	41	27	25	1.9%	-17.00 [-31.33 , -2.67]
Subtotal (95% CI)			**501**			**404**	**36.1%**	**-21.55 [-26.57 , -16.52]**
Heterogeneity: Tau² = 74.33; Chi² = 57.55, df = 17 (P < 0.00001); I² = 70%								
Test for overall effect: Z = 8.41 (P < 0.00001)								
Total (95% CI)			**1531**			**1215**	**100.0%**	**-15.20 [-18.24 , -12.16]**
Heterogeneity: Tau² = 75.83; Chi² = 181.55, df = 46 (P < 0.00001); I² = 75%								
Test for overall effect: Z = 9.80 (P < 0.00001)								
Test for subgroup differences: Chi² = 23.08, df = 2 (P < 0.00001), I² = 91.3%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.1. (Continued)

Heterogeneity: Tau² = 75.83; Chi² = 181.55, df = 46 (P < 0.00001); I² = 75%
Test for overall effect: Z = 9.80 (P < 0.00001)
Test for subgroup differences: Chi² = 23.08, df = 2 (P < 0.00001), I² = 91.3%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

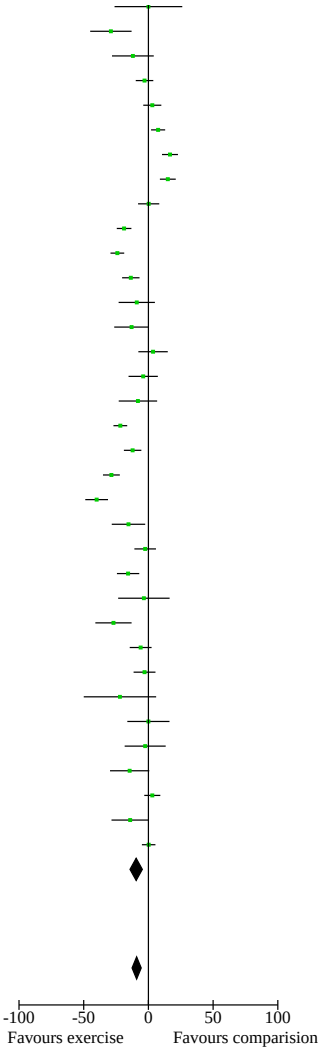
Analysis 3.2. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 2: Pain (/100): Earliest follow-up; Exercise vs. other conservative treatment

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
3.2.1 Moderate duration of symptoms (12 weeks - 3 years)								
Alayat 2014	37.1	13	28	56.5	10.4	10	1.3%	-19.40 [-27.45 , -11.35]
Alayat 2014	26.4	12.5	24	56.5	10.4	10	1.3%	-30.10 [-38.26 , -21.94]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.3%	-23.10 [-31.54 , -14.66]
Bronfort 1996	27	20	40	39	21	21	1.2%	-12.00 [-22.91 , -1.09]
Bronfort 1996	35	22	56	39	21	21	1.2%	-4.00 [-14.67 , 6.67]
Chan 2017	30	23	49	43	26	43	1.2%	-13.00 [-23.09 , -2.91]
Djavid 2007	43	16	18	44	20	8	1.1%	-1.00 [-16.71 , 14.71]
Djavid 2007	24	14	19	44	20	8	1.1%	-20.00 [-35.22 , -4.78]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.3%	-10.00 [-17.40 , -2.60]
Gur 2003	18	12	25	19	14	12	1.2%	-1.00 [-10.21 , 8.21]
Gur 2003	29	13	25	19	14	13	1.2%	10.00 [0.84 , 19.16]
Hansen 1993	33.33	21.25	44	33.33	21.25	59	1.3%	0.00 [-8.30 , 8.30]
Jaromi 2012	5.75	17.38	56	7.07	17.38	55	1.3%	-1.32 [-7.79 , 5.15]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.3%	-41.79 [-44.48 , -39.10]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.3%	-25.90 [-31.06 , -20.74]
Longo 2016	36.2	22	7	61.4	22.7	7	0.8%	-25.20 [-48.62 , -1.78]
Rantonen 2012	29	27	43	35	28	20	1.1%	-6.00 [-20.69 , 8.69]
Rantonen 2012	31	20	43	35	28	20	1.1%	-4.00 [-17.65 , 9.65]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	1.0%	-8.00 [-26.27 , 10.27]
Ritvanen 2007	28	14.11	28	23	14.11	33	1.3%	5.00 [-2.11 , 12.11]
Valenza 2017	39	17.67	27	49	17.67	27	1.2%	-10.00 [-19.43 , -0.57]
Weifen 2013	24.3	17.67	47	32.4	17.67	12	1.2%	-8.10 [-19.30 , 3.10]
Weifen 2013	22.5	17.67	141	32.4	17.67	12	1.2%	-9.90 [-20.31 , 0.51]
Weifen 2013	29.2	17.67	47	32.4	17.67	11	1.2%	-3.20 [-14.80 , 8.40]
Weifen 2013	30.6	17.67	38	32.4	17.67	12	1.2%	-1.80 [-13.27 , 9.67]
Subtotal (95% CI)			**973**			**582**	**29.8%**	**-10.57 [-18.12 , -3.03]**
Heterogeneity: Tau² = 338.40; Chi² = 488.03, df = 24 (P < 0.00001); I² = 95%								
Test for overall effect: Z = 2.75 (P = 0.006)								
3.2.2 Long duration of symptoms (> 3 years)								
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.3%	-24.98 [-29.09 , -20.87]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.2%	0.90 [-7.83 , 9.63]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.3%	-3.32 [-11.72 , 5.08]
Ferreira 2007	40	25	74	41	26	39	1.2%	-1.00 [-10.95 , 8.95]
Ferreira 2007	48	24	73	41	26	38	1.2%	7.00 [-2.93 , 16.93]
Goldby 2006	28.81	28.14	78	34.4	36.43	37	1.1%	-5.59 [-18.89 , 7.71]
Hemmilä 1997	30	22.63	35	25	22.63	34	1.2%	5.00 [-5.68 , 15.68]
Machado 2007	47	40	17	58	28	11	0.8%	-11.00 [-36.21 , 14.21]
Maul 2005	40	20	74	50	20	74	1.3%	-10.00 [-16.44 , -3.56]
McDonough 2013	45	27.08	39	39	25.24	17	1.1%	6.00 [-8.70 , 20.70]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.2%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	1.2%	-21.00 [-30.72 , -11.28]
Miyamoto 2018	33	23	69	56	26	23	1.2%	-23.00 [-34.93 , -11.07]
Miyamoto 2018	40	27	73	56	26	22	1.2%	-16.00 [-28.51 , -3.49]
Miyamoto 2018	32	24	67	56	26	23	1.2%	-24.00 [-36.08 , -11.92]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.3%	-10.90 [-17.43 , -4.37]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	1.2%	-2.60 [-13.31 , 8.11]
Soukup 1999	18	15.7	34	24	17	18	1.2%	-6.00 [-15.46 , 3.46]
Soukup 1999	23	16	38	24	17	17	1.2%	-1.00 [-10.55 , 8.55]
Torstensen 1998	37.2	25.3	70	39	28	33	1.2%	-1.80 [-13.04 , 9.44]
Torstensen 1998	50.4	27.2	71	39	28	34	1.2%	11.40 [0.06 , 22.74]
Williams 2005	10	11	20	21	23	24	1.2%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**1239**			**855**	**26.2%**	**-6.75 [-11.68 , -1.82]**
Heterogeneity: Tau² = 109.02; Chi² = 125.50, df = 21 (P < 0.00001); I² = 83%								
Test for overall effect: Z = 2.69 (P = 0.007)								
3.2.3 Duration of chronic symptoms not specified								
Alexandre 2001	6	27.2	14	37	35.61	14	0.8%	-31.00 [-54.47 , -7.53]
Barberini 2011	52	17.38	10	81	17.38	12	1.1%	-29.00 [-43.59 , -14.41]
Bellido-Fernandez 2018	30	18.52	9	30	22.22	5	0.9%	0.00 [-22.93 , 22.93]
Bellido-Fernandez 2018	30	22.22	9	30	22.22	4	0.8%	0.00 [-26.17 , 26.17]
Bendix 1995	27	23	31	56	29	17	1.1%	-29.00 [-44.99 , -13.01]
Bendix 1995	44	29	40	56	29	18	1.1%	-12.00 [-28.13 , 4.13]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.2. (Continued)

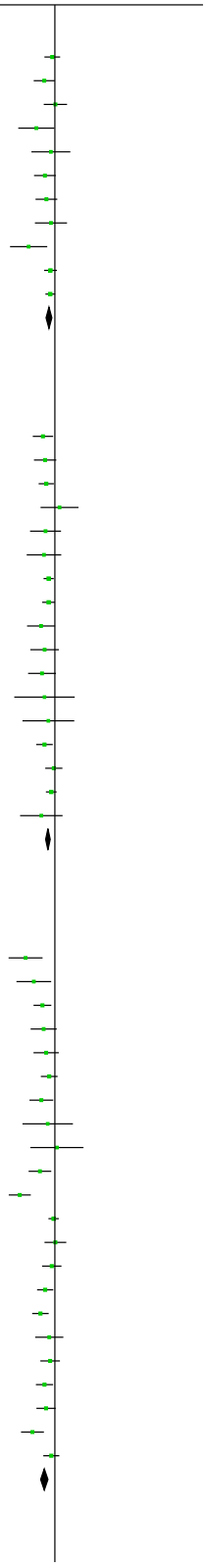
Bendix 1995	27	23	31	56	29	17	1.1%	-29.00 [-44.99 , -13.01]
Bendix 1995	44	29	40	56	29	18	1.1%	-12.00 [-28.13 , 4.13]
Bronfort 2011	26	21	93	29	19	50	1.3%	-3.00 [-9.78 , 3.78]
Bronfort 2011	32	22	91	29	19	49	1.3%	3.00 [-3.98 , 9.98]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.3%	7.50 [2.05 , 12.95]
Cecchi 2010	25	20	68	8.33	11.67	34	1.3%	16.67 [10.51 , 22.83]
Cecchi 2010	23.33	20	68	8.33	11.67	35	1.3%	15.00 [8.87 , 21.13]
Chown 2008	73.4	17.2	22	73.2	11.5	33	1.3%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016	38.1	12.1	53	56.9	16.3	48	1.3%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.3%	-24.00 [-29.30 , -18.70]
Diaz-Arribas 2009	23	18.5	63	36.6	20	63	1.3%	-13.60 [-20.33 , -6.87]
Fransoo 2006	11	17.67	19	20	17.67	9	1.1%	-9.00 [-23.01 , 5.01]
Fransoo 2006	7	17.67	20	20	17.67	10	1.1%	-13.00 [-26.41 , 0.41]
Hartvigsen 2010	45.6	19.82	36	42	21.3	20	1.2%	3.60 [-7.76 , 14.96]
Hartvigsen 2010	37.97	19.38	32	42	21.3	21	1.2%	-4.03 [-15.35 , 7.29]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	1.1%	-8.10 [-22.94 , 6.74]
Kim 2017	24	5.4	15	45.7	9	15	1.3%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.3%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.3%	-28.60 [-35.11 , -22.09]
Murtezani 2011	20	17	50	60	27	51	1.2%	-40.00 [-48.78 , -31.22]
Natour 2015	40.4	24.2	30	55.8	26.8	30	1.1%	-15.40 [-28.32 , -2.48]
Noori 2011	41.5	15.3	17	44	7.5	13	1.3%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.2%	-15.70 [-24.32 , -7.08]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.9%	-3.50 [-23.38 , 16.38]
Saper 2009	44	21	15	71	18	15	1.1%	-27.00 [-41.00 , -13.00]
Saper 2017	50	21	129	56	22	32	1.3%	-6.00 [-14.44 , 2.44]
Saper 2017	53	21	127	56	22	32	1.3%	-3.00 [-11.45 , 5.45]
Schenkman 2009	19	21	17	41	40	9	0.7%	-22.00 [-49.97 , 5.97]
Segal-Snir 2016	70	30	20	70	19	15	1.0%	0.00 [-16.29 , 16.29]
Shen 2009	23	17.67	15	25.5	17.67	7	1.1%	-2.50 [-18.35 , 13.35]
Shen 2009	11	17.67	15	25.5	17.67	8	1.1%	-14.50 [-29.66 , 0.66]
Sorensen 2010	49	22	91	46	21	94	1.3%	3.00 [-3.20 , 9.20]
Sousa 2009	33.5	24.8	27	47.6	28	25	1.1%	-14.10 [-28.52 , 0.32]
Ulger 2017	21	16	56	20.8	12	57	1.3%	0.20 [-5.02 , 5.42]
Subtotal (95% CI)			**1546**			**1100**	**44.0%**	**-9.46 [-14.22 , -4.70]**
Heterogeneity: Tau² = 185.36; Chi² = 394.94, df = 37 (P < 0.00001); I² = 91%								
Test for overall effect: Z = 3.90 (P < 0.0001)								
Total (95% CI)			**3758**			**2537**	**100.0%**	**-9.10 [-12.63 , -5.58]**
Heterogeneity: Tau² = 239.79; Chi² = 1168.81, df = 84 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 5.06 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.93, df = 2 (P = 0.63), I² = 0%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.3. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 3: Function (/100): Earliest follow-up; Exercise vs. no treatment or usual care

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
3.3.1 Moderate duration of symptoms (12 weeks - 3 years)								
Albaladejo 2010	31.25	24.69	100	33.33	21.6	109	2.9%	-2.08 [-8.39 , 4.23]
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.1%	-8.50 [-16.99 , -0.01]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	1.8%	0.43 [-8.93 , 9.79]
Kell 2009	24.2	13.27	9	39.1	13.27	5	0.9%	-14.90 [-29.41 , -0.39]
Kell 2009	35.9	13.27	9	39.1	13.27	4	0.8%	-3.20 [-18.83 , 12.43]
Kofotolis 2006	20	13.92	28	28	13.92	15	2.0%	-8.00 [-16.73 , 0.73]
Kofotolis 2006	21.2	13.92	28	28	13.92	15	2.0%	-6.80 [-15.53 , 1.93]
Little 2014	27.95	13.92	15	31.08	19.87	13	1.1%	-3.13 [-16.03 , 9.77]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	0.9%	-21.08 [-36.02 , -6.14]
Mazloum 2018	23.1	7.5	15	26.6	5	8	3.6%	-3.50 [-8.64 , 1.64]
Mazloum 2018	22.9	3.6	16	26.6	5	8	4.3%	-3.70 [-7.59 , 0.19]
Subtotal (95% CI)			**293**			**250**	**22.4%**	**-4.74 [-7.08 , -2.41]**
Heterogeneity: Tau² = 0.68; Chi² = 10.43, df = 10 (P = 0.40); I² = 4%								
Test for overall effect: Z = 3.98 (P < 0.0001)								
3.3.2 Long duration of symptoms (> 3 years)								
Costa 2009b	40	27.08	77	49.58	24.58	77	2.2%	-9.58 [-17.75 , -1.41]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	1.9%	-7.83 [-16.79 , 1.13]
Groessl 2017	30.62	19.06	75	37.54	19.06	75	3.0%	-6.92 [-13.02 , -0.82]
Harts 2008	25.42	25	20	21.67	16.25	9	0.8%	3.75 [-11.51 , 19.01]
Harts 2008	14.17	16.67	20	21.67	16.25	10	1.2%	-7.50 [-19.94 , 4.94]
Quinn 2011	22.5	19.06	15	31.25	19.06	14	1.0%	-8.75 [-22.63 , 5.13]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	4.2%	-5.00 [-9.08 , -0.92]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	3.6%	-5.06 [-10.19 , 0.07]
Smeets 2008	46.74	23.01	55	57.83	19.92	17	1.4%	-11.09 [-22.34 , 0.16]
Smeets 2008	49.6	24.56	55	57.83	19.92	17	1.3%	-8.23 [-19.71 , 3.25]
Smeets 2008	47.5	21.88	52	57.83	19.92	17	1.4%	-10.33 [-21.51 , 0.85]
Steele 2013	14.86	17.95	7	23.2	17.95	3	0.4%	-8.34 [-32.62 , 15.94]
Steele 2013	17.98	17.95	10	23.2	17.95	4	0.5%	-5.22 [-26.03 , 15.59]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	2.8%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	2.7%	-0.83 [-7.72 , 6.06]
Williams 2009	17.9	10.49	43	20.8	10.28	47	4.1%	-2.90 [-7.20 , 1.40]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	0.7%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**715**			**580**	**33.1%**	**-5.58 [-7.41 , -3.74]**
Heterogeneity: Tau² = 0.00; Chi² = 9.41, df = 16 (P = 0.90); I² = 0%								
Test for overall effect: Z = 5.96 (P < 0.00001)								
3.3.3 Duration of chronic symptoms not specified								
Cruz-Diaz 2017	19.83	20.42	34	43.37	23.33	15	1.0%	-23.54 [-37.20 , -9.88]
Cruz-Diaz 2017	26.46	22.08	34	43.37	23.33	15	1.0%	-16.91 [-30.86 , -2.96]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	2.6%	-10.00 [-17.14 , -2.86]
Dalichau 2000	40	13.92	21	49	13.92	10	1.5%	-9.00 [-19.48 , 1.48]
Dalichau 2000	42	13.92	21	49	13.92	11	1.6%	-7.00 [-17.15 , 3.15]
Hall 2011	29.21	21.49	80	33.75	21.87	80	2.7%	-4.54 [-11.26 , 2.18]
Highland 2018	18.46	19.25	33	29.33	18.96	29	1.8%	-10.87 [-20.40 , -1.34]
Keane 2017	25.8	20.1	10	31.56	18.24	5	0.5%	-5.76 [-26.03 , 14.51]
Keane 2017	33.2	18.84	10	31.56	18.24	4	0.5%	1.64 [-19.71 , 22.99]
Kofotolis 2016	33.25	10.17	36	45.25	16.17	14	1.9%	-12.00 [-21.10 , -2.90]
Kofotolis 2016	17.04	8.08	37	45.25	16.17	14	2.0%	-28.21 [-37.07 , -19.35]
Kuukkanen 2000	10	7.6	27	11	7.8	26	4.2%	-1.00 [-5.15 , 3.15]
Lee 2016a	7.5	8.75	15	7.08	6.67	3	2.0%	0.42 [-8.33 , 9.17]
Lee 2016a	4.58	3.75	15	7.08	6.67	3	2.3%	-2.50 [-10.28 , 5.28]
Morone 2011	18	12.9	41	25.8	14.1	29	2.9%	-7.80 [-14.28 , -1.32]
Morone 2012	13.68	9.39	25	25.28	14.03	25	2.8%	-11.60 [-18.22 , -4.98]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	1.4%	-4.50 [-15.82 , 6.82]
Patti 2016	13	8	19	16.8	15.6	19	2.3%	-3.80 [-11.68 , 4.08]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	2.7%	-8.34 [-15.23 , -1.45]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	2.3%	-7.10 [-14.82 , 0.62]
Shaughnessy 2004	26	14	20	44	16	21	1.9%	-18.00 [-27.19 , -8.81]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	2.9%	-2.87 [-9.31 , 3.57]
Subtotal (95% CI)			**656**			**448**	**44.5%**	**-8.50 [-11.44 , -5.56]**
Heterogeneity: Tau² = 27.91; Chi² = 55.27, df = 21 (P < 0.0001); I² = 62%								
Test for overall effect: Z = 5.67 (P < 0.00001)								



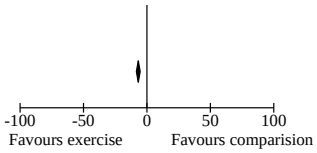
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.3. (Continued)

Test for overall effect: Z = 5.67 (P < 0.00001)

Total (95% CI)	**1664**	**1278**	**100.0%**	**-6.82 [-8.32 , -5.32]**

Heterogeneity: Tau² = 9.62; Chi² = 78.94, df = 49 (P = 0.004); I² = 38%
Test for overall effect: Z = 8.90 (P < 0.00001)
Test for subgroup differences: Chi² = 4.06, df = 2 (P = 0.13), I² = 50.7%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

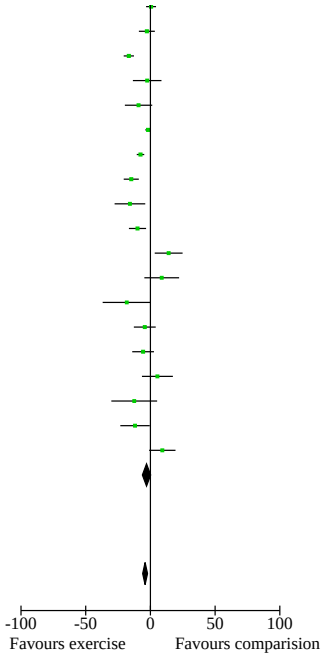
Analysis 3.4. Comparison 3: SUBGROUP ANALYSES (Symptom duration), Outcome 4: Function (/100): Earliest follow-up; Exercise vs. other conservative treatment

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
3.4.1 Moderate duration of symptoms (12 weeks - 3 years)								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.0%	-12.84 [-15.31 , -10.37]
Bronfort 1996	20.9	17	43	20.8	17.3	21	1.4%	0.10 [-8.88 , 9.08]
Bronfort 1996	15.1	17.4	56	20.8	17.3	21	1.4%	-5.70 [-14.39 , 2.99]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.7%	-6.80 [-12.74 , -0.86]
Djavid 2007	33.6	7.4	18	41.6	8.8	8	1.6%	-8.00 [-14.99 , -1.01]
Djavid 2007	48.2	10.4	19	41.6	8.8	8	1.5%	6.60 [-1.08 , 14.28]
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.7%	-4.90 [-11.09 , 1.29]
Gur 2003	22.91	13.33	25	27.5	12.08	12	1.4%	-4.59 [-13.19 , 4.01]
Gur 2003	26.25	14.58	25	27.5	12.08	13	1.4%	-1.25 [-9.96 , 7.46]
Little 2014	23.16	13.92	14	25.37	13.92	15	1.3%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.7%	-17.84 [-36.69 , 1.01]
Rantonen 2012	15	14	43	16	10	20	1.7%	-1.00 [-7.06 , 5.06]
Rantonen 2012	14	11	43	16	10	20	1.7%	-2.00 [-7.48 , 3.48]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	1.5%	-7.00 [-14.73 , 0.73]
Ritvanen 2007	17	13.27	28	12	13.27	33	1.6%	5.00 [-1.68 , 11.68]
Sherman 2005	12.92	15	36	26.67	21.67	15	1.1%	-13.75 [-25.76 , -1.74]
Sherman 2005	22.08	18.75	35	26.67	21.67	15	1.1%	-4.59 [-17.19 , 8.01]
Tilbrook 2011	23.75	20.31	156	32.17	20.31	136	1.8%	-8.42 [-13.09 , -3.75]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	1.8%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	1.8%	-10.33 [-15.25 , -5.41]
Valenza 2017	16.96	13.92	27	29.92	13.92	27	1.5%	-12.96 [-20.39 , -5.53]
Subtotal (95% CI)			**952**			**784**	**31.7%**	**-4.80 [-7.78 , -1.81]**
Heterogeneity: Tau² = 33.42; Chi² = 86.10, df = 20 (P < 0.00001); I² = 77%								
Test for overall effect: Z = 3.15 (P = 0.002)								
3.4.2 Long duration of symptoms (>3 years)								
Bid 2017	7.29	8.54	64	24.12	12	64	1.9%	-16.83 [-20.44 , -13.22]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	1.7%	-2.37 [-7.87 , 3.13]
Ferreira 2007	32.92	23.75	73	32.92	25	38	1.3%	0.00 [-9.64 , 9.64]
Ferreira 2007	40.42	26.25	74	32.92	25	39	1.3%	7.50 [-2.37 , 17.37]
Goldby 2006	31	17.07	78	28.1	17.34	37	1.6%	2.90 [-3.85 , 9.65]
Machado 2007	50	22.92	17	57.5	23.96	11	0.7%	-7.50 [-25.37 , 10.37]
Maul 2005	25	8.33	74	25	12.5	74	1.9%	0.00 [-3.42 , 3.42]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]
Miyamoto 2018	25.42	22.92	73	47.08	25.42	22	1.1%	-21.66 [-33.51 , -9.81]
Miyamoto 2018	32.5	21.67	69	47.08	25.42	23	1.1%	-14.58 [-26.16 , -3.00]
Miyamoto 2018	28.33	21.67	67	47.08	25.42	23	1.1%	-18.75 [-30.36 , -7.14]
Niemisto 2003	14.7	11.6	96	18.6	11	100	1.9%	-3.90 [-7.07 , -0.73]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.4%	-10.80 [-19.97 , -1.63]
Sherman 2011	18.74	16.28	92	28.52	20.61	23	1.4%	-9.78 [-18.84 , -0.72]
Sherman 2011	19.26	17.56	91	28.52	20.61	22	1.4%	-9.26 [-18.60 , 0.08]
Torstensen 1998	52.7	16.6	71	46.9	13.1	34	1.7%	5.80 [-0.06 , 11.66]
Torstensen 1998	46.2	13.1	70	46.9	13.1	33	1.7%	-0.70 [-6.12 , 4.72]
Vincent 2014a	34.17	22.92	17	26.25	17.5	7	0.8%	7.92 [-9.01 , 24.85]
Vincent 2014a	20.83	17.08	18	26.25	17.5	7	0.9%	-5.42 [-20.60 , 9.76]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.5%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**1282**			**794**	**29.4%**	**-4.87 [-8.52 , -1.23]**
Heterogeneity: Tau² = 51.73; Chi² = 111.12, df = 20 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 2.62 (P = 0.009)								
3.4.3 Duration of chronic symptoms not specified								
Bellido-Fernandez 2018	12	5.93	9	8	6.67	4	1.5%	4.00 [-3.60 , 11.60]
Bellido-Fernandez 2018	8	3.7	9	8	6.67	5	1.7%	0.00 [-6.33 , 6.33]
Bronfort 2011	16.96	20	91	21.3	21.74	49	1.6%	-4.34 [-11.68 , 3.00]
Bronfort 2011	18.7	19.57	92	21.3	21.74	50	1.6%	-2.60 [-9.83 , 4.63]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.8%	0.83 [-3.79 , 5.45]
Cecchi 2010	22.5	19.58	68	9.17	13.75	35	1.6%	13.33 [6.82 , 19.84]
Cecchi 2010	22.08	19.58	68	9.17	13.75	34	1.6%	12.91 [6.35 , 19.47]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Chown 2008	20.8	8.4	35	20.3	8	35	1.9%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.4. (Continued)

Cruz-Diaz 2016	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]
Fransoo 2006	11.25	13.92	19	13.75	13.92	9	1.2%	-2.50 [-13.54 , 8.54]
Fransoo 2006	4.58	13.92	20	13.75	13.92	10	1.2%	-9.17 [-19.74 , 1.40]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.0%	-1.80 [-3.96 , 0.36]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.0%	-7.66 [-10.55 , -4.77]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	1.7%	-14.80 [-20.62 , -8.98]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.1%	-15.83 [-27.67 , -3.99]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.6%	-9.97 [-16.60 , -3.34]
Okafor 2012	25	20.8	15	10.9	4.46	15	1.2%	14.10 [3.33 , 24.87]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.0%	8.75 [-4.68 , 22.18]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.7%	-18.27 [-36.90 , 0.36]
Saper 2017	49.13	22.17	129	53.48	21.74	32	1.4%	-4.35 [-12.80 , 4.10]
Saper 2017	47.83	21.3	127	53.48	21.74	32	1.4%	-5.65 [-14.04 , 2.74]
Schenkman 2009	15.81	10.4	17	10.4	16.64	9	1.1%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.7%	-12.50 [-30.19 , 5.19]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.2%	-11.88 [-23.25 , -0.51]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.3%	9.20 [-0.98 , 19.38]
Subtotal (95% CI)			**1300**			**892**	**38.9%**	**-3.00 [-6.20 , 0.20]**
Heterogeneity: Tau² = 52.65; Chi² = 173.43, df = 26 (P < 0.00001); I² = 85%								
Test for overall effect: Z = 1.84 (P = 0.07)								
Total (95% CI)			**3534**			**2470**	**100.0%**	**-4.11 [-5.98 , -2.24]**
Heterogeneity: Tau² = 44.47; Chi² = 382.69, df = 68 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 4.31 (P < 0.0001)								
Test for subgroup differences: Chi² = 0.83, df = 2 (P = 0.66), I² = 0%								



Comparison 4. SUBGROUP ANALYSES (Radicular symptoms)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Pain (/100): Earliest follow-up; Exercise vs. no treatment or usual care	35	2746	Mean Difference (IV, Random, 95% CI)	-15.23 [-18.26, -12.20]
4.1.1 Study populations excluded leg pain and/or neurologic symptoms	14	816	Mean Difference (IV, Random, 95% CI)	-16.95 [-22.61, -11.29]
4.1.2 Study populations included leg pain and/or neurologic symptoms	13	1382	Mean Difference (IV, Random, 95% CI)	-15.89 [-20.85, -10.94]
4.1.3 Leg pain and/or neurologic symptoms not specified	8	548	Mean Difference (IV, Random, 95% CI)	-11.23 [-14.65, -7.81]
4.2 Pain (/100): Earliest follow-up; Exercise vs. other conservative treatment	64	6295	Mean Difference (IV, Random, 95% CI)	-9.10 [-12.63, -5.57]
4.2.1 Study populations excluded leg pain and/or neurologic symptoms	19	1657	Mean Difference (IV, Random, 95% CI)	-9.51 [-14.59, -4.43]
4.2.2 Study populations included leg pain and/or neurologic symptoms	27	2766	Mean Difference (IV, Random, 95% CI)	-4.87 [-9.09, -0.65]
4.2.3 Leg pain and/or neurologic symptoms not specified	18	1872	Mean Difference (IV, Random, 95% CI)	-13.71 [-20.03, -7.39]
4.3 Function (/100): Earliest follow-up; Exercise vs. no treatment or usual care	38	2942	Mean Difference (IV, Random, 95% CI)	-6.80 [-8.30, -5.29]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.3.1 Study populations excluded leg pain and/or neurologic symptoms	15	983	Mean Difference (IV, Random, 95% CI)	-6.01 [-8.14, -3.87]
4.3.2 Study populations included leg pain and/or neurologic symptoms	17	1676	Mean Difference (IV, Random, 95% CI)	-6.63 [-8.37, -4.90]
4.3.3 Leg pain and/or neurologic symptoms not specified	6	283	Mean Difference (IV, Random, 95% CI)	-7.34 [-13.52, -1.16]
4.4 Function (/100): Earliest follow-up; Exercise vs. other conservative treatment	52	6004	Mean Difference (IV, Random, 95% CI)	-4.11 [-5.98, -2.23]
4.4.1 Study populations excluded leg pain and/or neurologic symptoms	19	2022	Mean Difference (IV, Random, 95% CI)	-5.85 [-8.85, -2.85]
4.4.2 Study populations included leg pain and/or neurologic symptoms	20	2789	Mean Difference (IV, Random, 95% CI)	-0.94 [-3.54, 1.66]
4.4.3 Leg pain and/or neurologic symptoms not specified	13	1193	Mean Difference (IV, Random, 95% CI)	-6.73 [-10.54, -2.92]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.1. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 1: Pain (/100): Earliest follow-up; Exercise vs. no treatment or usual care

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
4.1.1 Study populations excluded leg pain and/or neurologic symptoms								
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	2.2%	-9.50 [-21.15 , 2.15]
Cortell-Tormo 2018	15	15	11	44	14	8	2.0%	-29.00 [-42.14 , -15.86]
Cruz-Diaz 2017	17	14.1	34	49.6	13.1	15	2.6%	-32.60 [-40.75 , -24.45]
Cruz-Diaz 2017	21	13.6	34	49.6	13.1	15	2.6%	-28.60 [-36.65 , -20.55]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	2.5%	-14.90 [-23.47 , -6.33]
Dalichau 2000	21	17.67	21	55	17.67	10	2.0%	-34.00 [-47.31 , -20.69]
Dalichau 2000	31	17.67	21	55	17.67	11	2.0%	-24.00 [-36.89 , -11.11]
Garcia 2017	39.5	27.3	73	47	29.7	73	2.5%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	2.9%	-2.00 [-7.75 , 3.75]
Keane 2017	26	23.7	10	48.9	33.3	5	0.7%	-22.90 [-55.58 , 9.78]
Keane 2017	26.5	18.6	10	48.9	33.3	4	0.6%	-22.40 [-57.01 , 12.21]
Kell 2009	33	5	9	48	7	5	2.7%	-15.00 [-21.95 , -8.05]
Kell 2009	48	8	9	48	7	4	2.5%	0.00 [-8.62 , 8.62]
Masharawi 2013	16.8	8.2	16	38.8	15.4	20	2.6%	-22.00 [-29.85 , -14.15]
Mazloum 2018	48	11	15	69	16	8	2.1%	-21.00 [-33.41 , -8.59]
Mazloum 2018	30	9	20	69	16	8	2.2%	-39.00 [-50.77 , -27.23]
Steele 2013	16.43	25.76	10	25.91	14.89	4	1.2%	-9.48 [-31.11 , 12.15]
Steele 2013	25	10.97	7	25.91	14.89	3	1.4%	-0.91 [-19.62 , 17.80]
Williams 2009	24.3	17.9	43	36.9	19.81	47	2.6%	-12.60 [-20.39 , -4.81]
Yelland 2004	42	26.5	55	37	45.4	55	1.9%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**463**			**353**	**41.8%**	**-16.95 [-22.61 , -11.29]**
Heterogeneity: Tau² = 123.70; Chi² = 108.27, df = 19 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 5.87 (P < 0.00001)								
4.1.2 Study populations included leg pain and/or neurologic symptoms								
Albaladejo 2010	60	22.22	100	70	22.22	109	2.8%	-10.00 [-16.03 , -3.97]
Costa 2009b	46	28	77	56	26	77	2.5%	-10.00 [-18.53 , -1.47]
Hall 2011	34	20.31	80	47	20.54	80	2.8%	-13.00 [-19.33 , -6.67]
Highland 2018	27.5	24.3	33	33.5	17.9	29	2.3%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	2.6%	-5.00 [-13.44 , 3.44]
Morone 2011	45	23	41	76	21	29	2.3%	-31.00 [-41.39 , -20.61]
Morone 2012	42.4	21.85	25	72.08	17.19	25	2.3%	-29.68 [-40.58 , -18.78]
Muharram 2011	41.2	35.15	78	53.1	37.88	80	2.2%	-11.90 [-23.29 , -0.51]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	2.1%	-16.50 [-28.91 , -4.09]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	2.4%	-39.50 [-49.57 , -29.43]
Puppin 2011	24	27	30	41	27	25	1.9%	-17.00 [-31.33 , -2.67]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	2.5%	-15.60 [-24.89 , -6.31]
Smeets 2008	37.76	24.33	52	53.35	22.6	17	2.1%	-15.59 [-28.21 , -2.97]
Smeets 2008	42.31	25.56	55	53.35	22.6	17	2.0%	-11.04 [-23.73 , 1.65]
Smeets 2008	44.63	28.86	55	53.35	22.6	17	2.0%	-8.72 [-21.90 , 4.46]
Subtotal (95% CI)			**754**			**628**	**34.6%**	**-15.89 [-20.85 , -10.94]**
Heterogeneity: Tau² = 67.42; Chi² = 53.05, df = 14 (P < 0.00001); I² = 74%								
Test for overall effect: Z = 6.29 (P < 0.00001)								
4.1.3 Leg pain and/or neurologic symptoms not specified								
Cabak 2017	37	17.67	22	56	17.67	46	2.5%	-19.00 [-27.98 , -10.02]
Chen 2014	21.7	14.2	64	34.8	17.7	63	2.9%	-13.10 [-18.69 , -7.51]
Lee 2016a	22	11.3	15	35.8	17.2	3	1.3%	-13.80 [-34.09 , 6.49]
Lee 2016a	33.1	20	15	35.8	17.2	3	1.2%	-2.70 [-24.64 , 19.24]
Moseley 2002	19	15	29	31	15	28	2.6%	-12.00 [-19.79 , -4.21]
Oh 2014	34.4	14.11	9	38	14.11	3	1.5%	-3.60 [-22.04 , 14.84]
Oh 2014	14.4	21.23	9	38	14.11	3	1.2%	-23.60 [-44.75 , -2.45]
Oh 2014	11.1	14.11	10	38	14.11	3	1.5%	-26.90 [-45.10 , -8.70]
Quinn 2011	30.93	24.06	15	35.2	24.06	14	1.5%	-4.27 [-21.79 , 13.25]
Schinhan 2016	13.3	10.5	15	21	17.3	15	2.3%	-7.70 [-17.94 , 2.54]
Teut 2016	39.04	19.76	56	44.05	18.63	27	2.5%	-5.01 [-13.74 , 3.72]
Teut 2016	37.56	19.94	55	44.05	18.63	26	2.5%	-6.49 [-15.38 , 2.40]
Subtotal (95% CI)			**314**			**234**	**23.6%**	**-11.23 [-14.65 , -7.81]**
Heterogeneity: Tau² = 5.22; Chi² = 12.91, df = 11 (P = 0.30); I² = 15%								
Test for overall effect: Z = 6.43 (P < 0.00001)								
Total (95% CI)			**1531**			**1215**	**100.0%**	**-15.23 [-18.26 , -12.20]**
Heterogeneity: Tau² = 74.76; Chi² = 179.69, df = 46 (P < 0.00001); I² = 74%								
Test for overall effect: Z = 9.86 (P < 0.00001)								
Test for subgroup differences: Chi² = 4.01, df = 2 (P = 0.13), I² = 50.1%								

Mean Difference IV, Random, 95% CI

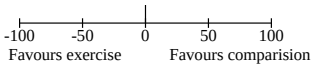
-100 -50 0 50 100

Favours exercise Favours comparision

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.1. (Continued)

Heterogeneity: Tau² = 74.76; Chi² = 179.69, df = 46 (P < 0.00001); I² = 74%
Test for overall effect: Z = 9.86 (P < 0.00001)
Test for subgroup differences: Chi² = 4.01, df = 2 (P = 0.13), I² = 50.1%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

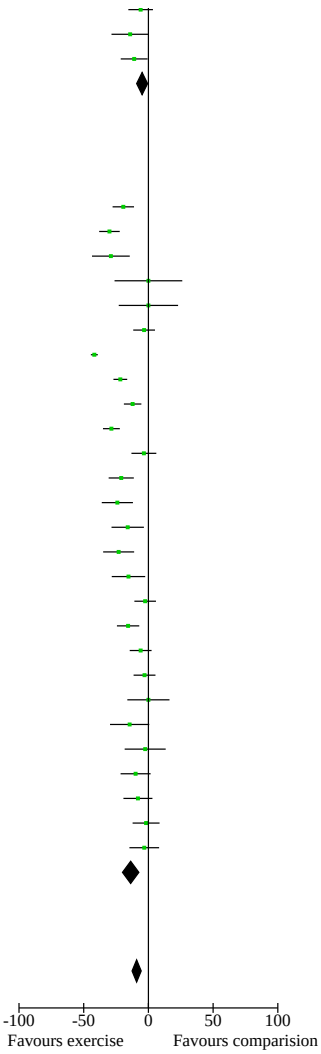
Analysis 4.2. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 2: Pain (/100): Earliest follow-up; Exercise vs. other conservative treatment

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
4.2.1 Study populations excluded leg pain and/or neurologic symptoms								
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.3%	-23.10 [-31.54 , -14.66]
Bendix 1995	44	29	31	56	29	17	1.0%	-12.00 [-29.15 , 5.15]
Bendix 1995	27	23	40	56	29	18	1.1%	-29.00 [-44.18 , -13.82]
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.3%	-24.98 [-29.09 , -20.87]
Bronfort 1996	35	22	40	39	21	21	1.2%	-4.00 [-15.28 , 7.28]
Bronfort 1996	27	20	56	39	21	21	1.2%	-12.00 [-22.40 , -1.60]
Chown 2008	73.4	17.2	22	73.2	11.5	33	1.3%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016	38.1	12.1	53	56.9	16.3	48	1.3%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.3%	-24.00 [-29.30 , -18.70]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.3%	-10.00 [-17.40 , -2.60]
Gur 2003	18	12	25	19	14	12	1.2%	-1.00 [-10.21 , 8.21]
Gur 2003	29	13	25	19	14	13	1.2%	10.00 [0.84 , 19.16]
Hansen 1993	33.33	21.25	44	33.33	21.25	59	1.3%	0.00 [-8.30 , 8.30]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	1.1%	-8.10 [-22.94 , 6.74]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.3%	-25.90 [-31.06 , -20.74]
Machado 2007	47	40	17	58	28	11	0.8%	-11.00 [-36.21 , 14.21]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.3%	-10.90 [-17.43 , -4.37]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	1.0%	-8.00 [-26.27 , 10.27]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	1.2%	-2.60 [-13.31 , 8.11]
Torstensen 1998	50.4	27.2	71	39	28	34	1.2%	11.40 [0.06 , 22.74]
Torstensen 1998	37.2	25.3	70	39	28	33	1.2%	-1.80 [-13.04 , 9.44]
Ulger 2017	21	16	56	20.8	12	57	1.3%	0.20 [-5.02 , 5.42]
Valenza 2017	39	17.67	27	49	17.67	27	1.2%	-10.00 [-19.43 , -0.57]
Subtotal (95% CI)			**917**			**740**	**27.5%**	**-9.51 [-14.59 , -4.43]**
Heterogeneity: Tau² = 125.99; Chi² = 191.57, df = 22 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 3.67 (P = 0.0002)								
4.2.2 Study populations included leg pain and/or neurologic symptoms								
Alexandre 2001	6	27.2	14	37	35.61	14	0.8%	-31.00 [-54.47 , -7.53]
Bronfort 2011	32	22	93	29	19	50	1.3%	3.00 [-3.91 , 9.91]
Bronfort 2011	26	21	91	29	19	49	1.3%	-3.00 [-9.85 , 3.85]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.3%	7.50 [2.05 , 12.95]
Cecchi 2010	25	20	68	8.33	11.67	35	1.3%	16.67 [10.54 , 22.80]
Cecchi 2010	23.33	20	68	8.33	11.67	34	1.3%	15.00 [8.84 , 21.16]
Chan 2017	30	23	49	43	26	43	1.2%	-13.00 [-23.09 , -2.91]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.2%	0.90 [-7.83 , 9.63]
Diaz-Arribas 2009	23	18.5	63	36.6	20	63	1.3%	-13.60 [-20.33 , -6.87]
Djavid 2007	43	16	19	44	20	8	1.1%	-1.00 [-16.62 , 14.62]
Djavid 2007	24	14	18	44	20	8	1.1%	-20.00 [-35.29 , -4.71]
Ferreira 2007	48	24	74	41	26	39	1.2%	7.00 [-2.82 , 16.82]
Ferreira 2007	40	25	73	41	26	38	1.2%	-1.00 [-11.06 , 9.06]
Fransoo 2006	11	17.67	20	20	17.67	10	1.1%	-9.00 [-22.41 , 4.41]
Fransoo 2006	7	17.67	19	20	17.67	9	1.1%	-13.00 [-27.01 , 1.01]
Goldby 2006	28.81	28.14	78	34.4	36.43	37	1.1%	-5.59 [-18.89 , 7.71]
Hartvigsen 2010	45.6	19.82	32	42	21.3	21	1.2%	3.60 [-7.81 , 15.01]
Hartvigsen 2010	37.97	19.38	36	42	21.3	20	1.2%	-4.03 [-15.31 , 7.25]
Hemmilä 1997	30	22.63	35	25	22.63	34	1.2%	5.00 [-5.68 , 15.68]
Jaromi 2012	5.75	17.38	56	7.07	17.38	55	1.3%	-1.32 [-7.79 , 5.15]
Longo 2016	36.2	22	7	61.4	22.7	7	0.8%	-25.20 [-48.62 , -1.78]
Maul 2005	40	20	74	50	20	74	1.3%	-10.00 [-16.44 , -3.56]
McDonough 2013	45	27.08	39	39	25.24	17	1.1%	6.00 [-8.70 , 20.70]
Murtezani 2011	20	17	50	60	27	51	1.2%	-40.00 [-48.78 , -31.22]
Rantonen 2012	31	20	43	35	28	20	1.1%	-4.00 [-17.65 , 9.65]
Rantonen 2012	29	27	43	35	28	20	1.1%	-6.00 [-20.69 , 8.69]
Ritvanen 2007	28	14.11	28	23	14.11	33	1.3%	5.00 [-2.11 , 12.11]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.9%	-3.50 [-23.38 , 16.38]
Saper 2009	44	21	15	71	18	15	1.1%	-27.00 [-41.00 , -13.00]
Schenkman 2009	19	21	17	41	40	9	0.7%	-22.00 [-49.97 , 5.97]
Sorensen 2010	49	22	91	46	21	94	1.3%	3.00 [-3.20 , 9.20]
Soukup 1999	23	16	34	24	17	18	1.2%	-1.00 [-10.52 , 8.52]
Soukup 1999	18	15.7	38	24	17	17	1.2%	-6.00 [-15.50 , 3.50]
Sousa 2009	33.5	24.8	27	47.6	28	25	1.1%	-14.10 [-28.52 , 0.32]
Williams 2005	10	11	20	21	23	24	1.2%	-11.00 [-21.39 , -0.61]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.2. (Continued)

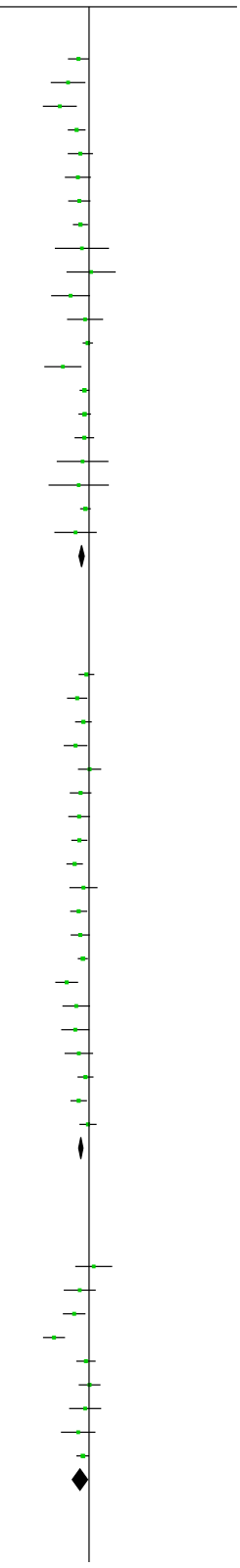
Sousa 2009	33.5	24.8	27	47.6	28	25	1.1%	-14.10 [-28.52 , 0.32]
Williams 2005	10	11	20	21	23	24	1.2%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**1595**			**1171**	**40.7%**	**-4.87 [-9.09 , -0.65]**
Heterogeneity: Tau² = 126.19; Chi² = 225.48, df = 34 (P < 0.00001); I² = 85%								
Test for overall effect: Z = 2.26 (P = 0.02)								
4.2.3 Leg pain and/or neurologic symptoms not specified								
Alayat 2014	37.1	13	24	56.5	10.4	10	1.3%	-19.40 [-27.68 , -11.12]
Alayat 2014	26.4	12.5	28	56.5	10.4	10	1.3%	-30.10 [-38.04 , -22.16]
Barberini 2011	52	17.38	10	81	17.38	12	1.1%	-29.00 [-43.59 , -14.41]
Bellido-Fernandez 2018	30	22.22	9	30	22.22	4	0.8%	0.00 [-26.17 , 26.17]
Bellido-Fernandez 2018	30	18.52	9	30	22.22	5	0.9%	0.00 [-22.93 , 22.93]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.3%	-3.32 [-11.72 , 5.08]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.3%	-41.79 [-44.48 , -39.10]
Kim 2017	24	5.4	15	45.7	9	15	1.3%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.3%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.3%	-28.60 [-35.11 , -22.09]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.2%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	1.2%	-21.00 [-30.72 , -11.28]
Miyamoto 2018	32	24	69	56	26	23	1.2%	-24.00 [-36.04 , -11.96]
Miyamoto 2018	40	27	73	56	26	22	1.2%	-16.00 [-28.51 , -3.49]
Miyamoto 2018	33	23	67	56	26	23	1.2%	-23.00 [-34.97 , -11.03]
Natour 2015	40.4	24.2	30	55.8	26.8	30	1.1%	-15.40 [-28.32 , -2.48]
Noori 2011	41.5	15.3	17	44	7.5	13	1.3%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.2%	-15.70 [-24.32 , -7.08]
Saper 2017	50	21	129	56	22	32	1.3%	-6.00 [-14.44 , 2.44]
Saper 2017	53	21	127	56	22	32	1.3%	-3.00 [-11.45 , 5.45]
Segal-Snir 2016	70	30	20	70	19	15	1.0%	0.00 [-16.29 , 16.29]
Shen 2009	11	17.67	15	25.5	17.67	8	1.1%	-14.50 [-29.66 , 0.66]
Shen 2009	23	17.67	15	25.5	17.67	7	1.1%	-2.50 [-18.35 , 13.35]
Weifen 2013	22.5	17.67	47	32.4	17.67	11	1.2%	-9.90 [-21.50 , 1.70]
Weifen 2013	24.3	17.67	47	32.4	17.67	12	1.2%	-8.10 [-19.30 , 3.10]
Weifen 2013	30.6	17.67	141	32.4	17.67	12	1.2%	-1.80 [-12.21 , 8.61]
Weifen 2013	29.2	17.67	38	32.4	17.67	12	1.2%	-3.20 [-14.67 , 8.27]
Subtotal (95% CI)			**1246**			**626**	**31.8%**	**-13.71 [-20.03 , -7.39]**
Heterogeneity: Tau² = 245.19; Chi² = 368.56, df = 26 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 4.25 (P < 0.0001)								
Total (95% CI)			**3758**			**2537**	**100.0%**	**-9.10 [-12.63 , -5.57]**
Heterogeneity: Tau² = 240.33; Chi² = 1171.16, df = 84 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 5.06 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.57, df = 2 (P = 0.06), I² = 64.1%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.3. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 3: Function (/100): Earliest follow-up; Exercise vs. no treatment or usual care

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
4.3.1 Study populations excluded leg pain and/or neurologic symptoms								
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.1%	-8.50 [-16.99 , -0.01]
Cruz-Diaz 2017	26.46	22.08	34	43.37	23.33	15	1.0%	-16.91 [-30.86 , -2.96]
Cruz-Diaz 2017	19.83	20.42	34	43.37	23.33	15	1.0%	-23.54 [-37.20 , -9.88]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	2.6%	-10.00 [-17.14 , -2.86]
Dalichau 2000	42	13.92	21	49	13.92	11	1.6%	-7.00 [-17.15 , 3.15]
Dalichau 2000	40	13.92	21	49	13.92	10	1.5%	-9.00 [-19.48 , 1.48]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	1.9%	-7.83 [-16.79 , 1.13]
Groessl 2017	30.62	19.06	75	37.54	19.06	75	3.0%	-6.92 [-13.02 , -0.82]
Keane 2017	25.8	20.1	10	31.56	18.24	4	0.4%	-5.76 [-27.55 , 16.03]
Keane 2017	33.2	18.84	10	31.56	18.24	5	0.5%	1.64 [-18.16 , 21.44]
Kell 2009	24.2	13.27	9	39.1	13.27	4	0.8%	-14.90 [-30.53 , 0.73]
Kell 2009	35.9	13.27	9	39.1	13.27	5	0.9%	-3.20 [-17.71 , 11.31]
Kuukkanen 2000	10	7.6	27	11	7.8	26	4.2%	-1.00 [-5.15 , 3.15]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	0.9%	-21.08 [-36.02 , -6.14]
Mazloum 2018	22.9	3.6	15	26.6	5	8	4.3%	-3.70 [-7.61 , 0.21]
Mazloum 2018	23.1	7.5	16	26.6	5	8	3.6%	-3.50 [-8.55 , 1.55]
Patti 2016	13	8	19	16.8	15.6	19	2.3%	-3.80 [-11.68 , 4.08]
Steele 2013	17.98	17.95	10	23.2	17.95	4	0.5%	-5.22 [-26.03 , 15.59]
Steele 2013	14.86	17.95	7	23.2	17.95	3	0.4%	-8.34 [-32.62 , 15.94]
Williams 2009	17.9	10.49	43	20.8	10.28	47	4.1%	-2.90 [-7.20 , 1.40]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	0.7%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**544**			**439**	**38.3%**	**-6.01 [-8.14 , -3.87]**
Heterogeneity: Tau² = 5.04; Chi² = 26.22, df = 20 (P = 0.16); I² = 24%								
Test for overall effect: Z = 5.51 (P < 0.00001)								
4.3.2 Study populations included leg pain and/or neurologic symptoms								
Albaladejo 2010	31.25	24.69	100	33.33	21.6	109	2.9%	-2.08 [-8.39 , 4.23]
Costa 2009b	40	27.08	77	49.58	24.58	77	2.2%	-9.58 [-17.75 , -1.41]
Hall 2011	29.21	21.49	80	33.75	21.87	80	2.7%	-4.54 [-11.26 , 2.18]
Highland 2018	18.46	19.25	33	29.33	18.96	29	1.8%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	1.8%	0.43 [-8.93 , 9.79]
Kofotolis 2006	21.2	13.92	28	28	13.92	15	2.0%	-6.80 [-15.53 , 1.93]
Kofotolis 2006	20	13.92	28	28	13.92	15	2.0%	-8.00 [-16.73 , 0.73]
Morone 2011	18	12.9	41	25.8	14.1	29	2.9%	-7.80 [-14.28 , -1.32]
Morone 2012	13.68	9.39	25	25.28	14.03	25	2.8%	-11.60 [-18.22 , -4.98]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	1.4%	-4.50 [-15.82 , 6.82]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	2.7%	-8.34 [-15.23 , -1.45]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	2.3%	-7.10 [-14.82 , 0.62]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	4.2%	-5.00 [-9.08 , -0.92]
Shaughnessy 2004	26	14	20	44	16	21	1.9%	-18.00 [-27.19 , -8.81]
Smeets 2008	47.5	21.88	55	57.83	19.92	17	1.4%	-10.33 [-21.43 , 0.77]
Smeets 2008	46.74	23.01	52	57.83	19.92	17	1.4%	-11.09 [-22.44 , 0.26]
Smeets 2008	49.6	24.56	55	57.83	19.92	17	1.3%	-8.23 [-19.71 , 3.25]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	2.9%	-2.87 [-9.31 , 3.57]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	2.8%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	2.7%	-0.83 [-7.72 , 6.06]
Subtotal (95% CI)			**932**			**744**	**46.0%**	**-6.63 [-8.37 , -4.90]**
Heterogeneity: Tau² = 1.11; Chi² = 20.46, df = 19 (P = 0.37); I² = 7%								
Test for overall effect: Z = 7.49 (P < 0.00001)								
4.3.3 Leg pain and/or neurologic symptoms not specified								
Harts 2008	25.42	25	20	21.67	16.25	10	0.9%	3.75 [-11.13 , 18.63]
Harts 2008	14.17	16.67	20	21.67	16.25	9	1.1%	-7.50 [-20.39 , 5.39]
Kofotolis 2016	33.25	10.17	36	45.25	16.17	14	1.9%	-12.00 [-21.10 , -2.90]
Kofotolis 2016	17.04	8.08	37	45.25	16.17	14	2.0%	-28.21 [-37.07 , -19.35]
Lee 2016a	4.58	3.75	15	7.08	6.67	3	2.3%	-2.50 [-10.28 , 5.28]
Lee 2016a	7.5	8.75	15	7.08	6.67	3	2.0%	0.42 [-8.33 , 9.17]
Little 2014	27.95	13.92	15	31.08	19.87	13	1.1%	-3.13 [-16.03 , 9.77]
Quinn 2011	22.5	19.06	15	31.25	19.06	14	1.0%	-8.75 [-22.63 , 5.13]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	3.6%	-5.06 [-10.19 , 0.07]
Subtotal (95% CI)			**188**			**95**	**15.8%**	**-7.34 [-13.52 , -1.16]**
Heterogeneity: Tau² = 61.46; Chi² = 30.16, df = 8 (P = 0.0002); I² = 73%								
Test for overall effect: Z = 2.33 (P = 0.02)								



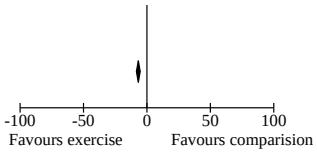
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.3. (Continued)

Test for overall effect: Z = 2.33 (P = 0.02)

Total (95% CI)	**1664**	**1278**	**100.0%**	**-6.80 [-8.30 , -5.29]**

Heterogeneity: Tau² = 9.62; Chi² = 78.93, df = 49 (P = 0.004); I² = 38%
Test for overall effect: Z = 8.87 (P < 0.00001)
Test for subgroup differences: Chi² = 0.29, df = 2 (P = 0.87), I² = 0%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

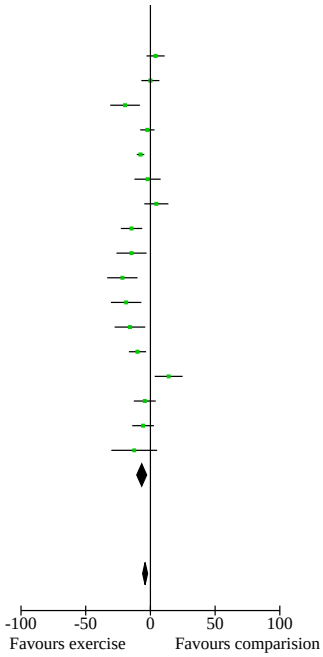
Analysis 4.4. Comparison 4: SUBGROUP ANALYSES (Radicular symptoms), Outcome 4: Function (/100): Earliest follow-up; Exercise vs. other conservative treatment

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
4.4.1 Study populations excluded leg pain and/or neurologic symptoms								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.0%	-12.84 [-15.31 , -10.37]
Bid 2017	7.29	8.54	64	24.12	12	64	1.9%	-16.83 [-20.44 , -13.22]
Bronfort 1996	15.1	17.4	43	20.8	17.3	21	1.4%	-5.70 [-14.74 , 3.34]
Bronfort 1996	20.9	17	56	20.8	17.3	21	1.4%	0.10 [-8.54 , 8.74]
Chown 2008	20.8	8.4	35	20.3	8	35	1.9%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.7%	-4.90 [-11.09 , 1.29]
Gur 2003	22.91	13.33	25	27.5	12.08	13	1.4%	-4.59 [-12.98 , 3.80]
Gur 2003	26.25	14.58	25	27.5	12.08	12	1.4%	-1.25 [-10.16 , 7.66]
Machado 2007	50	22.92	17	57.5	23.96	11	0.7%	-7.50 [-25.37 , 10.37]
Niemisto 2003	14.7	11.6	96	18.6	11	100	1.9%	-3.90 [-7.07 , -0.73]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	1.5%	-7.00 [-14.73 , 0.73]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.4%	-10.80 [-19.97 , -1.63]
Sherman 2005	22.08	18.75	36	26.67	21.67	15	1.1%	-4.59 [-17.15 , 7.97]
Sherman 2005	12.92	15	35	26.67	21.67	15	1.1%	-13.75 [-25.79 , -1.71]
Sherman 2011	18.74	16.28	91	28.52	20.61	22	1.4%	-9.78 [-19.02 , -0.54]
Sherman 2011	19.26	17.56	92	28.52	20.61	23	1.4%	-9.26 [-18.42 , -0.10]
Tilbrook 2011	23.75	20.31	156	32.17	20.31	136	1.8%	-8.42 [-13.09 , -3.75]
Torstensen 1998	46.2	13.1	70	46.9	13.1	33	1.7%	-0.70 [-6.12 , 4.72]
Torstensen 1998	52.7	16.6	71	46.9	13.1	34	1.7%	5.80 [-0.06 , 11.66]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.3%	9.20 [-0.98 , 19.38]
Valenza 2017	16.96	13.92	27	29.92	13.92	27	1.5%	-12.96 [-20.39 , -5.53]
Vincent 2014a	34.17	22.92	17	26.25	17.5	7	0.8%	7.92 [-9.01 , 24.85]
Vincent 2014a	20.83	17.08	18	26.25	17.5	7	0.9%	-5.42 [-20.60 , 9.76]
Subtotal (95% CI)			**1202**			**820**	**36.8%**	**-5.85 [-8.85 , -2.85]**
Heterogeneity: Tau² = 41.40; Chi² = 137.75, df = 24 (P < 0.00001); I² = 83%								
Test for overall effect: Z = 3.82 (P = 0.0001)								
4.4.2 Study populations included leg pain and/or neurologic symptoms								
Bronfort 2011	16.96	20	91	21.3	21.74	49	1.6%	-4.34 [-11.68 , 3.00]
Bronfort 2011	18.7	19.57	92	21.3	21.74	50	1.6%	-2.60 [-9.83 , 4.63]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.8%	0.83 [-3.79 , 5.45]
Cecchi 2010	22.08	19.58	68	9.17	13.75	35	1.6%	12.91 [6.40 , 19.42]
Cecchi 2010	22.5	19.58	68	9.17	13.75	34	1.6%	13.33 [6.77 , 19.89]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.7%	-6.80 [-12.74 , -0.86]
Djavid 2007	33.6	7.4	18	41.6	8.8	8	1.6%	-8.00 [-14.99 , -1.01]
Djavid 2007	48.2	10.4	19	41.6	8.8	8	1.5%	6.60 [-1.08 , 14.28]
Ferreira 2007	40.42	26.25	73	32.92	25	38	1.3%	7.50 [-2.47 , 17.47]
Ferreira 2007	32.92	23.75	74	32.92	25	39	1.3%	0.00 [-9.53 , 9.53]
Fransoo 2006	11.25	13.92	19	13.75	13.92	9	1.2%	-2.50 [-13.54 , 8.54]
Fransoo 2006	4.58	13.92	20	13.75	13.92	10	1.2%	-9.17 [-19.74 , 1.40]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.0%	-1.80 [-3.96 , 0.36]
Goldby 2006	31	17.07	78	28.1	17.34	37	1.6%	2.90 [-3.85 , 9.65]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.7%	-17.84 [-36.69 , 1.01]
Maul 2005	25	8.33	74	25	12.5	74	1.9%	0.00 [-3.42 , 3.42]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	1.7%	-14.80 [-20.62 , -8.98]
Rantonen 2012	15	14	43	16	10	20	1.7%	-1.00 [-7.06 , 5.06]
Rantonen 2012	14	11	43	16	10	20	1.7%	-2.00 [-7.48 , 3.48]
Ritvanen 2007	17	13.27	28	12	13.27	33	1.6%	5.00 [-1.68 , 11.68]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.0%	8.75 [-4.68 , 22.18]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.7%	-18.27 [-36.90 , 0.36]
Schenkman 2009	15.81	10.4	17	10.4	16.64	9	1.1%	5.41 [-6.53 , 17.35]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.2%	-11.88 [-23.25 , -0.51]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	1.8%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	1.8%	-10.33 [-15.25 , -5.41]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.5%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**1560**			**1229**	**40.1%**	**-0.94 [-3.54 , 1.66]**
Heterogeneity: Tau² = 31.53; Chi² = 112.38, df = 26 (P < 0.00001); I² = 77%								
Test for overall effect: Z = 0.71 (P = 0.48)								
4.4.3 Leg pain and/or neurologic symptoms not specified								
Bellido-Fernandez 2018	12	5.93	9	8	6.67	5	1.6%	4.00 [-3.01 , 11.01]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.4. (Continued)

4.4.3 Leg pain and/or neurologic symptoms not specified								
Bellido-Fernandez 2018	12	5.93	9	8	6.67	5	1.6%	4.00 [-3.01 , 11.01]
Bellido-Fernandez 2018	8	3.7	9	8	6.67	4	1.6%	0.00 [-6.97 , 6.97]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	1.7%	-2.37 [-7.87 , 3.13]
Kim 2017	21.27	3.37	15	28.93	4.61	15	1.9%	-7.66 [-10.55 , -4.77]
Little 2014	23.16	13.92	14	25.37	13.92	15	1.3%	-2.21 [-12.35 , 7.93]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]
Miyamoto 2018	32.5	21.67	67	47.08	25.42	23	1.1%	-14.58 [-26.19 , -2.97]
Miyamoto 2018	25.42	22.92	69	47.08	25.42	23	1.1%	-21.66 [-33.37 , -9.95]
Miyamoto 2018	28.33	21.67	73	47.08	25.42	22	1.1%	-18.75 [-30.48 , -7.02]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.1%	-15.83 [-27.67 , -3.99]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.6%	-9.97 [-16.60 , -3.34]
Okafor 2012	25	20.8	15	10.9	4.46	15	1.2%	14.10 [3.33 , 24.87]
Saper 2017	49.13	22.17	127	53.48	21.74	32	1.4%	-4.35 [-12.81 , 4.11]
Saper 2017	47.83	21.3	129	53.48	21.74	32	1.4%	-5.65 [-14.03 , 2.73]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.7%	-12.50 [-30.19 , 5.19]
Subtotal (95% CI)			**772**			**421**	**23.1%**	**-6.73 [-10.54 , -2.92]**
Heterogeneity: Tau² = 41.89; Chi² = 60.16, df = 16 (P < 0.00001); I² = 73%								
Test for overall effect: Z = 3.46 (P = 0.0005)								
Total (95% CI)			**3534**			**2470**	**100.0%**	**-4.11 [-5.98 , -2.23]**
Heterogeneity: Tau² = 44.57; Chi² = 383.17, df = 68 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 4.30 (P < 0.0001)								
Test for subgroup differences: Chi² = 8.72, df = 2 (P = 0.01), I² = 77.1%								



Comparison 5. SUBGROUP ANALYSES (Conservative comparisons)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Pain (/100): Earliest follow-up, Exercise vs. other conservative treatments	64	6295	Mean Difference (IV, Random, 95% CI)	-9.10 [-12.63, -5.57]
5.1.1 Exercise vs. Education	20	2039	Mean Difference (IV, Random, 95% CI)	-12.17 [-19.38, -4.97]
5.1.2 Exercise vs. Manual therapy	12	1632	Mean Difference (IV, Random, 95% CI)	0.99 [-3.14, 5.13]
5.1.3 Exercise vs. Electrotherapy	3	200	Mean Difference (IV, Random, 95% CI)	-10.34 [-23.86, 3.17]
5.1.4 Exercise vs. Psychological therapy	2	134	Mean Difference (IV, Random, 95% CI)	-18.67 [-30.75, -6.60]
5.1.5 Exercise vs. Mixed physical therapy	21	1869	Mean Difference (IV, Random, 95% CI)	-10.42 [-15.22, -5.62]
5.1.6 Exercise vs. Back school	3	241	Mean Difference (IV, Random, 95% CI)	-14.02 [-22.68, -5.35]
5.1.7 Exercise vs. Relaxation	1	68	Mean Difference (IV, Random, 95% CI)	-3.50 [-13.11, 6.11]
5.1.8 Exercise vs. Anti-inflammatories	2	112	Mean Difference (IV, Random, 95% CI)	-14.82 [-24.44, -5.20]

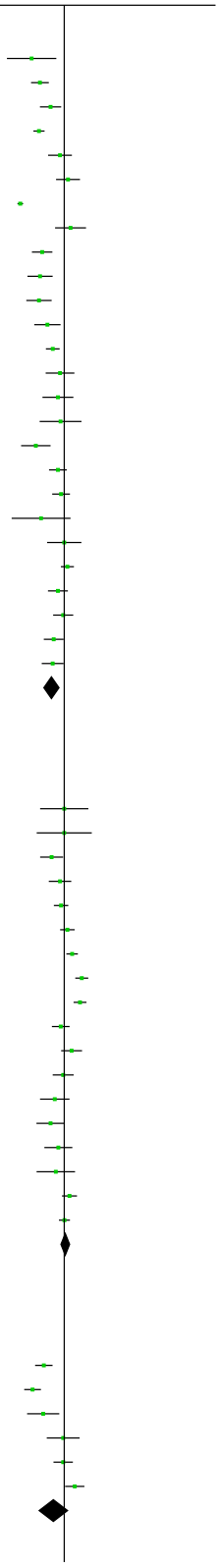
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.2 Function (/100): Earliest follow-up, Exercise vs. other conservative treatments	52	6004	Mean Difference (IV, Random, 95% CI)	-4.10 [-5.97, -2.24]
5.2.1 Exercise vs. Education	19	2598	Mean Difference (IV, Random, 95% CI)	-8.03 [-10.75, -5.30]
5.2.2 Exercise vs. Manual therapy	12	1637	Mean Difference (IV, Random, 95% CI)	1.46 [-1.47, 4.38]
5.2.3 Exercise vs. Electrotherapy	4	248	Mean Difference (IV, Random, 95% CI)	-3.81 [-10.07, 2.44]
5.2.4 Exercise vs. Psychological therapy	1	28	Mean Difference (IV, Random, 95% CI)	-7.50 [-25.37, 10.37]
5.2.5 Exercise vs. Mixed physical therapy	12	1074	Mean Difference (IV, Random, 95% CI)	-4.28 [-9.00, 0.43]
5.2.6 Exercise vs. Back school	2	209	Mean Difference (IV, Random, 95% CI)	-1.74 [-6.34, 2.85]
5.2.7 Exercise vs. Relaxation	1	68	Mean Difference (IV, Random, 95% CI)	4.59 [-4.72, 13.90]
5.2.8 Exercise vs. Anti-inflammatories	2	142	Mean Difference (IV, Random, 95% CI)	-13.78 [-21.98, -5.57]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.1. Comparison 5: SUBGROUP ANALYSES (Conservative comparisons), Outcome 1: Pain (/100): Earliest follow-up, Exercise vs. other conservative treatments

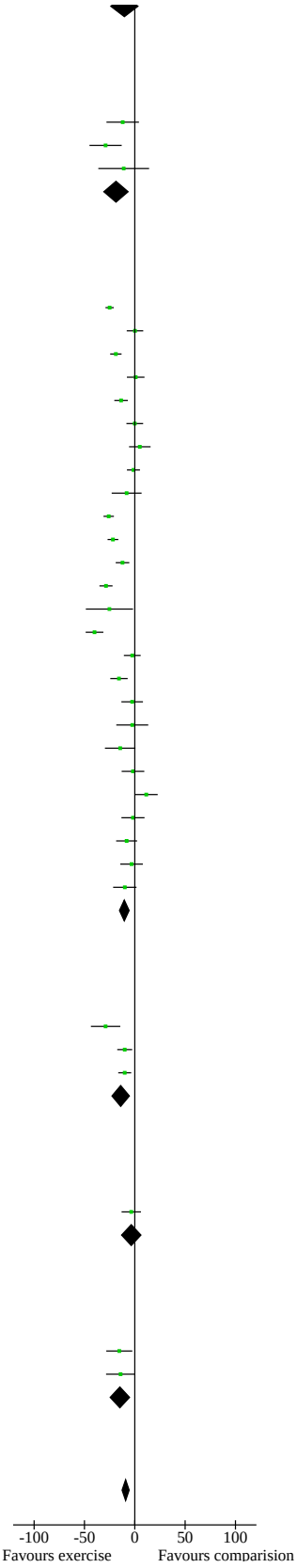
Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
5.1.1 Exercise vs. Education								
Alexandre 2001 (1)	6	27.2	14	37	35.61	14	0.8%	-31.00 [-54.47 , -7.53]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.3%	-23.10 [-31.54 , -14.66]
Chan 2017	30	23	49	43	26	43	1.2%	-13.00 [-23.09 , -2.91]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.3%	-24.00 [-29.30 , -18.70]
Hartvigsen 2010 (2)	37.97	19.38	32	42	21.3	21	1.2%	-4.03 [-15.35 , 7.29]
Hartvigsen 2010 (2)	45.6	19.82	36	42	21.3	20	1.2%	3.60 [-7.76 , 14.96]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.3%	-41.79 [-44.48 , -39.10]
McDonough 2013 (3)	45	27.08	39	39	25.24	17	1.1%	6.00 [-8.70 , 20.70]
Miyamoto 2013	31	23	43	52	23	43	1.2%	-21.00 [-30.72 , -11.28]
Miyamoto 2018 (2)	33	23	67	56	26	23	1.2%	-23.00 [-34.97 , -11.03]
Miyamoto 2018 (2)	32	24	69	56	26	23	1.2%	-24.00 [-36.04 , -11.96]
Miyamoto 2018 (2)	40	27	73	56	26	22	1.2%	-16.00 [-28.51 , -3.49]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.3%	-10.90 [-17.43 , -4.37]
Rantonen 2012 (2)	31	20	43	35	28	20	1.1%	-4.00 [-17.65 , 9.65]
Rantonen 2012 (2)	29	27	43	35	28	20	1.1%	-6.00 [-20.69 , 8.69]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.9%	-3.50 [-23.38 , 16.38]
Saper 2009	44	21	15	71	18	15	1.1%	-27.00 [-41.00 , -13.00]
Saper 2017 (2)	50	21	127	56	22	32	1.3%	-6.00 [-14.45 , 2.45]
Saper 2017 (2)	53	21	129	56	22	32	1.3%	-3.00 [-11.44 , 5.44]
Schenkman 2009 (4)	19	21	17	41	40	9	0.7%	-22.00 [-49.97 , 5.97]
Segal-Snir 2016	70	30	20	70	19	15	1.0%	0.00 [-16.29 , 16.29]
Sorensen 2010	49	22	91	46	21	94	1.3%	3.00 [-3.20 , 9.20]
Soukup 1999 (2)	18	15.7	34	24	17	18	1.2%	-6.00 [-15.46 , 3.46]
Soukup 1999 (2)	23	16	38	24	17	17	1.2%	-1.00 [-10.55 , 8.55]
Valenza 2017 (3)	39	17.67	27	49	17.67	27	1.2%	-10.00 [-19.43 , -0.57]
Williams 2005	10	11	20	21	23	24	1.2%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**1257**			**782**	**30.2%**	**-12.17 [-19.38 , -4.97]**
Heterogeneity: Tau² = 311.77; Chi² = 425.43, df = 25 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 3.31 (P = 0.0009)								
5.1.2 Exercise vs. Manual therapy								
Bellido-Fernandez 2018 (2)	30	18.52	9	30	22.22	5	0.9%	0.00 [-22.93 , 22.93]
Bellido-Fernandez 2018 (2)	30	22.22	9	30	22.22	4	0.8%	0.00 [-26.17 , 26.17]
Bronfort 1996 (5)	27	20	40	39	21	21	1.2%	-12.00 [-22.91 , -1.09]
Bronfort 1996 (5)	35	22	56	39	21	21	1.2%	-4.00 [-14.67 , 6.67]
Bronfort 2011 (2)	26	21	91	29	19	49	1.3%	-3.00 [-9.85 , 3.85]
Bronfort 2011 (2)	32	22	93	29	19	50	1.3%	3.00 [-3.91 , 9.91]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.3%	7.50 [2.05 , 12.95]
Cecchi 2010 (2)	25	20	68	8.33	11.67	34	1.3%	16.67 [10.51 , 22.83]
Cecchi 2010 (2)	23.33	20	68	8.33	11.67	35	1.3%	15.00 [8.87 , 21.13]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.3%	-3.32 [-11.72 , 5.08]
Ferreira 2007 (2)	48	24	73	41	26	38	1.2%	7.00 [-2.93 , 16.93]
Ferreira 2007 (2)	40	25	74	41	26	39	1.2%	-1.00 [-10.95 , 8.95]
Fransoo 2006 (5)	11	17.67	19	20	17.67	9	1.1%	-9.00 [-23.01 , 5.01]
Fransoo 2006 (5)	7	17.67	20	20	17.67	10	1.1%	-13.00 [-26.41 , 0.41]
Goldby 2006 (4)	28.81	28.14	78	34.4	36.43	37	1.1%	-5.59 [-18.89 , 7.71]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	1.0%	-8.00 [-26.27 , 10.27]
Ritvanen 2007 (3)	28	14.11	28	23	14.11	33	1.3%	5.00 [-2.11 , 12.11]
Ulger 2017	21	16	56	20.8	12	57	1.3%	0.20 [-5.02 , 5.42]
Subtotal (95% CI)			**974**			**658**	**21.1%**	**0.99 [-3.14 , 5.13]**
Heterogeneity: Tau² = 51.70; Chi² = 64.12, df = 17 (P < 0.00001); I² = 73%								
Test for overall effect: Z = 0.47 (P = 0.64)								
5.1.3 Exercise vs. Electrotherapy								
Alayat 2014 (2)	37.1	13	24	56.5	10.4	10	1.3%	-19.40 [-27.68 , -11.12]
Alayat 2014 (2)	26.4	12.5	28	56.5	10.4	10	1.3%	-30.10 [-38.04 , -22.16]
Djavid 2007 (2)	24	14	18	44	20	8	1.1%	-20.00 [-35.29 , -4.71]
Djavid 2007 (2)	43	16	19	44	20	8	1.1%	-1.00 [-16.62 , 14.62]
Gur 2003 (2)	18	12	25	19	14	12	1.2%	-1.00 [-10.21 , 8.21]
Gur 2003 (2)	29	13	25	19	14	13	1.2%	10.00 [0.84 , 19.16]
Subtotal (95% CI)			**139**			**61**	**7.1%**	**-10.34 [-23.86 , 3.17]**
Heterogeneity: Tau² = 252.99; Chi² = 53.64, df = 5 (P < 0.00001); I² = 91%								
Test for overall effect: Z = 1.50 (P = 0.13)								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.1. (Continued)

Heterogeneity: Tau² = 252.99; Chi² = 53.64, df = 5 (P < 0.00001); I² = 91%								
Test for overall effect: Z = 1.50 (P = 0.13)								
5.1.4 Exercise vs. Psychological therapy								
Bendix 1995 (2)	44	29	40	56	29	18	1.1%	-12.00 [-28.13 , 4.13]
Bendix 1995 (2)	27	23	31	56	29	17	1.1%	-29.00 [-44.99 , -13.01]
Machado 2007	47	40	17	58	28	11	0.8%	-11.00 [-36.21 , 14.21]
Subtotal (95% CI)			**88**			**46**	**2.9%**	**-18.67 [-30.75 , -6.60]**
Heterogeneity: Tau² = 27.33; Chi² = 2.61, df = 2 (P = 0.27); I² = 23%								
Test for overall effect: Z = 3.03 (P = 0.002)								
5.1.5 Exercise vs. Mixed physical therapy								
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.3%	-24.98 [-29.09 , -20.87]
Chown 2008 (4)	73.4	17.2	22	73.2	11.5	33	1.3%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016 (3)	38.1	12.1	53	56.9	16.3	48	1.3%	-18.80 [-24.45 , -13.15]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.2%	0.90 [-7.83 , 9.63]
Diaz-Arribas 2009 (3)	23	18.5	63	36.6	20	63	1.3%	-13.60 [-20.33 , -6.87]
Hansen 1993 (6)	33.33	21.25	44	33.33	21.25	59	1.3%	0.00 [-8.30 , 8.30]
Hemmilä 1997 (6)	30	22.63	35	25	22.63	34	1.2%	5.00 [-5.68 , 15.68]
Jaromi 2012 (3)	5.75	17.38	56	7.07	17.38	55	1.3%	-1.32 [-7.79 , 5.15]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	1.1%	-8.10 [-22.94 , 6.74]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.3%	-25.90 [-31.06 , -20.74]
Kim 2017	24	5.4	15	45.7	9	15	1.3%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.3%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.3%	-28.60 [-35.11 , -22.09]
Longo 2016	36.2	22	7	61.4	22.7	7	0.8%	-25.20 [-48.62 , -1.78]
Murtezani 2011	20	17	50	60	27	51	1.2%	-40.00 [-48.78 , -31.22]
Noori 2011	41.5	15.3	17	44	7.5	13	1.3%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.2%	-15.70 [-24.32 , -7.08]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	1.2%	-2.60 [-13.31 , 8.11]
Shen 2009 (5)	23	17.67	15	25.5	17.67	7	1.1%	-2.50 [-18.35 , 13.35]
Shen 2009 (5)	11	17.67	15	25.5	17.67	8	1.1%	-14.50 [-29.66 , 0.66]
Torstensen 1998 (5)	37.2	25.3	70	39	28	33	1.2%	-1.80 [-13.04 , 9.44]
Torstensen 1998 (5)	50.4	27.2	71	39	28	34	1.2%	11.40 [0.06 , 22.74]
Weifen 2013 (5)	30.6	17.67	38	32.4	17.67	12	1.2%	-1.80 [-13.27 , 9.67]
Weifen 2013 (5)	24.3	17.67	141	32.4	17.67	12	1.2%	-8.10 [-18.51 , 2.31]
Weifen 2013 (5)	29.2	17.67	47	32.4	17.67	12	1.2%	-3.20 [-14.40 , 8.00]
Weifen 2013 (5)	22.5	17.67	47	32.4	17.67	11	1.2%	-9.90 [-21.50 , 1.70]
Subtotal (95% CI)			**1087**			**782**	**31.5%**	**-10.42 [-15.22 , -5.62]**
Heterogeneity: Tau² = 129.98; Chi² = 220.92, df = 25 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 4.26 (P < 0.0001)								
5.1.6 Exercise vs. Back school								
Barberini 2011 (3)	52	17.38	10	81	17.38	12	1.1%	-29.00 [-43.59 , -14.41]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.3%	-10.00 [-17.40 , -2.60]
Maul 2005	40	20	74	50	20	74	1.3%	-10.00 [-16.44 , -3.56]
Subtotal (95% CI)			**120**			**121**	**3.7%**	**-14.02 [-22.68 , -5.35]**
Heterogeneity: Tau² = 37.08; Chi² = 5.87, df = 2 (P = 0.05); I² = 66%								
Test for overall effect: Z = 3.17 (P = 0.002)								
5.1.7 Exercise vs. Relaxation								
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.2%	-3.50 [-13.11 , 6.11]
Subtotal (95% CI)			**36**			**32**	**1.2%**	**-3.50 [-13.11 , 6.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.71 (P = 0.48)								
5.1.8 Exercise vs. Anti-inflammatories								
Natour 2015	40.4	24.2	30	55.8	26.8	30	1.1%	-15.40 [-28.32 , -2.48]
Sousa 2009	33.5	24.8	27	47.6	28	25	1.1%	-14.10 [-28.52 , 0.32]
Subtotal (95% CI)			**57**			**55**	**2.2%**	**-14.82 [-24.44 , -5.20]**
Heterogeneity: Tau² = 0.00; Chi² = 0.02, df = 1 (P = 0.90); I² = 0%								
Test for overall effect: Z = 3.02 (P = 0.003)								
Total (95% CI)			**3758**			**2537**	**100.0%**	**-9.10 [-12.63 , -5.57]**
Heterogeneity: Tau² = 240.09; Chi² = 1170.16, df = 84 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 5.06 (P < 0.00001)								
Test for subgroup differences: Chi² = 27.65, df = 7 (P = 0.0003), I² = 74.7%								



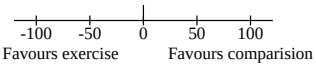
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.1. (Continued)

Heterogeneity: Tau² = 240.03; Chi² = 1170.10, df = 84 (P < 0.00001); I² = 93%

Test for overall effect: Z = 5.06 (P < 0.00001)

Test for subgroup differences: Chi² = 27.65, df = 7 (P = 0.0003), I² = 74.7%



Footnotes

(1) Study outcome data was excluded from meta-analysis due to implausible mean effect.
(2) Study has multiple exercise groups available: comparison group data divided.
(3) Standard deviation(s) required imputation due to missing or implausible value.
(4) Study has other comparison groups available for this meta-analysis.
(5) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(6) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

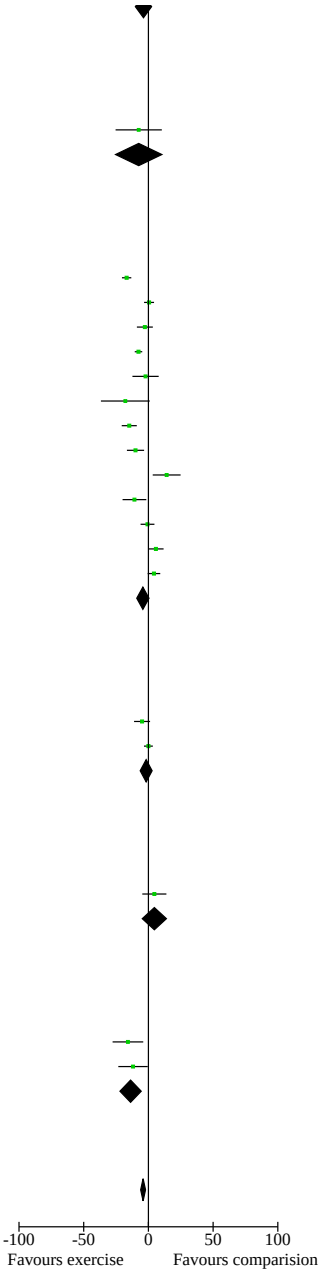
Analysis 5.2. Comparison 5: SUBGROUP ANALYSES (Conservative comparisons), Outcome 2: Function (/100): Earliest follow-up, Exercise vs. other conservative treatments

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
5.2.1 Exercise vs. Education								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.0%	-12.84 [-15.31 , -10.37]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.7%	-6.80 [-12.74 , -0.86]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.0%	-1.80 [-3.96 , 0.36]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]
Miyamoto 2018 (1)	28.33	21.67	69	47.08	25.42	23	1.1%	-18.75 [-30.33 , -7.17]
Miyamoto 2018 (1)	32.5	21.67	67	47.08	25.42	23	1.1%	-14.58 [-26.19 , -2.97]
Miyamoto 2018 (1)	25.42	22.92	73	47.08	25.42	22	1.1%	-21.66 [-33.51 , -9.81]
Niemisto 2003	14.7	11.6	96	18.6	11	100	1.9%	-3.90 [-7.07 , -0.73]
Rantonen 2012 (1)	15	14	43	16	10	20	1.7%	-1.00 [-7.06 , 5.06]
Rantonen 2012 (1)	14	11	43	16	10	20	1.7%	-2.00 [-7.48 , 3.48]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.0%	8.75 [-4.68 , 22.18]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.7%	-18.27 [-36.90 , 0.36]
Saper 2017 (1)	49.13	22.17	127	53.48	21.74	32	1.4%	-4.35 [-12.81 , 4.11]
Saper 2017 (1)	47.83	21.3	129	53.48	21.74	32	1.4%	-5.65 [-14.03 , 2.73]
Schenkman 2009 (2)	15.81	10.4	17	10.4	16.64	9	1.1%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.7%	-12.50 [-30.19 , 5.19]
Sherman 2005 (1)	22.08	18.75	36	26.67	21.67	15	1.1%	-4.59 [-17.15 , 7.97]
Sherman 2005 (1)	12.92	15	35	26.67	21.67	15	1.1%	-13.75 [-25.79 , -1.71]
Sherman 2011 (3)	18.74	16.28	91	28.52	20.61	22	1.4%	-9.78 [-19.02 , -0.54]
Sherman 2011 (3)	19.26	17.56	92	28.52	20.61	23	1.4%	-9.26 [-18.42 , -0.10]
Tilbrook 2011 (4)	23.75	20.31	156	32.17	20.31	136	1.8%	-8.42 [-13.09 , -3.75]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	1.8%	-10.33 [-15.25 , -5.41]
Valenza 2017 (4)	16.96	13.92	27	29.92	13.92	27	1.5%	-12.96 [-20.39 , -5.53]
Vincent 2014a (1)	34.17	22.92	17	26.25	17.5	7	0.8%	7.92 [-9.01 , 24.85]
Vincent 2014a (1)	20.83	17.08	18	26.25	17.5	7	0.9%	-5.42 [-20.60 , 9.76]
Subtotal (95% CI)			**1619**			**979**	**35.9%**	**-8.03 [-10.75 , -5.30]**
Heterogeneity: Tau² = 30.27; Chi² = 112.67, df = 25 (P < 0.00001); I² = 78%								
Test for overall effect: Z = 5.77 (P < 0.00001)								
5.2.2 Exercise vs. Manual therapy								
Bellido-Fernandez 2018 (1)	8	3.7	9	8	6.67	5	1.7%	0.00 [-6.33 , 6.33]
Bellido-Fernandez 2018 (1)	12	5.93	9	8	6.67	4	1.5%	4.00 [-3.60 , 11.60]
Bronfort 1996 (3)	20.9	17	56	20.8	17.3	21	1.4%	0.10 [-8.54 , 8.74]
Bronfort 1996 (3)	15.1	17.4	43	20.8	17.3	21	1.4%	-5.70 [-14.74 , 3.34]
Bronfort 2011 (1)	18.7	19.57	91	21.3	21.74	49	1.6%	-2.60 [-9.90 , 4.70]
Bronfort 2011 (1)	16.96	20	92	21.3	21.74	50	1.6%	-4.34 [-11.62 , 2.94]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.8%	0.83 [-3.79 , 5.45]
Cecchi 2010 (1)	22.5	19.58	68	9.17	13.75	35	1.6%	13.33 [6.82 , 19.84]
Cecchi 2010 (1)	22.08	19.58	68	9.17	13.75	34	1.6%	12.91 [6.35 , 19.47]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	1.7%	-2.37 [-7.87 , 3.13]
Ferreira 2007 (1)	40.42	26.25	74	32.92	25	39	1.3%	7.50 [-2.37 , 17.37]
Ferreira 2007 (1)	32.92	23.75	73	32.92	25	38	1.3%	0.00 [-9.64 , 9.64]
Fransoo 2006 (3)	11.25	13.92	20	13.75	13.92	10	1.2%	-2.50 [-13.07 , 8.07]
Fransoo 2006 (3)	4.58	13.92	19	13.75	13.92	9	1.2%	-9.17 [-20.21 , 1.87]
Goldby 2006 (2)	31	17.07	78	28.1	17.34	37	1.6%	2.90 [-3.85 , 9.65]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	1.5%	-7.00 [-14.73 , 0.73]
Ritvanen 2007 (4)	17	13.27	28	12	13.27	33	1.6%	5.00 [-1.68 , 11.68]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.3%	9.20 [-0.98 , 19.38]
Subtotal (95% CI)			**977**			**660**	**27.0%**	**1.46 [-1.47 , 4.38]**
Heterogeneity: Tau² = 24.43; Chi² = 46.81, df = 17 (P = 0.0001); I² = 64%								
Test for overall effect: Z = 0.98 (P = 0.33)								
5.2.3 Exercise vs. Electrotherapy								
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Djavid 2007 (1)	48.2	10.4	18	41.6	8.8	8	1.5%	6.60 [-1.16 , 14.36]
Djavid 2007 (1)	33.6	7.4	19	41.6	8.8	8	1.6%	-8.00 [-14.95 , -1.05]
Gur 2003 (1)	26.25	14.58	25	27.5	12.08	13	1.4%	-1.25 [-9.96 , 7.46]
Gur 2003 (1)	22.91	13.33	25	27.5	12.08	12	1.4%	-4.59 [-13.19 , 4.01]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.5%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**147**			**101**	**8.6%**	**-3.81 [-10.07 , 2.44]**
Heterogeneity: Tau² = 42.51; Chi² = 16.97, df = 5 (P = 0.005); I² = 71%								
Test for overall effect: Z = 1.19 (P = 0.23)								

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.2. (Continued)

Heterogeneity: Tau² = 42.51; Chi² = 16.97, df = 5 (P = 0.005); I² = 71%								
Test for overall effect: Z = 1.19 (P = 0.23)								
5.2.4 Exercise vs. Psychological therapy								
Machado 2007	50	22.92	17	57.5	23.96	11	0.7%	-7.50 [-25.37 , 10.37]
Subtotal (95% CI)			**17**			**11**	**0.7%**	**-7.50 [-25.37 , 10.37]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.82 (P = 0.41)								
5.2.5 Exercise vs. Mixed physical therapy								
Bid 2017	7.29	8.54	64	24.12	12	64	1.9%	-16.83 [-20.44 , -13.22]
Chown 2008 (2)	20.8	8.4	35	20.3	8	35	1.9%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016 (4)	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.0%	-7.66 [-10.55 , -4.77]
Little 2014 (4)	23.16	13.92	14	25.37	13.92	15	1.3%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.7%	-17.84 [-36.69 , 1.01]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	1.7%	-14.80 [-20.62 , -8.98]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.6%	-9.97 [-16.60 , -3.34]
Okafor 2012 (5)	25	20.8	15	10.9	4.46	15	1.2%	14.10 [3.33 , 24.87]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.4%	-10.80 [-19.97 , -1.63]
Torstensen 1998 (3)	46.2	13.1	71	46.9	13.1	34	1.8%	-0.70 [-6.05 , 4.65]
Torstensen 1998 (3)	52.7	16.6	70	46.9	13.1	33	1.7%	5.80 [-0.12 , 11.72]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	1.8%	4.29 [-0.57 , 9.15]
Subtotal (95% CI)			**548**			**526**	**20.5%**	**-4.28 [-9.00 , 0.43]**
Heterogeneity: Tau² = 61.42; Chi² = 110.42, df = 12 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 1.78 (P = 0.08)								
5.2.6 Exercise vs. Back school								
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.7%	-4.90 [-11.09 , 1.29]
Maul 2005	25	8.33	74	25	12.5	74	1.9%	0.00 [-3.42 , 3.42]
Subtotal (95% CI)			**103**			**106**	**3.6%**	**-1.74 [-6.34 , 2.85]**
Heterogeneity: Tau² = 5.50; Chi² = 1.85, df = 1 (P = 0.17); I² = 46%								
Test for overall effect: Z = 0.74 (P = 0.46)								
5.2.7 Exercise vs. Relaxation								
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Subtotal (95% CI)			**36**			**32**	**1.4%**	**4.59 [-4.72 , 13.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.97 (P = 0.33)								
5.2.8 Exercise vs. Anti-inflammatories								
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.1%	-15.83 [-27.67 , -3.99]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.2%	-11.88 [-23.25 , -0.51]
Subtotal (95% CI)			**87**			**55**	**2.3%**	**-13.78 [-21.98 , -5.57]**
Heterogeneity: Tau² = 0.00; Chi² = 0.22, df = 1 (P = 0.64); I² = 0%								
Test for overall effect: Z = 3.29 (P = 0.0010)								
Total (95% CI)			**3534**			**2470**	**100.0%**	**-4.10 [-5.97 , -2.24]**
Heterogeneity: Tau² = 44.42; Chi² = 382.34, df = 68 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 4.31 (P < 0.0001)								
Test for subgroup differences: Chi² = 31.36, df = 7 (P < 0.0001), I² = 77.7%								



Footnotes

(1) Study has multiple exercise groups available: comparison group data divided.
(2) Study has other comparison groups available for this meta-analysis.
(3) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(4) Standard deviation(s) required imputation due to missing or implausible value.
(5) Large baseline differences.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Comparison 6. SENSITIVITY ANALYSES (No assumptions about data)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 Pain (/100): Earliest follow-up, no imputation of SD	71	6345	Mean Difference (IV, Random, 95% CI)	-12.08 [-15.49, -8.66]
6.1.1 Exercise vs. No treatment or usual care	24	1717	Mean Difference (IV, Random, 95% CI)	-16.65 [-20.81, -12.50]
6.1.2 Exercise vs. Other conservative treatment	47	4628	Mean Difference (IV, Random, 95% CI)	-9.85 [-14.40, -5.31]
6.2 Function (/100): Earliest follow-up, no imputation of SD	68		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.2.1 Exercise vs. No treatment or usual care	28	2076	Mean Difference (IV, Random, 95% CI)	-7.38 [-9.33, -5.44]
6.2.2 Exercise vs. Other conservative treatment	41	4743	Mean Difference (IV, Random, 95% CI)	-4.12 [-6.28, -1.97]
6.3 Function (/100): Earliest follow-up, RMDQ measurement only	53		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.3.1 Exercise vs. No treatment or usual care	22	2146	Mean Difference (IV, Random, 95% CI)	-7.83 [-10.21, -5.45]
6.3.2 Exercise vs. Other conservative treatment	33	4082	Mean Difference (IV, Random, 95% CI)	-5.27 [-8.05, -2.49]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

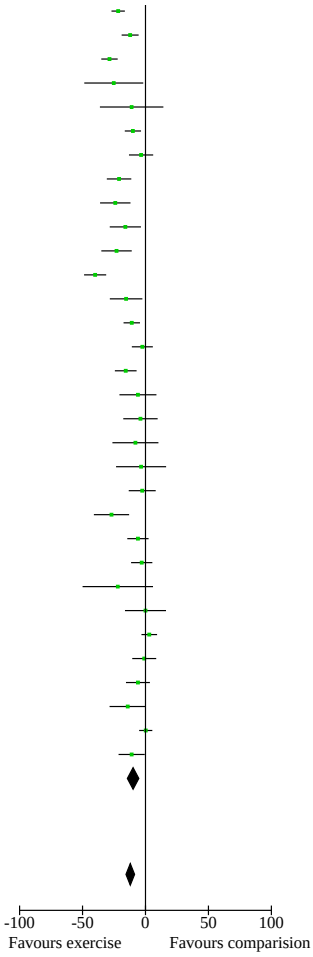
Analysis 6.1. Comparison 6: SENSITIVITY ANALYSES (No assumptions about data), Outcome 1: Pain (/100): Earliest follow-up, no imputation of SD

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
6.1.1 Exercise vs. No treatment or usual care								
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	1.1%	-9.50 [-21.15 , 2.15]
Chen 2014	21.7	14.2	64	34.8	17.7	63	1.3%	-13.10 [-18.69 , -7.51]
Cortell-Tormo 2018	15	15	11	44	14	8	1.1%	-29.00 [-42.14 , -15.86]
Costa 2009b	46	28	77	56	26	77	1.2%	-10.00 [-18.53 , -1.47]
Cruz-Diaz 2017	17	14.1	34	49.6	13.1	15	1.2%	-32.60 [-40.75 , -24.45]
Cruz-Diaz 2017	21	13.6	34	49.6	13.1	15	1.2%	-28.60 [-36.65 , -20.55]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	1.2%	-14.90 [-23.47 , -6.33]
Garcia 2017	39.5	27.3	73	47	29.7	73	1.2%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	1.3%	-2.00 [-7.75 , 3.75]
Highland 2018	27.5	24.3	33	33.5	17.9	29	1.2%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	1.2%	-5.00 [-13.44 , 3.44]
Keane 2017	26	23.7	10	48.9	33.3	5	0.6%	-22.90 [-55.58 , 9.78]
Keane 2017	26.5	18.6	10	48.9	33.3	4	0.6%	-22.40 [-57.01 , 12.21]
Lee 2016a	22	11.3	15	35.8	17.2	3	0.9%	-13.80 [-34.09 , 6.49]
Lee 2016a	33.1	20	15	35.8	17.2	3	0.8%	-2.70 [-24.64 , 19.24]
Masharawi 2013	16.8	8.2	20	38.8	15.4	20	1.2%	-22.00 [-29.65 , -14.35]
Mazloum 2018	48	11	15	69	16	8	1.1%	-21.00 [-33.41 , -8.59]
Mazloum 2018	30	9	16	69	16	8	1.1%	-39.00 [-50.93 , -27.07]
Morone 2011	45	23	41	76	21	29	1.2%	-31.00 [-41.39 , -20.61]
Morone 2012	42.4	21.85	25	72.08	17.19	25	1.1%	-29.68 [-40.58 , -18.78]
Moseley 2002	19	15	29	31	15	28	1.2%	-12.00 [-19.79 , -4.21]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	1.1%	-16.50 [-28.91 , -4.09]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	1.2%	-39.50 [-49.57 , -29.43]
Puppin 2011	24	27	30	41	27	25	1.1%	-17.00 [-31.33 , -2.67]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	1.2%	-15.60 [-24.89 , -6.31]
Schinhan 2016	13.3	10.5	15	21	17.3	15	1.2%	-7.70 [-17.94 , 2.54]
Smeets 2008	37.76	24.33	55	53.35	22.6	17	1.1%	-15.59 [-28.11 , -3.07]
Smeets 2008	44.63	28.86	55	53.35	22.6	17	1.1%	-8.72 [-21.90 , 4.46]
Smeets 2008	42.31	25.56	52	53.35	22.6	17	1.1%	-11.04 [-23.83 , 1.75]
Yelland 2004	42	26.5	55	37	45.4	55	1.1%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**977**			**740**	**32.9%**	**-16.65 [-20.81 , -12.50]**
Heterogeneity: Tau² = 97.39; Chi² = 137.34, df = 29 (P < 0.00001); I² = 79%								
Test for overall effect: Z = 7.85 (P < 0.00001)								
6.1.2 Exercise vs. Other conservative treatment								
Alayat 2014	26.4	12.5	28	56.5	10.4	10	1.2%	-30.10 [-38.04 , -22.16]
Alayat 2014	37.1	13	24	56.5	10.4	10	1.2%	-19.40 [-27.68 , -11.12]
Alexandre 2001	6	27.2	14	37	35.61	14	0.8%	-31.00 [-54.47 , -7.53]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.2%	-23.10 [-31.54 , -14.66]
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.3%	-24.98 [-29.09 , -20.87]
Bronfort 2011	32	22	93	29	19	50	1.2%	3.00 [-3.91 , 9.91]
Bronfort 2011	26	21	91	29	19	49	1.2%	-3.00 [-9.85 , 3.85]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.3%	7.50 [2.05 , 12.95]
Cecchi 2010	25	20	68	8.33	11.67	34	1.2%	16.67 [10.51 , 22.83]
Cecchi 2010	23.33	20	68	8.33	11.67	35	1.2%	15.00 [8.87 , 21.13]
Chan 2017	30	23	49	43	26	43	1.2%	-13.00 [-23.09 , -2.91]
Chown 2008	73.4	17.2	22	73.2	11.5	33	1.2%	0.20 [-7.99 , 8.39]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.2%	0.90 [-7.83 , 9.63]
Djavid 2007	43	16	19	44	20	8	1.0%	-1.00 [-16.62 , 14.62]
Djavid 2007	24	14	18	44	20	8	1.0%	-20.00 [-35.29 , -4.71]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.2%	-3.32 [-11.72 , 5.08]
Ferreira 2007	40	25	74	41	26	39	1.2%	-1.00 [-10.95 , 8.95]
Ferreira 2007	48	24	73	41	26	38	1.2%	7.00 [-2.93 , 16.93]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.2%	-10.00 [-17.40 , -2.60]
Goldby 2006	28.81	28.14	78	34.4	36.43	37	1.1%	-5.59 [-18.89 , 7.71]
Gur 2003	29	13	25	19	14	13	1.2%	10.00 [0.84 , 19.16]
Gur 2003	18	12	25	19	14	12	1.2%	-1.00 [-10.21 , 8.21]
Hartvigsen 2010	37.97	19.38	32	42	21.3	21	1.1%	-4.03 [-15.35 , 7.29]
Hartvigsen 2010	45.6	19.82	36	42	21.3	20	1.1%	3.60 [-7.76 , 14.96]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.3%	-41.79 [-44.48 , -39.10]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	1.0%	-8.10 [-22.94 , 6.74]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.3%	-25.90 [-31.06 , -20.74]
Kim 2017	24	5.4	15	45.7	9	15	1.3%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.2%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.2%	-28.60 [-35.11 , -22.09]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 6.1. (Continued)

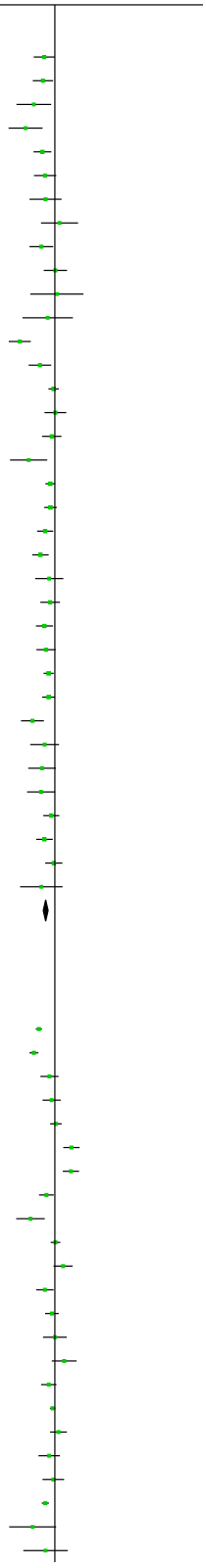
Kim 2017	24	5.4	15	45.7	9	15	1.3%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.2%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.2%	-28.60 [-35.11 , -22.09]
Longo 2016	36.2	22	7	61.4	22.7	7	0.8%	-25.20 [-48.62 , -1.78]
Machado 2007	47	40	17	58	28	11	0.8%	-11.00 [-36.21 , 14.21]
Maul 2005	40	20	74	50	20	74	1.2%	-10.00 [-16.44 , -3.56]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.2%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	1.2%	-21.00 [-30.72 , -11.28]
Miyamoto 2018	32	24	67	56	26	23	1.1%	-24.00 [-36.08 , -11.92]
Miyamoto 2018	40	27	69	56	26	23	1.1%	-16.00 [-28.39 , -3.61]
Miyamoto 2018	33	23	73	56	26	22	1.1%	-23.00 [-35.08 , -10.92]
Murtezani 2011	20	17	50	60	27	51	1.2%	-40.00 [-48.78 , -31.22]
Natour 2015	40.4	24.2	30	55.8	26.8	30	1.1%	-15.40 [-28.32 , -2.48]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.2%	-10.90 [-17.43 , -4.37]
Noori 2011	41.5	15.3	17	44	7.5	13	1.2%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.2%	-15.70 [-24.32 , -7.08]
Rantonen 2012	29	27	43	35	28	20	1.0%	-6.00 [-20.69 , 8.69]
Rantonen 2012	31	20	43	35	28	20	1.1%	-4.00 [-17.65 , 9.65]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	0.9%	-8.00 [-26.27 , 10.27]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.9%	-3.50 [-23.38 , 16.38]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	1.1%	-2.60 [-13.31 , 8.11]
Saper 2009	44	21	15	71	18	15	1.1%	-27.00 [-41.00 , -13.00]
Saper 2017	50	21	127	56	22	32	1.2%	-6.00 [-14.45 , 2.45]
Saper 2017	53	21	129	56	22	32	1.2%	-3.00 [-11.44 , 5.44]
Schenkman 2009	19	21	17	41	40	9	0.7%	-22.00 [-49.97 , 5.97]
Segal-Snir 2016	70	30	20	70	19	15	1.0%	0.00 [-16.29 , 16.29]
Sorensen 2010	49	22	91	46	21	94	1.2%	3.00 [-3.20 , 9.20]
Soukup 1999	23	16	34	24	17	18	1.2%	-1.00 [-10.52 , 8.52]
Soukup 1999	18	15.7	38	24	17	17	1.2%	-6.00 [-15.50 , 3.50]
Sousa 2009	33.5	24.8	27	47.6	28	25	1.1%	-14.10 [-28.52 , 0.32]
Ulger 2017	21	16	56	20.8	12	57	1.3%	0.20 [-5.02 , 5.42]
Williams 2005	10	11	20	21	23	24	1.2%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**2703**			**1925**	**67.1%**	**-9.85 [-14.40 , -5.31]**
Heterogeneity: Tau² = 283.26; Chi² = 1038.23, df = 58 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 4.25 (P < 0.0001)								
Total (95% CI)			**3680**			**2665**	**100.0%**	**-12.08 [-15.49 , -8.66]**
Heterogeneity: Tau² = 234.44; Chi² = 1179.18, df = 88 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 6.93 (P < 0.00001)								
Test for subgroup differences: Chi² = 4.69, df = 1 (P = 0.03), I² = 78.7%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 6.2. Comparison 6: SENSITIVITY ANALYSES (No assumptions about data), Outcome 2: Function (/100): Earliest follow-up, no imputation of SD

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
6.2.1 Exercise vs. No treatment or usual care								
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.8%	-8.50 [-16.99 , -0.01]
Costa 2009b	40	27.08	77	49.58	24.58	77	3.0%	-9.58 [-17.75 , -1.41]
Cruz-Diaz 2017	26.46	22.08	34	43.37	23.33	15	1.5%	-16.91 [-30.86 , -2.96]
Cruz-Diaz 2017	19.83	20.42	34	43.37	23.33	15	1.5%	-23.54 [-37.20 , -9.88]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	3.4%	-10.00 [-17.14 , -2.86]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	2.7%	-7.83 [-16.79 , 1.13]
Harts 2008	14.17	16.67	20	21.67	16.25	9	1.7%	-7.50 [-20.39 , 5.39]
Harts 2008	25.42	25	20	21.67	16.25	10	1.3%	3.75 [-11.13 , 18.63]
Highland 2018	18.46	19.25	33	29.33	18.96	29	2.5%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	2.6%	0.43 [-8.93 , 9.79]
Keane 2017	33.2	18.84	10	31.56	18.24	4	0.7%	1.64 [-19.71 , 22.99]
Keane 2017	25.8	20.1	10	31.56	18.24	5	0.8%	-5.76 [-26.03 , 14.51]
Kofotolis 2016	17.04	8.08	36	45.25	16.17	14	2.7%	-28.21 [-37.08 , -19.34]
Kofotolis 2016	33.25	10.17	37	45.25	16.17	14	2.6%	-12.00 [-21.08 , -2.92]
Kuukkanen 2000	10	7.6	27	11	7.8	26	4.8%	-1.00 [-5.15 , 3.15]
Lee 2016a	7.5	8.75	15	7.08	6.67	3	2.8%	0.42 [-8.33 , 9.17]
Lee 2016a	4.58	3.75	15	7.08	6.67	3	3.1%	-2.50 [-10.28 , 5.28]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	1.3%	-21.08 [-36.02 , -6.14]
Mazloum 2018	22.9	3.6	15	26.6	5	8	5.0%	-3.70 [-7.61 , 0.21]
Mazloum 2018	23.1	7.5	16	26.6	5	8	4.4%	-3.50 [-8.55 , 1.55]
Morone 2011	18	12.9	41	25.8	14.1	29	3.7%	-7.80 [-14.28 , -1.32]
Morone 2012	13.68	9.39	25	25.28	14.03	25	3.6%	-11.60 [-18.22 , -4.98]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	2.0%	-4.50 [-15.82 , 6.82]
Patti 2016	13	8	19	16.8	15.6	19	3.1%	-3.80 [-11.68 , 4.08]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	3.5%	-8.34 [-15.23 , -1.45]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	3.1%	-7.10 [-14.82 , 0.62]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	4.9%	-5.00 [-9.08 , -0.92]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	4.3%	-5.06 [-10.19 , 0.07]
Shaughnessy 2004	26	14	20	44	16	21	2.6%	-18.00 [-27.19 , -8.81]
Smeets 2008	49.6	24.56	52	57.83	19.92	17	1.9%	-8.23 [-19.82 , 3.36]
Smeets 2008	47.5	21.88	55	57.83	19.92	17	2.1%	-10.33 [-21.43 , 0.77]
Smeets 2008	46.74	23.01	55	57.83	19.92	17	2.0%	-11.09 [-22.34 , 0.16]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	3.7%	-2.87 [-9.31 , 3.57]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	3.6%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	3.5%	-0.83 [-7.72 , 6.06]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	1.1%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**1203**			**873**	**100.0%**	**-7.38 [-9.33 , -5.44]**
Heterogeneity: Tau² = 15.92; Chi² = 72.41, df = 35 (P = 0.0002); I² = 52%								
Test for overall effect: Z = 7.43 (P < 0.00001)								
6.2.2 Exercise vs. Other conservative treatment								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.6%	-12.84 [-15.31 , -10.37]
Bid 2017	7.29	8.54	64	24.12	12	64	2.5%	-16.83 [-20.44 , -13.22]
Bronfort 2011	16.96	20	92	21.3	21.74	50	2.1%	-4.34 [-11.62 , 2.94]
Bronfort 2011	18.7	19.57	91	21.3	21.74	49	2.1%	-2.60 [-9.90 , 4.70]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	2.4%	0.83 [-3.79 , 5.45]
Cecchi 2010	22.5	19.58	68	9.17	13.75	34	2.2%	13.33 [6.77 , 19.89]
Cecchi 2010	22.08	19.58	68	9.17	13.75	35	2.2%	12.91 [6.40 , 19.42]
Chan 2017	17.6	12.6	47	24.4	15.8	43	2.2%	-6.80 [-12.74 , -0.86]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.5%	-19.58 [-31.05 , -8.11]
Chown 2008	20.8	8.4	35	20.3	8	35	2.5%	0.50 [-3.34 , 4.34]
Djavid 2007	48.2	10.4	19	41.6	8.8	8	2.0%	6.60 [-1.08 , 14.28]
Djavid 2007	33.6	7.4	18	41.6	8.8	8	2.1%	-8.00 [-14.99 , -1.01]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	2.3%	-2.37 [-7.87 , 3.13]
Ferreira 2007	32.92	23.75	74	32.92	25	39	1.8%	0.00 [-9.53 , 9.53]
Ferreira 2007	40.42	26.25	73	32.92	25	38	1.7%	7.50 [-2.47 , 17.47]
Frost 1995	16.3	10.3	29	21.2	14.2	32	2.2%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.6%	-1.80 [-3.96 , 0.36]
Goldby 2006	31	17.07	78	28.1	17.34	37	2.1%	2.90 [-3.85 , 9.65]
Gur 2003	22.91	13.33	25	27.5	12.08	12	1.9%	-4.59 [-13.19 , 4.01]
Gur 2003	26.25	14.58	25	27.5	12.08	13	1.9%	-1.25 [-9.96 , 7.46]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.6%	-7.66 [-10.55 , -4.77]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.9%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	0.9%	-7.50 [-25.37 , 10.37]



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

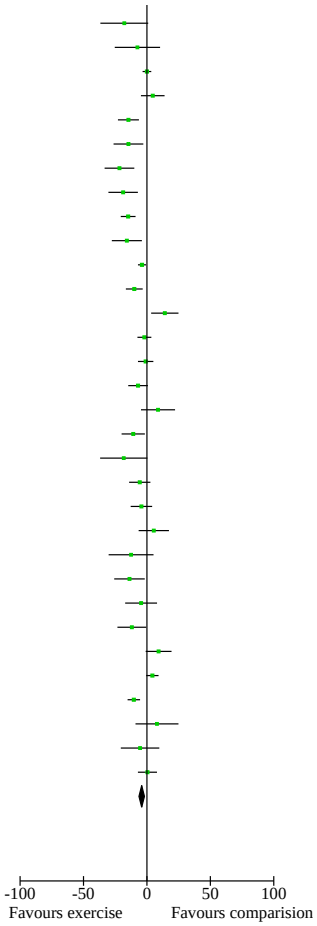
Analysis 6.2. (Continued)

Longo 2016	16.08	14.92	7	33.92	20.62	7	0.9%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	0.9%	-7.50 [-25.37 , 10.37]
Maul 2005	25	8.33	74	25	12.5	74	2.5%	0.00 [-3.42 , 3.42]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.8%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.9%	-14.58 [-22.85 , -6.31]
Miyamoto 2018	32.5	21.67	73	47.08	25.42	22	1.5%	-14.58 [-26.31 , -2.85]
Miyamoto 2018	25.42	22.92	69	47.08	25.42	23	1.5%	-21.66 [-33.37 , -9.95]
Miyamoto 2018	28.33	21.67	67	47.08	25.42	23	1.5%	-18.75 [-30.36 , -7.14]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	2.3%	-14.80 [-20.62 , -8.98]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.5%	-15.83 [-27.67 , -3.99]
Niemisto 2003	14.7	11.6	96	18.6	11	100	2.6%	-3.90 [-7.07 , -0.73]
Noori 2011	21.41	10.48	17	31.38	8.06	13	2.2%	-9.97 [-16.60 , -3.34]
Okafor 2012	25	20.8	15	10.9	4.46	15	1.6%	14.10 [3.33 , 24.87]
Rantonen 2012	14	11	43	16	10	20	2.3%	-2.00 [-7.48 , 3.48]
Rantonen 2012	15	14	43	16	10	20	2.2%	-1.00 [-7.06 , 5.06]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	2.0%	-7.00 [-14.73 , 0.73]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.3%	8.75 [-4.68 , 22.18]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.8%	-10.80 [-19.97 , -1.63]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.9%	-18.27 [-36.90 , 0.36]
Saper 2017	47.83	21.3	129	53.48	21.74	32	1.9%	-5.65 [-14.03 , 2.73]
Saper 2017	49.13	22.17	127	53.48	21.74	32	1.9%	-4.35 [-12.81 , 4.11]
Schenkman 2009	15.81	10.4	17	10.4	16.64	9	1.5%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	1.0%	-12.50 [-30.19 , 5.19]
Sherman 2005	12.92	15	35	26.67	21.67	15	1.5%	-13.75 [-25.79 , -1.71]
Sherman 2005	22.08	18.75	36	26.67	21.67	15	1.4%	-4.59 [-17.15 , 7.97]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.5%	-11.88 [-23.25 , -0.51]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.7%	9.20 [-0.98 , 19.38]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	2.4%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	2.4%	-10.33 [-15.25 , -5.41]
Vincent 2014a	34.17	22.92	17	26.25	17.5	7	1.0%	7.92 [-9.01 , 24.85]
Vincent 2014a	20.83	17.08	18	26.25	17.5	7	1.2%	-5.42 [-20.60 , 9.76]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	2.0%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**2744**			**1999**	**100.0%**	**-4.12 [-6.28 , -1.97]**

Heterogeneity: Tau² = 44.85; Chi² = 305.97, df = 52 (P < 0.00001); I² = 83%

Test for overall effect: Z = 3.75 (P = 0.0002)

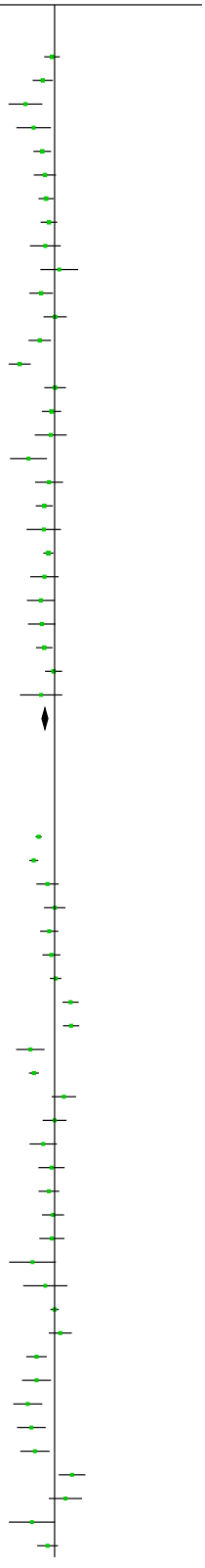
Test for subgroup differences: Chi² = 4.83, df = 1 (P = 0.03), I² = 79.3%



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 6.3. Comparison 6: SENSITIVITY ANALYSES (No assumptions about data), Outcome 3: Function (/100): Earliest follow-up, RMDQ measurement only

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
6.3.1 Exercise vs. No treatment or usual care								
Albaladejo 2010	31.25	24.69	100	33.33	21.6	109	5.1%	-2.08 [-8.39 , 4.23]
Costa 2009b	40	27.08	77	49.58	24.58	77	4.1%	-9.58 [-17.75 , -1.41]
Cruz-Diaz 2017	19.83	20.42	34	43.37	23.33	15	2.2%	-23.54 [-37.20 , -9.88]
Cruz-Diaz 2017	26.46	22.08	34	43.37	23.33	15	2.1%	-16.91 [-30.86 , -2.96]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	4.7%	-10.00 [-17.14 , -2.86]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	3.8%	-7.83 [-16.79 , 1.13]
Groessl 2017	30.62	19.06	75	37.54	19.06	75	5.3%	-6.92 [-13.02 , -0.82]
Hall 2011	29.21	21.49	80	33.75	21.87	80	4.9%	-4.54 [-11.26 , 2.18]
Harts 2008	14.17	16.67	20	21.67	16.25	10	2.5%	-7.50 [-19.94 , 4.94]
Harts 2008	25.42	25	20	21.67	16.25	9	1.9%	3.75 [-11.51 , 19.01]
Highland 2018	18.46	19.25	33	29.33	18.96	29	3.5%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	3.6%	0.43 [-8.93 , 9.79]
Kofotolis 2016	33.25	10.17	37	45.25	16.17	14	3.7%	-12.00 [-21.08 , -2.92]
Kofotolis 2016	17.04	8.08	36	45.25	16.17	14	3.8%	-28.21 [-37.08 , -19.34]
Lee 2016a	7.5	8.75	15	7.08	6.67	3	3.9%	0.42 [-8.33 , 9.17]
Lee 2016a	4.58	3.75	15	7.08	6.67	3	4.3%	-2.50 [-10.28 , 5.28]
Little 2014	27.95	13.92	15	31.08	19.87	13	2.4%	-3.13 [-16.03 , 9.77]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	1.9%	-21.08 [-36.02 , -6.14]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	2.9%	-4.50 [-15.82 , 6.82]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	4.8%	-8.34 [-15.23 , -1.45]
Quinn 2011	22.5	19.06	15	31.25	19.06	14	2.2%	-8.75 [-22.63 , 5.13]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	6.5%	-5.00 [-9.08 , -0.92]
Smeets 2008	49.6	24.56	55	57.83	19.92	17	2.8%	-8.23 [-19.71 , 3.25]
Smeets 2008	46.74	23.01	55	57.83	19.92	17	2.9%	-11.09 [-22.34 , 0.16]
Smeets 2008	47.5	21.88	52	57.83	19.92	17	2.9%	-10.33 [-21.51 , 0.85]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	4.9%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	4.8%	-0.83 [-7.72 , 6.06]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	1.6%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**1191**			**955**	**100.0%**	**-7.83 [-10.21 , -5.45]**
Heterogeneity: Tau² = 18.24; Chi² = 53.02, df = 27 (P = 0.002); I² = 49%								
Test for overall effect: Z = 6.45 (P < 0.00001)								
6.3.2 Exercise vs. Other conservative treatment								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.9%	-12.84 [-15.31 , -10.37]
Bid 2017	7.29	8.54	64	24.12	12	64	2.8%	-16.83 [-20.44 , -13.22]
Bronfort 1996	15.1	17.4	43	20.8	17.3	21	2.3%	-5.70 [-14.74 , 3.34]
Bronfort 1996	20.9	17	56	20.8	17.3	21	2.3%	0.10 [-8.54 , 8.74]
Bronfort 2011	16.96	20	91	21.3	21.74	49	2.5%	-4.34 [-11.68 , 3.00]
Bronfort 2011	18.7	19.57	92	21.3	21.74	50	2.5%	-2.60 [-9.83 , 4.63]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	2.7%	0.83 [-3.79 , 5.45]
Cecchi 2010	22.08	19.58	68	9.17	13.75	35	2.5%	12.91 [6.40 , 19.42]
Cecchi 2010	22.5	19.58	68	9.17	13.75	34	2.5%	13.33 [6.77 , 19.89]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	2.0%	-19.58 [-31.05 , -8.11]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	2.8%	-16.67 [-20.63 , -12.71]
Ferreira 2007	40.42	26.25	74	32.92	25	39	2.2%	7.50 [-2.37 , 17.37]
Ferreira 2007	32.92	23.75	73	32.92	25	38	2.2%	0.00 [-9.64 , 9.64]
Fransoo 2006	4.58	13.92	19	13.75	13.92	9	2.0%	-9.17 [-20.21 , 1.87]
Fransoo 2006	11.25	13.92	20	13.75	13.92	10	2.1%	-2.50 [-13.07 , 8.07]
Gur 2003	22.91	13.33	25	27.5	12.08	13	2.3%	-4.59 [-12.98 , 3.80]
Gur 2003	26.25	14.58	25	27.5	12.08	12	2.3%	-1.25 [-10.16 , 7.66]
Little 2014	23.16	13.92	14	25.37	13.92	15	2.1%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	1.3%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	1.3%	-7.50 [-25.37 , 10.37]
Maul 2005	25	8.33	74	25	12.5	74	2.8%	0.00 [-3.42 , 3.42]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	2.2%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	2.3%	-14.58 [-22.85 , -6.31]
Miyamoto 2018	32.5	21.67	73	47.08	25.42	22	1.9%	-14.58 [-26.31 , -2.85]
Miyamoto 2018	25.42	22.92	67	47.08	25.42	23	1.9%	-21.66 [-33.41 , -9.91]
Miyamoto 2018	28.33	21.67	69	47.08	25.42	23	2.0%	-18.75 [-30.33 , -7.17]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.9%	-15.83 [-27.67 , -3.99]
Okafor 2012	25	20.8	15	10.9	4.46	15	2.1%	14.10 [3.33 , 24.87]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.8%	8.75 [-4.68 , 22.18]
Saper 2009	35.67	30.01	15	53.94	21.32	15	1.3%	-18.27 [-36.90 , 0.36]
Saper 2017	47.83	21.3	127	53.48	21.74	32	2.3%	-5.65 [-14.04 , 2.74]



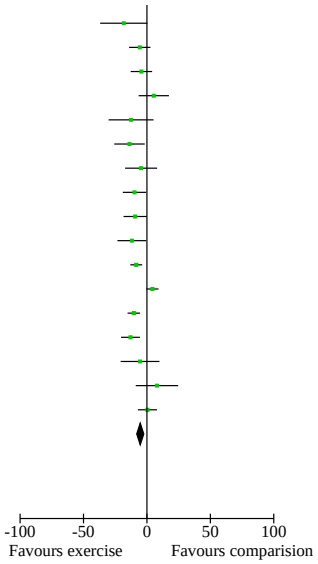
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 6.3. (Continued)

Saper 2009	35.67	30.01	15	53.94	21.32	15	1.3%	-18.27 [-36.90 , 0.36]
Saper 2017	47.83	21.3	127	53.48	21.74	32	2.3%	-5.65 [-14.04 , 2.74]
Saper 2017	49.13	22.17	129	53.48	21.74	32	2.3%	-4.35 [-12.80 , 4.10]
Schenkman 2009	15.81	10.4	17	10.4	16.64	9	1.9%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	1.3%	-12.50 [-30.19 , 5.19]
Sherman 2005	12.92	15	36	26.67	21.67	15	1.9%	-13.75 [-25.76 , -1.74]
Sherman 2005	22.08	18.75	35	26.67	21.67	15	1.8%	-4.59 [-17.19 , 8.01]
Sherman 2011	18.74	16.28	91	28.52	20.61	22	2.2%	-9.78 [-19.02 , -0.54]
Sherman 2011	19.26	17.56	92	28.52	20.61	23	2.2%	-9.26 [-18.42 , -0.10]
Sousa 2009	22.12	19.96	57	34	25.83	25	2.0%	-11.88 [-23.25 , -0.51]
Tilbrook 2011	23.75	20.31	156	32.17	20.31	136	2.7%	-8.42 [-13.09 , -3.75]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	2.7%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	2.7%	-10.33 [-15.25 , -5.41]
Valenza 2017	16.96	13.92	27	29.92	13.92	27	2.4%	-12.96 [-20.39 , -5.53]
Vincent 2014a	20.83	17.08	17	26.25	17.5	7	1.6%	-5.42 [-20.72 , 9.88]
Vincent 2014a	34.17	22.92	18	26.25	17.5	7	1.4%	7.92 [-8.82 , 24.66]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	2.4%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**2478**			**1604**	**100.0%**	**-5.27 [-8.05 , -2.49]**

Heterogeneity: Tau² = 67.86; Chi² = 280.64, df = 45 (P < 0.00001); I² = 84%
Test for overall effect: Z = 3.72 (P = 0.0002)

Test for subgroup differences: Chi² = 1.88, df = 1 (P = 0.17), I² = 46.7%



Comparison 7. SENSITIVITY ANALYSES (Low risk of bias)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 Pain (/100): Earliest follow-up	81	7680	Mean Difference (IV, Random, 95% CI)	-11.18 [-14.11, -8.26]
7.1.1 Exercise vs. No treatment or usual care	28	2220	Mean Difference (IV, Random, 95% CI)	-15.34 [-18.91, -11.76]
7.1.2 Exercise vs. Other conservative treatment	53	5460	Mean Difference (IV, Random, 95% CI)	-9.06 [-12.99, -5.13]
7.2 Function (/100): Earliest follow-up	72		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.2.1 Exercise vs. No treatment or usual care	29	2447	Mean Difference (IV, Random, 95% CI)	-7.27 [-9.04, -5.49]
7.2.2 Exercise vs. Other conservative treatment	45	5433	Mean Difference (IV, Random, 95% CI)	-4.39 [-6.45, -2.34]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

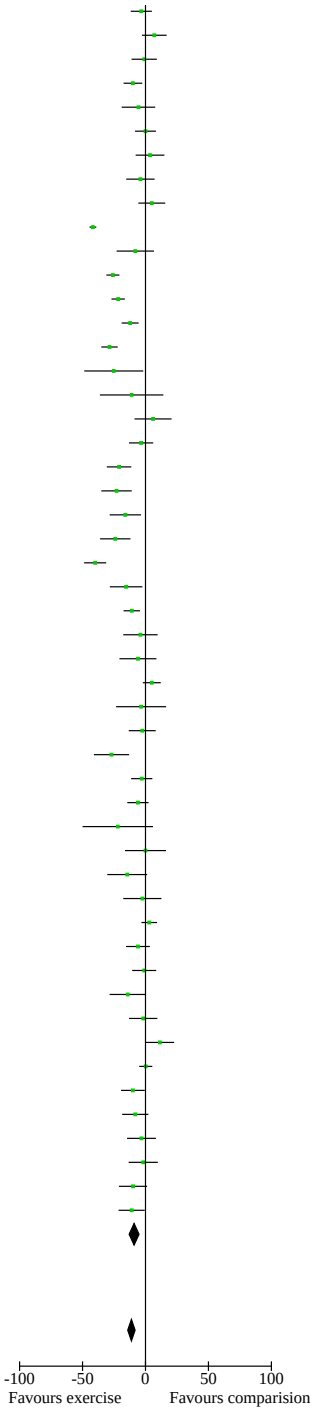
Analysis 7.1. Comparison 7: SENSITIVITY ANALYSES (Low risk of bias), Outcome 1: Pain (/100): Earliest follow-up

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
7.1.1 Exercise vs. No treatment or usual care								
Albaladejo 2010	60	22.22	100	70	22.22	109	1.0%	-10.00 [-16.03 , -3.97]
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	0.9%	-9.50 [-21.15 , 2.15]
Cabak 2017	37	17.67	22	56	17.67	46	1.0%	-19.00 [-27.98 , -10.02]
Cortell-Tormo 2018	15	15	11	44	14	8	0.9%	-29.00 [-42.14 , -15.86]
Costa 2009b	46	28	77	56	26	77	1.0%	-10.00 [-18.53 , -1.47]
Cruz-Diaz 2017	17	14.1	34	49.6	13.1	15	1.0%	-32.60 [-40.75 , -24.45]
Cruz-Diaz 2017	21	13.6	34	49.6	13.1	15	1.0%	-28.60 [-36.65 , -20.55]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	1.0%	-14.90 [-23.47 , -6.33]
Garcia 2017	39.5	27.3	73	47	29.7	73	1.0%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	1.0%	-2.00 [-7.75 , 3.75]
Hall 2011	34	20.31	80	47	20.54	80	1.0%	-13.00 [-19.33 , -6.67]
Highland 2018	27.5	24.3	33	33.5	17.9	29	1.0%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	1.0%	-5.00 [-13.44 , 3.44]
Keane 2017	26.5	18.6	10	48.9	33.3	4	0.4%	-22.40 [-57.01 , 12.21]
Keane 2017	26	23.7	10	48.9	33.3	5	0.5%	-22.90 [-55.58 , 9.78]
Masharawi 2013	16.8	8.2	20	38.8	15.4	20	1.0%	-22.00 [-29.65 , -14.35]
Mazloum 2018	30	9	16	69	16	8	0.9%	-39.00 [-50.93 , -27.07]
Mazloum 2018	48	11	15	69	16	8	0.9%	-21.00 [-33.41 , -8.59]
Morone 2011	45	23	41	76	21	29	1.0%	-31.00 [-41.39 , -20.61]
Morone 2012	42.4	21.85	25	72.08	17.19	25	0.9%	-29.68 [-40.58 , -18.78]
Moseley 2002	19	15	29	31	15	28	1.0%	-12.00 [-19.79 , -4.21]
Oh 2014	14.4	21.23	9	38	14.11	3	0.7%	-23.60 [-44.75 , -2.45]
Oh 2014	11.1	14.11	9	38	14.11	3	0.8%	-26.90 [-45.34 , -8.46]
Oh 2014	34.4	14.11	10	38	14.11	3	0.8%	-3.60 [-21.80 , 14.60]
Phattharasupharerk 2019a	14	20.5	33	53.5	20.9	32	1.0%	-39.50 [-49.57 , -29.43]
Quinn 2011	30.93	24.06	15	35.2	24.06	14	0.8%	-4.27 [-21.79 , 13.25]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	1.0%	-15.60 [-24.89 , -6.31]
Schinhan 2016	13.3	10.5	15	21	17.3	15	1.0%	-7.70 [-17.94 , 2.54]
Smeets 2008	37.76	24.33	55	53.35	22.6	17	0.9%	-15.59 [-28.11 , -3.07]
Smeets 2008	44.63	28.86	55	53.35	22.6	17	0.9%	-8.72 [-21.90 , 4.46]
Smeets 2008	42.31	25.56	52	53.35	22.6	17	0.9%	-11.04 [-23.83 , 1.75]
Steele 2013	16.43	25.76	10	25.91	14.89	4	0.7%	-9.48 [-31.11 , 12.15]
Steele 2013	25	10.97	7	25.91	14.89	3	0.8%	-0.91 [-19.62 , 17.80]
Teut 2016	39.04	19.76	56	44.05	18.63	27	1.0%	-5.01 [-13.74 , 3.72]
Teut 2016	37.56	19.94	55	44.05	18.63	26	1.0%	-6.49 [-15.38 , 2.40]
Williams 2009	24.3	17.9	43	36.9	19.81	47	1.0%	-12.60 [-20.39 , -4.81]
Yelland 2004	42	26.5	55	37	45.4	55	0.9%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**1237**			**983**	**33.5%**	**-15.34 [-18.91 , -11.76]**
Heterogeneity: Tau² = 85.79; Chi² = 156.43, df = 36 (P < 0.00001); I² = 77%								
Test for overall effect: Z = 8.41 (P < 0.00001)								
7.1.2 Exercise vs. Other conservative treatment								
Alayat 2014	37.1	13	28	56.5	10.4	10	1.0%	-19.40 [-27.45 , -11.35]
Alayat 2014	26.4	12.5	24	56.5	10.4	10	1.0%	-30.10 [-38.26 , -21.94]
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.0%	-23.10 [-31.54 , -14.66]
Bellido-Fernandez 2018	30	18.52	9	30	22.22	5	0.7%	0.00 [-22.93 , 22.93]
Bellido-Fernandez 2018	30	22.22	9	30	22.22	4	0.6%	0.00 [-26.17 , 26.17]
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.1%	-24.98 [-29.09 , -20.87]
Bronfort 1996	27	20	56	39	21	21	1.0%	-12.00 [-22.40 , -1.60]
Bronfort 1996	35	22	40	39	21	21	0.9%	-4.00 [-15.28 , 7.28]
Bronfort 2011	32	22	91	29	19	49	1.0%	3.00 [-3.98 , 9.98]
Bronfort 2011	26	21	93	29	19	50	1.0%	-3.00 [-9.78 , 3.78]
Cecchi 2010	23.33	20	68	8.33	11.67	34	1.0%	15.00 [8.84 , 21.16]
Cecchi 2010	25	20	68	8.33	11.67	35	1.0%	16.67 [10.54 , 22.80]
Chan 2017	30	23	49	43	26	43	1.0%	-13.00 [-23.09 , -2.91]
Chown 2008	73.4	17.2	22	73.2	11.5	33	1.0%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016	38.1	12.1	53	56.9	16.3	48	1.0%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.1%	-24.00 [-29.30 , -18.70]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.0%	0.90 [-7.83 , 9.63]
Diaz-Arribas 2009	23	18.5	63	36.6	20	63	1.0%	-13.60 [-20.33 , -6.87]
Djavid 2007	43	16	18	44	20	8	0.8%	-1.00 [-16.71 , 14.71]
Djavid 2007	24	14	19	44	20	8	0.8%	-20.00 [-35.22 , -4.78]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.0%	-3.32 [-11.72 , 5.08]
Ferreira 2007	48	24	74	41	26	39	1.0%	7.00 [-2.82 , 16.82]
Ferreira 2007	40	25	73	41	26	38	1.0%	-1.00 [-11.06 , 9.06]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 7.1. (Continued)

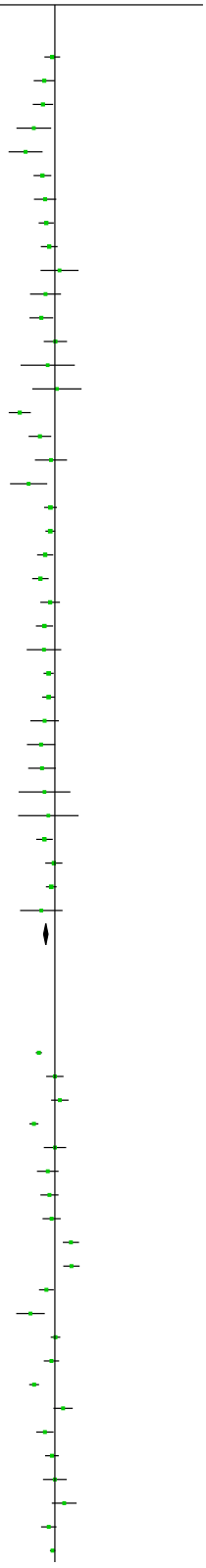
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.0%	-3.32 [-11.72 , 5.08]
Ferreira 2007	48	24	74	41	26	39	1.0%	7.00 [-2.82 , 16.82]
Ferreira 2007	40	25	73	41	26	38	1.0%	-1.00 [-11.06 , 9.06]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.0%	-10.00 [-17.40 , -2.60]
Goldby 2006	28.81	28.14	78	34.4	36.43	37	0.9%	-5.59 [-18.89 , 7.71]
Hansen 1993	33.33	21.25	44	33.33	21.25	59	1.0%	0.00 [-8.30 , 8.30]
Hartvigsen 2010	45.6	19.82	32	42	21.3	21	0.9%	3.60 [-7.81 , 15.01]
Hartvigsen 2010	37.97	19.38	36	42	21.3	20	0.9%	-4.03 [-15.31 , 7.25]
Hemmilä 1997	30	22.63	35	25	22.63	34	1.0%	5.00 [-5.68 , 15.68]
Jaromi 2018	7.51	8.3	67	49.3	7.73	70	1.1%	-41.79 [-44.48 , -39.10]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	0.9%	-8.10 [-22.94 , 6.74]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.1%	-25.90 [-31.06 , -20.74]
Kim 2017	24	5.4	15	45.7	9	15	1.1%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.0%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.0%	-28.60 [-35.11 , -22.09]
Longo 2016	36.2	22	7	61.4	22.7	7	0.6%	-25.20 [-48.62 , -1.78]
Machado 2007	47	40	17	58	28	11	0.6%	-11.00 [-36.21 , 14.21]
McDonough 2013	45	27.08	39	39	25.24	17	0.9%	6.00 [-8.70 , 20.70]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	1.0%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	1.0%	-21.00 [-30.72 , -11.28]
Miyamoto 2018	33	23	73	56	26	22	0.9%	-23.00 [-35.08 , -10.92]
Miyamoto 2018	40	27	67	56	26	23	0.9%	-16.00 [-28.44 , -3.56]
Miyamoto 2018	32	24	69	56	26	23	0.9%	-24.00 [-36.04 , -11.96]
Murtezani 2011	20	17	50	60	27	51	1.0%	-40.00 [-48.78 , -31.22]
Natour 2015	40.4	24.2	30	55.8	26.8	30	0.9%	-15.40 [-28.32 , -2.48]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.0%	-10.90 [-17.43 , -4.37]
Rantonen 2012	31	20	43	35	28	20	0.9%	-4.00 [-17.65 , 9.65]
Rantonen 2012	29	27	43	35	28	20	0.9%	-6.00 [-20.69 , 8.69]
Ritvanen 2007	28	14.11	28	23	14.11	33	1.0%	5.00 [-2.11 , 12.11]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.7%	-3.50 [-23.38 , 16.38]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	1.0%	-2.60 [-13.31 , 8.11]
Saper 2009	44	21	15	71	18	15	0.9%	-27.00 [-41.00 , -13.00]
Saper 2017	53	21	129	56	22	32	1.0%	-3.00 [-11.44 , 5.44]
Saper 2017	50	21	127	56	22	32	1.0%	-6.00 [-14.45 , 2.45]
Schenkman 2009	19	21	17	41	40	9	0.5%	-22.00 [-49.97 , 5.97]
Segal-Snir 2016	70	30	20	70	19	15	0.8%	0.00 [-16.29 , 16.29]
Shen 2009	11	17.67	15	25.5	17.67	7	0.8%	-14.50 [-30.35 , 1.35]
Shen 2009	23	17.67	15	25.5	17.67	8	0.8%	-2.50 [-17.66 , 12.66]
Sorensen 2010	49	22	91	46	21	94	1.0%	3.00 [-3.20 , 9.20]
Soukup 1999	18	15.7	34	24	17	18	1.0%	-6.00 [-15.46 , 3.46]
Soukup 1999	23	16	38	24	17	17	1.0%	-1.00 [-10.55 , 8.55]
Sousa 2009	33.5	24.8	27	47.6	28	25	0.9%	-14.10 [-28.52 , 0.32]
Torstensen 1998	37.2	25.3	70	39	28	33	0.9%	-1.80 [-13.04 , 9.44]
Torstensen 1998	50.4	27.2	71	39	28	34	0.9%	11.40 [0.06 , 22.74]
Ulger 2017	21	16	56	20.8	12	57	1.1%	0.20 [-5.02 , 5.42]
Valenza 2017	39	17.67	27	49	17.67	27	1.0%	-10.00 [-19.43 , -0.57]
Weifen 2013	24.3	17.67	141	32.4	17.67	12	1.0%	-8.10 [-18.51 , 2.31]
Weifen 2013	29.2	17.67	38	32.4	17.67	12	0.9%	-3.20 [-14.67 , 8.27]
Weifen 2013	30.6	17.67	47	32.4	17.67	11	0.9%	-1.80 [-13.40 , 9.80]
Weifen 2013	22.5	17.67	47	32.4	17.67	12	0.9%	-9.90 [-21.10 , 1.30]
Williams 2005	10	11	20	21	23	24	1.0%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**3309**			**2151**	**66.5%**	**-9.06 [-12.99 , -5.13]**
Heterogeneity: Tau² = 251.14; Chi² = 1048.31, df = 70 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 4.52 (P < 0.00001)								
Total (95% CI)			**4546**			**3134**	**100.0%**	**-11.18 [-14.11 , -8.26]**
Heterogeneity: Tau² = 204.12; Chi² = 1205.02, df = 107 (P < 0.00001); I² = 91%								
Test for overall effect: Z = 7.49 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.37, df = 1 (P = 0.02), I² = 81.4%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 7.2. Comparison 7: SENSITIVITY ANALYSES (Low risk of bias), Outcome 2: Function (/100): Earliest follow-up

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
7.2.1 Exercise vs. No treatment or usual care								
Albaladejo 2010	31.25	24.69	100	33.33	21.6	109	3.9%	-2.08 [-8.39 , 4.23]
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	2.8%	-8.50 [-16.99 , -0.01]
Costa 2009b	40	27.08	77	49.58	24.58	77	2.9%	-9.58 [-17.75 , -1.41]
Cruz-Diaz 2017	26.46	22.08	34	43.37	23.33	15	1.3%	-16.91 [-30.86 , -2.96]
Cruz-Diaz 2017	19.83	20.42	34	43.37	23.33	15	1.4%	-23.54 [-37.20 , -9.88]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	3.5%	-10.00 [-17.14 , -2.86]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	2.6%	-7.83 [-16.79 , 1.13]
Groessl 2017	30.62	19.06	75	37.54	19.06	75	4.1%	-6.92 [-13.02 , -0.82]
Hall 2011	29.21	21.49	80	33.75	21.87	80	3.7%	-4.54 [-11.26 , 2.18]
Harts 2008	25.42	25	20	21.67	16.25	9	1.2%	3.75 [-11.51 , 19.01]
Harts 2008	14.17	16.67	20	21.67	16.25	10	1.6%	-7.50 [-19.94 , 4.94]
Highland 2018	18.46	19.25	33	29.33	18.96	29	2.4%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	2.5%	0.43 [-8.93 , 9.79]
Keane 2017	25.8	20.1	10	31.56	18.24	4	0.6%	-5.76 [-27.55 , 16.03]
Keane 2017	33.2	18.84	10	31.56	18.24	5	0.7%	1.64 [-18.16 , 21.44]
Kofotolis 2016	17.04	8.08	36	45.25	16.17	14	2.6%	-28.21 [-37.08 , -19.34]
Kofotolis 2016	33.25	10.17	37	45.25	16.17	14	2.6%	-12.00 [-21.08 , -2.92]
Little 2014	27.95	13.92	15	31.08	19.87	13	1.5%	-3.13 [-16.03 , 9.77]
Masharawi 2013	38.79	24.17	20	59.87	24.04	20	1.2%	-21.08 [-36.02 , -6.14]
Mazloum 2018	23.1	7.5	15	26.6	5	8	4.7%	-3.50 [-8.64 , 1.64]
Mazloum 2018	22.9	3.6	16	26.6	5	8	5.7%	-3.70 [-7.59 , 0.19]
Morone 2011	18	12.9	41	25.8	14.1	29	3.8%	-7.80 [-14.28 , -1.32]
Morone 2012	13.68	9.39	25	25.28	14.03	25	3.8%	-11.60 [-18.22 , -4.98]
Patti 2016	13	8	19	16.8	15.6	19	3.1%	-3.80 [-11.68 , 4.08]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	3.6%	-8.34 [-15.23 , -1.45]
Quinn 2011	22.5	19.06	15	31.25	19.06	14	1.4%	-8.75 [-22.63 , 5.13]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	5.6%	-5.00 [-9.08 , -0.92]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	4.7%	-5.06 [-10.19 , 0.07]
Smeets 2008	49.6	24.56	55	57.83	19.92	17	1.8%	-8.23 [-19.71 , 3.25]
Smeets 2008	46.74	23.01	52	57.83	19.92	17	1.9%	-11.09 [-22.44 , 0.26]
Smeets 2008	47.5	21.88	55	57.83	19.92	17	1.9%	-10.33 [-21.43 , 0.77]
Steele 2013	14.86	17.95	10	23.2	17.95	4	0.7%	-8.34 [-29.15 , 12.47]
Steele 2013	17.98	17.95	7	23.2	17.95	3	0.5%	-5.22 [-29.50 , 19.06]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	3.7%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	3.6%	-0.83 [-7.72 , 6.06]
Williams 2009	17.9	10.49	43	20.8	10.28	47	5.4%	-2.90 [-7.20 , 1.40]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	0.9%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**1351**			**1096**	**100.0%**	**-7.27 [-9.04 , -5.49]**
Heterogeneity: Tau² = 10.39; Chi² = 60.22, df = 36 (P = 0.007); I² = 40%								
Test for overall effect: Z = 8.03 (P < 0.00001)								
7.2.2 Exercise vs. Other conservative treatment								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.2%	-12.84 [-15.31 , -10.37]
Bellido-Fernandez 2018	8	3.7	9	8	6.67	4	1.8%	0.00 [-6.97 , 6.97]
Bellido-Fernandez 2018	12	5.93	9	8	6.67	5	1.8%	4.00 [-3.01 , 11.01]
Bid 2017	7.29	8.54	64	24.12	12	64	2.2%	-16.83 [-20.44 , -13.22]
Bronfort 1996	20.9	17	43	20.8	17.3	21	1.6%	0.10 [-8.88 , 9.08]
Bronfort 1996	15.1	17.4	56	20.8	17.3	21	1.6%	-5.70 [-14.39 , 2.99]
Bronfort 2011	16.96	20	92	21.3	21.74	50	1.8%	-4.34 [-11.62 , 2.94]
Bronfort 2011	18.7	19.57	91	21.3	21.74	49	1.8%	-2.60 [-9.90 , 4.70]
Cecchi 2010	22.08	19.58	68	9.17	13.75	35	1.9%	12.91 [6.40 , 19.42]
Cecchi 2010	22.5	19.58	68	9.17	13.75	34	1.9%	13.33 [6.77 , 19.89]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.9%	-6.80 [-12.74 , -0.86]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.4%	-19.58 [-31.05 , -8.11]
Chown 2008	20.8	8.4	35	20.3	8	35	2.1%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016	16.55	15.78	53	19.29	15.78	48	1.9%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	2.1%	-16.67 [-20.63 , -12.71]
Djavid 2007	48.2	10.4	18	41.6	8.8	8	1.7%	6.60 [-1.16 , 14.36]
Djavid 2007	33.6	7.4	19	41.6	8.8	8	1.8%	-8.00 [-14.95 , -1.05]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	2.0%	-2.37 [-7.87 , 3.13]
Ferreira 2007	32.92	23.75	74	32.92	25	39	1.6%	0.00 [-9.53 , 9.53]
Ferreira 2007	40.42	26.25	73	32.92	25	38	1.5%	7.50 [-2.47 , 17.47]
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.9%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.3%	-1.80 [-3.96 , 0.36]



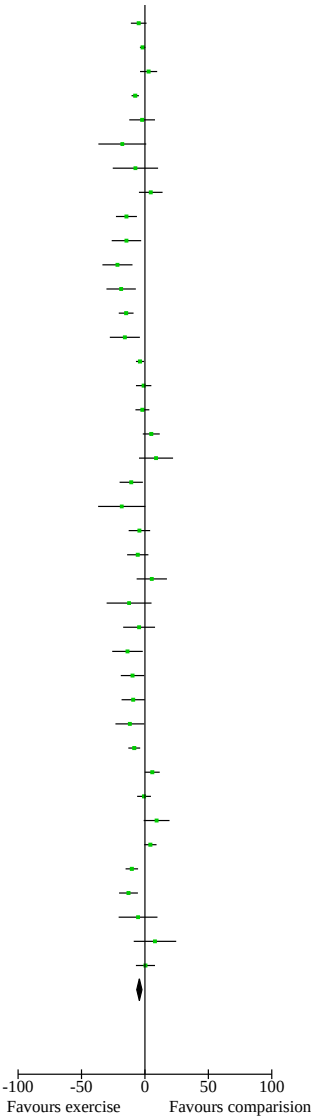
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 7.2. (Continued)

Frost 1995	16.3	10.3	29	21.2	14.2	32	1.9%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.3%	-1.80 [-3.96 , 0.36]
Goldby 2006	31	17.07	78	28.1	17.34	37	1.9%	2.90 [-3.85 , 9.65]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.2%	-7.66 [-10.55 , -4.77]
Little 2014	23.16	13.92	14	25.37	13.92	15	1.5%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.8%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	0.8%	-7.50 [-25.37 , 10.37]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.6%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.7%	-14.58 [-22.85 , -6.31]
Miyamoto 2018	32.5	21.67	67	47.08	25.42	23	1.3%	-14.58 [-26.19 , -2.97]
Miyamoto 2018	25.42	22.92	73	47.08	25.42	22	1.3%	-21.66 [-33.51 , -9.81]
Miyamoto 2018	28.33	21.67	69	47.08	25.42	23	1.3%	-18.75 [-30.33 , -7.17]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	2.0%	-14.80 [-20.62 , -8.98]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.3%	-15.83 [-27.67 , -3.99]
Niemisto 2003	14.7	11.6	96	18.6	11	100	2.2%	-3.90 [-7.07 , -0.73]
Rantonen 2012	15	14	43	16	10	20	1.9%	-1.00 [-7.06 , 5.06]
Rantonen 2012	14	11	43	16	10	20	2.0%	-2.00 [-7.48 , 3.48]
Ritvanen 2007	17	13.27	28	12	13.27	33	1.9%	5.00 [-1.68 , 11.68]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.2%	8.75 [-4.68 , 22.18]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.6%	-10.80 [-19.97 , -1.63]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.8%	-18.27 [-36.90 , 0.36]
Saper 2017	49.13	22.17	127	53.48	21.74	32	1.7%	-4.35 [-12.81 , 4.11]
Saper 2017	47.83	21.3	129	53.48	21.74	32	1.7%	-5.65 [-14.03 , 2.73]
Schenkman 2009	15.81	10.4	17	10.4	16.64	9	1.3%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.9%	-12.50 [-30.19 , 5.19]
Sherman 2005	22.08	18.75	36	26.67	21.67	15	1.2%	-4.59 [-17.15 , 7.97]
Sherman 2005	12.92	15	35	26.67	21.67	15	1.3%	-13.75 [-25.79 , -1.71]
Sherman 2011	18.74	16.28	91	28.52	20.61	22	1.6%	-9.78 [-19.02 , -0.54]
Sherman 2011	19.26	17.56	92	28.52	20.61	23	1.6%	-9.26 [-18.42 , -0.10]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.4%	-11.88 [-23.25 , -0.51]
Tilbrook 2011	23.75	20.31	156	32.17	20.31	136	2.1%	-8.42 [-13.09 , -3.75]
Torstensen 1998	52.7	16.6	71	46.9	13.1	34	2.0%	5.80 [-0.06 , 11.66]
Torstensen 1998	46.2	13.1	70	46.9	13.1	33	2.0%	-0.70 [-6.12 , 4.72]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.5%	9.20 [-0.98 , 19.38]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	2.1%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	2.1%	-10.33 [-15.25 , -5.41]
Valenza 2017	16.96	13.92	27	29.92	13.92	27	1.8%	-12.96 [-20.39 , -5.53]
Vincent 2014a	20.83	17.08	17	26.25	17.5	7	1.0%	-5.42 [-20.72 , 9.88]
Vincent 2014a	34.17	22.92	18	26.25	17.5	7	0.9%	7.92 [-8.82 , 24.66]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.8%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**3235**			**2198**	**100.0%**	**-4.39 [-6.45 , -2.34]**

Heterogeneity: Tau² = 47.46; Chi² = 353.06, df = 59 (P < 0.00001); I² = 83%
Test for overall effect: Z = 4.18 (P < 0.0001)

Test for subgroup differences: Chi² = 4.30, df = 1 (P = 0.04), I² = 76.7%



Comparison 8. SENSITIVITY ANALYSES (Excluding outlying means)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.1 Pain (/100): Earliest follow-up	87	8017	Mean Difference (IV, Random, 95% CI)	-8.68 [-10.82, -6.54]
8.1.1 Exercise vs. No treatment or usual care	30	2403	Mean Difference (IV, Random, 95% CI)	-11.49 [-13.82, -9.17]
8.1.2 Exercise vs. Other conservative treatment	57	5614	Mean Difference (IV, Random, 95% CI)	-7.25 [-10.15, -4.34]
8.2 Function (/100): Earliest follow-up	84		Mean Difference (IV, Random, 95% CI)	Subtotals only

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.2.1 Exercise vs. No treatment or usual care	35	2703	Mean Difference (IV, Random, 95% CI)	-5.35 [-6.47, -4.23]
8.2.2 Exercise vs. Other conservative treatment	51	5727	Mean Difference (IV, Random, 95% CI)	-3.61 [-5.49, -1.73]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

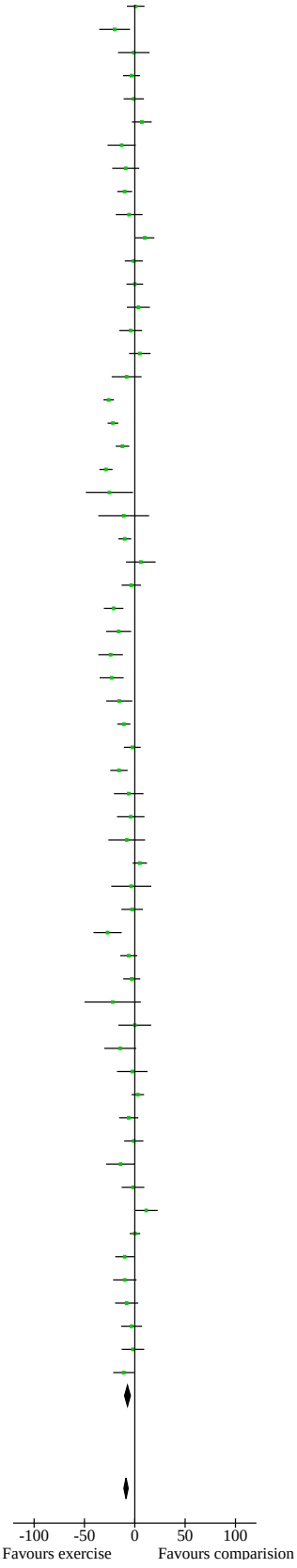
Analysis 8.1. Comparison 8: SENSITIVITY ANALYSES (Excluding outlying means), Outcome 1: Pain (/100): Earliest follow-up

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
8.1.1 Exercise vs. No treatment or usual care								
Albaladejo 2010 (1)	60	22.22	100	70	22.22	109	1.1%	-10.00 [-16.03 , -3.97]
Arampatzis 2017	29.8	18.5	20	39.3	19.1	20	0.9%	-9.50 [-21.15 , 2.15]
Cabak 2017 (2)	37	17.67	22	56	17.67	46	1.0%	-19.00 [-27.98 , -10.02]
Chen 2014	21.7	14.2	64	34.8	17.7	63	1.1%	-13.10 [-18.69 , -7.51]
Cortell-Tormo 2018	15	15	11	44	14	8	0.8%	-29.00 [-42.14 , -15.86]
Costa 2009b	46	28	77	56	26	77	1.0%	-10.00 [-18.53 , -1.47]
Cuesta-Vargas 2012	18	10.3	25	32.9	18.9	24	1.0%	-14.90 [-23.47 , -6.33]
Garcia 2017	39.5	27.3	73	47	29.7	73	1.0%	-7.50 [-16.75 , 1.75]
Gladwell 2006	22	9	20	24	8	14	1.1%	-2.00 [-7.75 , 3.75]
Hall 2011 (2)	34	20.31	80	47	20.54	80	1.1%	-13.00 [-19.33 , -6.67]
Highland 2018	27.5	24.3	33	33.5	17.9	29	0.9%	-6.00 [-16.54 , 4.54]
Jensen 2012	45	21	42	50	19	45	1.0%	-5.00 [-13.44 , 3.44]
Keane 2017 (3)	26.5	18.6	10	48.9	33.3	4	0.3%	-22.40 [-57.01 , 12.21]
Keane 2017 (3)	26	23.7	10	48.9	33.3	5	0.3%	-22.90 [-55.58 , 9.78]
Kell 2009 (4)	33	5	9	48	7	4	1.0%	-15.00 [-22.60 , -7.40]
Kell 2009 (4)	48	8	9	48	7	5	1.0%	0.00 [-8.06 , 8.06]
Lee 2016a (3)	22	11.3	15	35.8	17.2	3	0.6%	-13.80 [-34.09 , 6.49]
Lee 2016a (3)	33.1	20	15	35.8	17.2	3	0.5%	-2.70 [-24.64 , 19.24]
Masharawi 2013	16.8	8.2	20	38.8	15.4	20	1.0%	-22.00 [-29.65 , -14.35]
Morone 2012 (5)	42.4	21.85	25	72.08	17.19	25	0.9%	-29.68 [-40.58 , -18.78]
Moseley 2002	19	15	29	31	15	28	1.0%	-12.00 [-19.79 , -4.21]
Muharram 2011 (2)	41.2	35.15	78	53.1	37.88	80	0.9%	-11.90 [-23.29 , -0.51]
Nassif 2011	27.6	20.6	32	44.1	27.4	28	0.8%	-16.50 [-28.91 , -4.09]
Oh 2014 (4)	34.4	14.11	10	38	14.11	3	0.6%	-3.60 [-21.80 , 14.60]
Oh 2014 (4)	14.4	21.23	9	38	14.11	3	0.5%	-23.60 [-44.75 , -2.45]
Oh 2014 (4)	11.1	14.11	9	38	14.11	3	0.6%	-26.90 [-45.34 , -8.46]
Puppin 2011	24	27	30	41	27	25	0.8%	-17.00 [-31.33 , -2.67]
Quinn 2011 (2)	30.93	24.06	15	35.2	24.06	14	0.7%	-4.27 [-21.79 , 13.25]
Rydeard 2006a	18.3	14.66	21	33.9	14.85	18	1.0%	-15.60 [-24.89 , -6.31]
Schinhan 2016	13.3	10.5	15	21	17.3	15	0.9%	-7.70 [-17.94 , 2.54]
Smeets 2008 (3)	44.63	28.86	55	53.35	22.6	17	0.8%	-8.72 [-21.90 , 4.46]
Smeets 2008 (3)	42.31	25.56	52	53.35	22.6	17	0.8%	-11.04 [-23.83 , 1.75]
Smeets 2008 (3)	37.76	24.33	55	53.35	22.6	17	0.8%	-15.59 [-28.11 , -3.07]
Steele 2013 (4)	16.43	25.76	10	25.91	14.89	4	0.5%	-9.48 [-31.11 , 12.15]
Steele 2013 (4)	25	10.97	7	25.91	14.89	3	0.6%	-0.91 [-19.62 , 17.80]
Teut 2016 (4)	37.56	19.94	56	44.05	18.63	27	1.0%	-6.49 [-15.25 , 2.27]
Teut 2016 (4)	39.04	19.76	55	44.05	18.63	26	1.0%	-5.01 [-13.87 , 3.85]
Williams 2009 (2)	24.3	17.9	43	36.9	19.81	47	1.0%	-12.60 [-20.39 , -4.81]
Yelland 2004	42	26.5	55	37	45.4	55	0.8%	5.00 [-8.89 , 18.89]
Subtotal (95% CI)			**1316**			**1087**	**32.6%**	**-11.49 [-13.82 , -9.17]**
Heterogeneity: Tau² = 22.74; Chi² = 71.99, df = 38 (P = 0.0007); I² = 47%								
Test for overall effect: Z = 9.68 (P < 0.00001)								
8.1.2 Exercise vs. Other conservative treatment								
Areeudomwong 2017	15.4	15.6	21	38.5	12.1	21	1.0%	-23.10 [-31.54 , -14.66]
Barberini 2011 (2)	52	17.38	10	81	17.38	12	0.8%	-29.00 [-43.59 , -14.41]
Bellido-Fernandez 2018 (3)	30	22.22	9	30	22.22	5	0.5%	0.00 [-24.29 , 24.29]
Bellido-Fernandez 2018 (3)	30	18.52	9	30	22.22	4	0.5%	0.00 [-24.91 , 24.91]
Bid 2017	5.62	8.7	61	30.6	14.2	64	1.1%	-24.98 [-29.09 , -20.87]
Bronfort 1996 (4)	27	20	40	39	21	21	0.9%	-12.00 [-22.91 , -1.09]
Bronfort 1996 (4)	35	22	56	39	21	21	0.9%	-4.00 [-14.67 , 6.67]
Bronfort 2011 (3)	26	21	93	29	19	50	1.0%	-3.00 [-9.78 , 3.78]
Bronfort 2011 (3)	32	22	91	29	19	49	1.0%	3.00 [-3.98 , 9.98]
Cambron 2006	22.1	20.4	86	14.6	17.67	108	1.1%	7.50 [2.05 , 12.95]
Cecchi 2010 (3)	25	20	68	8.33	11.67	35	1.1%	16.67 [10.54 , 22.80]
Cecchi 2010 (3)	23.33	20	68	8.33	11.67	34	1.1%	15.00 [8.84 , 21.16]
Chan 2017	30	23	49	43	26	43	0.9%	-13.00 [-23.09 , -2.91]
Chown 2008 (5)	73.4	17.2	22	73.2	11.5	33	1.0%	0.20 [-7.99 , 8.39]
Cruz-Diaz 2016 (2)	38.1	12.1	53	56.9	16.3	48	1.1%	-18.80 [-24.45 , -13.15]
Cruz-Diaz 2018	19.5	8.08	32	43.5	12.58	30	1.1%	-24.00 [-29.30 , -18.70]
Deyo 1990	26.5	24.6	62	25.6	24.6	60	1.0%	0.90 [-7.83 , 9.63]
Djavid 2007 (3)	24	14	18	44	20	8	0.7%	-20.00 [-35.29 , -4.71]
Djavid 2007 (3)	43	16	19	44	20	8	0.7%	-1.00 [-16.62 , 14.62]

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.1. (Continued)

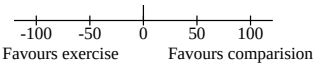
Djavid 2007 (3)	24	14	18	44	20	8	0.7%	-20.00 [-35.29 , -4.71]
Djavid 2007 (3)	43	16	19	44	20	8	0.7%	-1.00 [-16.62 , 14.62]
Dougherty 2014	38.1	30.88	89	41.42	26.52	92	1.0%	-3.32 [-11.72 , 5.08]
Ferreira 2007 (3)	40	25	73	41	26	38	0.9%	-1.00 [-11.06 , 9.06]
Ferreira 2007 (3)	48	24	74	41	26	39	0.9%	7.00 [-2.82 , 16.82]
Fransoo 2006 (4)	7	17.67	19	20	17.67	9	0.8%	-13.00 [-27.01 , 1.01]
Fransoo 2006 (4)	11	17.67	20	20	17.67	10	0.8%	-9.00 [-22.41 , 4.41]
Frost 1995	12.1	9.9	36	22.1	20.1	35	1.0%	-10.00 [-17.40 , -2.60]
Goldby 2006 (5)	28.81	28.14	78	34.4	36.43	37	0.8%	-5.59 [-18.89 , 7.71]
Gur 2003 (3)	29	13	25	19	14	12	1.0%	10.00 [0.58 , 19.42]
Gur 2003 (3)	18	12	25	19	14	13	1.0%	-1.00 [-9.95 , 7.95]
Hansen 1993 (1)	33.33	21.25	44	33.33	21.25	59	1.0%	0.00 [-8.30 , 8.30]
Hartvigsen 2010 (3)	45.6	19.82	32	42	21.3	21	0.9%	3.60 [-7.81 , 15.01]
Hartvigsen 2010 (3)	37.97	19.38	36	42	21.3	20	0.9%	-4.03 [-15.31 , 7.25]
Hemmilä 1997 (1)	30	22.63	35	25	22.63	34	0.9%	5.00 [-5.68 , 15.68]
Kankaanpaa 1999	36.8	28.8	30	44.9	26.7	24	0.7%	-8.10 [-22.94 , 6.74]
Kim 2015a	26.7	8.9	27	52.6	10.2	26	1.1%	-25.90 [-31.06 , -20.74]
Kim 2017	24	5.4	15	45.7	9	15	1.1%	-21.70 [-27.01 , -16.39]
Kumar 2009a	12.5	14.28	51	24.7	20	51	1.0%	-12.20 [-18.94 , -5.46]
Kumar 2009b	14.7	9.9	15	43.3	8.2	15	1.1%	-28.60 [-35.11 , -22.09]
Longo 2016	36.2	22	7	61.4	22.7	7	0.5%	-25.20 [-48.62 , -1.78]
Machado 2007	47	40	17	58	28	11	0.4%	-11.00 [-36.21 , 14.21]
Maul 2005	40	20	74	50	20	74	1.1%	-10.00 [-16.44 , -3.56]
McDonough 2013 (2)	45	27.08	39	39	25.24	17	0.8%	6.00 [-8.70 , 20.70]
Michalsen 2016	37.3	18.2	36	40.8	21.8	32	0.9%	-3.50 [-13.11 , 6.11]
Miyamoto 2013	31	23	43	52	23	43	0.9%	-21.00 [-30.72 , -11.28]
Miyamoto 2018 (3)	40	27	69	56	26	23	0.8%	-16.00 [-28.39 , -3.61]
Miyamoto 2018 (3)	32	24	73	56	26	22	0.8%	-24.00 [-36.18 , -11.82]
Miyamoto 2018 (3)	33	23	67	56	26	23	0.9%	-23.00 [-34.97 , -11.03]
Natour 2015	40.4	24.2	30	55.8	26.8	30	0.8%	-15.40 [-28.32 , -2.48]
Niemisto 2003	25.2	23.3	96	36.1	23.3	100	1.1%	-10.90 [-17.43 , -4.37]
Noori 2011	41.5	15.3	17	44	7.5	13	1.0%	-2.50 [-10.84 , 5.84]
Okafor 2012	33.2	12.2	15	48.9	11.9	15	1.0%	-15.70 [-24.32 , -7.08]
Rantonen 2012 (3)	29	27	43	35	28	20	0.8%	-6.00 [-20.69 , 8.69]
Rantonen 2012 (3)	31	20	43	35	28	20	0.8%	-4.00 [-17.65 , 9.65]
Rasmussen-Barr 2003	14	16.18	17	22	33.82	16	0.6%	-8.00 [-26.27 , 10.27]
Ritvanen 2007 (2)	28	14.11	28	23	14.11	33	1.0%	5.00 [-2.11 , 12.11]
Ryan 2010	19.1	18.9	15	22.6	30.8	12	0.6%	-3.50 [-23.38 , 16.38]
Salavati 2016	26.1	19.6	20	28.7	14.6	20	0.9%	-2.60 [-13.31 , 8.11]
Saper 2009	44	21	15	71	18	15	0.8%	-27.00 [-41.00 , -13.00]
Saper 2017 (3)	50	21	127	56	22	32	1.0%	-6.00 [-14.45 , 2.45]
Saper 2017 (3)	53	21	129	56	22	32	1.0%	-3.00 [-11.44 , 5.44]
Schenkman 2009 (5)	19	21	17	41	40	9	0.4%	-22.00 [-49.97 , 5.97]
Segal-Snir 2016	70	30	20	70	19	15	0.7%	0.00 [-16.29 , 16.29]
Shen 2009 (4)	11	17.67	15	25.5	17.67	7	0.7%	-14.50 [-30.35 , 1.35]
Shen 2009 (4)	23	17.67	15	25.5	17.67	8	0.7%	-2.50 [-17.66 , 12.66]
Sorensen 2010	49	22	91	46	21	94	1.1%	3.00 [-3.20 , 9.20]
Soukup 1999 (3)	18	15.7	38	24	17	17	1.0%	-6.00 [-15.50 , 3.50]
Soukup 1999 (3)	23	16	34	24	17	18	1.0%	-1.00 [-10.52 , 8.52]
Sousa 2009	33.5	24.8	27	47.6	28	25	0.8%	-14.10 [-28.52 , 0.32]
Torstensen 1998 (4)	37.2	25.3	70	39	28	33	0.9%	-1.80 [-13.04 , 9.44]
Torstensen 1998 (4)	50.4	27.2	71	39	28	34	0.9%	11.40 [0.06 , 22.74]
Ulger 2017	21	16	56	20.8	12	57	1.1%	0.20 [-5.02 , 5.42]
Valenza 2017 (2)	39	17.67	27	49	17.67	27	1.0%	-10.00 [-19.43 , -0.57]
Weifen 2013 (4)	22.5	17.67	47	32.4	17.67	11	0.9%	-9.90 [-21.50 , 1.70]
Weifen 2013 (4)	24.3	17.67	38	32.4	17.67	12	0.9%	-8.10 [-19.57 , 3.37]
Weifen 2013 (4)	29.2	17.67	141	32.4	17.67	12	0.9%	-3.20 [-13.61 , 7.21]
Weifen 2013 (4)	30.6	17.67	47	32.4	17.67	12	0.9%	-1.80 [-13.00 , 9.40]
Williams 2005	10	11	20	21	23	24	0.9%	-11.00 [-21.39 , -0.61]
Subtotal (95% CI)			**3385**			**2229**	**67.4%**	**-7.25 [-10.15 , -4.34]**
Heterogeneity: Tau² = 133.08; Chi² = 562.32, df = 75 (P < 0.00001); I² = 87%								
Test for overall effect: Z = 4.89 (P < 0.00001)								
Total (95% CI)			**4701**			**3316**	**100.0%**	**-8.68 [-10.82 , -6.54]**
Heterogeneity: Tau² = 101.63; Chi² = 645.11, df = 114 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 7.96 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.01, df = 1 (P = 0.03), I² = 80.0%								



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.1. (Continued)

Heterogeneity: Tau² = 101.83; Chi² = 645.11, df = 114 (P < 0.00001); I² = 82%
Test for overall effect: Z = 7.96 (P < 0.00001)
Test for subgroup differences: Chi² = 5.01, df = 1 (P = 0.03), I² = 80.0%



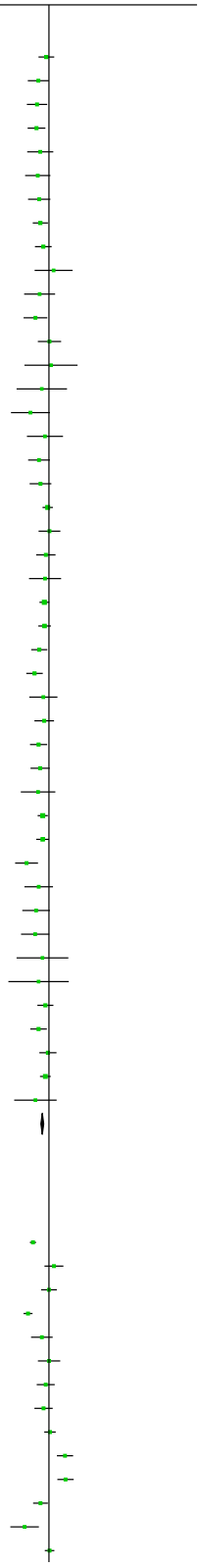
Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Standard deviation(s) required imputation due to missing or implausible value.
(3) Study has multiple exercise groups available: comparison group data divided.
(4) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(5) Study has other comparison groups available for this meta-analysis.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 8.2. Comparison 8: SENSITIVITY ANALYSES (Excluding outlying means), Outcome 2: Function (/100): Earliest follow-up

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Comparison Mean	Comparison SD	Comparison Total	Weight	Mean Difference IV, Random, 95% CI
8.2.1 Exercise vs. No treatment or usual care								
Albaladejo 2010 (1)	31.25	24.69	100	33.33	21.6	109	3.1%	-2.08 [-8.39 , 4.23]
Cortell-Tormo 2018	6	6.1	11	14.5	11.1	8	1.7%	-8.50 [-16.99 , -0.01]
Costa 2009b	40	27.08	77	49.58	24.58	77	1.9%	-9.58 [-17.75 , -1.41]
Cuesta-Vargas 2012	11.25	7.5	25	21.25	16.25	24	2.5%	-10.00 [-17.14 , -2.86]
Dalichau 2000 (2)	42	13.92	21	49	13.92	10	1.1%	-7.00 [-17.48 , 3.48]
Dalichau 2000 (2)	40	13.92	21	49	13.92	11	1.2%	-9.00 [-19.15 , 1.15]
Garcia 2017	33.21	26.33	73	41.04	28.87	73	1.6%	-7.83 [-16.79 , 1.13]
Groessl 2017 (3)	30.62	19.06	75	37.54	19.06	75	3.4%	-6.92 [-13.02 , -0.82]
Hall 2011 (3)	29.21	21.49	80	33.75	21.87	80	2.8%	-4.54 [-11.26 , 2.18]
Harts 2008 (4)	25.42	25	20	21.67	16.25	9	0.5%	3.75 [-11.51 , 19.01]
Harts 2008 (4)	14.17	16.67	20	21.67	16.25	10	0.8%	-7.50 [-19.94 , 4.94]
Highland 2018	18.46	19.25	33	29.33	18.96	29	1.4%	-10.87 [-20.40 , -1.34]
Jensen 2012	48.26	23.48	42	47.83	20.87	45	1.4%	0.43 [-8.93 , 9.79]
Keane 2017 (4)	33.2	18.84	10	31.56	18.24	4	0.3%	1.64 [-19.71 , 22.99]
Keane 2017 (4)	25.8	20.1	10	31.56	18.24	5	0.3%	-5.76 [-26.03 , 14.51]
Kell 2009 (2)	24.2	13.27	9	39.1	13.27	4	0.5%	-14.90 [-30.53 , 0.73]
Kell 2009 (2)	35.9	13.27	9	39.1	13.27	5	0.6%	-3.20 [-17.71 , 11.31]
Kofotolis 2006 (2)	20	13.92	28	28	13.92	15	1.6%	-8.00 [-16.73 , 0.73]
Kofotolis 2006 (2)	21.2	13.92	28	28	13.92	15	1.6%	-6.80 [-15.53 , 1.93]
Kuukkanen 2000	10	7.6	27	11	7.8	26	7.3%	-1.00 [-5.15 , 3.15]
Lee 2016a (4)	7.5	8.75	15	7.08	6.67	3	1.6%	0.42 [-8.33 , 9.17]
Lee 2016a (4)	4.58	3.75	15	7.08	6.67	3	2.1%	-2.50 [-10.28 , 5.28]
Little 2014 (3)	27.95	13.92	15	31.08	19.87	13	0.8%	-3.13 [-16.03 , 9.77]
Mazloum 2018 (4)	22.9	3.6	15	26.6	5	8	8.1%	-3.70 [-7.61 , 0.21]
Mazloum 2018 (4)	23.1	7.5	16	26.6	5	8	4.9%	-3.50 [-8.55 , 1.55]
Morone 2011	18	12.9	41	25.8	14.1	29	3.0%	-7.80 [-14.28 , -1.32]
Morone 2012 (5)	13.68	9.39	25	25.28	14.03	25	2.8%	-11.60 [-18.22 , -4.98]
Nassif 2011	40.62	20.83	32	45.12	23.54	28	1.0%	-4.50 [-15.82 , 6.82]
Patti 2016	13	8	19	16.8	15.6	19	2.0%	-3.80 [-11.68 , 4.08]
Phattharasupharerk 2019a	8.33	15	33	16.67	13.33	32	2.6%	-8.34 [-15.23 , -1.45]
Puppin 2011	18.4	13.3	30	25.5	15.5	25	2.1%	-7.10 [-14.82 , 0.62]
Quinn 2011 (3)	22.5	19.06	15	31.25	19.06	14	0.6%	-8.75 [-22.63 , 5.13]
Rydeard 2006a	8.33	5.73	21	13.33	7.07	18	7.5%	-5.00 [-9.08 , -0.92]
Schinhan 2016	6.4	5.36	15	11.46	8.6	15	4.7%	-5.06 [-10.19 , 0.07]
Shaughnessy 2004	26	14	20	44	16	21	1.5%	-18.00 [-27.19 , -8.81]
Smeets 2008 (4)	49.6	24.56	55	57.83	19.92	17	0.9%	-8.23 [-19.71 , 3.25]
Smeets 2008 (4)	47.5	21.88	55	57.83	19.92	17	1.0%	-10.33 [-21.43 , 0.77]
Smeets 2008 (4)	46.74	23.01	52	57.83	19.92	17	1.0%	-11.09 [-22.44 , 0.26]
Steele 2013 (2)	17.98	17.95	10	23.2	17.95	4	0.3%	-5.22 [-26.03 , 15.59]
Steele 2013 (2)	14.86	17.95	7	23.2	17.95	3	0.2%	-8.34 [-32.62 , 15.94]
Stephan 2011	4.1	7.98	58	6.97	12.45	16	3.0%	-2.87 [-9.31 , 3.57]
Vollenbroek-Hutten 2004	45.83	20.83	72	54.17	20.83	79	2.8%	-8.34 [-14.99 , -1.69]
Weiner 2008	31.67	19.17	50	32.5	15.83	50	2.6%	-0.83 [-7.72 , 6.06]
Williams 2009 (3)	17.9	10.49	43	20.8	10.28	47	6.8%	-2.90 [-7.20 , 1.40]
Yelland 2004	36.97	41.33	55	47.85	49.59	55	0.4%	-10.88 [-27.94 , 6.18]
Subtotal (95% CI)			**1503**			**1200**	**100.0%**	**-5.35 [-6.47 , -4.23]**
Heterogeneity: Tau² = 0.00; Chi² = 39.09, df = 44 (P = 0.68); I² = 0%								
Test for overall effect: Z = 9.38 (P < 0.00001)								
8.2.2 Exercise vs. Other conservative treatment								
Areeudomwong 2017	7.04	3.58	21	19.88	4.54	21	2.1%	-12.84 [-15.31 , -10.37]
Bellido-Fernandez 2018 (4)	12	5.93	9	8	6.67	4	1.6%	4.00 [-3.60 , 11.60]
Bellido-Fernandez 2018 (4)	8	3.7	9	8	6.67	5	1.7%	0.00 [-6.33 , 6.33]
Bid 2017	7.29	8.54	64	24.12	12	64	2.0%	-16.83 [-20.44 , -13.22]
Bronfort 1996 (2)	15.1	17.4	56	20.8	17.3	21	1.5%	-5.70 [-14.39 , 2.99]
Bronfort 1996 (2)	20.9	17	43	20.8	17.3	21	1.4%	0.10 [-8.88 , 9.08]
Bronfort 2011 (4)	18.7	19.57	92	21.3	21.74	50	1.6%	-2.60 [-9.83 , 4.63]
Bronfort 2011 (4)	16.96	20	91	21.3	21.74	49	1.6%	-4.34 [-11.68 , 3.00]
Cambron 2006	12.08	15.55	87	11.25	17.48	110	1.9%	0.83 [-3.79 , 5.45]
Cecchi 2010 (4)	22.08	19.58	68	9.17	13.75	35	1.7%	12.91 [6.40 , 19.42]
Cecchi 2010 (4)	22.5	19.58	68	9.17	13.75	34	1.7%	13.33 [6.77 , 19.89]
Chan 2017	17.6	12.6	47	24.4	15.8	43	1.8%	-6.80 [-12.74 , -0.86]
Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Chown 2008 (5)	20.8	8.4	35	20.3	8	35	2.0%	0.50 [-3.34 , 4.34]



Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

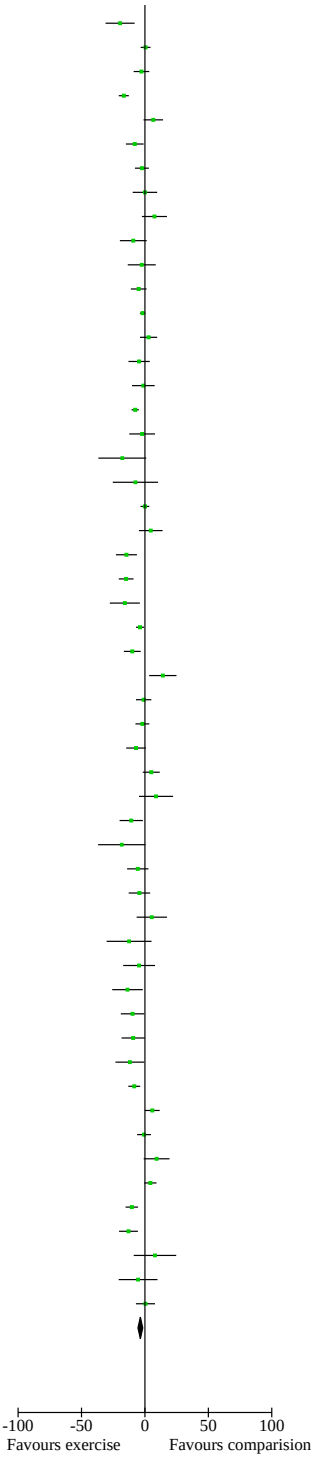
Analysis 8.2. (Continued)

Chatzitheodorou 2007	40	10.83	10	59.58	15	10	1.2%	-19.58 [-31.05 , -8.11]
Chown 2008 (5)	20.8	8.4	35	20.3	8	35	2.0%	0.50 [-3.34 , 4.34]
Cruz-Diaz 2016 (3)	16.55	15.78	53	19.29	15.78	48	1.7%	-2.74 [-8.90 , 3.42]
Cruz-Diaz 2018	20.83	8.18	32	37.5	7.74	30	1.9%	-16.67 [-20.63 , -12.71]
Djavid 2007 (4)	48.2	10.4	19	41.6	8.8	8	1.6%	6.60 [-1.08 , 14.28]
Djavid 2007 (4)	33.6	7.4	18	41.6	8.8	8	1.6%	-8.00 [-14.99 , -1.01]
Dougherty 2014	28.14	20.19	89	30.51	17.42	92	1.8%	-2.37 [-7.87 , 3.13]
Ferreira 2007 (4)	32.92	23.75	73	32.92	25	38	1.4%	0.00 [-9.64 , 9.64]
Ferreira 2007 (4)	40.42	26.25	74	32.92	25	39	1.3%	7.50 [-2.37 , 17.37]
Fransoo 2006 (2)	4.58	13.92	20	13.75	13.92	10	1.3%	-9.17 [-19.74 , 1.40]
Fransoo 2006 (2)	11.25	13.92	19	13.75	13.92	9	1.2%	-2.50 [-13.54 , 8.54]
Frost 1995	16.3	10.3	29	21.2	14.2	32	1.7%	-4.90 [-11.09 , 1.29]
Frost 2004	18.47	9.34	144	20.27	9.29	142	2.1%	-1.80 [-3.96 , 0.36]
Goldby 2006 (5)	31	17.07	78	28.1	17.34	37	1.7%	2.90 [-3.85 , 9.65]
Gur 2003 (4)	22.91	13.33	25	27.5	12.08	13	1.5%	-4.59 [-12.98 , 3.80]
Gur 2003 (4)	26.25	14.58	25	27.5	12.08	12	1.4%	-1.25 [-10.16 , 7.66]
Kim 2017	21.27	3.37	15	28.93	4.61	15	2.0%	-7.66 [-10.55 , -4.77]
Little 2014 (3)	23.16	13.92	14	25.37	13.92	15	1.3%	-2.21 [-12.35 , 7.93]
Longo 2016	16.08	14.92	7	33.92	20.62	7	0.7%	-17.84 [-36.69 , 1.01]
Machado 2007	50	22.92	17	57.5	23.96	11	0.7%	-7.50 [-25.37 , 10.37]
Maul 2005	25	8.33	74	25	12.5	74	2.0%	0.00 [-3.42 , 3.42]
Michalsen 2016	26.67	20.42	36	22.08	18.75	32	1.4%	4.59 [-4.72 , 13.90]
Miyamoto 2013	15	14.17	43	29.58	23.75	43	1.5%	-14.58 [-22.85 , -6.31]
Murtezani 2011	15.8	12.7	50	30.6	16.9	51	1.8%	-14.80 [-20.62 , -8.98]
Natour 2015	28.29	22.25	30	44.12	24.5	30	1.2%	-15.83 [-27.67 , -3.99]
Niemisto 2003	14.7	11.6	96	18.6	11	100	2.0%	-3.90 [-7.07 , -0.73]
Noori 2011	21.41	10.48	17	31.38	8.06	13	1.7%	-9.97 [-16.60 , -3.34]
Okafor 2012 (6)	25	20.8	15	10.9	4.46	15	1.3%	14.10 [3.33 , 24.87]
Rantonen 2012 (4)	15	14	43	16	10	20	1.7%	-1.00 [-7.06 , 5.06]
Rantonen 2012 (4)	14	11	43	16	10	20	1.8%	-2.00 [-7.48 , 3.48]
Rasmussen-Barr 2003	6	5.88	17	13	14.71	16	1.6%	-7.00 [-14.73 , 0.73]
Ritvanen 2007 (3)	17	13.27	28	12	13.27	33	1.7%	5.00 [-1.68 , 11.68]
Ryan 2010	26.67	17.92	15	17.92	17.5	12	1.0%	8.75 [-4.68 , 22.18]
Salavati 2016	21.13	11.82	20	31.93	17.26	20	1.4%	-10.80 [-19.97 , -1.63]
Saper 2009	35.67	30.01	15	53.94	21.32	15	0.7%	-18.27 [-36.90 , 0.36]
Saper 2017 (4)	47.83	21.3	127	53.48	21.74	32	1.5%	-5.65 [-14.04 , 2.74]
Saper 2017 (4)	49.13	22.17	129	53.48	21.74	32	1.5%	-4.35 [-12.80 , 4.10]
Schenkman 2009 (5)	15.81	10.4	17	10.4	16.64	9	1.1%	5.41 [-6.53 , 17.35]
Segal-Snir 2016	45.83	26.67	20	58.33	26.25	15	0.7%	-12.50 [-30.19 , 5.19]
Sherman 2005 (4)	22.08	18.75	36	26.67	21.67	15	1.1%	-4.59 [-17.15 , 7.97]
Sherman 2005 (4)	12.92	15	35	26.67	21.67	15	1.1%	-13.75 [-25.79 , -1.71]
Sherman 2011 (2)	18.74	16.28	91	28.52	20.61	22	1.4%	-9.78 [-19.02 , -0.54]
Sherman 2011 (2)	19.26	17.56	92	28.52	20.61	23	1.4%	-9.26 [-18.42 , -0.10]
Sousa 2009	22.12	19.96	57	34	25.83	25	1.2%	-11.88 [-23.25 , -0.51]
Tilbrook 2011 (3)	23.75	20.31	156	32.17	20.31	136	1.9%	-8.42 [-13.09 , -3.75]
Torstensen 1998 (2)	52.7	16.6	71	46.9	13.1	34	1.8%	5.80 [-0.06 , 11.66]
Torstensen 1998 (2)	46.2	13.1	70	46.9	13.1	33	1.8%	-0.70 [-6.12 , 4.72]
Ulger 2017	47	28.4	56	37.8	26.8	57	1.3%	9.20 [-0.98 , 19.38]
Underwood 2004	24.62	20.75	117	20.33	20.83	176	1.9%	4.29 [-0.57 , 9.15]
Underwood 2004	18.5	19.33	146	28.83	21.62	125	1.9%	-10.33 [-15.25 , -5.41]
Valenza 2017 (3)	16.96	13.92	27	29.92	13.92	27	1.6%	-12.96 [-20.39 , -5.53]
Vincent 2014a (4)	34.17	22.92	18	26.25	17.5	7	0.8%	7.92 [-8.82 , 24.66]
Vincent 2014a (4)	20.83	17.08	17	26.25	17.5	7	0.9%	-5.42 [-20.72 , 9.88]
Weiner 2008	33.33	19.58	50	32.92	18.75	50	1.6%	0.41 [-7.10 , 7.92]
Subtotal (95% CI)			**3325**			**2402**	**100.0%**	**-3.61 [-5.49 , -1.73]**

Heterogeneity: Tau² = 43.11; Chi² = 366.26, df = 65 (P < 0.00001); I² = 82%

Test for overall effect: Z = 3.76 (P = 0.0002)

Test for subgroup differences: Chi² = 2.44, df = 1 (P = 0.12), I² = 59.0%



Footnotes

(1) Standard deviation(s) required imputation due to missing or implausible value. Study has other comparison groups available for this meta-analysis.
(2) Standard deviation(s) required imputation due to missing or implausible value. Study has multiple exercise groups available: comparison group data divided.
(3) Standard deviation(s) required imputation due to missing or implausible value.
(4) Study has multiple exercise groups available: comparison group data divided.
(5) Study has other comparison groups available for this meta-analysis.
(6) Large baseline differences.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

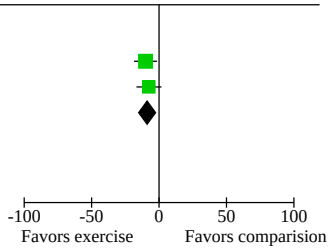
Comparison 9. SENSITIVITY ANALYSES (Placebo only comparisons)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9.1 Pain (/100): Earliest follow-up	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.1.1 Exercise vs. Placebo treatment	2	300	Mean Difference (IV, Random, 95% CI)	-8.85 [-15.12, -2.58]
9.2 Function (/100): Earliest follow-up	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.2.1 Exercise vs. Placebo treatment	3	400	Mean Difference (IV, Random, 95% CI)	-5.60 [-11.19, -0.01]

Analysis 9.1. Comparison 9: SENSITIVITY ANALYSES (Placebo only comparisons), Outcome 1: Pain (/100): Earliest follow-up

	Exercise			Comparison				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
9.1.1 Exercise vs. Placebo treatment									
Costa 2009b	46	28	77	56	26	77	54.0%	-10.00 [-18.53 , -1.47]	
Garcia 2017	39.5	27.3	73	47	29.7	73	46.0%	-7.50 [-16.75 , 1.75]	
Subtotal (95% CI)			**150**			**150**	**100.0%**	**-8.85 [-15.12 , -2.58]**	

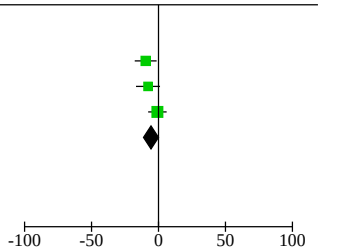
Heterogeneity: Tau² = 0.00; Chi² = 0.15, df = 1 (P = 0.70); I² = 0%
Test for overall effect: Z = 2.77 (P = 0.006)



Analysis 9.2. Comparison 9: SENSITIVITY ANALYSES (Placebo only comparisons), Outcome 2: Function (/100): Earliest follow-up

	Exercise			Comparison				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
9.2.1 Exercise vs. Placebo treatment									
Costa 2009b	40	27.08	77	49.58	24.58	77	32.0%	-9.58 [-17.75 , -1.41]	
Garcia 2017	33.21	26.33	73	41.04	28.87	73	28.1%	-7.83 [-16.79 , 1.13]	
Weiner 2008	31.67	19.17	50	32.5	15.83	50	39.9%	-0.83 [-7.72 , 6.06]	
Subtotal (95% CI)			**200**			**200**	**100.0%**	**-5.60 [-11.19 , -0.01]**	

Heterogeneity: Tau² = 8.07; Chi² = 2.98, df = 2 (P = 0.23); I² = 33%
Test for overall effect: Z = 1.96 (P = 0.05)



ADDITIONAL TABLES

Table 1. Study population and exercise treatment characteristic definitions

Characteristic	Definitions

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 1. Study population and exercise treatment characteristic definitions *(Continued)*

Population LBP duration	**Acute** = Low back pain episode less than 6 weeks in duration **Subacute** = Low back pain episode between 6 weeks and 3 months in duration **Chronic** = Low back pain episode 3 months or more in duration **Chronic recurrent** = Two episodes in a year, lasting more than 24 hours, with more than 30 days pain-free between **Mixed** = If population was mixed, we recorded the proportion of the population that had chronic low back pain
Population source	**Healthcare** = Primary care, secondary/tertiary (referred), physiotherapy, occupational therapy, etc. **Occupational** = Population selected by workplace, workers' compensation claims, etc. **General** = Population selected to represent all residents living in a geographic area (non-clinical) **Mixed** = Two or more of the above categories
Exercise type	**Strengthening** = Strengthening exercises were those that were designed to increase an individual's ability to exert or resist force using load-bearing or resistance exercises including weights, weight machines, or elastic cords. Strengthening exercises were categorised as whole-body where they involved a broad set of exercises that were not specific only to the low back region (i.e. directed at muscles connecting to the lumbar spine and/or pelvis). **Stretching** = Stretching exercises were those that intended to elongate soft tissues and poses held for varying periods of time. For the purpose of this review, stretching exercises were self-administered (but could use aids). Depending on the location and types of stretches, stretches could be classified as whole-body or back-specific or both. **Core strengthening** = Core strengthening exercises were those that were designed to improve the control, co-ordination, and strength of the core, lumbar, and pelvic region muscles. These included motor control exercises, core stability or sling exercises, and Pilates, but excluded those that were generalised (whole-body) stability exercises. We categorised core strengthening exercises as back-specific. **Flexibility/mobilising** = Neuromuscular techniques involved repeated movements within the patient's range of motion and within the patient's control. Typically, such exercises were geared toward improving the range of motion of joints related to low back pain and thereby were classified as low back-specific. **Aerobic** = Aerobic exercises were those that were designed to improve cardiorespiratory fitness and used large muscle groups, as such, we categorised it as a whole-body exercise. **Functional restoration** = Functional restoration treats chronic pain and is focussed on improving/maximizing patient function and activities through a multi-pronged strategy of psychosocial support and physical activities. We categorised it as whole-body. **McKenzie** = McKenzie involves individualised assessment, specific loading exercises, and postural training and education including managing pain through self-management. Because McKenzie was developed for and focusses on managing low back pain, we categorised it as low back-specific. **Yoga** = Yoga is a practice with many branches and styles, but which may involve physical poses designed for strengthening and stretching (asanas), controlled breathing (pranayama), and often medication (dhyana). Given the generalised physical nature of yoga practices, we categorised yoga as a whole-body exercise. **Mixed** (3 or more types) = Mixed exercises incorporated three or more types of exercise in which one did not clearly predominate/form most of the exercise intervention. Mixed exercises could be whole-body, low back-specific, or both, depending on the types and combination of exercises involved.
Exercise specificity	**Whole-body** = Exercises involved many muscle groups from around the body, generally seeking to improve overall fitness. **Back-specific** = Exercises were concentrated around muscle groups that support the back (core, trunk, abdominal, back, pelvic, upper leg, etc.). **Both** = Aspects of the intervention were focussed on both overall fitness and on the back specifically. **Unclear** = Not enough description of the exercise(s) was provided to determine what part of the body it was working on.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 1. Study population and exercise treatment characteristic definitions *(Continued)*

	Other = Exercises were not working on the body overall, nor focussed on the back, but a different specific part of the body.
Programme individualisation	**Individually designed** = Treating therapist completed a clinical history and physical examination and delivered an exercise treatment specifically designed for the individual participant. **Partially individually designed** = Exercise treatment included the same type of exercises for all participants, but varied in intensity, duration or both. **Standard design** = A fixed exercise treatment was delivered to all participants.
Delivery mode	**Independent** = Home exercises only, in which the participants met initially with therapists, then participated in the exercise treatment with no supervision or follow-up **Independent with follow-up** = Supervised home exercises in which the participants met initially with therapists, participated in the exercise treatment, and had follow-up with their therapist at least every six weeks **Group** = Group supervision, in which participants attended exercise treatment sessions with two or more participants **Individual** = Individual supervision, in which participants received one-on-one intervention or supervision
Non-exercise comparisons	**Usual care/no treatment** = Participants received no intervention and received the same standard medical care they would have without being recruited to the study or participants who received a placebo or sham treatment. **Education** = Verbal or written guidance, online or in person, related to their back pain (a formal intervention – not just advice) **Electrotherapy** = e.g. TENS, laser, ultrasound, PENS **Manual therapy** = Manual therapist, physiotherapist, or other professional delivers manual therapy; these can be passive techniques, mobilisation. **Physical therapy** = Mixed physical therapy interventions, including more than one non-modality (e.g. heat + mobilisation; manual therapy + TENS + heat); does not include exercise as a main treatment component of intervention. **Psychological therapy** = A psychological intervention, delivered by a healthcare professional, including Cognitive Behavioural Therapy (psychologist, psychological therapist, psychiatrist, counsellor, etc.) **Back school** = A programme containing both educational and active training components, delivered by a therapist or other professional. Topics cover a range, including spinal anatomy, causes of LBP, ergonomic advice, exercises to deal with LBP, causes of LBP, posture, relaxation, LBP prevention. **Anti-inflammatory/analgesics** = Participants in this group were given analgesics or nonsteroidal anti-inflammatory drugs to deal with LBP symptoms. **Relaxation** = Relaxation programme/intervention delivered by a healthcare professional

LBP: low back pain
PENS: Percutaneous electrical nerve stimulation
TENS: Transcutaneous electrical nerve stimulation

Table 2. Sources of risk of bias

Bias domain	Source of bias	Possible answers
Selection	(1) Was the method of randomisation adequate?	Yes/no/unsure

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 2. Sources of risk of bias *(Continued)*

Selection	(2) Was the treatment allocation concealed?	Yes/no/unsure
Performance	(3) Was the patient blinded to the intervention?	Yes/no/unsure
Performance	(4) Was the care provider blinded to the intervention?	Yes/no/unsure
Detection	(5) Was the outcome assessor blinded to the intervention?	Yes/no/unsure
Attrition	(6) Was there missing outcome data (dropout rate described and acceptable)?	Yes/no/unsure
Performance	(7) Were all randomised participants analysed in the group to which they were allocated?	Yes/no/unsure
Reporting	(8) Are reports of the study free of suggestion of selective outcome reporting?	Yes/no/unsure
Selection	(9) Were the groups similar at baseline regarding the most important prognostic indicators?	Yes/no/unsure
Other (adherence)	(10) Were co-interventions avoided or similar?	Yes/no/unsure
Other (adherence)	(11) Was the compliance acceptable in all groups?	Yes/no/unsure
Detection	(12) Was the timing of the outcome assessment similar in all groups?	Yes/no/unsure
Other	(13) Are other sources of potential bias unlikely?	Yes/no/unsure

See Furlan 2015

Table 3. Criteria for a judgement of ''yes'' for the sources of risk of bias

1	A random (unpredictable) assignment sequence. Examples of adequate methods are coin toss (for studies with two groups), rolling a dice (for studies with two or more groups), drawing of balls of different colours, drawing of ballots with the study group labels from a dark bag, computer-generated random sequence, pre-ordered sealed envelopes, sequentially-ordered vials, telephone call to a central office, and preordered list of treatment assignments. Examples of inadequate methods are: alternation, birth date, social insurance/security number, date in which they are invited to participate in the study, and hospital registration number.
2	Assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the persons included in the trial and has no influence on the assignment sequence or on the decision about eligibility of the patient.
3	Index and control groups are indistinguishable for the patients or if the success of blinding was tested among the patients and it was successful
4	Index and control groups are indistinguishable for the care providers or if the success of blinding was tested among the care providers and it was successful
5	Adequacy of blinding should be assessed for each primary outcome separately. This item should be scored ''yes'' if the success of blinding was tested among the outcome assessors and it was successful or: - for patient-reported outcomes in which the patient is the outcome assessor (e.g. pain, functional limitations): the blinding procedure is adequate for outcome assessors if participant blinding is scored ''yes'' - for outcome criteria assessed during scheduled visit and that supposes a contact between participants and outcome assessors (e.g. clinical examination): the blinding procedure is adequate if pa-

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 3. Criteria for a judgement of ''yes'' for the sources of risk of bias *(Continued)*

	tients are blinded, and the treatment or adverse effects of the treatment cannot be noticed during clinical examination - for outcome criteria that do not suppose a contact with participants (e.g. radiography, magnetic resonance imaging): the blinding procedure is adequate if the treatment or adverse effects of the treatment cannot be noticed when assessing the main outcome - for outcome criteria that are clinical or therapeutic events that will be determined by the interaction between patients and care providers (e.g. co-interventions, hospitalisation length, treatment failure), in which the care provider is the outcome assessor: the blinding procedure is adequate for outcome assessors if item ''4'' (caregivers) is scored ''yes'' - for outcome criteria that are assessed from data of the medical forms: the blinding procedure is adequate if the treatment or adverse effects of the treatment cannot be noticed from the extracted data
6	The number of participants who were included in the study but did not complete the observation period or were not included in the analysis must be described and reasons given. If the percentage of withdrawals and dropouts does not exceed 20% for short-term follow-up and 30% for long-term follow-up and does not lead to substantial bias, a ''yes'' is scored. (N.B. these percentages are arbitrary, not supported by literature)
7	All randomised patients are reported/analysed in the group they were allocated to by randomisation for the most important moments of effect measurement (minus missing values) irrespective of noncompliance and co-interventions.
8	All the results from all prespecified outcomes have been adequately reported in the published report of the trial. This information is either obtained by comparing the protocol and the report, or in the absence of the protocol, assessing that the published report includes enough information to make this judgement.
9	Groups have to be similar at baseline regarding demographic factors, duration and severity of complaints, percentage of patients with radicular symptoms, and value of main outcome measure(s).
10	If there were no co-interventions or they were similar between the index and control groups
11	The reviewer determines if the compliance with the interventions is acceptable, based on the reported intensity, duration, number and frequency of sessions for both the index intervention and control intervention(s). For example, physiotherapy treatment is usually administered for several sessions; therefore it is necessary to assess how many sessions each patient attended. For single-session interventions (e.g. surgery), this item is irrelevant.
12	Timing of outcome assessment should be identical for all intervention groups and for all primary outcome measures.
13	Other types of biases. For example: - When the outcome measures were not valid. There should be evidence from a previous or present scientific study that the primary outcome can be considered valid in the context of the present. - Industry-sponsored trials. The conflict of interest (COI) statement should explicitly state that the researchers have had full possession of the trial process from planning to reporting without funders with potential COI having any possibility of interfering in the process. If, for example, the statistical analyses have been done by a funder with a potential COI, usually ''unsure'' is scored.

See Furlan 2015
COI: Conflict of interest

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

APPENDICES

Appendix 1. Search strategies

WILEY CENTRAL

Last searched 7 Dec 2020

1	MeSH descriptor: [Back Pain] explode all trees
2	dorsalgia
3	backache or "back pain"
4	lumb* pain
5	coccyx or coccydynia or sciatica or spondylosis
6	MeSH descriptor: [Sciatic Neuropathy] explode all trees
7	MeSH descriptor: [Spine] explode all trees
8	MeSH descriptor: [Spinal Diseases] explode all trees
9	lumbago
10	(disc or disk) NEAR/1 herniat*
11	spinal fusion
12	facet NEAR/1 joint*
13	MeSH descriptor: [Intervertebral Disc] explode all trees
14	postlaminectomy
15	arachnoiditis
16	failed NEAR/1 back
17	MeSH descriptor: [Cauda Equina] explode all trees
18	lumb* near/1 vertebra*
19	spinal near/1 stenosis
20	slipped near/1 (disc* or disk*)
21	degenerat* near/1 (disc* or disk*)
22	stenosis near/1 (spine or root or spinal)
23	displace* near/1 (disc* or disk*)
24	prolap* near/1 (disc* or disk*)
25	MeSH descriptor: [Low Back Pain] explode all trees

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

26	#1 OR #2 OR #3 OR #4 OR #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 #25
27	exercis*
28	MeSH descriptor: [Exercise] explode all trees
29	MeSH descriptor: [Exercise Therapy] explode all trees
30	MeSH descriptor: [Exercise Movement Techniques] explode all trees
31	MeSH descriptor: [Physical Therapy Modalities] explode all trees
32	McKenzie
33	Alexander
34	William
35	Feldenkrais
36	MeSH descriptor: [Yoga] explode all trees
37	MeSH descriptor: [Recreation] explode all trees
38	MeSH descriptor: [Physical Fitness] explode all trees
39	yoga or pilates
40	Tai Chi or Tai Ji or Taiji or Taijiquan
41	#27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #34 OR #33 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40
42	MeSH descriptor: [Alexander Disease] explode all trees
43	MeSH descriptor: [Williams Syndrome] explode all trees
44	#43 OR #42
45	#41 NOT #44
46	#45 AND #26

OVID MEDLINE

Last searched 7 Dec 2020

1	randomized controlled trial.pt.
2	controlled clinical trial.pt.
3	pragmatic clinical trial.pt.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

4	random*.ti,ab.
5	placebo.ab,ti.
6	drug therapy.fs.
7	trial.ab,ti.
8	groups.ab,ti.
9	or/1-8
10	(animals not (humans and animals)).sh.
11	9 not 10
12	dorsalgia.tw,kf.
13	exp Back Pain/
14	(backache or back pain).tw,kf.
15	(lumb* adj pain).tw,kf.
16	coccyx.tw,kf.
17	coccydynia.tw,kf.
18	sciatica.tw,kf.
19	exp sciatic neuropathy/
20	spondylosis.tw,kf.
21	lumbago.tw,kf.
22	exp low back pain/
23	or/12-22
24	exp Exercise/
25	exercis*.tw,kf.
26	exp Exercise Therapy/
27	exp Exercise Movement Techniques/
28	exp Physical Therapy Modalities/
29	McKenzie.tw,kf.
30	Alexander.tw,kf.
31	William.tw,kf.

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

32	Feldenkrais.tw,kf.
33	exp Yoga/
34	exp Recreation/
35	exp Physical Fitness/
36	(yoga or pilates).tw,kf.
37	(Tai Chi or Tai Ji or Taiji or Taijiquan).tw,kf.
38	or/24-37
39	exp Alexander Disease/
40	exp Williams Syndrome/
41	39 or 40
42	38 not 41
43	11 and 23 and 42

Embase

Last searched 7 Dec 2020

1	'randomized controlled trial'/de
2	'controlled clinical trial'/exp
3	'controlled study'/de
4	'double blind procedure'/de
5	'single blind procedure'/de
6	'crossover procedure'/de
7	'placebo'/de
8	'randomization'/de
9	random*:ti,ab
10	placebo$:ti,ab
11	allocat*:ti,ab
12	assign*:ti,ab
13	blind*:ti,ab

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

14	('cross-over' OR crossover):ti,ab
15	(compare OR compared OR comparing OR comparison OR comparative):ti,ab
16	(controlled NEAR/7 (study OR design OR trial)):ti,ab
17	((singl* OR doubl* OR trebl* OR tripl*) NEAR/7 (blind* OR mask*)):ti,ab
18	trial:ti,ab
19	#1 OR #2 OR #3 OR #4 OR #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
20	'animal'/exp OR 'invertebrate'/exp OR 'animal experiment'/de OR 'animal model'/de OR 'animal tissue'/de OR 'animal cell'/de OR 'nonhuman'/de
21	'human'/de OR 'normal human'/de OR 'human cell'/de
22	#20 AND #21
23	#20 NOT #22
24	#19 NOT #23
25	dorsalgia:ti,ab,kw
26	'back pain':ti,ab,kw
27	'backache'/exp
28	(lumb* NEAR/1 pain):ti,ab,kw
29	coccyx:ti,ab,kw
30	coccydynia:ti,ab,kw
31	sciatica:ti,ab,kw
32	'ischialgia'/exp
33	spondylosis:ti,ab,kw
34	lumbago:ti,ab,kw
35	'low back pain'/exp
36	'back disorder*':ti,ab,kw
37	#25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36
38	'exercise'/exp
39	exercis*:ti,ab,kw
40	'kinesiotherapy'/exp

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

41	McKenzie:ti,ab,kw
42	'Alexander technique'/exp
43	Alexander:ti,ab,kw
44	William:ti,ab,kw
45	'Feldenkrais method'/exp
46	Feldenkrais:ti,ab,kw
47	'yoga'/exp
48	yoga:ti,ab,kw
49	'fitness'/exp
50	pilates'/exp
51	'pilates'/exp
52	'Tai Chi'/exp
53	('Tai Chi' or 'Tai Ji' or Taiji or Taijiquan):ti,ab,kw
54	#38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53
55	'Alexander disease':ti,ab,kw OR 'Alexander disease'/exp
56	'williams beuren syndrome':ti,ab,kw OR 'williams beuren syndrome'/exp
57	#55 OR #56
58	#54 NOT #57
59	#24 AND #37 AND #58

CINAHL

Last searched 7 Dec 2020

39	back pain
40	"coccydynia"
41	(MH "Lumbar Vertebrae")
42	lumbar N2 vertebra
43	S36 or S37 or S38 or S39 or S40 or S41 or S42

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

44	(MH "Thoracic Vertebrae")
45	(MH "Spondylolisthesis") OR (MH "Spondylolysis")
46	"lumbago"
47	S44 or S45 or S46
48	S35 or S43 or S47
49	S28 and S48
50	(MH "Exercise+")
51	(MH "Physical Activity")
52	(MH "Physical Fitness+")
53	(MH "Physical Therapy+")
54	(MH "Therapeutic Exercise+")
55	McKenzie or Alexander or William or Feldenkrais
56	(MH "Structural-Functional-Movement Integration+")
57	(MH "Alexander Technique")
58	(MH "Feldenkrais Method")
59	("yoga") or (MH "Yoga")
60	(MH "Recreation+")
61	(MH "Pilates") or pilates
62	(MH "Tai Chi") or Tai Chi or Tai Ji or Taiji or Taijiquan
63	exercise
64	S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR S62 OR S63
65	S49 AND S64
66	(MH "Williams Syndrome")
67	"alexander disease"
68	S66 OR S67
69	S65 not S68

PsycINFO

Last searched 7 Dec 2020

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1	DE "Clinical Trials"
2	placebo#
3	control*
4	random*
5	blind*
6	DE "Treatment" OR DE "Addiction Treatment" OR DE "Adjunctive Treatment" OR DE "Adventure Therapy" OR DE "Aftercare" OR DE "Alternative Medicine" OR DE "Anxiety Management" OR DE "Behavior Modification" OR DE "Bibliotherapy" OR DE "Caregiving" OR DE "Client Transfer" OR DE "Client Treatment Matching" OR DE "Cognitive Behavior Therapy" OR DE "Cognitive Stimulation Therapy" OR DE "Cognitive Techniques" OR DE "Computer Assisted Therapy" OR DE "Counseling" OR DE "Creative Arts Therapy" OR DE "Cross Cultural Treatment" OR DE "Disease Management" OR DE "Habilitation" OR DE "Health Care Services" OR DE "Horticulture Therapy" OR DE "Hospice" OR DE "Human Potential Movement" OR DE "Human Services" OR DE "Hydrotherapy" OR DE "Institutionalization" OR DE "Integrated Services" OR DE "Interdisciplinary Treatment Approach" OR DE "Intervention" OR DE "Involuntary Treatment" OR DE "Language Therapy" OR DE "Life Sustaining Treatment" OR DE "Maintenance Therapy" OR DE "Medical Treatment (General)" OR DE "Mental Health Programs" OR DE "Milieu Therapy" OR DE "Mind Body Therapy" OR DE "Mindfulness-Based Interventions" OR DE "Movement Therapy" OR DE "Multimodal Treatment Approach" OR DE "Multisystemic Therapy" OR DE "Outpatient Treatment" OR DE "Pain Management" OR DE "Partial Hospitalization" OR DE "Personal Therapy" OR DE "Physical Treatment Methods" OR DE "Private Practice" OR DE "Psychoeducation" OR DE "Psychotherapy" OR DE "Rehabilitation" OR DE "Relaxation Therapy" OR DE "Respite Care" OR DE "Self-Help Techniques" OR DE "Sex Therapy" OR DE "Social Casework" OR DE "Sociotherapy" OR DE "Speech Therapy" OR DE "Spiritual Care" OR DE "Symptoms Based Treatment" OR DE "Therapeutic Processes" OR DE "Trauma-Informed Care" OR DE "Trauma Treatment" OR DE "Treatment Guidelines" OR DE "Treatment Outcomes" OR DE "Treatment Planning" OR DE "Video-Based Interventions"
7	S1 or S2 or S3 or S4 or S5 or S6
8	DE "Back Pain"
9	dorsalgia
10	(backache or back pain)
11	(lumb* N1 pain)
12	(low N1 back N1 pain)
13	sciatica
14	lumbago
15	DE "Spinal Nerves"
16	DE "Lumbar Spinal Cord"
17	((disc or disk) N1 degenerat*)
18	((disc or disk) N1 prolapse*)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

19	((disc or disk) N1 herniat*)
20	S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19
21	S7 AND S20
22	DE "Exercise" OR DE "Aerobic Exercise" OR DE "Weightlifting" OR DE "Yoga"
23	exercise
24	DE "Physical Fitness"
25	DE "Physical Activity"
26	DE "Physical Education"
27	DE "Movement Therapy"
28	(Feldenkrais or McKenzie)
29	(Alexander or William)
30	DE "Yoga"
31	(pilates or yoga)
32	(Tai Chi or Tai Ji or Taiji or Taijiquan)
33	S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32
34	S21 AND S33
35	williams syndrome
36	alexander disease
	-
37	DE "Williams Syndrome"
38	S35 OR S36 OR S37
39	S34 NOT S38

SportDiscus

Last searched 7 Dec 2020

1	placebo
2	clinical trials
3	SU clinical trials

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

4	SU randomized controlled trial
5	random allocation
6	single blind
7	random*
8	double blind
9	controlled clinical trial
10	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9
11	DE "BACKACHE" or DE "SCIATICA" or DE "LUMBAR vertebrae" or DE "LUMBOSACRAL region"
12	back pain
13	backache or sciatica or dorsalgia or lumbago
14	lumb* pain
15	S11 OR S12 OR S13 OR S14
16	S10 AND S15
17	DE "EXERCISE" or DE "BACK exercises" or DE "EXERCISE therapy" or DE "PHYSICAL education & training" or DE "PHYSICAL fitness"
18	exercise
19	pilates or yoga
20	DE "PILATES method"
21	DE "YOGA"
22	McKenzie or Alexander or William or Feldenkrais
23	DE "TAI chi"
24	Tai Chi or Tai Ji or Taiji or Taijiquan
25	S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24
26	S16 AND S25
27	SU Alexander Disease
28	SU Williams syndrome
29	SU "alexander disease" or "williams syndrome"
30	"alexander disease" or "williams syndrome"
31	S27 OR S28 OR S29 OR S30

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

32	S26 NOT S31

PEDro

Last searched 6 Aug 2019

1	Therapy: Fitness training
	AND
2	Problem: pain
	AND
3	Body part: Lumbar spine, sacroiliac joint or pelvis
	AND
4	Method: blank
	AND
5	New Records added since: 18/01/2019

ClinicalTrials.gov

Last searched 6 Aug 2019

1	Other terms: "back pain" AND exercise ; "back pain" AND (yoga OR pilates OR tai chi)
2	Limit to intervention studies

First posted from 01/18/2019 to 08/06/2019

WHO ICTRP

Last searched 6 Aug 2019. Selected studies from 2019

1	back pain AND exercise
2	back pain AND yoga
3	back pain AND pilates
4	back pain AND tai chi

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CBN trials register in CRS web

Last searched 6 Aug 2019

1	(back pain AND (exercise OR pilates OR yoga OR tai chi)) AND INREGISTER 1433
2	2019:YR AND INREGISTER 149
3	#1 AND #2 20

Appendix 2. Definitions of the GRADE domains used for evidence synthesis

The certainty of evidence was assessed using the GRADE Handbook framework, categorised as:

- High (⊙⊙⊙⊙): we are very confident that the true effect lies close to that of the estimate of effect.
- Moderate (⊙⊙⊙○): we are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different.
- Low (⊙⊙○○): our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of effect.
- Very low (⊙○○○): we have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

The recommended five domains were assessed to judge the certainty of evidence.

1. Limitations in study design or execution. Certainty of evidence was downgraded if studies had limitations that were likely to bias the estimate of effect. We assessed 12 items describing five types of bias: selection bias (random sequence generation, allocation concealment, group similarities at baseline), performance bias (blinding of participants, blinding of personnel or care providers, co-interventions, compliance), attrition bias (incomplete outcome data, intention-to-treat analysis), detection bias (blinding of outcome assessors, timing of outcome assessments), reporting bias (selective reporting).

We did not downgrade the certainty of evidence if 75% or more of participant information was from studies at low risk of bias and plausible bias was unlikely to seriously alter the results. We downgraded the certainty of evidence by one level if more than 25% of participant information was from studies rated as having high risk of bias and plausible bias raised serious doubt about the results. We downgraded the certainty of evidence by two levels if more than 50% of participant information was from studies rated as having high risk of bias resulting in limitations that substantially lowered our certainty of the estimate of effect.

2. Inconsistency. We judged the heterogeneity of evidence considering the variation in effect size across studies, including with exploration of heterogeneity in subgroup analyses. The I^2 statistic was used according to 'rough guidance' provided in the *Cochrane Handbook for Systematic Reviews of Interventions*. We downgraded the certainty of evidence by one level if I^2 statistic values of 50% or greater were unexplained, and by two levels if I^2 statistic values of 75% or greater were unexplained.

3. Indirectness. We judged the generalisability of evidence considering important differences in the populations, interventions and outcomes of interest available in the included studies. We did not downgrade evidence for indirectness as the broad inclusion criteria in this review resulted in satisfaction of this criteria.

4. Imprecision. We judged the extent to which confidence in the estimate of the treatment effect was adequate to support decision-making. We downgraded the certainty of evidence if wide confidence intervals around the point estimate resulted in uncertainty about treatment effect interpretation. We downgraded by one level if the point estimate and 95% confidence interval boundaries did not exclude the possibility of a clinically important and no/small treatment effect. A clinically important treatment effect was defined as 15 points difference in pain intensity and 10 points difference in functional limitations based on the observed smallest worthwhile effect for this type of treatment compared to no treatment.

5. Publication (small study) bias. We judged potential publication bias by considering the distribution of small and positive/negative studies included in our review using a funnel plot and Egger's test. We downgraded the certainty of evidence by one level if publication bias was suggested (visual inspection of the funnel plot and significant Egger's test, $P < 0.05$).

WHAT'S NEW

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Date	Event	Description
1 October 2021	Amended	Additional detail about geographic location of trials provided in text; number 'awaiting assessment' corrected in text

HISTORY

Protocol first published: Issue 4, 2012
Review first published: Issue 9, 2021

CONTRIBUTIONS OF AUTHORS

Study design: JH, MvT, AM

Search strategy development and searching: Cochrane Back and Neck Group

Trial selection and data extraction: JH, RO, JE

Data analysis: JH, JE

Manuscript development: JH, RO, JE

Critical revision of the manuscript for important intellectual content: JH, RO, JE, AM, MvT

Final approval of the manuscript: JH, RO, JE, AM, MvT

DECLARATIONS OF INTEREST

Jill Hayden has received peer-reviewed funding from the Nova Scotia Health Research Foundation (now Research Nova Scotia), the Canadian Institutes of Health Research, Cochrane Strategic Methods Fund and Cochrane Methods Innovation Fund, the QEII Foundation, and the Nova Scotia Health Authority. She previously held a Canadian Chiropractic Research Foundation/Dalhousie University Research Professorship in Epidemiology.

Jenna Ellis: Research funded by grant from the Canadian Institute of Health Research Nova Scotia Health Research Foundation (presently Research Nova Scotia).

Rachel Ogilvie: Canadian Institutes of Health Research, Nova Scotia Health Research Foundation (presently Research Nova Scotia)

Antti Malmivaara: none known

Maurits van Tulder: No competing interest; all research funding comes from non-profit, governmental funding agencies (Danish Occupational Therapy Association, European Pain Federation, Netherlands Health Care Institute and the Netherlands Organisation for Health Research and Development) and all funding including travel and stay expenses were paid to the VU University Amsterdam.

Maurits van Tulder was formerly Co-ordinating Editor of the Cochrane Back and Neck Review Group; he and Antti Malmivaara are currently on the Editorial Board, and Jill Hayden is on the Associate Editorial Board. Editors are required to conduct at least one Cochrane Review, to ensure that editors are aware of the processes and commitment needed to conduct reviews. This involvement does not seem to be a source of conflict of interest in the Cochrane Back and Neck Review Group. Any editor who is a review author is excluded from editorial decisions on the review in which they are contributors.

SOURCES OF SUPPORT

Internal sources

- Cochrane Back Review Group, Canada

  Editorial and administrative support, conducting the search strategy

External sources

- Department of Community Health & Epidemiology, Dalhousie University, Canada

  Administrative support and office space
- Research Nova Scotia, Canada

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Funding support over 3 years to carry out analyses related to Cochrane Review update - meta-regression analyses and IPD meta-analysis

- Canadian Chiropractic Research Foundation/Dalhousie University, Canada

Research Professorship Award to Jill Hayden

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

· We use the language of functional limitations instead of disability.

· We use the language of exercise treatment instead of exercise therapy.

· We use the comparison group label 'no treatment' instead of 'no treatment (including placebo and sham treatment)'. This group includes no treatment, usual care and placebo comparisons.

· We reduced the minimal clinically important difference for the pain outcome from 20 points to 15 points.

· We added interpretation of smaller differences in effectiveness of exercise treatment compared to other conservative treatments as 'probably meaningful' when the 95% confidence interval was entirely on one side of the no effect line. This is relevant given similar costs, inconveniences and adverse effects for comparison treatments considered in this review.

· We created a new exercise type of ‘yoga’. The stability training category was renamed core strengthening. A mixed category was added.

· We did not contact content experts for additional trials outside our team.

· For screening studies, we included a preliminary title screen in addition to title abstract and full text.

· We did not assess the clinical relevance of trials.

· We did not assess global improvement as a primary outcome (moved to secondary outcome).

· We did not describe the secondary outcomes satisfaction with treatment and reduction in frequency of analgesic use or number of medications used due to limited and inconsistent reporting.

· We described secondary outcomes reported in trials, but did not synthesise these outcomes.

· We did not conduct a fatal flaws assessment of individual trials.

· We added a comprehensive research integrity assessment of studies.

· We added an exploration of specific types of conservative treatment comparisons.

· We used an approach to assess risk of bias recommended by the Cochrane Back & Neck group as planned in the protocol; however, we used an approach modified from the new Cochrane RoB 2 tool for the most recent set of studies assessed after release of this new tool.

· The meta-analyses were restricted to VAS and NRS for pain outcome measures and did not include the McGill pain score.

· We added a sensitivity analysis to explore improbable or outlying effect sizes.

· We added a post hoc sensitivity analysis restricting to only placebo-controlled trials.

· The related IPD meta-analysis mentioned in the protocol was published separately (Hayden 2020b).

· We did not investigate the treatment effect for specific types of exercise and characteristics in this publication due to the extensive amount of information available and presented; a separate publication presents results of network meta-regression analyses (Hayden 2021b).

INDEX TERMS

Medical Subject Headings (MeSH)

Exercise; Exercise Therapy; *Low Back Pain [therapy]; Mind-Body Therapies; *Musculoskeletal Manipulations

MeSH check words

Adult; Female; Humans

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.