標題：透過動作控制及核心肌群訓練課程改善慢性下背痛患者之症狀

投稿者：張堯舜／臺北市立萬芳醫院物理治療

前言：

根據中央健康保險局統計，在門、住診合計醫療點數前 10 名的疾病中，下背痛排名第 6，可見國人有慢性下背痛困擾比例之高。大部分慢性下背痛患者至各醫療院所的復健科或骨科就診並接受物理治療時仍以腰椎牽引為主要之治療模式。然而 Wegner 等學者於 2013 年所發表探討腰椎牽引對治療下背痛之效果的系統性回顧文獻結論顯示，腰椎牽引對治療下背痛並無顯著效果。

實證文獻內容：

為探究不同的治療方式以提升慢性下背痛患者之治療成效，我們參考了 Cochrane Database of Systematic Reviews 資料庫中，Geneen LJ 等學者於 2017 年所發表探討運動對成年人慢性疼痛之效益的回顧文獻，文中回顧了 21 篇對慢性疼痛進行運動介入的系統性回顧文獻，其中慢性下背痛佔了 3 篇。該文獻歸納出慢性下背痛的運動介入方式以肌力訓練、動作控制訓練(核心肌群)為主，單次訓練時間 40 至 60 分鐘，訓練頻率大多為一周 2 次，訓練時程介於單堂課程至 30 個月。成效評估包含:疼痛程度、身體活動功能及心理健康。

作者結論表示，透過運動訓練減低慢性疼痛患者疼痛程度方面的證據力是有限的，但對於改善身體活動功能、心理健康及生活品質仍有部分的證據力，且重要的是運動訓練對受試者皆沒有造成傷害。

實證文獻應用:

我們參考該文獻所建議的運動計畫，對慢性下背痛患者建構了一系列的肌力與動作控制訓練課程。此課程共招募 12 位患者參與，每堂課 60 分鐘，每周 2 堂，為期八周共 16 堂課，由物理治療師進行指導教學。並於課程結束後進行滿意度、疼痛指數、生活功能之問卷調查。

課程內容以訓練核心肌群的肌力及骨盆、腰椎的動作控制為主，包含:腹式呼吸訓練、骨盆動作控制、深層腹肌誘發訓練、橋式(bridging exercise)、貓和駱駝(cat and camel)等。運動強度隨著課程進展漸進式增加，且在學員不會疼痛的動作範圍內進行訓練。

經過 16 堂的訓練後，參加學員的疼痛指數(NAS)平均降低了 2.25±0.97，整體生活改善程度為 4.17±0.69 分(滿分 5 分)，且對課程整體滿意度為 4.86±0.12 分(滿分 5 分)。

呼吸訓練

骨盆動作控制

貓與駱駝

橋式

深層腹肌誘發

推廣價值：

我們實行了 Cochrane 實證文獻的建議，在慢性下背痛患者進行動作控制及核心肌群的訓練後，確實改善患者們的疼痛程度及生活功能，且患者們對課程也相當滿意。因此我們將持續於臨床推廣此運動訓練計畫，提供慢性下背痛患者另一種安全且有效的治療選擇。

參考文獻：

1. 主要參考文獻:

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database of Systematic Reviews 2017, Issue 4. Art. No.: CD011279. DOI: 10.1002/14651858.CD011279.pub3.

2. 佐證文獻

Wegner, I., Widyahening, I.S., van Tulder, M.W., Blomberg, S.E.I., de Vet, H.C.W., Brønfort, G. et al, Traction for low－back pain with or without sciatica. Cochrane Database Syst Rev. 2013;8：Article No.：CD003010.

Cochrane Library

Cochrane Database of Systematic Reviews

Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews (Review)

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH

Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH.
Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews.
Cochrane Database of Systematic Reviews 2017, Issue 4. Art. No.: CD011279.
DOI: 10.1002/14651858.CD011279.pub3.

www.cochranelibrary.com

TABLE OF CONTENTS

[Overview of Reviews]

Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews

Louise J Geneen[1], R Andrew Moore[2], Clare Clarke[3], Denis Martin[4], Lesley A Colvin[5], Blair H Smith[1]

[1]Division of Population Health Sciences, University of Dundee, Dundee, UK. [2]Pain Research and Nuffield Department of Clinical Neurosciences (Nuffield Division of Anaesthetics), University of Oxford, Oxford, UK. [3]Division of Population Health Sciences, University of Dundee, Dundee, UK. [4]Institute of Health and Social Care, Teesside University, Middlesbrough, UK. [5]Anaesthesia & Pain Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK

Contact address: Louise J Geneen, Division of Population Health Sciences, University of Dundee, Dundee, UK. louise.geneen@dph.ox.ac.uk.

Editorial group: Cochrane Pain, Palliative and Supportive Care Group.
Publication status and date: Edited (no change to conclusions), published in Issue 4, 2017.

Citation: Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. *Cochrane Database of Systematic Reviews* 2017, Issue 4. Art. No.: CD011279. DOI: 10.1002/14651858.CD011279.pub3.

ABSTRACT

Background

Chronic pain is defined as pain lasting beyond normal tissue healing time, generally taken to be 12 weeks. It contributes to disability, anxiety, depression, sleep disturbances, poor quality of life, and healthcare costs. Chronic pain has a weighted mean prevalence in adults of 20%.

For many years, the treatment choice for chronic pain included recommendations for rest and inactivity. However, exercise may have specific benefits in reducing the severity of chronic pain, as well as more general benefits associated with improved overall physical and mental health, and physical functioning.

Physical activity and exercise programmes are increasingly being promoted and offered in various healthcare systems, and for a variety of chronic pain conditions. It is therefore important at this stage to establish the efficacy and safety of these programmes, and furthermore to address the critical factors that determine their success or failure.

Objectives

To provide an overview of Cochrane Reviews of adults with chronic pain to determine (1) the effectiveness of different physical activity and exercise interventions in reducing pain severity and its impact on function, quality of life, and healthcare use; and (2) the evidence for any adverse effects or harm associated with physical activity and exercise interventions.

Methods

We searched the *Cochrane Database of Systematic Reviews* (CDSR) on the Cochrane Library (CDSR 2016, Issue 1) for systematic reviews of randomised controlled trials (RCTs), after which we tracked any included reviews for updates, and tracked protocols in case of full review publication until an arbitrary cut-off date of 21 March 2016 (CDSR 2016, Issue 3). We assessed the methodological quality of the reviews using the AMSTAR tool, and also planned to analyse data for each painful condition based on quality of the evidence.

We extracted data for (1) self-reported pain severity, (2) physical function (objectively or subjectively measured), (3) psychological function, (4) quality of life, (5) adherence to the prescribed intervention, (6) healthcare use/attendance, (7) adverse events, and (8) death.

Due to the limited data available, we were unable to directly compare and analyse interventions, and have instead reported the evidence qualitatively.

Main results

We included 21 reviews with 381 included studies and 37,143 participants. Of these, 264 studies (19,642 participants) examined exercise versus no exercise/minimal intervention in adults with chronic pain and were used in the qualitative analysis.

Pain conditions included rheumatoid arthritis, osteoarthritis, fibromyalgia, low back pain, intermittent claudication, dysmenorrhoea, mechanical neck disorder, spinal cord injury, postpolio syndrome, and patellofemoral pain. None of the reviews assessed 'chronic pain' or 'chronic widespread pain' as a general term or specific condition. Interventions included aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi.

Reviews were well performed and reported (based on AMSTAR), and included studies had acceptable risk of bias (with inadequate reporting of attrition and reporting biases). However the quality of evidence was low due to participant numbers (most included studies had fewer than 50 participants in total), length of intervention and follow-up (rarely assessed beyond three to six months). We pooled the results from relevant reviews where appropriate, though results should be interpreted with caution due to the low quality evidence.

Pain severity: several reviews noted favourable results from exercise: only three reviews that reported pain severity found no statistically significant changes in usual or mean pain from any intervention. However, results were inconsistent across interventions and follow-up, as exercise did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point.

Physical function: was the most commonly reported outcome measure. Physical function was significantly improved as a result of the intervention in 14 reviews, though even these statistically significant results had only small-to-moderate effect sizes (only one review reported large effect sizes).

Psychological function and quality of life: had variable results: results were either favourable to exercise (generally small and moderate effect size, with two reviews reporting significant, large effect sizes for quality of life), or showed no difference between groups. There were no negative effects.

Adherence to the prescribed intervention: could not be assessed in any review. However, risk of withdrawal/dropout was slightly higher in the exercising group (82.8/1000 participants versus 81/1000 participants), though the group difference was non-significant.

Healthcare use/attendance: was not reported in any review.

Adverse events, potential harm, and death: only 25% of included studies (across 18 reviews) actively reported adverse events. Based on the available evidence, most adverse events were increased soreness or muscle pain, which reportedly subsided after a few weeks of the intervention. Only one review reported death separately to other adverse events: the intervention was protective against death (based on the available evidence), though did not reach statistical significance.

Authors' conclusions

The quality of the evidence examining physical activity and exercise for chronic pain is low. This is largely due to small sample sizes and potentially underpowered studies. A number of studies had adequately long interventions, but planned follow-up was limited to less than one year in all but six reviews.

There were some favourable effects in reduction in pain severity and improved physical function, though these were mostly of small-to-moderate effect, and were not consistent across the reviews. There were variable effects for psychological function and quality of life.

The available evidence suggests physical activity and exercise is an intervention with few adverse events that may improve pain severity and physical function, and consequent quality of life. However, further research is required and should focus on increasing participant numbers, including participants with a broader spectrum of pain severity, and lengthening both the intervention itself, and the follow-up period.

PLAIN LANGUAGE SUMMARY

Physical activity and exercise for chronic pain in adults - an overview of Cochrane Reviews

Background

Chronic (long-term) pain is pain that has lasted beyond the body's usual healing time. It is often described as pain that has lasted for at least three months. Chronic pain causes many problems, beyond the pain itself, including fatigue, anxiety, depression, and a poor quality of life.

In the past, people with chronic pain were told to rest. However, general advice now is to keep active - whether to affect the pain directly or to combat the other problems associated with it. Therefore, research studies have attempted to examine the effect of physical activity in people with chronic pain.

This overview aimed to bring together and analyse any reviews published by Cochrane that looked at physical activity and exercise studies in any chronic pain condition, including arthritis, back and neck pain, and menstrual (period) pain.

Key results and quality of the evidence

In January 2016, we identified 21 Cochrane Reviews which covered 10 different diagnoses (osteoarthritis (a joint disease), rheumatoid arthritis (joint pain and swelling), fibromyalgia (widespread pain condition), low back pain, intermittent claudication (cramping pain in the legs), dysmenorrhoea (period pain), mechanical neck disorders (neck pain), spinal cord injury, postpolio syndrome (a condition occurring in people who have had polio), patellofemoral pain (pain at the front of the knee)). The physical activity or exercise programme used in the trials ranged in frequency, intensity, and type, including land- and water-based activities, those focusing on building strength, endurance, flexibility and range of motion, and muscle activation exercises.

The quality of the evidence was low. This was mostly due to the small numbers of people with chronic pain who participated in each reviewed study. Ideally, a study should have hundreds of people assigned to each group, whereas most of the studies included in the review process here had fewer than 50 people in total.

There was evidence that physical activity reduced the severity of pain, improved physical function, and had a variable effect on both psychological function and quality of life. However, these results were not found in all studies. The inconsistency could be due to the quality of the studies or because of the mix of different types of physical activity tested in the studies. Additionally, participants had predominantly mild-to-moderate pain, not moderate-to-severe pain.

Conclusions

According to the available evidence (only 25% of included studies reported on possible harm or injury from the intervention), physical activity did not cause harm. Muscle soreness that sometimes occurs with starting a new exercise subsided as the participants adapted to the new activities. This is important as it shows physical activity in general is acceptable and unlikely to cause harm in people with chronic pain, many of whom may have previously feared it would increase their pain further.

Future studies should focus on increasing participant numbers, including a wider range of severity of pain (more people with more severe pain), and lengthening both the intervention (exercise programme) itself, and the follow-up period. This pain is chronic in nature, and so a long-term intervention, with longer periods of recovery or follow-up, may be more effective.

BACKGROUND

Description of the condition

Chronic pain has been defined as pain lasting beyond normal tissue healing time, generally taken to be 12 weeks (International Association for the Study of Chronic Pain; Merskey 2011). It contributes to disability, anxiety and depression, sleep disturbances, poor quality of life, and healthcare costs (Leadley 2014; Moore 2014a; Park 2012).

Chronic pain has a weighted mean prevalence in adults of 20% (Breivik 2006; Moore 2014a), which increases as the population ages (32% of adults aged 25 to 34 years, 62% of adults over 75

years; Abdulla 2013; Elliott 1999). This is a greater proportion than people with asthma (To 2012) or diabetes (IDF 2012) in the same population (van Hecke 2013a). The World Health Organization (WHO) recognises chronic pain as a public health problem throughout the world, with one systematic review assessing the growing evidence that the prevalence of chronic pain in the general population is high internationally (34% in low-income countries and 30% in high-income countries; Elzahaf 2012). Chronic painful conditions comprise four of the 10 highest ranking conditions for years lived with disability in 2013 (Vos 2015), and are responsible for considerable loss of quality of life and employment, and increased healthcare costs (Moore 2014b). Despite this, the term 'chronic pain' was only added as a MeSH term in MEDLINE in January 2012 (National Library of Medicine), highlighting the relatively small proportion of specific research dedicated to this population.

Certain factors can contribute to an increased risk of chronic pain (female gender, older age, lower socioeconomic status, geographical and cultural background, and genetics; Smith 2007; van Hecke 2013b). Other factors associated with chronic pain conditions are modifiable, such as smoking status, alcohol intake, nutrition, obesity, comorbidities, employment status and occupational factors, and physical activity level (Smith 2007; van Hecke 2013a).

A review of current issues in the treatment of chronic pain strongly suggests that health professionals traditionally focus on biomedical views of pain, utilising pharmacology first and foremost, and sometimes not addressing potential non-pharmacological approaches such as physical activity and changing attitudes towards chronic pain (Schofield 2011). Guidance often suggests that lifestyle advice is important: for example, the National Institute for Health and Care Excellence (NICE) osteoarthritis guidelines state that "exercise should be a core treatment ... irrespective of age, comorbidity, pain severity and disability. Exercise should include: local muscle strengthening [and] general aerobic fitness" (NICE 2014). Non-pharmacological treatments have been developed, investigated, and implemented, with Cochrane Reviews and protocols evaluating the available evidence for psychological, physical, and other non-medical interventions (e.g. cognitive behavioural and behavioural therapy, Eccleston 2014; Williams 2012; TENS, Nnoaham 2008; low-impact/intensity movement/exercise therapy, Wieland 2013; dietary, Straube 2015; and patient education, Engers 2008; Gross 2009). While evidence for the effectiveness of these interventions is of variable quantity and quality, the 2013 Scottish Intercollegiate Guideline Network (SIGN) guidelines on the management of chronic pain made strong recommendations on the use of exercise, based on evidence drawn from randomised controlled trials (RCTs), stating: "exercise and exercise therapies, regardless of their form, are recommended in the management of patients with chronic pain" (SIGN 2013).

Description of the interventions

Physical activity has been defined by the WHO as "any bodily movement produced by skeletal muscles that requires energy expenditure, including activities undertaken while working, playing, carrying out household chores, travelling, and engaging in recreational pursuits" (WHO 2015). WHO also states that "exercise ... is a sub-category of physical activity that is planned, structured, repetitive, and aims to improve or maintain one or more components of physical fitness" (WHO 2015).

Physical activity for health can take many different forms: it can be structured exercise, such as in classes, gym-based, or a DVD or programme performed at home; or unstructured and involve adding just a few small activities each day (activities of daily living). Physical activity and exercise can also vary in intensity, duration, and type: aerobic (such as walking) or more focused on increasing flexibility, strength, or balance. Physical activity and exercise can also be taught (or led) by another individual such as an exercise professional, or initiated and maintained through the person's own initiative and motivation.

Both physical activity and exercise can be performed on land or in the water, and can range from whole-body to localised (body site-specific) training. Most forms of exercise can also be modified to be performed where there is restricted movement (e.g. in a chair, a bed, or another assistive device).

How the intervention might work

Physical activity and exercise can be adapted for an individual, and is something people can do to help themselves. It is likely to be associated with minimal adverse effects, such as interactions with medication and potential for abuse in adults with chronic pain, when compared to pharmaceutical and surgical interventions. It is therefore an attractive option to help manage an individual's pain if the systematic reviews show benefit. However, current evidence suggests that simply giving an individual advice to exercise is insufficient to bring about significant change (SIGN 2013), and a badly prescribed intervention that does not consider the individual's conditions and present state of health and fitness, such as one that does not incorporate pacing or gradual progression, may bring about adverse events such as pain 'flare-ups', or lead to cardiac or respiratory events (American College of Sports Medicine 2007). This suggests that supervised or structured interventions may be more fruitful, though this is currently unconfirmed.

Since the 1980s, primary care physician advice for treating pain has changed, moving away from "rest", to minimising or eliminating bedrest and instead remaining active (back pain, Waddell 1987). Exercise may have specific benefits in reducing the severity of chronic pain, as well as more general benefits associated with improved overall physical and mental health, and physical functioning of people with chronic pain, as depression (Finan 2013), deconditioning (Bousema 2007), and obesity are commonly observed in these people (headache/migraine, Bigal 2012; fibromyalgia, Ursini 2011). For example, studies have revealed that a sin-

gle bout of exercise increases the production of endogenous opioids, leading to transient anti-nociception in both animals and humans, and repeated exercise produces long-lasting anti-nociception in otherwise untreated animals (Stagg 2011). Aerobic exercise is also strongly linked to weight loss (Messier 2013), which in turn has implications for the management of chronic pain as the pressure on joints is reduced. Alternatively, resistance exercise, or other forms of strength training, can improve the person's capacity to support bone and cartilage through improved musculature supporting movement around a joint, with potential to relieve stiffness (Mayer 2008) and bringing about some pain relief. Resistance training through repetitive full range-of-motion exercise around the lumbar spine (in chronic low back pain) may affect disc metabolism itself, with the possibility that the exercise programme could improve metabolic exchange in the lumbar discs and aid in repair (Mooney 2006). Training to improve balance and flexibility also has benefits as it reduces the risk of falls, and the potential for further pain or injury (Harvard 2013).

Why it is important to do this overview

If physical activity and exercise interventions are shown to effectively and safely reduce pain intensity or frequency (or both), they are likely to be a preferable alternative or adjunct therapy to pharmacological/surgical treatments for chronic pain. The interventions could promote personal involvement of individuals in the management of their pain, thus increasing self-efficacy and the ability to self-manage. In turn this could lead to an increase in overall quality of life and a consequent reduction in healthcare use. In addition, exercise is of great importance for cardiovascular (Vigorito 2014) and bone health (Sakuma 2012). Reduced physical function and consequent lack of mobility in people with chronic pain is associated with increased all-cause and cardiovascular mortality (Nüesch 2011), with other studies linking severe chronic pain to general increased all-cause mortality (Moore 2014a; Torrance 2010).

Physical activity and exercise programmes are increasingly being promoted and offered in various healthcare systems (American College of Sports Medicine (ACSM) 'Exercise is Medicine' global pledge at the Inaugural World Congress 2010) and for a variety of chronic pain conditions, including arthritis (Fransen 2014; Silva 2010), fibromyalgia (Busch 2013), and dysmenorrhoea (Brown 2010). At this stage it is important to establish the efficacy and safety of these programmes, and furthermore to address the critical factors that determine their success or failure.

It is therefore important to identify whether (and how) exercise interventions can be effectively and safely applied in people with chronic pain.

With a number of systematic reviews published by Cochrane evaluating the effectiveness of exercise in various painful conditions, it is timely and important to bring together all relevant published information to evaluate the current evidence, and identify the availability and quality of evidence-based exercise interventions. This overview will determine the extent to which the published systematic reviews have accurately assessed the evidence for exercise in chronic pain conditions/syndromes, which will help to direct future guidelines and identify current research gaps.

OBJECTIVES

To provide an overview of Cochrane Reviews of adults with chronic pain to determine (1) the effectiveness of different physical activity and exercise interventions in reducing pain severity and its impact on function, quality of life, and healthcare use; and (2) the evidence for any adverse effects or harm associated with physical activity and exercise interventions.

METHODS

Criteria for considering reviews for inclusion

We included only systematic reviews of RCTs of physical activity and exercise in participants with chronic pain, and published in the *Cochrane Database of Systematic Reviews*. The included reviews had to fulfil the following criteria:

Participants

Adults (aged 18 years and over) reporting chronic non-cancer pain, including persistent (e.g. chronic back pain, fibromyalgia) and intermittent (e.g. migraine, dysmenorrhoea) pain, for at least three months (12 weeks) in any body site.

Intervention

Reviews of RCTs assessing physical activity or exercise as the intervention (any reviews where that assessed physical activity or exercise as a stand-alone intervention). This included physical activity interventions that could be initially taught by an exercise professional, or involve periodical/ongoing supervision.

Exclusions

Interventions not deemed physical activity or exercise using the WHO definition, such as manipulation, mobilisation, or passive movement. Any multi-modal interventions were excluded if physical activity/exercise could not be assessed for effect (the effect of exercise must have been measured distinctly).

Comparison

Usual care, waiting list control, placebo/sham treatment, other treatment, or a combination of treatments (as long as the effect of exercise could be measured distinctly).

Primary outcome

- self-reported pain (severity).

This could be presented and analysed as change on a continuous scale, the proportion of participants who 'responded', or, ideally, in a dichotomised format as the proportion of participants in each group who achieved a predetermined threshold of improvement (e.g. outcome in individual participants of at least 50% pain intensity reduction, or no worse than mild pain, at the end of the trial, with at least 30% pain intensity reduction as a secondary outcome, or recovery; Moore 2013).

Secondary outcomes

- Physical function (objectively or subjectively measured).
- Psychological function.
- Quality of life.
- Adherence to the prescribed intervention.
- Healthcare use/attendance.
- Adverse events (not death).
- Death.

Reviews may not always report specifically on activity or exercise for chronic pain in adults. We anticipated two possible circumstances which might have arisen.

- A review included some interventions of interest or reported only some outcomes of interest. In this case we extracted the interventions and outcomes of interest, but we did not include interventions or outcomes outside the scope of this overview.
- Reviews occasionally included papers that included children and adults together, but the results for adults were not reported or analysed separately in the included papers or the review. In this case we made a judgement as to whether the review could be included based on the proportion of adults. Our intention was to include only those reviews where more than 80% of participants were adults.

Search methods for identification of reviews

We searched the *Cochrane Database of Systematic Reviews* (CDSR), 2016, Issue 1, on the Cochrane Library for relevant reviews using the search strategy: (*pain or migraine or headache*) and (*exercise or activity or physical*). We did not seek non-Cochrane reviews.

Data collection and analysis

Two overview authors (LG, CC) independently carried out searches and selected reviews for inclusion. Disagreements were resolved through discussion, and a third overview author (RAM) acted as arbitrator where necessary.

Two overview authors (independently carried out assessment of methodological quality (LG, CC), and extracted data (LG, RAM). Any disagreements were resolved through discussion, or involving a third overview author if necessary (DM).

One overview author (LG) tracked results of the search for the most up to date version of each review and protocol that fulfilled the inclusion criteria.

Selection of reviews

Included reviews assessed RCTs of the effects of exercise for pain management in adults (as defined by individual reviews), compared with any of the listed comparators, and included:

- a clearly defined clinical question;
- details of inclusion and exclusion criteria;
- details of databases searched and relevant search strategies;
- participant-reported pain severity (primary outcome measure);
- summary results for at least one other desired outcome.

Data extraction and management

Two overview authors (LG, RAM) independently extracted data from the included review using a standardised data extraction form and checked for agreement prior to entry into Microsoft Excel for Windows. We did not extract data from reports included in the reviews again, neither did we undertake any re-analysis of data from reviews. Data were not entered for analysis into Cochrane's statistical software due to the lack of relevant and comparable data (RevMan 2014).

We collected the following information (where available) from the reviews:

- number of included studies and participants;
- intervention (exercise or activity type) and dose (frequency/intensity);
- comparator;
- condition treated;
- time of assessment;
- duration of follow-up;
- relevant outcomes.

Where possible we extracted risk ratio (RR), number needed to treat for an additional beneficial outcome (NNTB), mean difference (MD), and standardised mean difference (SMD), and other relevant statistical data for the primary and secondary outcomes. This included:

- obtaining 50% pain relief (participant-reported);

- obtaining any other measure of 'improvement' (participant-reported);
- adverse events;
- death;
- withdrawals.

Assessment of methodological quality of included reviews

Quality of included reviews

Two overview authors (LG, CC) independently assessed each included review to see if it satisfied the criteria specified in the 'assessment of multiple systematic reviews' (AMSTAR) measurement tool (Shea 2007), for rigorous methodological quality. Arbitration by a third overview author (DM) was necessary for some fields. High quality reviews were required to fulfil each of the established AMSTAR criteria (further criteria to fulfil each field is listed in Table 1).

For each review we also planned to assess the likelihood of publication bias by calculating the number of participants in studies with zero effect (relative benefit of one) that would be needed to give an NNTB too high to be clinically relevant (Moore 2008). In this case we would have considered an NNTB of 10 or greater for the outcome of participant-reported pain relief of 30% or greater to be the cut-off for clinical relevance. This method is used as statistical tests for the presence of publication bias have been shown to be unhelpful (Thornton 2000). However, assessment of publication bias was not possible due to the lack of specificity of the populations included within the reviews, and so we were unable to extract comparable data.

Quality of evidence in included reviews

We planned to use two main indicators for the quality of evidence: all included reviews must have used only primary studies that were both randomised and double-blind, so minimising the risk of bias from these items; and all included reviews must have included only people with at least moderate pain intensity at baseline (visual analogue scale greater than 30/100, categorical rating scale greater than 1/3, and numerical rating scale greater than 3/10, Collins 1997), providing a sensitive assay of intervention efficacy. Subsequently, we planned to analyse data for each painful condition in three tiers, according to outcome and freedom from known sources of bias.

- The first tier used data meeting current best standards, where studies reported the outcome of at least 50% pain intensity reduction from baseline (where 50% was the cut-off for a dichotomous (yes/no) outcome: was a 50% reduction in pain observed?), or its equivalent, without using last observation carried forward (LOCF) or other imputation method for dropouts, reported an intention-to-treat (ITT) analysis, lasted eight or more weeks, had a parallel-group design, and had at least 200 participants (preferably at least 400) in the comparison (Moore 2010). These top-tier results were usually reported first.
- The second tier used any available data, but where one or more of these conditions were not met, for example reporting at least 30% pain intensity reduction, using LOCF or a completer analysis, lasting four to eight weeks, and where the numbers of participants were at least 200.
- A third tier of evidence related to small amounts of data (fewer than 200 participants), or short studies of less than four weeks, or where there was obvious major heterogeneity between studies, or where there were other shortcomings in allocation concealment, considerable attrition, and incomplete outcome data. For this third tier of evidence, no data synthesis was reasonable, and may have been misleading, but an indication of beneficial effects might be possible.

This overview examined the quality of all included reviews according to current best standards for reporting in pain. These included the attempt and ability of the reviews to identify studies/interventions with the maximum evidence of effectiveness, and minimum risk of bias, including the reporting of the following.

- Outcomes in trials of the proportion of participants obtaining at least 50% pain intensity reduction, or no worse than mild pain, at the end of the trial (with at least 30% pain intensity reduction as a secondary outcome). We did not consider the use of mean changes in pain scores as high quality because responses to pain interventions are not Gaussian, and few people have the mean response.
- Duration of included studies of eight weeks or longer.
- Imputation method of baseline observation carried forward (BOCF), LOCF, or worst observation carried forward (WOCF) if adverse event withdrawals were similar in active and control groups.
- At least 200 participants per treatment group in included studies, with at least two trials, as a minimum criterion for trustworthiness of any analysis. Pooled analysis of small studies may be considered good quality if at least 400 participants were involved, but we regarded these as being potentially subject to bias.

We extracted the 'Risk of bias' as assessed by the original review authors from included reviews. Counts of low risk of bias were extracted from relevant studies in the included reviews and tabulated under the following headings to evaluate the proportion of studies achieving a low risk of bias for each:

- random sequence generation (selection bias);
- allocation concealment (selection bias);
- blinding of participants and personnel (performance bias);
- blinding of outcome assessment (detection bias);
- incomplete outcome data (attrition bias);
- selective reporting (reporting bias);
- sample size;

- any other biases.

Data synthesis

Additional quantitative analyses were not required, since we only considered results from properly conducted (Cochrane) reviews. The aim was to concentrate on specific outcomes such as the proportion of participants with at least 50% pain relief, all-cause or adverse event discontinuations, or serious adverse events, and to explore how these can be compared across different treatments for the same condition. We planned to compare only like with like (where possible); for example in study duration, which can be an additional source of bias if insufficient in length (Moore 2010).

However due to the limited data available, we were unable to directly compare and analyse interventions, and have instead reported the evidence qualitatively only. We had also planned to employ subgroup analyses assessing age, condition, and intervention type/intensity, though this was not feasible using the available data from included reviews. For this reason we have also been unable to include a 'Summary of findings' table as planned and stated in the protocol.

Importantly, we have tried to highlight issues of low trial quality, inadequate size, and whether trials were truly valid for the particular condition in making between-therapy comparisons.

We approached each review with four main questions/focus, and extracted data accordingly.

- Did they report exercise versus non-exercise studies?
- Did the review or studies included in the review (or both) have low risk of bias?
- Did they have our main outcome?
- What were the actual intervention/s included in the review?

RESULTS

We included 21 reviews with 381 included studies, totalling 37,143 participants. Of these, 264 studies (19,642 participants) examined exercise versus no exercise/minimal intervention in adults with chronic pain (the focus of this overview) and so were used in the qualitative analysis.

Description of included reviews

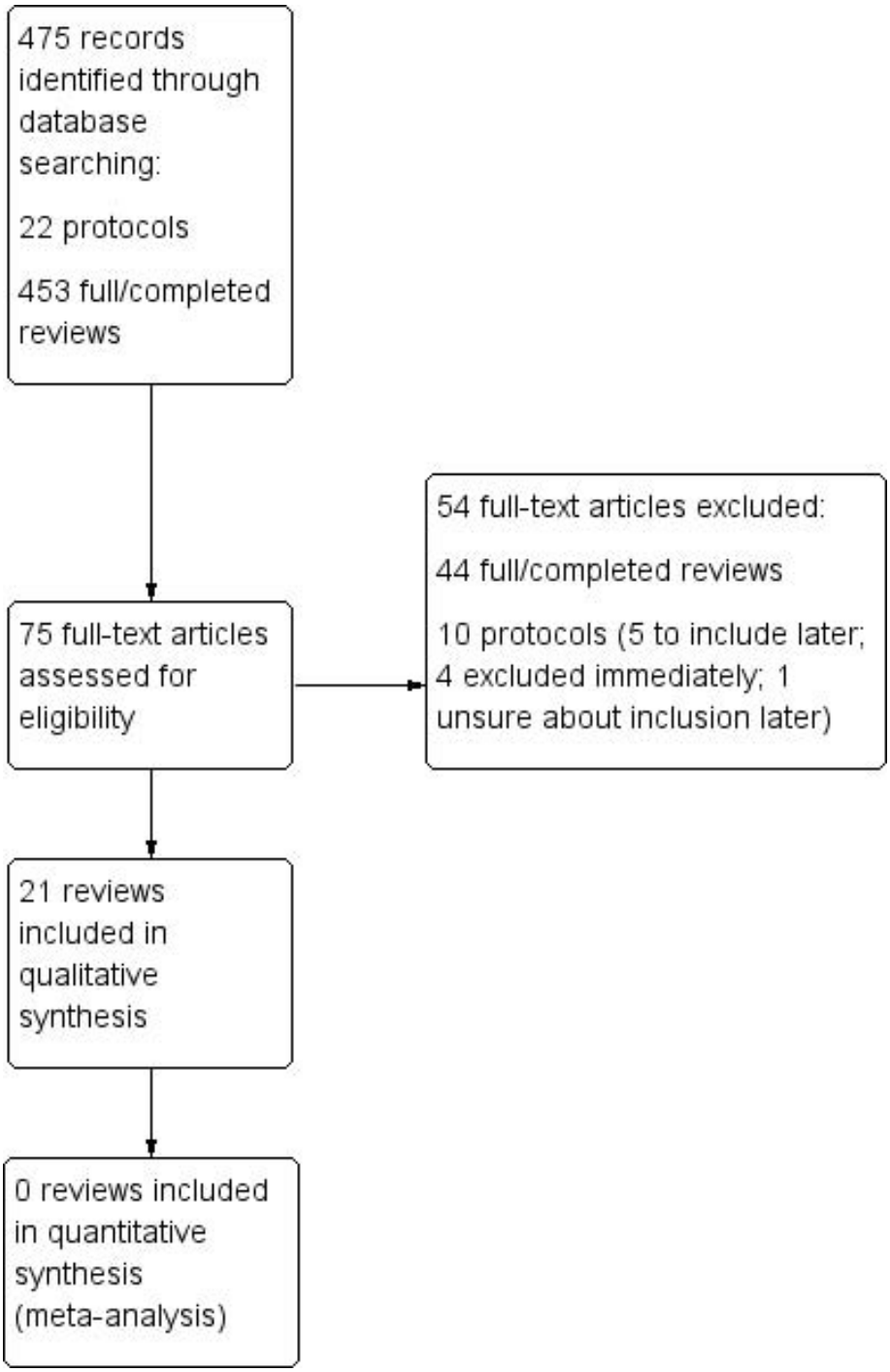
The search strategy was performed in the Cochrane Library only, and revealed 475 potentially relevant titles, of which 75 were assessed as full papers.

The search was undertaken on 31 January 2016 (CDSR 2016, Issue 1), after which any included reviews were tracked for updates, and protocols were followed in case of full review publication until 21 March 2016 (CDSR 2016, Issue 3).

All extracted data and methodological quality assessment were taken from the most recent published version of the full review.

Ultimately, of the 75 titles requiring further assessment, 10 were reviews at protocol stage only (five of which have potential to be included once published as a full review, one which was unclear, and four that were excluded based on information within the protocol). Hence, we excluded 54 titles (10 protocols and 44 full reviews; Figure 1), reasons for which are listed in Table 2.

Figure 1. Study flow diagram.



Detailed information about the included reviews is available in Table 3. Trial and participant number, age, and gender distribution is reported in Table 4.

Specificity of chronic pain condition of included reviews

Following abstract and full paper assessment, 21 reviews fulfilled the inclusion criteria: four in rheumatoid arthritis (Cramp 2013; Han 2004; Hurkmans 2009; Silva 2010), four in osteoarthritis (Bartels 2007; Fransen 2014; Fransen 2015; Regnaux 2015), three in fibromyalgia (Bidonde 2014; Busch 2007; Busch 2013), three in low back pain (Hayden 2005; Saragiotto 2016; Yamato 2015), two in intermittent claudication (Lane 2014; Lauret 2014), one in dysmenorrhoea (Brown 2010), one in mechanical neck disorder (Gross 2015a), one in spinal cord injury (Boldt 2014), one in postpolio syndrome (Koopman 2015), and one in patellofemoral pain (van der Heijden 2015). None of the included reviews assessed 'chronic pain' or 'chronic widespread pain' as a general term or specific condition.

The 21 included reviews were published by five different Cochrane Review groups: 11 from the Cochrane Musculoskeletal Group (Bartels 2007; Bidonde 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Han 2004; Hurkmans 2009; Regnaux 2015; Silva 2010); four from the Cochrane Neck and Back Group previously the Cochrane Back Group) (Gross 2015a; Hayden 2005; Saragiotto 2016; Yamato 2015); two from the Cochrane Peripheral Vascular Diseases Group (Lane 2014; Lauret 2014); one from the Cochrane Menstrual Disorders and Subfertility Group (Brown 2010); one from the Cochrane Injuries Group (Boldt 2014); one from the Cochrane Neuromuscular Group (Koopman 2015); and one from the Cochrane Bone, Joint and Muscle Trauma Group (van der Heijden 2015).

Protocols that may be included in updates of this overview focus on osteoarthritis (Østerås 2013 from the Cochrane Musculoskeletal Group), migraine (Brønfort 2015 from the Cochrane Pain, Palliative and Supportive Care Group), chronic low back pain (Hayden 2012 from the Cochrane Back Group), ankylosing spondylitis (Regnaux 2014 from the Cochrane Musculoskeletal Group), and temporomandibular disorders (Craane 2006 from the Cochrane Oral Health Group).

Exercise and physical activity interventions implemented in the included reviews

Interventions assessed included: any specified style of land-based exercise or physical activity such as one designed to improve strength, range of movement, aerobic capacity, or a combination of these (Boldt 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hurkmans 2009; Koopman 2015; Regnaux 2015; van der Heijden 2015); a single style of land-based exercise only (tai chi only: Han 2004, walking only: Lauret 2014, walking or jogging only: Brown 2010; Lane 2014, balance training only: Silva 2010, motor control exercise only: Saragiotto 2016, Pilates method only: Yamato 2015); any pool-based or aquatic therapy (Bartels 2007; Bidonde 2014; Cramp 2013), or "any exercise therapy" (Hayden 2005).

Aquatic exercise

Any exercise performed in water. This can include swimming, though many studies will be referring to exercises performed vertically in the water (not horizontally), either using the water to support the body through the exercise, or as resistance against the body.

Range of motion and flexibility exercise

Can be performed in water or on land. The intention is to increase the range of motion around a joint through progressive stretching and mobilising of the muscles around and crossing the joint. For the purposes of this overview, we only included active movement where the movement was brought about by the participant, and not passively moved by an external force such as a therapist.

Aerobic exercise

Can be performed in water or on land. Exercise usually performed continuously to raise the heart rate and breathing rate for a prolonged period. Examples include walking, jogging, running, cycling, and swimming. Often presented as a percentage of the participant's heart rate max (HRmax) - the highest heart rate reached when performing at their absolute maximum. Similarly it may be presented as a percentage of VO_2max or VO_2peak (a proportion of the maximum amount of oxygen the muscle can take up per minute), or as an absolute value (mL/kg/minute).

Strength/resistance exercise

Can be performed in water or on land. Exercise performed against a progressive resistance with the intention of improving muscle strength, muscle endurance, muscle power, or a combination of these. Resistance can come from fixed or free weights, elastic bands, body weight (against gravity), and water resistance. It may also involve static or isometric strength (holding a position or weight without moving against it). Often presented as a percentage of the participant's one repetition maximum (1-RM) - the maximum weight they can lift/move if they only have to do it once.

Motor control exercise

Can be performed in water or on land. Exercise to bring about activation of the deep trunk muscles, targeting the restoration of control and co-ordination of these 'core muscles' (Saragiotto 2016).

Balance (proprioceptive) training

Can be performed in water or on land (water may be used initially for support). Exercise emphasises the maintenance of balance during visual and perturbation challenges with eyes open or closed, range of motion, and maintaining stability over reduced areas of support and unstable surface (Silva 2010), that is improving balance in increasingly unstable situations.

Tai chi

An ancient Chinese discipline developed from martial arts, involving a continuous series of very controlled (and usually slow) movements designed to improve physical and mental wellbeing.

Yoga

Arising out of Hindu philosophy. Exercise includes breath control, simple meditation, and the adoption of specific bodily postures. It is widely practised for health, relaxation, and control (physically and mentally). Incorporates stretching and flexibility training with isometric strength training (holding certain poses, with no movement against a resistance).

Pilates

Developed by Joseph Pilates in the 20th Century, it is a system of exercises (often using special apparatus) designed to improve physical strength, flexibility, and posture, and enhance mental awareness.

Duration and dose (frequency/intensity) of the exercise and physical activity interventions

A detailed breakdown of each review can be seen in Table 5.

Duration of intervention

Interventions assessed by the included reviews varied in length from a single session (Fransen 2015) to 30 months (Fransen 2015). Only five reviews enforced a minimum intervention period to reduce risk of bias, and were able to attribute any effects to the intervention (Brown 2010; Busch 2013; Gross 2015a; Hurkmans 2009; Silva 2010).

Frequency

There was large variation in the exercise or physical activity intervention being implemented, ranging from just once a week (Bidonde 2014; Busch 2007; Fransen 2014; Fransen 2015; Han 2004; Saragiotto 2016), to twice a day (Boldt 2014), and some performing a short series of exercises (two-minute duration) every 15 minutes during the day (Gross 2015a). However, when reported, most included studies in the reviews implemented the programme twice a week (or stated at least twice a week).

Intensity

Few studies quantified the intensity of each session. Baseline intensity was often accepted as low/moderate, with the aim to progress over the intervention period to 70% to 85% of HRmax or heart rate reserve (HRR) for aerobic interventions (Brown 2010; Cramp 2013; Hurkmans 2009), 70% to 80% of an individual's 1-RM, or 50% to 70% maximum voluntary contraction (Koopman 2015) in strength/resistance training programmes (Busch 2013; Hurkmans 2009). In other reviews, intensity was described more loosely as "variable" or "low intensity (very light) to maximum effort (vigorous)" (Bidonde 2014; Fransen 2014; Lane 2014; Regnaux 2015), "low intensity" (Fransen 2014; Gross 2015a; Han 2004; Silva 2010), or "moderate or moderate-to-high" (Cramp 2013; Fransen 2015).

Duration (per session)

Individual sessions varied in length from two minutes (Gross 2015a), to 90 minutes (Busch 2013; Cramp 2013; Han 2004) or 120 minutes (Boldt 2014), but mostly situated around 45 to 60 minutes. However, it is important to note that the shorter sessions were often performed more regularly than longer sessions. With more information it would have been possible to calculate total volume of exercise or physical activity (session duration × frequency per week × number of weeks), for a more accurate and detailed analysis.

Intervention specificity for chronic pain in the included reviews

The focus of this overview was exercise versus no-exercise interventions with the intention of answering the original question: is exercise beneficial, detrimental, or ineffective for people with chronic pain when compared to inactivity? Two of the 21 reviews did not include/locate any studies that examined simply exercise versus no exercise (Lauret 2014; Silva 2010). However, many of the included reviews compared varying exercise modality, duration, intensity, and frequency. The "no-exercise" intervention referred to the control group where there was a minimal intervention (such as sham exercise or education) or wait-list control/no treatment (see Table 3 for more information on control group activity).

Time points reported

Four of the 19 reviews that reported data, reported results at a single time point only ('post-intervention': Bidonde 2014; Busch 2007; Cramp 2013; Han 2004). Reviews also analysed outcome measures immediately post-intervention and at one or more follow-up points. Each review defined short-, intermediate-, and long-term follow-up according to their own assessment, so when the time period was not mentioned explicitly, we grouped the reviews according to the review authors' own classification only, and where a time period (weeks, month, years) was explicitly listed but not defined by the authors, we grouped them as short-term (follow-up as under six months), intermediate-term (six to 12 months), and long-term (longer than 12 months): short-term: Busch 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Lane 2014; Regnaux 2015; Saragiotto 2016; intermediate-term: Bartels 2007; Fransen 2015; Gross 2015a; Hayden 2005; Lane 2014; Regnaux 2015; Saragiotto 2016; long-term: Gross 2015a; Hayden 2005; Regnaux 2015; Saragiotto 2016. Five reviews did not report "post-intervention" but at short-term, mid/intermediate-term, and long-term postrandomisation (short, mid, and long term: Boldt 2014; short and intermediate term: Koopman 2015; Yamato 2015; short and long-term: Hurkmans 2009; van der Heijden 2015). One review assessed participants in an ongoing fashion "over three menstrual cycles" (Brown 2010).

Long-term follow-up

Of the seven reviews claiming to report "long term" follow-up, one classed long-term as longer than six weeks (intermediate term as one to six weeks' follow-up) (Boldt 2014). The remaining six reviews defined long-term follow up as over 12 months (one year) post-intervention (Gross 2015a; Hayden 2005; Hurkmans 2009; Regnaux 2015; Saragiotto 2016; van der Heijden 2015).

Methodological quality of included reviews

AMSTAR quality assessment of included reviews

No review achieved a perfect score of 11/11, though five achieved 10/11 (Boldt 2014; Busch 2013; Hayden 2005; Koopman 2015; Regnaux 2015) and eight scored 9/11 (Cramp 2013; Gross 2015a; Hurkmans 2009; Lane 2014; Lauret 2014; Saragiotto 2016; van der Heijden 2015; Yamato 2015). The lowest score was 6/11 (Silva 2010) though five categories were not applicable (n/a) due to there being no included studies. Quality assessment results for each individual review are presented in Table 6.

All reviews except one (Bidonde 2014) fulfilled the basic criteria (questions one to three of Table 1); to follow an 'a priori' design as Cochrane implements a system of protocol publication before undertaking the full reviews, where it also specifies dual study selection and data extraction from a comprehensive literature search. One review did not fulfil the 'a priori' design as this was an update and separation from a broader review series, and so the criteria had not been explicitly listed prior to publication for this specific title (Bidonde 2014).

Criteria which scored badly using the AMSTAR tool were characteristics of included studies (question six of Table 1), reporting of publication bias (question 10 of Table 1), and conflict of interest declarations (question 11 of Table 1).

- Included study characteristics were limited, often reporting the "inclusion criteria" used to recruit participants in the study instead of the characteristics of actual included participants, and excluding information such as participants' age, gender split, ethnicity, and disease status.

- Assessment of publication bias was omitted entirely in five reviews (Bartels 2007; Fransen 2014; Fransen 2015; Han 2004; Hurkmans 2009), and when it was assessed, it was reported using only a simple statement (with no test values, analyses used, or diagrams to demonstrate the result; Busch 2007; Koopman 2015). Two reviews mentioned in the methods as planned analyses, though was not mentioned again (Brown 2010; van der Heijden 2015), and a third review mentioned it in the methods, but appeared to use it interchangeably with reporting bias causing great confusion (Bidonde 2014).

- Conflicts of interest were sufficiently reported in only three out of 21 of the included reviews (Hayden 2005; Koopman 2015; Silva 2010). In the remaining reviews, a cursory statement was commonly made regarding the review authors' conflicts of interests, however, fulfilling the AMSTAR criteria also requires a statement to be made regarding any conflict of interest for any of the included studies.

Risk of bias in included reviews

The original review authors assessed risk of bias (see Table 7). The table shows the number of studies assessed as low risk of bias only, and excluded those that were assessed as unclear or high risk of bias.

Selection bias (randomisation and allocation concealment)

Selection bias had the largest proportion of included studies with low risk of bias (63% and 42% of studies adequately undertaking and reporting the methods used).

Performance and detection bias (blinding participants, personnel, outcome assessors)

With any exercise or physical activity intervention it is very difficult to blind both participants and personnel to the allocation, though some studies included in reviews attempted to by offering sham exercise.

Due to the difficulty of blinding participants to their group allocation, review authors assessed the risk of bias in different ways,

which may cause confusion: whereas the majority declared this lack of possible blinding to be high risk of bias or unclear, two reviews labelled such cases as low risk of bias in order not to exclude these studies unnecessarily from their analysis (Lane 2014; Lauret 2014). Without these two reviews, only a small percentage (7.8% or 18/229) of the included studies would have scored low risk of performance bias (blinding of participants and personnel), but by including them (all 35 studies from those two reviews assessed as low risk of bias) the overall proportion of studies assessed as having low risk of bias was closer to 20% (53/264).

Attrition (incomplete outcome data, withdrawals/dropouts)

About 55% (144/264) of the studies included in these reviews showed low risk of bias.

Reporting bias (selective reporting)

Reporting bias was classed as low risk in only 46% of included studies. However, it is important to note this was not due to the remainder having high risk of bias, but instead 'unclear', as trial protocols were not always published or accessible to the review authors to accurately assess/interpret.

Study/sample/group size

Sample size was not always included within the risk of bias assessment. It was therefore extracted directly from each review's table of included study characteristics by a single overview author (LG), and assessed as being low risk of bias when there was a minimum of 50 participants per arm, or 100 in total. Numbers were then separated for the proportion of studies with greater than 100 participants per arm (or 200 in total), and 200 participants per arm (or 400 in total), as this could then be considered higher tiered evidence.

Only 26 out of 264 included studies (10%) across the 21 reviews reported over 100 participants in total (or 50 per arm), a further 6% (15/264) included over 200 participants per arm. The remaining 223 studies (84%) had fewer than 50 participants per arm (or sample size was not reported), often not reaching 50 in total.

Other bias

The format for reporting bias has changed, and therefore some earlier reviews (that are yet to be updated) did not assess bias using the same format. Others reported additional criteria as 'other bias' including the similarity of baseline characteristics, and similarity of timing points.

Interpretation of results/conclusions by original review authors

For conclusions made by the original review authors, see Table 8. We assessed whether these conclusions/interpretations of the results accurately reflected the information provided within the review, and if any further information should have been included. This final assessment of the review is an important stage in determining any author bias within the review process, as many readers, funders, and policy makers will focus on the author conclusions without a full appraisal of the actual presented data.

Eleven of the 21 reviews reported appropriate conclusions based on the data available in the context of the quality of evidence (Bidonde 2014; Boldt 2014; Busch 2007; Busch 2013; Fransen 2015; Gross 2015a; Koopman 2015; Regnaux 2015; Saragiotto 2016; Silva 2010; Yamato 2015); five reviews had appropriate conclusions, did not mention quality of the evidence in the conclusion, but did discuss it in detail earlier in the review (Bartels 2007; Cramp 2013; Han 2004; Hayden 2005; Lauret 2014); two reviews had appropriate conclusions but had only limited discussion of quality or did not adequately consider the quality of the evidence in the interpretation of the results (Hurkmans 2009; Lane 2014); and three reviews needed further comment as the strength of the conclusions were not appropriate based on the available data (Brown 2010; Fransen 2014), or we were unable to agree with their interpretation due to difficulty in extracting the data (van der Heijden 2015).

Effect of interventions

We have interpreted results using data reported in the reviews, and did not return to the original studies. Where data have been reported as MDs or as an absolute or relative change score we have used the appropriate scales (where possible) to determine whether this was clinically significant. When data have only been presented as SMD, with or without 95% confidence intervals (CI), with or without level of significance (P value), we have cautiously used the interpretation by Cohen 1988 who defined effect size using the SMD as small (SMD 0.2 to 0.5), moderate (SMD 0.5 to 0.8), or large (SMD greater than 0.8).

For the purposes of clarity, we have used the term 'intervention' to refer to the exercise or physical activity intervention, and 'control' to refer to the included comparison group which did not involve any exercise or physical activity element.

Primary outcome

Self-reported pain (severity)

Part of the inclusion criteria for this overview was for pain severity to be listed as an outcome measure.

Two of the 21 reviews did not include/identify any studies that examined intervention versus control (Lauret 2014; Silva 2010). Of the remaining reviews that did report studies examining intervention versus control (no physical activity or exercise, or minimal intervention), two did not report pain as an absolute or relative score of severity, intensity, or change as a result of the intervention (Brown 2010; Han 2004), and one review assessed pain-free time and distance during exercise (they did not assess pain using a mean/usual pain scale; Lane 2014). We could not extract relevant data for one review as they compared two different exercise interventions and a control but did not report the data compared to the control (Regnaux 2015).

The remaining 15 reviews reported a mean or usual pain score for exercise (intervention) and no-exercise (control) groups (Bartels 2007; Bidonde 2014; Boldt 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Hurkmans 2009; Koopman 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015).

Reported baseline pain score

Of the 15 reviews that were able to assess pain (Table 9), only three reviews reported actual baseline pain scores (Bidonde 2014; Boldt 2014; Hayden 2005). Three reviews reported change data (Bartels 2007; Busch 2007; Busch 2013), but we were able to use control group baseline and earliest control group scores as assumed or approximate baseline measures for the intervention groups in nine reviews (Bartels 2007; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Koopman 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015). Overall, only three reviews that assessed pain did not provide baseline or control group scores for comparison (Busch 2013; Cramp 2013; Hurkmans 2009).

Intervention group at baseline	Control group at baseline	Control group at earliest follow-up
Median pain score 70.9/100 (based on 7 studies, n = 382; Bidonde 2014)	WOMAC 9.1/20 (2 studies, n = 380) VAS ~ 55/100 (3 studies, n = 117) HAQ 1.05/3 (1 study, n = 249) (Bartels 2007)	Mean pain score ~ 29/100 (9 studies, n = 549; Fransen 2014)
11.05 to 22.6 on a 0 to 150 WUSPI score (1 study, n = 35; Boldt 2014)	VAS 35/100 to 61/100 (4 studies, n = 204; Busch 2007)	44/100 (44 studies, n = 3537; Fransen 2015)
Mean pain score 46/100 (95% CI 41 to 50) (8 studies, n = 370; Hayden 2005)	-	40/100 to 60/100 (2 studies, n = 147; Gross 2015a)
-	-	44/100 SD 24 (1 study, n = 55; Koopman 2015)
-	-	range 25/100 to 56/100 (4 studies, n = 291; Saragiotto 2016)
-	-	2.1/10 to 6.0/10 (2 studies, n = 41; van der Heijden 2015)
-	-	range 18/100 to 52/100 (6 studies, n = 148; Yamato 2015)
Range: 46 to 70.9 on a 0 to 100 scale 16 studies, n = 787	**Range: 35 to 55 on a 0 to 100 scale 10 studies, n = 950**	**Range: 18 to 60 on a 0 to 100 scale 68 studies, n = 4768**

CI: confidence interval; HAQ: Health Assessment Questionnaire; n: number of participants; SD: standard deviation; VAS: visual analogue score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; WUSPI: Wheelchair User's Shoulder Pain Index

HAQ: mean of different category scores, 0 or 1 (mild to moderate disability), up to 2 or 3 (severe to very severe disability); WOMAC pain score: 5 items summed to 0 (no pain) to 20 (worst pain ever); WUSPI: 15 items of 0 to 10 VAS scores, summed to form total

(*Continued*)

of 0 (no pain) to 150 (worst pain ever)

This suggests the majority of participants reviewed had mild-to-moderate pain (only one review reported a mean of severe pain (aquatic exercise for fibromyalgia, Bidonde 2014) at the commencement of each intervention (less than 30/100 mild pain, 30/100 to 60/100 moderate pain, more than 60/100 severe pain; Collins 1997), though labelling the majority as having only mild-to-moderate pain should be interpreted with caution due to the lack of specific data available - the baseline data of the intervention group would have been preferable to the proxies we have had to use.

Quality judgement/ tiered quality (first, second, third tier evidence)

Our assessment criteria stated that we would accept the information as graded evidence when reported as the number of participants achieving a 50% (first tier evidence) or 30% (second tier evidence) reduction in pain, but none of the included reviews reported results in this way, and so instead we used the reported absolute and relative change values.

None of the included reviews fulfilled the requirements for first tier evidence (at least 50% pain reduction from baseline, study duration longer than eight weeks, and more than 200 participants per arm).

Second tier evidence (at least 30% pain reduction from baseline, study duration between four and eight weeks, and more than 200 participants in total or 100 participants per arm) was also lacking in these reviews; three reviews found at least 30% reduction in pain from baseline (Busch 2007; Busch 2013; van der Heijden 2015), one of which also used long enough exercise programmes (eight to 21 weeks' intervention, Busch 2013) but totalled only 81 participants across two studies. The other two reviews did not fulfil the study duration criteria (interventions from 2.5 weeks, Busch 2007; and three weeks, van der Heijden 2015) or study size criteria.

Consequently results from relevant reviews have been pooled (all tier three quality) where appropriate, though results should be interpreted with caution due to the low quality evidence.

Treatment effect

Data that could be extracted for pain can be seen in Table 9 for all reviews. Only three reviews found no statistically significant changes in usual or mean pain from any intervention (Cramp 2013; Hurkmans 2009; Koopman 2015 (assumed due to lack of presented data)). The remaining reviews reported a statistically significant effect of the intervention at one or more time points, in at least one subgroup.

Three reviews found at least 30% pain reduction from baseline (post-intervention - strength training: Busch 2007; Busch 2013, at short-term follow-up: van der Heijden 2015). Additionally, seven reviews reported clinically significant results (minimally important difference: reduction in pain from baseline of at least 10 points on a 0 to 100 scale or an absolute improvement of at least 10% to 20%, Dworkin 2008) as a result of the exercise intervention (1.3/10 from aerobic training, Busch 2007; 12/100 (95% CI 10 to 15), Fransen 2015,; 14.9/100 (95% CI 7.39 to 22.40), Gross 2015a; 10.2/100 (95% CI 1.31 to 19.09), Hayden 2005; 2.5/10 (95% CI 1.52 to 3.48), Boldt 2014; 10.01/100 (95% CI 4.35 to 15.67), Saragiotto 2016; 14.05/100 (95% CI 9.19 to 18.91), Yamato 2015). Three reviews found statistically significant improvements as a result of the intervention, but they did not reach clinical significance (post-intervention, P = 0.02, Bartels 2007; "small to moderate" benefit post-intervention and at six-month follow-up, P < 0.001, Fransen 2014; "moderate effect" of 7% (95% CI 3 to 11) benefit post-intervention, Bidonde 2014).

Overall, results were inconsistent across interventions and follow-up (see Table 9), as exercise did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point.

Secondary outcomes

Physical function (objectively or subjectively measured)

Measures of physical function were the primary outcome measure in eight out of 21 reviews (Busch 2013; Han 2004; Hayden 2005; Hurkmans 2009; Koopman 2015; Lane 2014; Lauret 2014; Silva 2010), and a reported (non-primary) outcome measure in nine more reviews (Bartels 2007; Bidonde 2014; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Regnaux 2015; Saragiotto 2016; van der Heijden 2015, plus some which assessed disability; Cramp 2013; Saragiotto 2016; Yamato 2015). Only Boldt 2014 and Brown 2010 did not list physical function (or disability, or activity limitation) as a potential outcome measure.

Treatment effect

Data that could be extracted for physical function are shown in Table 10. Two reviews which reported physical function had no

data to extract (Lauret 2014; Silva 2010), and for one review we were unable to extract the relevant data (Regnaux 2015). Two reviews found no significant difference in physical function between the intervention and control groups (Han 2004; Hurkmans 2009, both rheumatoid arthritis, 8 studies, n = 240). The remaining 14 reviews showed that the intervention produced a statistically significant benefit over the control at a minimum of one reported time point (Bartels 2007; Bidonde 2014; Busch 2007; Busch 2013; Cramp 2013; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Koopman 2015; Lane 2014; Saragiotto 2016; van der Heijden 2015; Yamato 2015; 129 studies, n greater than 9559 (exact number unknown due to some participant numbers not being reported)).

Many of these statistically significant results were of small or moderate effect size (as reported by the review authors, or using the definition by Cohen 1988 if unreported; small effect size: Bartels 2007; Bidonde 2014; Fransen 2014; Fransen 2015; Gross 2015a; Koopman 2015; Saragiotto 2016; Yamato 2015, moderate effect size: Busch 2007; Fransen 2015; Yamato 2015).

Only one review reported statistical significance and large effect size (both short-term and long-term follow-up: SMD 1.10 (95% CI 0.58 to 1.63) and 1.62 (95% CI 0.31 to 2.94), van der Heijden 2015). However, the original review authors highlighted the low to very low quality of the evidence as many studies had high or unclear risk of bias across multiple domains (van der Heijden 2015).

Psychological function

Only five out of 21 reviews assessed psychological function as mental health (Bartels 2007; Bidonde 2014; Busch 2013), anxiety (Cramp 2013), and depression (Boldt 2014; Busch 2013; Cramp 2013).

Treatment effect

Data that could be extracted for psychological function can be seen in Table 11. There were significant effects in favour of the intervention for mental health (Bartels 2007) and depression (Busch 2013) scores, and "variable effect" for depression (Cramp 2013). However, there was also no effect or no differences between control and intervention groups reported for mental health (Bidonde 2014; Busch 2013), anxiety (Cramp 2013), and depression (Boldt 2014).

Quality of life

A version of quality of life assessment was reported in nine reviews. Six were termed quality of life or health-related quality of life (HRQoL) (Bartels 2007; Boldt 2014; Fransen 2014; Fransen 2015; Gross 2015a; Lauret 2014).

Other reviews assessed global perceived effect (Gross 2015a), global wellbeing (Busch 2007), global assessment (Hayden 2005), global impression of recovery (Saragiotto 2016; Yamato 2015), health assessment questionnaire (Silva 2010), multi-dimensional function (Bidonde 2014; Busch 2013), and work status (Hayden 2005). These have been reported separately to quality of life (Table 12).

Treatment effect

Data that could be extracted for quality of life can be seen in Table 12. Four reviews found no significant difference between intervention and control groups in health-related quality of life post-intervention (9 studies, n = 556) (HRQoL: Boldt 2014; Fransen 2014; Gross 2015a, global assessment: Bidonde 2014; Gross 2015a)), three reviews did not or were unable to report any data (HRQoL: Lauret 2014, global assessment: Hayden 2005, other assessment: Silva 2010), and seven reviews found a significant improvement as a result of the intervention (34 studies, n = 2700) (HRQoL: Bartels 2007, Fransen 2015, global assessment: Busch 2007; Saragiotto 2016; Yamato 2015, other assessment: Bidonde 2014; Busch 2013).

Two reviews assessing strength/resistance training interventions found significantly large effect sizes (SMD greater than 0.8, as defined by Cohen 1988) in favour of the intervention (global wellbeing measure, SMD 1.43 (95% CI 0.76 to 2.10), Busch 2007; Fibromyalgia Impact Questionnaire, SMD 1.27 (95% CI 0.72 to 1.83), Busch 2013). Other statistically significant changes reported in the included reviews were of small-to-moderate effect size (SMD 0.2 to 0.8, Cohen 1988).

Adherence to the prescribed intervention

Only one review reported adherence to the intervention as an outcome measure (Regnaux 2015), but the authors were unable to perform an analysis on attendance as most studies did not clearly report attendance or compliance (Regnaux 2015). However, five reviews assessed withdrawals or dropouts (Bidonde 2014; Fransen 2014; Han 2004; Regnaux 2015; Saragiotto 2016), one reported all-cause attrition (Busch 2013), and another reported the discontinuation rate (Silva 2010).

Data that could be extracted for adherence, withdrawals, and attrition can be seen in Table 13. Pooling all available data for withdrawals/dropout/attrition gave an RR of 1.02 (95% CI 0.94 to 1.12) in favour of the control group (6 reviews, 30 studies, n = 2256, control withdrawal 81/1000, intervention withdrawal 82.8/1000).

One clinically controlled trial (CCT) in one review reported statistically significant improvement in enjoyment of exercise/rest (P = 0.0002) and self-reported benefit from exercise/rest (P = 0.006) at both post-intervention (end of therapy, 10 weeks) and follow-up (four months later) (n = 95, Han 2004).

Healthcare use/attendance

None of the reviews reported healthcare use/attendance.

Adverse events (not death)

Eighteen out of 21 reviews reported adverse effects (three reviews did not report adverse events as an outcome measure due to lack of studies or other undisclosed reasons; Brown 2010; Lauret 2014; Silva 2010). Two reviews only assessed a specific adverse event ("amputation" Lane 2014; "motor unit survival" Koopman 2015), one review observed "safety - pain and radiological damage" (Hurkmans 2009), and another referred to any "side-effects" (Han 2004).

Data that could be extracted for adverse events (not death) can be seen in Table 14. The total number of reported adverse events (not death) was 137 events across 39 studies out of 61 studies that had adverse events as an outcome measure (over one-third of all trials that reported them found no adverse events related to the intervention): six reviews reported no adverse events from the included trials (Bartels 2007; Busch 2013; Cramp 2013; Hurkmans 2009; Koopman 2015; Yamato 2015) though the authors questioned whether this was due to lack of reporting by the trial authors, or whether there were no adverse events.

Adverse events were largely reported as a total number per trial, though one review separately reported results for the intervention group versus the control group (Saragiotto 2016), and two others reported adverse events for the intervention group only (Boldt 2014; Regnaux 2015). Only one review calculated an RR for the adverse events, showing a reduced risk for amputation in the intervention group (two amputations in the usual care/control group: RR 0.20, 95% CI 0.01 to 4.15, based on one study in one review, Lane 2014).

Death

Only one out of 21 reviews reported death separately to other adverse events (Lane 2014). Based on five studies within the review, death had an RR of 0.71 (95% CI 0.28 to 1.78) in favour of exercise as being protective, though was not statistically significant (P = 0.47).

DISCUSSION

Specificity of the condition: despite the heterogeneous nature of chronic pain, in this overview we have combined several painful conditions covering a number of conditions and diagnoses. Regardless of aetiology, the impact of chronic pain is broadly similar across many conditions.

Summary of main results

Pain severity: there were favourable results in a number of reviews as a result of exercise: only three reviews found no statistically significant changes in usual or mean pain from any intervention. However, results were inconsistent across interventions and follow-up, as the intervention did not consistently bring about a change (positive or negative) in self-reported pain scores at any single point. The exercise or physical activity interventions did not have a negative effect on the outcome (did not worsen the pain). A factor in the lack of statistical and clinically significant result may be the baseline pain severity of participants. The majority of the included population had an assumed mild-to-moderate pain severity score (assumed only due to lack of exact group data at baseline). This is often the desired outcome (post-intervention) of many drug therapies for pain, and it may therefore be difficult to show a clinically significant improvement in these people.

Physical function: physical function/disability was the most commonly reported outcome measure, and was the primary measure in eight out of the 21 reviews. Physical function was significantly (statistically) improved as a result of the intervention in 14 reviews, though even these statistically significant results had only small-to-moderate effect sizes in all but one review.

Psychological function and quality of life: there were variable results for psychological function and quality of life: results were either favourable to exercise (two reviews reporting significantly large effect sizes for quality of life), or showed no difference between groups. There were no negative effects.

Adherence to the prescribed intervention: could not be assessed in any included review. However, risk of withdrawal/dropout was slightly higher in the exercising group (82.8/1000 participants versus 81/1000 participants), though the group difference was not significant.

Healthcare use/attendance: not reported in any included review.

Adverse events, potential harm, and death: importantly, exercise caused no actual harm, with most adverse events being increased soreness or muscle pain, which reportedly subsided after several weeks of the intervention. One review reported a non-significant reduction in risk of death as a result of the intervention.

Overall completeness and applicability of evidence

Of the 21 included reviews, seven could be considered out of date as they were most recently assessed as up-to-date prior to 2010 such that any recent controlled trials assessing pain severity have not been included in this overview (Cochrane recommends updating reviews every two years) (Bartels 2007; Brown 2010; Busch 2007; Han 2004; Hayden 2005; Hurkmans 2009; Silva 2010). We included these reviews in the overview, but they may not be as relevant now due to the elapsed time since they were updated. One protocol that had potential to be included was published in

2006 with no full review available yet (Craane 2006).

Available data suggest that participants in the included reviews and studies would generally be characterised as having mild-moderate pain (moderate greater than 30/100 or 3/10) with only one review reporting moderate-severe pain (severe greater than 60/100 or 6/10). Therefore whether the evidence of change or no change seen here as a result of each intervention is applicable to people further along on the pain spectrum (with higher pain scores/worse pain) is debatable. However, it can be argued that those people are more likely to be assigned medical or surgical interventions than physical activity and exercise alone (where available), and as a group they may be less able to engage in exercise, and may therefore be more difficult to recruit into exercise-only studies. Having said this, the labelling of participants as having mild-moderate pain was a cautious one within this overview due to the lack of specific data available at baseline assessment; only three reviews included baseline pain scores in the intervention group, and two further reviews provided control group baseline scores.

There are still gaps in the available literature, and therefore also within this overview. None of the included reviews examined generalised or widespread chronic pain as a global condition, each instead examined specific conditions that included chronic pain as a symptom or result of the ongoing condition (rheumatoid arthritis, osteoarthritis, fibromyalgia, low back pain, intermittent claudication, dysmenorrhoea, mechanical neck disorder, spinal cord injury, postpolio syndrome, and patellofemoral pain). The pain in these cases can occur secondary to other symptoms such as fatigue, muscle stiffness, difficulty sleeping, and depression, all of which could separately (and more effectively) be influenced by the intervention. Additionally, only 25% of included studies actively reported adverse events. This may affect the completeness of the evidence as conclusions have been drawn based on the available data. The included reviews did not discuss the possible impact of this non-reporting by the original trials, and this may lead to underestimating possible adverse events from an intervention, or overestimating its safety.

The exercise interventions examined in the included reviews were broad; including aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi. Many of these interventions can be accessed in the community by the general public and people with chronic pain, either individually or in classes (yoga, Pilates, tai chi). Other exercise intervention programmes, such as the motor control exercise and proprioceptive (balance) training, requires at least initial supervision by a therapist to teach the correct techniques and provide feedback for progression.

Quality of the evidence

In assessing the quality of the evidence, we employed the AMSTAR tool to examine the reviews, extracted data on risk of bias to examine the available primary evidence, and evaluated the authors' conclusions to ensure that they were appropriate based on the available data.

The AMSTAR tool is useful in assessing the reporting of a systematic review, though it does not inform us of the actual undertaking or conduct of the review process. All 21 included reviews scored well across the AMSTAR assessment, though this is likely due to the stringent reporting guidelines implemented by Cochrane prior to publication. However, it may be necessary or advisable for the Cochrane guidelines to be further expanded and detailed with regards to reporting study characteristics, publication bias, and conflicts of interest, as these areas often did not meet the requirements laid out in the AMSTAR criteria (Table 1).

Data extracted from the reviews regarding their assessment of bias (risk of bias) showed moderate level scores at best across all included studies within the included reviews. Other than issues surrounding blinding (which are problematic in exercise intervention studies due to the nature of the intervention), the trials did not consistently and adequately report potential attrition and reporting biases, with less than half of studies within these reviews at low risk of bias.

However, the most prominent issue with regards to bias in these exercise and physical activity intervention studies is the sample size used. This subcategory is not used as standard in the assessment of bias in Cochrane Reviews, despite the increasing volume of research available suggesting that small studies of fewer than 100 participants per arm (Moore 2010; Nüesch 2010) are at increased risk of succumbing to the random effects in estimating both direction and magnitude of treatment effects (Moore 1998; Turner 2013) due to greater heterogeneity within and between small studies (IntHout 2015).

Studies within the included reviews here were very small (often fewer than 50 participants in total). For greater quality and a more reliable effect, at least 100 participants per arm should be analysed for a study to potentially be classed as tier two evidence (200 per arm for tier one); small studies are known to overestimate the treatment effect by up to 32% in comparison with larger studies (Deschartes 2013).

Assessing studies for risk of bias based on study size (total number or per arm) should be included in any review or meta-analysis in future, to adequately assess the influence of small trials on the estimated treatment effect (Nüesch 2010). Inclusion in the standard assessment process may in turn influence the design and undertaking of future research trials to increase the sample size, and produce more consistent clinically and statistically accurate results.

Of the 21 included reviews, 12 used a pain measure as their primary outcome (Bartels 2007; Boldt 2014; Brown 2010; Busch 2007; Fransen 2014; Fransen 2015; Gross 2015a; Hayden 2005; Regnaux 2015; Saragiotto 2016; van der Heijden 2015; Yamato 2015), and the remaining nine reviews included the measure as a secondary outcome only. Other outcomes were shared, including physical and psychological function, and quality of life. Likewise,

each review team will have included studies that did not use their chosen outcome measures as the primary measure, and that were therefore powered according to a different primary outcome. On collating the evidence, some studies may appear underpowered for the outcome(s) of interest to us (Turner 2013), yet were adequately powered for the studies' primary measure. To increase the power of the results of this overview, and the intermediary reviews we have included, intervention studies that focus on painful conditions should include pain intensity as the primary outcome, or at least as a prominent secondary outcome; alternatively review authors should seek to include only those studies that were adequately powered for pain intensity as a primary outcome measure.
Intervention length ranged from a single session to regular sessions over a period of 30 months, though the majority were between eight and 12 weeks. Durations of this length are common among exercise and physical activity intervention studies to allow for physiological adaptation and familiarisation. In contrast, the follow-up period was often inadequate, as many reviews reported only a single follow-up point (immediately post-intervention), or repeated measures over the short-term (less than six months): only six of the 21 reviews planned to assess participants over the long term (over 12 months: Gross 2015a; Hayden 2005; Hurkmans 2009; Regnaux 2015; Saragiotto 2016; van der Heijden 2015). With chronic conditions, it would be advisable to include longer follow-up periods (beyond 12 months post-randomisation) as long-term solutions may be more relevant to their control or pain management. It is also possible that initial adaptation and potential benefits as a result of an exercise intervention may take longer to manifest in comparison to a 'healthy' person due to the possible limitations in exercise intensity and progression (a training threshold) beyond which any additional physical training may be detrimental to the underlying pathophysiological mechanisms (Daenen 2015) or simply be additional physical stress with no additional physical benefit (Benton 2011).
We grouped outcome measurement points in this overview into short term (less than six months), intermediate term (six to 12 months), and long term (longer than 12 months). The broad time window for 'short term' outcomes (less than six months) is a potential source of heterogeneity as the early period is the one where time of measurement is most likely to result in variable outcomes. These initial problems could be overcome by use of standard reporting periods in exercise intervention studies (suggested four-weekly within the 'short term' period to assess both neural adaptation and other physiological changes). This would allow review authors to use the data gathered closest to the time point they are assessing, for more accurate analyses. Additionally, by extending the follow-up period beyond one year (long-term follow-up), heterogeneity may be reduced further.
Reviews generally did not enforce a minimum exercise requirement for inclusion in their review. Additionally, not all exercise sessions were supervised or baseline fitness/physical ability was assessed subjectively, and consequently it was not reported whether the intervention was fulfilled as described, or whether the dose was enough to elicit a physiological response. Studies often rely on the self-report of participants as to the actual physical activity and exercise being undertaken, which can lead to a greater risk of bias, and reduced study quality as it is questionable as to whether the effect can be truly attributed to the intervention. This was examined in a previous review, where it was concluded that non-subjective physical assessment should be performed where possible (Perruchoud 2014), though these still have challenges regarding implementation.
In summary, the quality of the evidence was low (third tier): within this overview we found no tier one or tier two evidence. This is largely due to the small sample sizes and potentially underpowered studies. A number of studies within the reviews had adequately long interventions, but planned follow-up was limited to less than one year (12 months) in all but six reviews.
Interpretation of the available data, and conclusions drawn by the review authors, were appropriate, although the conclusions were sometimes stronger than warranted by the available data. Occasionally results were not discussed with regards to the quality of the evidence or risk of bias: it is important to discuss the findings in the context of the quality of the evidence, with complete transparency, as this may affect future research, and implications for patients, funders, and policy makers.

Potential biases in the overview process

While we have attempted to include all relevant reviews in the overview process, we do concede that by only searching the Cochrane Library, and including only current Cochrane Reviews we may have missed some key literature. However previous publications have referred to the higher quality grading (high AMSTAR score) in Cochrane Reviews due to the basic criteria necessary for publication at any stage (protocol or full review) suggesting they may be the most reliable source of evidence (O'Connell 2013).

Agreements and disagreements with other studies or reviews

This is a summary overview of current Cochrane Reviews, we are not aware of any overviews or reviews summarising non-Cochrane reviews.

AUTHORS' CONCLUSIONS

There is limited evidence of improvement in pain severity as a result of exercise. There is some evidence of improved physical function and a variable effect on both psychological function and quality of life. However, results are inconsistent and the evidence is

low quality (tier three). Promisingly however, none of the physical and activity interventions assessed appeared to cause harm to the participants.

Implications for practice

For clinicians and people with chronic pain

The evidence in this overview suggests that the broad spectrum of physical activity and exercise interventions assessed here (aerobic, strength, flexibility, range of motion, and core or balance training programmes, as well as yoga, Pilates, and tai chi) are potentially beneficial, though the evidence for benefit is low quality and inconsistent. The most commonly reported adverse events were increased soreness or muscle pain, which subsided after several weeks of the intervention.

Physical activity and exercise may improve pain severity as well as physical function and quality of life.

For policy makers

The evidence showed variable results, though in some reviews there was a clinical and statistical benefit in pain relief and physical function (based on low quality evidence). The evidence suggests that physical activity or exercise is an acceptable intervention in people with chronic pain, with minimal negative adverse effects. However based on this low quality evidence, we cannot provide direction to the content of an exercise programme should clinicians decide to implement one.

Implications for research

There is a clear need for further research into exercise and physical activity for chronic pain in adults.

General implications

- Future research should report baseline values for outcome measures in both intervention and control groups, together with detailed relevant information about the participants. Knowing the baseline value is relevant to interpreting any change observed as a result of the intervention, and understanding the broader value of the intervention.
- Where possible, pain results should be reported as the number of people achieving 50%, 30%, and 10% pain relief, and the number who did not meet that point (dichotomous outcome). These are clinically important cut-offs in pain intervention research, and reporting in this way allows readers to observe the clinical effect more effectively.
- Reporting should include median and range as well as mean and standard deviation (SD) of results. This will allow readers to review the effects of any outliers that may have skewed the data, which often goes unnoticed in the reporting of mean and SD alone.
- The importance of clear intervention reporting is underestimated: often studies report both intervention and control programmes simply, where other researchers and clinicians alike are unable to replicate the trial or intervention. Recommendations for reporting are based on the Consolidated Standards of Reporting Trials (CONSORT) statement (www.consort-statement.org/), but this alone does not detail the extent of necessary intervention and control programmes reporting. The template for intervention description and replication (TIDieR) approach (Hoffman 2014) is intended as an extension to CONSORT item 5 ("The interventions for each group with sufficient details to allow replication, including how and when they were actually administered") and is a checklist for detailing the programmes using: why (rationale), what (materials and procedures), who, how, where, when, and how much.

Design

- One previous review highlighted the increased bias often present in questionnaires and other self-report measures of physical activity in people with chronic pain, and as a result made the recommendation to use objective measures instead, such as accelerometers, or the use of direct and indirect calorimetry, where possible (Perruchoud 2014), though these still have challenges regarding implementation. This would allow direct and exact comparison and analyses of actual energy expenditure and treatment effect.

Population/participants/sample

- There needs to be a focus on participants with generalised and/or widespread chronic pain, instead of (or as well as) condition-specific populations.
- Studies should include people with higher pain severity (greater than 50/100 on a 100-point visual analogue scale) at baseline. People with mild-moderate pain should still be included, but it would be advisable to separate the results for analysis, ensuring the study is adequately powered to allow this subgroup analysis in advance. This way we could determine if exercise has benefit overall, or affects one group more than another, and tailor exercise programmes according to the individual needs.
- It has been previously suggested that for 20% to 25% of participants undertaking an exercise programme there is little to no favourable response (Timmons 2014), while a small percentage (5% to 10%) have adverse events (Bouchard 2012). It is therefore vitally important that much larger sample sizes are used: ideally *more than 200 participants per arm*, though even this number in total would increase the quality of the evidence in the

first instance. In this way we may be able to learn to identify individuals who will benefit, and those who will require further intervention.

Interventions

- Different forms of exercise should be researched in detail. For the purposes of this overview, we combined all physical activity and exercise interventions under one banner to determine if there was any effect. However a number of reviews separately analysed resistance (strength) training, aerobic (endurance), and combination programmes. It is important to continue to examine different modalities, but currently there is not enough high quality evidence to exclude or prioritise one specific mode (resistance, endurance, stability) or medium (land/water based), or the proportion of a combination programme to be assigned to each, as all may have individual benefits for people with chronic pain.
- Intensity of exercise, duration of individual sessions, and frequency should be investigated. It is this dose alongside duration (of the entire intervention) and adherence that may determine the actual efficacy.
- More reviews and trials should attempt to minimise intervention heterogeneity by implementing minimum and maximum requirements. Only this way will the research community be able to determine more accurately the direction and magnitude of effect of a specific programme or intervention. Many of these important restrictions can be implemented as subgroup analyses, though if this is the case it is important to have adequate study numbers (ideally 200 participants per arm or subgroup).
- Due to the chronicity and long-term nature of the condition, physiological and psychological changes may take longer to manifest. It is widely accepted that there is a delay in muscular hypertrophy as a result of exercise, and initial gains within the first few weeks of any training programme will be as a result of neural factors (Enoka 1997); this is also in line with the grading of evidence (tier two evidence or higher requires a minimum of a four-week intervention). This suggests that longer interventions may be necessary (eight weeks for tier one evidence), though assessing participants at regular intervals, including at four weeks, would be beneficial to examine the effect of the neural adaptation alone.

Measurement (end-points)

- Randomised controlled trials with long-term follow-up are needed. Chronic pain is defined by its chronic nature, and therefore long-term follow-up of results is equally important as the initial short-term effect (if not more so): outcomes should be assessed beyond one year after randomisation. In turn this will inform the direct effect of the intervention, as well as the proportion of the population who maintains the programme of exercise employed in the intervention, or something else under the guise of physical activity as a result of participation.
- The broad time window for 'short term' outcomes (less than six months) is a potential source of heterogeneity as the early period is the one where time of measurement is most likely to result in variable outcomes. These initial problems could be overcome by use of standard reporting periods in exercise intervention studies (suggested four-weekly assessment within the 'short term' period to assess both neural adaptation and other physiological changes). This would allow review authors to use the data recorded closest to the time point they are assessing, for more accurate and comparable analyses.
- Outcome measures used by researchers should be standardised across trials and studies. Recommendations for selecting the most appropriate and important outcome measures to those who live with chronic pain have previously been published (Initiatives on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) Consensus Recommendations, Dworkin 2005; Turk 2003).

Other

- It would be of interest in future research to determine the reasons for non-participation in regular physical activity or non-compliance to a prescribed exercise intervention in people with chronic pain, and how to overcome these barriers.
- Future Cochrane Reviews could include: exercise for chronic pain or chronic widespread pain (and not specific conditions such as osteoarthritis, fibromyalgia, etc.), and exercise for neuropathic pain. These areas have not been covered by Cochrane with an exercise or physical activity intervention.

ACKNOWLEDGEMENTS

Thanks to Dr Amanda Williams, Professor Chris Eccleston, and Anna Erskine at the Cochrane Pain, Palliative and Supportive Care Review Group (PaPaS), for their support and rigour in the development of the concept of this overview, title proposal, and protocol. Thanks also to peer reviewers Dr Neil O'Connell and Dr Clare Stevinson, and consumer referees Jennifer Patrick and Dr Shailen Mehta.

Cochrane Review Group funding acknowledgement: the National Institute for Health Research (NIHR) is the largest single funder of the Cochrane PaPaS Group.

Disclaimer: the views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, National Health Service (NHS), or the Department of Health.

External funding support from the National Institute of Academic Anaesthesia (BJA RCoA NIAA) and the MRC Lifelong Health

and Wellbeing Grant to complete the overview as part of two larger projects.

REFERENCES

References to included reviews

Bartels EM, Lund H, Hagen KB, Dagfinrud H, Christensen R, Danneskiold-Samsøe B. Aquatic exercise for the treatment of knee and hip osteoarthritis. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD005523.pub2

Bidonde J, Busch AJ, Webber SC, Schachter CL, Danyliw A, Overend TJ, et al. Aquatic exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011336

Boldt I, Eriks-Hoogland I, Brinkhof MWG, de Bie R, Joggi D, von Elm E. Non-pharmacological interventions for chronic pain in people with spinal cord injury. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD009177.pub2

Brown J, Brown S. Exercise for dysmenorrhoea. *Cochrane Database of Systematic Reviews* 2010, Issue 2. [DOI: 10.1002/14651858.CD004142.pub2

Busch AJ, Barber KA, Overend TJ, Peloso PMJ, Schachter CL. Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD003786.pub2

Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010884

Cramp F, Hewlett S, Almeida C, Kirwan JR, Choy EHS, Chalder T, et al. Non-pharmacological interventions for fatigue in rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD008322.pub2

Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2008, Issue 4. [DOI: 10.1002/14651858.CD004376.pub2

Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: 10.1002/14651858.CD007912.pub2

Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD004376.pub3

Gross A, Kay TM, Paquin JP, Blanchette S, Lalonde P, Christie T, et al. Exercises for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD004250.pub5

Han A, Judd M, Welch V, Wu T, Tugwell P, Wells GA. Tai chi for treating rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2004, Issue 3. [DOI: 10.1002/14651858.CD004849

Hayden J, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database of Systematic Reviews* 2005, Issue 3. [DOI: 10.1002/14651858.CD000335.pub2

Hurkmans E, van der Giesen FJ, Vliet Vlieland TPM, Schoones J, Van den Ende ECHM. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2009, Issue 4. [DOI: 10.1002/14651858.CD006853.pub2

Kay TM, Gross A, Goldsmith CH, Rutherford S, Voth S, Hoving JL, et al. Exercises for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 2012, Issue 8. [DOI: 10.1002/14651858.CD004250.pub4

Koopman FS, Beelen A, Gilhus NE, de Visser M, Nollet F. Treatment for postpolio syndrome. *Cochrane Database of Systematic Reviews* 2015, Issue 5. [DOI: 10.1002/14651858.CD007818.pub3

Lane R, Ellis B, Watson L, Leng GC. Exercise for intermittent claudication. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD000990.pub3

Lauret GJ, Fakhry F, Fokkenrood HJP, Hunink MGM, Teijink JAW, Spronk S. Modes of exercise training for intermittent claudication. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD009638.pub2

Regnaux JP, Lefevre-Colau MM, Trinquart L, Nguyen C, Boutron I, Brosseau L, et al. High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis. *Cochrane Database of Systematic Reviews* 2015, Issue 10. [DOI: 10.1002/14651858.CD010203.pub2

Saragiotto BT, Maher CG, Yamato TP, Costa LOP, Menezes Costa LC, Ostelo RWJG, et al. Motor control exercise for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2016, Issue 1. [DOI: 10.1002/14651858.CD012004

Silva KNG, Mizusaki Imoto A, Almeida GJM, Atallah ÁN, Peccin MS, Fernandes Moça Trevisani V. Balance training (proprioceptive training) for patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2010, Issue 5. [DOI: 10.1002/14651858.CD007648.pub2

van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SMA, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2015, Issue 1. [DOI: 10.1002/14651858.CD010387.pub2

Yamato TP, Maher CG, Saragiotto BT, Hancock MJ,

Ostelo RWJG, Cabral CMN, et al. Pilates for low back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 7. [DOI: 10.1002/14651858.CD010265.pub2

References to excluded reviews

Aggarwal VR, Lovell K, Peters S, Javidi H, Joughin A, Goldthorpe J. Psychosocial interventions for the management of chronic orofacial pain. *Cochrane Database of Systematic Reviews* 2011, Issue 11. [DOI: 10.1002/14651858.CD008456.pub2

Bierma-Zeinstra SMA, Brinks A, Verhagen AP, Van Rijn RM, Koes BW, Verhaar JAN. Interventions for lateral hip pain (tendinopathy or bursitis). *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.CD008924

Brønfort G, Haas M, Evans RL, Goldsmith CH, Assendelft WJJ, Bouter LM. Non-invasive physical treatments for chronic/recurrent headache. *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD001878.pub3

Brønfort G, Evans RL, Goldsmith CH, Haas M, Leininger B, Levin M, et al. Spinal rehabilitative exercise and manual treatment for the prevention of migraine attacks in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 8. [DOI: 10.1002/14651858.CD011848

Choi BKL, Verbeek JH, Tam WWS, Jiang JY. Exercises for prevention of recurrences of low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD006555.pub2

Craane B, De Laat A, Dijkstra PU, Stappaerts K, Stegenga B. Physical therapy for the management of patients with temporomandibular disorders and related pain. *Cochrane Database of Systematic Reviews* 2006, Issue 1. [DOI: 10.1002/14651858.CD005621

Dagfinrud H, Hagen KB, Kvien TK. Physiotherapy interventions for ankylosing spondylitis. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD002822.pub3

Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. *Cochrane Database of Systematic Reviews* 2010, Issue 6. [DOI: 10.1002/14651858.CD007612.pub2

Dal Bello-Haas V, Florence JM. Therapeutic exercise for people with amyotrophic lateral sclerosis or motor neuron disease. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD005229.pub3

de Souza RF, Lovato da Silva CH, Nasser M, Fedorowicz Z, Al-Muharraqi MA. Interventions for the management of temporomandibular joint osteoarthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD007261.pub2

Fokkenrood HJP, Bendermacher BLW, Lauret GJ, Willigendael EM, Prins MH, Teijink JAW. Supervised exercise therapy versus non-supervised exercise therapy for intermittent claudication. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD005263.pub3

Franke H, Fryer G, Ostelo RWJG, Kamper SJ. Muscle energy technique for non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: 10.1002/14651858.CD009852.pub2

Green S, Buchbinder R, Hetrick SE. Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD004258

Gross A, Aker P, Goldsmith CH, Peloso PMJ. Physical medicine modalities for mechanical neck disorders. *Cochrane Database of Systematic Reviews* 1998, Issue 2. [DOI: 10.1002/14651858.CD000961

Gross A, Forget M, St George K, Fraser MMH, Graham N, Perry L, et al. Patient education for neck pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. [DOI: 10.1002/14651858.CD005106.pub4

Gross A, Langevin P, Burnie SJ, Bédard-Brochu MS, Empey B, Dugas E, et al. Manipulation and mobilisation for neck pain contrasted against an inactive control or another active treatment. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD004249.pub4

Hayden JA, Cartwright J, van Tulder MW, Malmivaara A. Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD009790

Heintjes EM, Berger M, Bierma-Zeinstra SMA, Bernsen RMD, Verhaar JAN, Koes BW. Exercise therapy for patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2003, Issue 4. [DOI: 10.1002/14651858.CD003472

Henschke N, Ostelo RWJG, van Tulder MW, Vlaeyen JWS, Morley S, Assendelft WJJ, et al. Behavioural treatment for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2010, Issue 7. [DOI: 10.1002/14651858.CD002014.pub3

Heymans MW, van Tulder MW, Esmail R, Bombardier C, Koes BW. Back schools for non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD000261.pub2

Hilde G, Hagen KB, Jamtvedt G, Winnem M. Advice to stay active as a single treatment for low-back pain and sciatica. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD003632.pub2

Hoving JL, Lacaille D, Urquhart DM, Hannu TJ, Sluiter JK, Frings-Dresen MHW. Non-pharmacological interventions for preventing job loss in workers with inflammatory arthritis. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD010208.pub2

Hurley M, Dickson K, Walsh N, Hauari H, Grant R, Cumming J, et al. Exercise interventions and patient beliefs for people with chronic hip and knee pain: a mixed methods review. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010842

IJzelenberg W, Rubinstein SM, Hayden J, Koes BW, van Tulder MW. Exercise therapy for acute non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2011,

Issue 10. [DOI: 10.1002/14651858.CD009365
Jones G, Crotty M, Brooks P. Interventions for treating psoriatic arthritis. *Cochrane Database of Systematic Reviews* 2000, Issue 3. [DOI: 10.1002/14651858.CD000212
Jordan JL, Holden MA, Mason EEJ, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD005956.pub2
Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RWJG, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. *Cochrane Database of Systematic Reviews* 2014, Issue 9. [DOI: 10.1002/14651858.CD000963.pub3
Karjalainen KA, Malmivaara A, van Tulder MW, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary rehabilitation for fibromyalgia and musculoskeletal pain in working age adults. *Cochrane Database of Systematic Reviews* 1999, Issue 3. [DOI: 10.1002/14651858.CD001984
Karjalainen KA, Malmivaara A, van Tulder MW, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD002194
Larun L, Brurberg KG, Odgaard-Jensen J, Price JR. Exercise therapy for chronic fatigue syndrome. *Cochrane Database of Systematic Reviews* 2016, Issue 2. [DOI: 10.1002/14651858.CD003200.pub4
Liddle SD, Pennick V. Interventions for preventing and treating low-back and pelvic pain during pregnancy. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: 10.1002/14651858.CD001139.pub4
Liu H, Yao K, Zhang J, Li L, Wu T, Brox JI, et al. Sling exercise therapy for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 9. [DOI: 10.1002/14651858.CD010689
Miller J, Gross A, Kay TM, Graham N, Burnie SJ, Goldsmith CH, et al. Manual therapy with exercise for neck pain. *Cochrane Database of Systematic Reviews* 2014, Issue 7. [DOI: 10.1002/14651858.CD011225
Moi JHY, Sriranganathan MK, Edwards CJ, Buchbinder R. Lifestyle interventions for chronic gout. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD010039.pub2
O'Brien K, Nixon S, Glazier R, Tynan AM. Progressive resistive exercise interventions for adults living with HIV/AIDS. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD004248.pub2
O'Connell NE, Wand BM, McAuley J, Marston L, Moseley GL. Interventions for treating pain and disability in adults with complex regional pain syndrome- an overview of systematic reviews. *Cochrane Database of Systematic Reviews* 2013, Issue 4. [DOI: 10.1002/14651858.CD009416.pub2
Page MJ, O'Connor D, Pitt V, Massy-Westropp N. Exercise and mobilisation interventions for carpal tunnel syndrome. *Cochrane Database of Systematic Reviews* 2012, Issue 6. [DOI: 10.1002/14651858.CD009899
Page MJ, Green S, Kramer S, Johnston RV, McBain B, Chau M, et al. Manual therapy and exercise for adhesive capsulitis (frozen shoulder). *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD011275
Peters S, Page MJ, Coppieters MW, Ross M, Johnston V. Rehabilitation following carpal tunnel release. *Cochrane Database of Systematic Reviews* 2013, Issue 6. [DOI: 10.1002/14651858.CD004158.pub2
Preston NJ, Seers K, Mortimer PS. Physical therapies for reducing and controlling lymphoedema of the limbs. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD003141.pub2
Proctor M, Murphy PA, Pattison HM, Suckling JA, Farquhar C. Behavioural interventions for dysmenorrhoea. *Cochrane Database of Systematic Reviews* 2007, Issue 3. [DOI: 10.1002/14651858.CD002248.pub3
Radner H, Ramiro S, Buchbinder R, Landewé RBM, van der Heijde D, Aletaha D. Pain management for inflammatory arthritis (rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis and other spondyloarthritis) and gastrointestinal or liver comorbidity. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.CD008951.pub2
Regnaux JP, Trinquart L, Boutron I, Nguyen C, Brosseau L, Ravaud P. High-intensity versus low-intensity physical activity or exercise in patients with hip or knee osteoarthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD010203
Regnaux JP, Lefevre-Colau MM, Palazzo C, Poiraudeau S, Ravaud P, Boutron I. Exercise therapy for ankylosing spondylitis. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011321
Richards BL, Whittle SL, Buchbinder R. Neuromodulators for pain management in rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.CD008921.pub2
Riemsma RP, Kirwan JR, Taal E, Rasker HJJ. Patient education for adults with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD003688
Schaafsma FG, Whelan K, van der Beek AJ, van der Es-Lambeek LC, Ojajärvi A, Verbeek JH. Physical conditioning as part of a return to work strategy to reduce sickness absence for workers with back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: 10.1002/14651858.CD001822.pub3
Steultjens EEMJ, Dekker JJ, Bouter LM, Schaardenburg DD, Kuyk MAMAH, Van den Ende ECHM. Occupational therapy for rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2004, Issue 1. [DOI: 10.1002/14651858.CD003114.pub2
Stones W, Cheong YC, Howard FM, Singh S. Interventions for treating chronic pelvic pain in women. *Cochrane Database of Systematic Reviews* 2005, Issue 2. [DOI: 10.1002/14651858.CD000387
Takken T, Van Brussel M, Engelbert RH, van der Net

JJ, Kuis W, Helders PPJM. Exercise therapy in juvenile idiopathic arthritis. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD005954.pub2
van der Heijden RA, Lankhorst NE, van Linschoten R, Bierma-Zeinstra SMA, van Middelkoop M. Exercise for treating patellofemoral pain syndrome. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010387
van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD011142.pub2
White CM, Pritchard J, Turner-Stokes L. Exercise for people with peripheral neuropathy. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD003904.pub2
Zammit GV, Menz HB, Munteanu SE, Landorf KB, Gilheany MF. Interventions for treating osteoarthritis of the big toe joint. *Cochrane Database of Systematic Reviews* 2010, Issue 9. [DOI: 10.1002/14651858.CD007809.pub2
Østerås N, Kjeken I, Smedslund G, Moe RH, Slatkowsky-Christensen B, Uhlig T, et al. Exercises for hand osteoarthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010388

Additional references

Abdulla 2013
Abdulla A, Adams N, Bone M, Elliott AM, Gaffin J, Jones D, et al. British Geriatric Society. Guidance on the management of pain in older people. *Age and Ageing* 2013;**42**(Suppl 1):i1–57. [DOI: 10.1093/ageing/afs200; PUBMED: 23420266

American College of Sports Medicine 2007
American College of Sports Medicine, American Heart Association. Exercise and acute cardiovascular events: placing the risks into perspective. *Medicine and Science in Sport and Exercise* 2007;**39**(5):886–97. [DOI: 10.1249/mss.0b013e3180574e0e; PUBMED: 17468590

Benton 2011
Benton MJ, Kasper MJ, Raab SA, Waggener GT, Swan PD. Short-term effects of resistance training frequency on body composition and strength in middle-aged women. *Journal of Strength and Conditioning Research/National Strength & Conditioning Association* 2011;**25**(11):3142–9. [DOI: 10.1519/JSC.0b013e31820f505f

Bigal 2012
Bigal ME, Rapoport AM. Obesity and chronic daily headache. *Current Pain and Headache Reports* 2012;**16**(1):101–9. [DOI: 10.1007/s11916-011-0232-0; PUBMED: 22076673

Bouchard 2012
Bouchard C, Blair SN, Church TS, Earnest CP, Hagberg JM, Häkkinen K, et al. Adverse metabolic response to regular exercise: is it a rare or common occurrence?. *PLoS One* 2012;**7**(5):e37887. [DOI: 10.1371/journal.pone.0037887; PUBMED: 22666405

Bousema 2007
Bousema EJ, Verbunt JA, Seelen HA, Vlaeyen JW, Knottnerus JA. Disuse and physical deconditioning in the first year after the onset of back pain. *Pain* 2007;**130**(3):279–86. [PUBMED: 17467902]

Breivik 2006
Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *European Journal of Pain* 2006;**10**(4):287–333. [PUBMED: 16095934]

Cohen 1988
Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edition. Hillsdale, NJ: L. Erlbaum Associates, 1988. [ISBN 0805802835]

Collins 1997
Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? . *Pain* 1997;**72**(1-2):95–7. [DOI: 10.1016/S0304-3959(97)00005-5; PUBMED: 9272792

Daenen 2015
Daenen L, Varkey E, Kellmann M, Nijs J. Exercise, not to exercise, or how to exercise in patients with chronic pain? Applying science to practice. *Clinical Journal of Pain* 2015;**31**(2):108–14. [DOI: 10.1097/AJP.0000000000000099; PUBMED: 24662498

Deschartes 2013
Deschartes A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ* 2013;**346**:f2304. [DOI: 10.1136/bmj.f2304

Dworkin 2005
Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;**113**:9–19. [PUBMED: 15621359]

Dworkin 2008
Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Consensus Statement: Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *Journal of Pain* 2008;**9**(2):105–21. [DOI: 10.1016/j.jpain.2007.09.005; PUBMED: 18055266

Eccleston 2014
Eccleston C, Fisher E, Craig L, Duggan GB, Rosser BA, Keogh E. Psychological therapies (Internet-delivered) for the management of chronic pain in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD010152.pub2

Elliott 1999
Elliott AM, Smith BH, Penny KI, Smith WC, Chambers WA. The epidemiology of chronic pain in the community. *Lancet* 1999;**354**:1248–52. [PUBMED: 10520633]

Elzahaf 2012
Elzahaf RA, Tashani OA, Unsworth BA, Johnson MI. The prevalence of chronic pain with an analysis of countries with a Human Development Index less than 0.9: a systematic review without meta-analysis. *Current Medical Research & Opinion* 2012;**28**(7):1221–9. [DOI: 10.1185/03007995.2012.703132

Engers 2008
Engers A, Jellema P, Wensing M, van der Windt DA, Grol R, van Tulder MW. Individual patient education for low back pain. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD004057.pub3; PUBMED: 18254037

Enoka 1997
Enoka RM. Neural adaptations with chronic physical activity. *Journal of Biomechanics* 1997;**30**(5):447–55. [PUBMED: 9109556]

Exercise is Medicine
American College of Sports Medicine (ACSM). Exercise is Medicine. http://www.exerciseismedicine.org/ (accessed 27 June 2014).

Finan 2013
Finan PH, Smith MT. The comorbidity of insomnia, chronic pain, and depression: dopamine as a putative mechanism. *Sleep Medicine Reviews* 2013;**17**(3):173–83. [DOI: 10.1016/j.smrv.2012.03.003; PUBMED: 22748562

Gross 2009
Gross A, Forget M, St George K, Fraser MM, Graham N, Perry L, et al. Patient education for neck pain. *Cochrane Database of Systematic Reviews* 2012, Issue 3. [DOI: 10.1002/14651858.CD005106.pub4; PUBMED: 22419306

Harvard 2013
Better balance: mental and physical fitness are both essential: daily physical activity, targeted balance exercise, and steps to maintain your brain all help prevent falls. *Harvard Men's Health Watch* 2013;**17**(11):5. [PUBMED: 24000423]

Hoffman 2014
Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**. [DOI: 10.1136/bmj.g1687; PUBMED: 24609605

IDF 2012
Guariguata L. By the numbers: new estimates from the IDF Diabetes Atlas Update for 2012. *Diabetes Research and Clinical Practice* 2012;**98**(3):524–5. [DOI: 10.1016/j.diabres.2012.11.006; PUBMED: 23217268

IntHout 2015
IntHout J, Ioannidis JPA, Borm GF, Goeman JJ. Small studies are more heterogeneous than large ones: a meta-meta-analysis. *Journal of Clinical Epidemiology* 2015;**68**(8):860–9. [DOI: 10.1016/j.jclinepi.2015.03.017; PUBMED: 25959635

Leadley 2014
Leadley RM, Armstrong N, Reid KJ, Allen A, Misso KV, Kleijnen J. Healthy aging in relation to chronic pain and quality of life in Europe. *Pain Practice* 2014;**14**(6):547–58. [DOI: 10.1111/papr.12125; PUBMED: 24138082

Mayer 2008
Mayer J, Mooney V, Dagenais S. Evidence-informed management of low back pain with lumbar extensor strengthening exercises. *The Spine Journal* 2008;**8**(1):96–113. [DOI: 10.1016/j.spinee.2007.09.008; PUBMED: 18164458

Merskey 2011
Merskey H, Bogduk N. Part III: Pain terms: a current list with definitions and notes on usage. *Classification of Chronic Pain*. 2nd Edition. Seattle: IASP Task Force on Taxonomy, IASP Press, 2011 [update from 1994].

Messier 2013
Messier SP, Mihalko SL, Legault C, Miller GD, Nicklas BJ, DeVita P, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. *JAMA* 2013;**310**(12):1263–73. [DOI: 10.1001/jama.2013.277669; PUBMED: 24065013

Mooney 2006
Mooney V, Verna J, Morris C. Clinical management of chronic, disabling low back syndromes. In: Morris C editor (s). *Rehabilitation of the Spine: a Practitioner's Manual*. New York: McGraw-Hill, 2006.

Moore 1998
Moore RA, Gavaghanb D, Tramer MR, Collins SL, McQuay HJ. Size is everything - large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. *Pain* 1998;**78**:209–16. [PUBMED: 9870574]

Moore 2008
Moore RA, Barden J, Derry S, McQuay HJ. Managing potential publication bias. In: McQuay HJ, Kalso E, Moore RA editor(s). *Systematic Reviews in Pain Research: Methodology Refined*. Seattle: IASP Press, 2008:15–23. [ISBN: 978–0–931092–69–5]

Moore 2010
Moore RA, Moore OA, Derry S, Peloso PM, Gammaitono AR, Wang H. Responder analysis for pain relief and numbers needed to treat in a meta-analysis of etoricoxib osteoarthritis trials: bridging a gap between clinical trials and clinical practice. *Annals of Rheumatic Diseases* 2010;**69**(2):374–9. [DOI: 10.1136/ard.2009.107805; PUBMED: 19364730

Moore 2013
Moore RA, Straube S, Aldington D. Pain measures and cut-offs - 'no worse than mild pain' as a simple, universal outcome. *Anaesthesia* 2013;**68**(4):400–12. [DOI: 10.1111/anae.12148; PUBMED: 23347230

Moore 2014a

Moore RA, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. *Pain Practice* 2014;**14**(1):79–94. [PUBMED: 23464879]

Moore 2014b

Moore RA, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. Pain Practice 2014; Vol. 14, issue 1:79–94. [DOI: 10.1111/papr.12050

National Library of Medicine

National Library of Medicine. Medical Subject Headings. search "chronic pain" at www.nlm.nih.gov/mesh/MBrowser.html (accessed 22 August 2014). [MeSH unique ID: D059350]

NICE 2014

National Institute for Health and Care Excellence (NICE). Osteoarthritis. Care and management in adults (CG177), February 2014. www.nice.org.uk/guidance/cg177 (accessed 4 December 2015). [ISBN: 978–1–4731–0426–6]

Nnoaham 2008

Nnoaham KE, Kumbang J. Transcutaneous electrical nerve stimulation (TENS) for chronic pain. *Cochrane Database of Systematic Reviews* 2008, Issue 3. [DOI: 10.1002/14651858.CD003222.pub2; PUBMED: 18646088

Nüesch 2010

Nüesch E, Trelle S, Reichenbach S, Rutjes AW, Tschannen B, Altman DG, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;**341**:c3515. [PUBMED: 20639294]

Nüesch 2011

Nüesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Jüni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011;**342**:d1165. [10.1136/bmj.d1165]

Park 2012

Park J, Hughes AK. Non-pharmacological approaches to the management of chronic pain in community-dwelling older adults: a review of empirical evidence. *Journal of the American Geriatric Society* 2012;**60**(3):555–68. [DOI: 10.1111/j.1532-5415.2011.03846.x; PUBMED: 22288789

Perruchoud 2014

Perruchoud C, Buchser E, Johanek LM, Aminian K, Paraschiv-Ionescu A, Taylor RS. Assessment of physical activity of patients with chronic pain. *Neuromodulation* 2014;**17**:42–7. [DOI: 10.1111/ner.12036

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, Cochrane. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, Cochrane, 2014.

Sakuma 2012

Sakuma M, Endo N. Space flight/bedrest immobilization and bone. Exercise training for osteoporosis. *Clinical Calcium* 2012;**22**(12):1903–7. [DOI: CliCa121219031907; PUBMED: 23187084

Schofield 2011

Schofield P, Clarke A, Jones D, Martin D, McNamee P, Smith BH. Chronic pain in later life: a review of current issues and challenges. *Aging Health* 2011;**7**(4):551–6. [DOI: 10.2217/ahe.11.41

Shea 2007

Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007;**7**:10. [DOI: 10.1186/1471-2288-7-10; PUBMED: 17302989

SIGN 2013

Scottish Intercollegiate Guidelines Network (SIGN). *Management of chronic pain. SIGN publication no. 136*. 1st Edition. Edinburgh, UK: SIGN, 2013. [ISBN: 978–1–909103–17–7]

Smith 2007

Smith BH, Macfarlane GJ, Torrance N. Epidemiology of chronic pain, from the laboratory to the bus stop: time to add understanding of biological mechanisms to the study of risk factors in population-based research?. *Pain* 2007;**127**(1-2):5–10. [DOI: 10.1016/j.pain.2006.11.001; PUBMED: 17140732

Stagg 2011

Stagg NJ, Mata HP, Ibrahim MM, Henriksen EJ, Porreca F, Vanderah TW, et al. Regular exercise reverses sensory hypersensitivity in a rat neuropathic pain model: role of endogenous opioids. *Anaesthesiology* 2011;**114**(4):940–8. [DOI: 10.1097/ALN.0b013e318210f880; PUBMED: 21386701

Straube 2015

Straube S, Derry S, Moore RA, McQuay HJ. Vitamin D for the treatment of chronic painful conditions in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 5. [DOI: 10.1002/14651858.CD007771.pub3; PUBMED: 25946084

Thornton 2000

Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *Journal of Clinical Epidemiology* 2000;**53**(2):207–16. [DOI: 10.1016/S0895-4356(99)00161-4

Timmons 2014

Timmon J, Phillips SM. Do we have a valid public health strategy for exercise prescription?. *Sport and Exercise Scientist* 2014;**39**:26–7.

To 2012

To T, Stanojevic S, Moores G, Gershon AS, Bateman ED, Cruz AA, et al. Global asthma prevalence in adults: findings from the cross-sectional world health study. *BMC Public Health* 2012;**12**:204. [DOI: 10.1186/1471-2458-12-204; PUBMED: 22429515

Torrance 2010

Torrance N, Elliott AM, Lee AJ, Smith BH. Severe chronic pain is associated with increased 10 year mortality. A cohort

record linkage study. *European Journal of Pain* 2010;**14**(4): 380–6. [DOI: 10.1016/j.ejpain.2009.07.006

Turk 2003
Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003; **106**:337–45. [PUBMED: 14659516]

Turner 2013
Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane Reviews. *PLoS One* 2013;**8**(3):e59202. [DOI: 10.1371/journal.pone.0059202; PUBMED: 23544056

Ursini 2011
Ursini F, Naty S, Grembiale RD. Fibromyalgia and obesity: the hidden link. *Rheumatology International* 2011;**31**(11): 1403–8. [DOI: 10.1007/s00296-011-1885-z; PUBMED: 21476098

van Hecke 2013a
van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology - where do lifestyle factors fit in?. British Journal of Pain 2013; Vol. 7, issue 4:209–17. [DOI: 10.1177/2049463713493264

van Hecke 2013b
van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology and its clinical relevance. *British Journal of Anaesthesia* 2013;**111**(1):13–8. [DOI: 10.1093/bja/aet123; PUBMED: 23794640

Vigorito 2014
Vigorito C, Giallauria F. Effects of exercise on cardiovascular performance in the elderly. *Frontiers in Physiology* 2014;**5**: 51. [DOI: 10.3389/fphys.2014.00051

Vos 2015
Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;**386**(9995):743–800. [DOI: 10.1016/S0140-6736(15)60692-4; PUBMED: 26063472

Waddell 1987
Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine* 1987;**12**(7):632–44. [PUBMED: 2961080]

WHO 2015
World Health Organization. Physical activity. Fact sheet 385. www.who.int/mediacentre/factsheets/fs385/en/ (accessed 4 December 2015).

Wieland 2013
Wieland LS, Skoetz N, Manheimer E, Pilkington K, Vempati R, Berman BM. Yoga treatment for chronic non-specific low-back pain. *Cochrane Database of Systematic Reviews* 2013, Issue 7. [DOI: 10.1002/14651858.CD010671

Williams 2012
Williams ACDC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD007407.pub3

References to other published versions of this review

Geneen 2014
Geneen L, Smith B, Clarke C, Martin D, Colvin LA, Moore RA. Physical activity and exercise for chronic pain in adults: an overview of Cochrane reviews. *Cochrane Database of Systematic Reviews* 2014, Issue 8. [DOI: 10.1002/14651858.CD011279

* *Indicates the major publication for the study*

ADDITIONAL TABLES

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews

Criteria	Specific requirements (possible answers: yes, no, cannot answer, not applicable)
1. Was an 'a priori' design used?	The research question and inclusion criteria should be established before the conduct of the review *Note: need to refer to a protocol, ethics approval, or predetermined/a priori published research objectives to score a "yes."*
2. Was there duplicate study selection and data extraction?	There should be at least 2 independent data extractors and a consensus procedure for disagreements should be in place *Note: 2 people do study selection, 2 people do data extraction, consensus process or 1 person checks the other person's work.*

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews (*Continued*)

3. Was a comprehensive literature search performed?	At least 2 electronic sources should be searched. The report must include years and databases used (e.g. CENTRAL, MEDLINE, and Embase). Keywords or MeSH terms (or both) must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialised registers, or experts in the particular field of study, and by reviewing the references in the studies found *Note: if at least 2 sources + 1 supplementary strategy used, select "yes" (Cochrane register/ CENTRAL counts as 2 sources; a grey literature search counts as supplementary).*
4. Was the status of the publication (i.e. grey literature) used as inclusion criteria?	The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language, etc *Note: if review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished literature.*
5. Was a list of studies (included and excluded) provided?	A list of included and excluded studies should be provided. *Note: acceptable if the excluded studies were referenced. If there was an electronic link to the list but the link is no longer active, select "no. "*
6. Were the characteristics of the included studies provided?	In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions, and outcomes. The ranges of characteristics in all the studies analysed, e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported *Note: acceptable if not in table format as long as they are described as above.*
7. Was the scientific quality of the included studies assessed and documented?	'A priori' methods of assessment should be provided (e.g. for effectiveness studies if the author(s) chose to include only randomised, double-blind, placebo-controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant *Note: can include use of a quality scoring tool or checklist, e.g. Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some type of result for EACH study ("low" or "high" is acceptable, as long as it is clear which studies scored "low" and which scored "high;" a summary score/range for all studies is not acceptable).*
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations

Table 1. AMSTAR tool to assess the methodological quality of systematic reviews (*Continued*)

	Note: might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.
9. Were the methods used to combine findings of studies appropriate?	For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi^2 test for homogeneity, I^2 statistic). If heterogeneity exists, a random-effects model should be used or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?), or both *Note: indicate "yes" if they mention or describe heterogeneity, i.e. if they explain that they cannot pool because of heterogeneity/variability between interventions.*
10. Was the likelihood of publication bias assessed?	An assessment of publication bias should include a combination of graphical aids (e.g. funnel plot, other available tests) or statistical tests (e.g. Egger regression test), or both *Note: if no test values or funnel plot included, score "no." Score "yes" if they mention that publication bias could not be assessed because there were fewer than 10 included studies.*
11. Was the conflict of interest stated?	Potential sources of support should be clearly acknowledged in both the systematic review and the included studies *Note: to get a "yes," must indicate source of funding or support for the systematic review AND for each of the included studies.*

Table 2. Reasons for exclusion

Review	Reason for exclusion from overview
Aggarwal 2011	Not exercise/physical activity
Brønfort 2015	Protocol stage only - possibly include when published as full review
Bierma-Zeinstra 2011	Protocol stage only - exclude when published as full review
Brønfort 2014	Withdrawn from the Cochrane Library
Choi 2010	Not chronic using definition of > 3 months
Craane 2006	Protocol stage only - possibly include when published as full review
Dagfinrud 2008	Physiotherapy - required therapist to perform intervention
Dahm 2010	Acute pain, not chronic. Intervention was advice
Dal Bello-Haas 2013	Malignant condition

Table 2. Reasons for exclusion (*Continued*)

de Souza 2012	Drug- and surgery-based interventions
Fokkenrood 2013	Did not include RCTs (excluded studies with control groups)
Franke 2015	Not exercise/physical activity
Green 2003	Physiotherapy - required therapist to perform intervention
Gross 1998	Withdrawn from the Cochrane Library
Gross 2012	Not exercise/physical activity
Gross 2015b	Not exercise/physical activity
Hayden 2012	Protocol stage only - possibly include when published as full review
Heintjes 2003	Withdrawn from the Cochrane Library January 2015
Henschke 2010	Not exercise/physical activity
Heymans 2004	Exercise could not be assessed as stand-alone intervention
Hilde 2006	Withdrawn from the Cochrane Library
Hoving 2014	No exercise intervention, and no pain outcome measure
Hurley 2013	Protocol stage only - exclude when published as full review
IJzelenberg 2011	Protocol stage only - exclude when published as full review
Jones 2000	Drug-based interventions
Jordan 2010	Intervention to improve adherence to exercise, not exercise itself
Kamper 2014	Exercise could not be assessed as stand-alone intervention
Karjalainen 1999	Exercise could not be assessed as stand-alone intervention
Karjalainen 2003	Exercise could not be assessed as stand-alone intervention
Larun 2016	Chronic fatigue, not chronic pain
Liddle 2015	Pain in pregnancy only, not chronic pain
Liu 2013	Protocol stage only - unsure about inclusion when published as full review

Table 2. Reasons for exclusion (*Continued*)

Miller 2014	Protocol stage only - exclude when published as full review
Moi 2013	Exercise could not be assessed as stand-alone intervention
O'Brien 2004	No pain outcome measure
O'Connell 2013	Overview of reviews, not systematic review
Østerås 2013	Protocol stage only - possibly include when published as full review
Page 2012	No pain outcome measure
Page 2014	Manual therapy - required therapist to perform intervention
Peters 2013	Exercise could not be assessed as stand-alone intervention
Preston 2004	No pain outcome measure
Proctor 2007	Exercise could not be assessed as stand-alone intervention
Radner 2012	Drug-based interventions
Regnaux 2014	Protocol stage only - possibly include when published as full review
Richards 2012	Not exercise/physical activity
Riemsma 2003	Not exercise/physical activity
Schaafsma 2013	No pain outcome measure
Steultjens 2004	Occupational therapy - exercise could not be assessed as stand-alone intervention
Stones 2005	Exercise cannot be assessed as stand-alone intervention
Takken 2008	Aged < 18 years - not adults
van Dessel 2014	Not chronic pain and no specific pain outcome measure
White 2004	No pain outcome measure
Williams 2012	Not exercise/physical activity
Zammit 2010	Surgery or required therapist to perform intervention

RCT: randomised controlled trial.

Table 3. Characteristics of included reviews

Review and Cochrane Review Group	Assessed as up to date	Chronic pain condition	Duration of pain/ diagnosis	Intervention description	Control description	Outcomes with data reported	Time points reported
Bartels 2007 Cochrane Musculoskeletal Group	Aug 2007	Hip or knee OA	Not reported	All types of exercises developed in the therapeutic/heated indoor pool (ROM, dynamics, aerobics, etc.) were permitted	No treatment or other treatment.	Function, quality of life, mental health, pain, adverse events	Post-intervention (immediate), 6-month follow-up
Bidonde 2014 Cochrane Musculoskeletal Group	Oct 2013	Fibromyalgia	12 yr (range 6 to 24)	Aquatic exercise training intervention defined as "exercise conducted in a vertical standing position."	Treatment as usual, physical activity as usual, wait list control, placebo or sham, education-only, water immersion-only, and attention only	Multi-dimensional function (wellness), self-reported physical function (wellness), pain (symptoms), stiffness (symptoms), muscle strength (physical fitness), submaximal cardiorespiratory function (physical fitness), withdrawals (safety and acceptability), adverse effects (safety and acceptability)	Post-intervention (4 to 32 wk)
Boldt 2014 Cochrane Injuries Group	Mar 2011	Spinal cord injury	Mean 66 months, and 1 to 24 yr when reported	"Exercise": stretching and strengthening exercises aimed at mobilising	Wait list control or no intervention.	Pain, depression, quality of life, adverse effects	Short term (within 24 hours of last intervention, i.e. post-intervention)

Table 3. Characteristics of included reviews (*Continued*)

				painful shoulder joint			and intermediate term (1 to 6 wk post-intervention) and long term (> 6 wk post-intervention)
Brown 2010 Cochrane Menstrual Disorders and Subfertility Group	Aug 2009	Primary dysmenorrhoea in the majority (≥ 50%) of cycles	Ongoing/not appropriate	12-wk walk or jog training programme at an intensity of 70% to 85% of the HR range. Training for 3 days/wk and duration of aerobic phase was 30 minutes with 15-minute warm-up and cool-down periods	Asked not to exercise during the experimental period.	Pain: menstrual disorders questionnaire (MDQ) score	Ongoing - over 3 menstrual cycles
Busch 2007 Cochrane Musculoskeletal Group	Aug 2007	Fibromyalgia	Not reported	Exercise-only interventions included aerobic-only training, strength-only training, flexibility-only training, or mixed exercise-only interventions	"Untreated."	Pain, global wellbeing, objectively measured physical function	Post-intervention (strength exercise 21 wk, aerobic exercise 6 to 23 wk)
Busch 2013 Cochrane Musculoskeletal Group	Mar 2013	Fibromyalgia	mean range from 4 yrs (SD 3.1) to 12 yrs (SD 4)	Defined resistance training as exercise performed against a progressive resistance on a minimum of 2 days/wk (on non-consecutive days) with the in-	Untreated control conditions (treatment as usual, activity as usual, wait list control, and placebo), other types of exercise or physical activity in-	Multidimensional function, self-reported physical function, pain, tenderness, muscle strength, adverse effects, all-cause attrition	Post-intervention, follow-up (12 wk) in 1 study only

Table 3. Characteristics of included reviews (*Continued*)

				tention of improving muscle strength, muscle endurance, muscle power, or a combination of these	terventions (e.g. aerobic, flexibility), and other resistance training interventions (head-to-head comparisons)		
Cramp 2013 Cochrane Musculoskeletal Group	Oct 2012	Rheumatoid arthritis	Not reported	Included pool-based therapy (twice/wk, moderate intensity, music-paced), yoga (6 wk, twice/wk, 1.5-hour sessions), dynamic strength training (home-based after inpatient programme, all main muscle groups using dumbbells and elastic bands), stationary cycling (70% HRmax, 5 minute excluding: 1-minute of rest, increased duration), low-impact aerobics (class at fitness centre and video at home, individual HR targets), tai chi (1-hour group sessions)	"Could have been placebo, an alternative intervention (pharmacological or non-pharmacological) or usual care."	Fatigue, pain, anxiety, depression, disability, tender and swollen joints, adverse events	Post-intervention (only a single time point analysed)

Table 3. Characteristics of included reviews (*Continued*)

Fransen 2014 Cochrane Musculoskeletal Group	May 2013	Hip OA	Not reported	Any land-based therapeutic exercise regimens aiming to relieve the symptoms of hip OA, regardless of content, duration, frequency, or intensity. This included any exercise designed to improve muscle strength, range of joint movement or aerobic capacity (or combinations of the three). Programmes could be designed and supervised by physiotherapists or other professionals, or provided as a home programme with minimal monitoring	Wait-list control, usual care, GP education.	Self-reported pain, physical function, quality of life, withdrawal or dropouts, adverse events	post-intervention (immediate in 9/10 studies) follow-up 3 to 6 months
Fransen 2015 Cochrane Musculoskeletal Group	May 2013	Knee OA	Often not reported: some less than 1yr, others over 10yr	"land-based therapeutic exercise." Along with delivery mode and content, treatment 'dosage' (duration, frequency, intensity) varied widely be-	No exercise: active (any no-exercise intervention) or no treatment (including waiting list)	Knee pain, self-reported physical function, quality of life	Immediately at the end of treatment (post-treatment), 2 to 6 months after cessation of monitored study treatment and longer than six months

Table 3. Characteristics of included reviews (*Continued*)

				tween studies			after cessation of monitored study treatment
Gross 2015a Cochrane Back Group	May 2014	Mechanical neck disorders	"Chronic" (not subacute or acute)	Cervical stretch/ROM exercises + cervical/ scapulothoracic strengthening + static/ dynamic cervical/shoulder stabilisation	Wait list control.	Pain intensity, function, quality of life, global perceived effect, adverse effects	Immediately post-treatment (≤ 1 day), short-term follow-up (1 day to 3 months), intermediate-term follow-up (3 months up to, but not including, 1 yr), and long-term follow-up (≥ 1 yr)
Han 2004 Cochrane Musculoskeletal Group	Apr 2004	Rheumatoid arthritis	Not reported	Only trials of exercise programmes with tai chi instruction or incorporating principles of tai chi philosophy	Not reported.	Function, tender and swollen joints, ROM, strength, enjoyment, withdrawals, adverse effects	Post-intervention (8 to 10 wk)
Hayden 2005 Cochrane Back Group	Sep 2004	Non-specific low back pain	Chronic, i.e. longer than 12 wk: 5.6 yr (95% CI 3.4 to 7.8)	Exercise therapy defined as "a series of specific movements with the aim of training or developing the body by a routine practice or as physical training to promote good physical health;"	No exercise: no treatment or placebo treatment, other conservative therapy, or another exercise group	Pain, functional ability, work status, global assessment, adverse events	Earliest, 6 wk, 6 months, 12 months

Table 3. Characteristics of included reviews (*Continued*)

				only 54% adequately described the exercise intervention			
Hurkmans 2009 Cochrane Musculoskeletal Group	Jun 2009	Rheumatoid arthritis	5 to 14 yr	Dynamic exercise programmes - aerobic capacity and muscle strength training; short-term muscle strength training (high quality); short-term dynamic exercise to improve aerobic capacity (not high methodological quality); exercise frequency of at least 20 minutes twice a week. Duration of exercise programme at least 6 wk (duration < 3 months was considered short-term; duration > 3 months was considered long-term). Exercise programme performed under supervision Aerobic exercise intensity at least 55% of	Not reported	Functional ability, aerobic capacity, muscle strength, safety (pain and radiological damage)	Follow-up (12 wk and 24 months)

Table 3. Characteristics of included reviews (*Continued*)

				the maximum HR; or intensity starting at 40% to 50% of the maximum oxygen uptake reserve or HR maximum reserve. Furthermore, the intensity was increased up to 85% during the intervention. Progressively strengthening exercise loads starting at 30% to 50% and increasing to 80% of maximum (defined as the percentage of either 1 repetition maximum, 1 MVC, maximum speed, or as maximal subjective exertion)			
Koopman 2015 Cochrane Neuromuscular Group	Jul 2014	Postpolio syndrome (PPS)	Not reported	Exercise therapy (e.g. aerobic exercise, muscle strengthening exercise, respiratory muscle training, warm climate training, hydro training)	Placebo, usual care or no treatment.	Self-perceived activity limitations, muscle strength, muscle endurance, fatigue, pain, adverse events (minor and serious)	3 and 6 months

Table 3. Characteristics of included reviews (*Continued*)

Lane 2014 Cochrane Peripheral Vascular Diseases Group	Sep-2013	intermittent claudication	not reported	Any exercise programme used in the treatment of intermittent claudication was included, such as walking, skipping and running. Inclusion of trials was not affected by the duration, frequency or intensity of the exercise programme but these issues were taken into account in the meta-analysis	Exercise was compared to six different modes of treatment, the most common being usual care or placebo. Two early trials compared exercise with placebo tablets but in more recent studies usual care was used as the control comparator. Exercise was compared with the following drug therapies: antiplatelet agents pentoxifylline, iloprost, and vitamin E. One study compared exercise with pneumatic foot and calf compression	maximal walking time, pain-free walking time, pain-free walking distance, maximum walking distance, ankle brachial index (ABI), peak exercise calf blood flow, mortality, amputation	Post-intervention, 3-month follow up, six-month follow up
Lauret 2014 Cochrane Peripheral Vascular Diseases Group	Jul 2013	Intermittent claudication	Not reported	Supervised walking programme needed to be supervised at least twice a week for a consecutive 6 wk of training	Alternative exercise.	Maximum walking distance (METs), pain-free walking distance (METs), health-related quality of life and functional impairment	n/a

Table 3. Characteristics of included reviews (*Continued*)

Regnaux 2015 Cochrane Musculoskeletal Group	Jun 2014	Hip or knee OA	> 6 months	High-intensity physical activity or exercise programme.	Low-intensity physical activity or exercise programme and control (no-exercise) group in 1 study.	Pain, physical function, quality of life, adverse effects (related to intervention), severe adverse events or withdrawal (due to intervention)	Post-intervention, intermediate term (6 to 12 months), long-term (over 12 months) follow-up
Saragiotto 2016 Cochrane Back and Neck Group	Apr 2015	Low back pain	> 12 wk	MCE: activation of the deep trunk muscles, targeting the restoration of control and co-ordination of these muscles	Placebo, no treatment, another active treatment, or when MCE was added as a supplement to other interventions. When MCE was used in addition to other treatments, it had to represent at least 50% of the total treatment programme to be included	Pain intensity and disability, function, quality of life, global impression of recovery, return to work, adverse events and recurrence	Post-intervention, short term (4 to 10 wk), intermediate term (3 to 6 months), long term (12 to 36 months)
Silva 2010 Cochrane Musculoskeletal Group	Jun 2009	Rheumatoid arthritis	No studies found	Balance training (proprioceptive training).	No intervention or other intervention.	ACR-50, pain, disease activity score (DAS), Health Assessment Questionnaire (HAQ for function), gait, adverse effects, discontinuation rate	n/a
van der Heijden 2015 Cochrane	May 2014	Adolescents and adults with	3 wk to 8 months (as minimum re-	Exercise therapy for patellofemoral	No treatment, placebo, or waiting list	Pain during activity, usual pain, func-	4- to 12-wk follow-up (short

Table 3. Characteristics of included reviews (*Continued*)

Bone, Joint and Muscle Trauma Group		patellofemoral pain	quirement) ; reported pain 4 wk to 9 yr	pain syndrome; exercises could be performed at home or under supervision of a therapist - various descriptions in the included trials, including knee exercises, hip and knee exercises, home exercises, supervised exercises, closed kinetic chain, open kinetic chain	controls. This also included 'exercise therapy + another intervention (e.g. taping) versus the other intervention alone (e.g. taping).'	tional ability, recovery	term) and 16 wk to 12 months (long term)
Yamato 2015 Cochrane Back Group	Mar 2014	Low back pain	Acute, subacute, chronic (i.e. no minimum)	Explicitly stated as based on Pilates principles, or the therapists who provided the interventions had previous training in Pilates exercises or the therapists were described as certified Pilates instructors	No intervention, placebo, or other interventions.	Pain intensity, disability, global impression of recovery, quality of life, return to work, adverse effects	Short term (4 to 8 wk), intermediate term (3 to 6 months)

ACR: American College of Rheumatology; GP: general practitioner; HR: heart rate; MCE: motor control exercise; MET: metabolic equivalents; n/a: not applicable; OA: osteoarthritis; ROM: range of motion; wk: week; yr: year.

Table 4. Further characteristics of included reviews

Review	Number of trials included	Total number of participants	Gender distribution	Participants ages
Bartels 2007	6 (4 exercise vs no exercise)	800 (674 exercise vs no exercise)	50% to 86% Female	Means ranged from 66 to 71 yr

Table 4. Further characteristics of included reviews (*Continued*)

Bidonde 2014	16 (9 exercise vs no exercise)	881 (519 exercise vs no exercise)	513 female, 6 male	Means ranged from 46.3 to 48.3 yr
Boldt 2014	16 (3 exercise vs no exercise)	616 (149 exercise vs no exercise)	115 male, 41 female across 3 studies	Range 19 to 65 yr and mean 35 to 45 yr
Brown 2010	1	36	100% female	Not reported
Busch 2007	34 (in meta-analysis - strength training vs control: 2; aerobic training vs control: 4)	2276 total (in meta-analysis - strength: 47, aerobic: 269)	96.4% female when reported (in 2197 participants)	Range reported as 27.5 to 60.2 yr
Busch 2013	5 studies as 7 publications (exercise vs control: 3 publications, 2 studies)	219 with fibromyalgia (exercise vs control: 81)	100% female	Not reported
Cramp 2013	24 (only 6 using physical activity interventions)	2882 (physical activity interventions: 371)	"A higher percentage of females"… when reported	"Mainly within the fifth decade"
Fransen 2014	10	> 549	75% to 80% female when reported	58 to 70 yr (means) when reported
Fransen 2015	54	5362	When reported 55% to 100% female	When reported mean age 60 to 70 yr
Gross 2015a	27 (16 chronic pain)	2485	Not reported	Not reported
Han 2004	4 (3 RCTs). Pain not reported in any included study	206 total; pain not reported in any included study	Not reported	Range 38 to 72 yr
Hayden 2005	61 (43 chronic low back pain)	6390 (3907 chronic low back pain)	Chronic: 46% male (95% CI 39 to 52)	Chronic: 42 yr (95% CI 40 to 44)
Hurkmans 2009	8 RCTs (5 exercise vs no-exercise)	575	"Mainly female"	52 yr
Koopman 2015	13 (2 exercise vs no exercise)	675 (68 exercise vs no exercise) - 1 study used 3 arms (no treatment in cold, exercise in cold, exercise in warm; we have excluded the warm exercise arm as cannot compare directly to the con-	~ 25% male	Mean 58 and 65 yr

Table 4. Further characteristics of included reviews (*Continued*)

		trol)		
Lane 2014	30	1822 total	Not reported	Mean > 65 yr
Lauret 2014	5 (0 for exercise vs no exercise)	184 (0 for exercise vs no exercise)	n/a	n/a
Regnaux 2015	6 (1 for exercise vs no exercise) only 1 study that had a no exercise control	656 (102 for exercise vs no exercise)	79 female	62.6 yr
Saragiotto 2016	29 (7 for exercise vs no exercise/minimal intervention)	2431 (671 for exercise vs no exercise)	"Mixed"	Median 40.9 yr (IQR 11.2) (range 20.8 to 54.8)
Silva 2010	None	None	n/a	n/a
van der Heijden 2015	31 (10 for exercise vs control)	1690	0% to 100% female; equally distributed across range	Mean 25 to 50 yr
Yamato 2015	10 (6 exercise vs minimal intervention (control))	478 (265 exercise vs control)	2 trials were all female, the others included both genders	Mean 38 yr (range 22 to 50)

CI: confidence interval; GP: general practitioner; IQR: interquartile range; OA: osteoarthritis; RCT: randomised controlled trial; ROM: range of motion; wk: week; yr: year.

Table 5. Dose and duration of exercise interventions in included reviews

Review	Duration	Frequency (sessions per day/wk/month)	Intensity	Duration (per session)	Other description
Bartels 2007	Not reported	Not reported	"Muscle maintenance" and "range of motion"	Not reported	No minimum requirement for inclusion. Actual intervention only reported by 2 of 6 included studies
Bidonde 2014	17 wk (range 4 to 32)	1 to 4/wk	Very light (< 57% HRmax) to vigorous (95% HRmax), self-selected, and not specified	45 minutes (range 30 to 70)	No minimum requirement for inclusion. None of the studies met the ACSM exercise guidelines

Table 5. Dose and duration of exercise interventions in included reviews (*Continued*)

					specified for aerobic or strength training. Only 1 study met the ACSM guidelines for flexibility training
Boldt 2014	12 wk to 9 months	2/day to 2/wk	Not reported	Reported for 1 study only (90 to 120 minutes)	No minimum requirement for inclusion. Stretching and strengthening exercises aimed at mobilising painful shoulder joint
Brown 2010	≥ 12 wk	3/wk	70% to 85% HRR	1 hour	No minimum requirement for inclusion.
Busch 2007	3 wk to 6 months	1 to 5/wk	Not reported	Not reported	No minimum requirement for inclusion. Assessed as whether they "met ACSM recommendations."
Busch 2013	8 to 21 wk (median 16 wk)	≥ 2/wk	> 4/10 RPE rating progressing to 70% to 80% 1RM	40 to 90 minutes	Assessed as whether they "met ACSM recommendations."
Cramp 2013	6 wk (when reported)	2/wk	"Low impact", "moderate", and 70% HRmax	1 to 1.5 hours, when reported	No minimum requirement for inclusion.
Fransen 2014	6 to 12 wk (median 8)	1 to 3/wk	"Low intensity" to "max effort"	30 to 60 minutes	No minimum requirement for inclusion. Intensity only reported in 2 of 10 studies.
Fransen 2015	single session to 30 months	1 to 5/wk	"Moderate to moderately high intensity"	15 to 60 minutes	No minimum requirement for inclusion. Varied in dose and duration.

Table 5. Dose and duration of exercise interventions in included reviews (*Continued*)

Gross 2015a	2 wk to 3 months	5/wk to every 15 minutes/day	Low intensity	2 to 20 minutes	-
Han 2004	8 to 10 wk (when reported)	1 to 7/wk (median 1/wk)	Tai chi = low intensity	1 to 1.5 hours	No minimum requirement for inclusion.
Hayden 2005	Not reported	Not reported	Not reported	Not reported	No minimum requirement for inclusion. Could not extract actual data.
Hurkmans 2009	≥ 6 wk	2/wk	Aerobic: ≥ 55% HRmax increasing to 85% HRmax strength: start 30% 1RM increasing to 80% 1RM	20 minutes	-
Koopman 2015	4 to 12 wk	Daily to 3/wk	Reported in 1 study: 50% to 70% MVC	45 minutes	No minimum requirement for inclusion. 1 study: supervised progressive resistance training consisting of 3 sets of 8 isometric contractions of the thumb muscles 1 study: combination of individual and group therapy with daily treatment in a swimming pool (45 minutes), physiotherapy, individually adapted training programme
Lane 2014	3 to 12 months	≥ 2/wk	"Variable"	~ 60 minutes	No minimum requirement for inclusion.
Lauret 2014	≥ 6 wk	≥ 2/wk	Not reported	Not reported	No minimum requirement for inclusion. Must be supervised.

Table 5. Dose and duration of exercise interventions in included reviews (*Continued*)

Regnaux 2015	8 wk	3/wk	Compared high vs low intensity vs control	30 to 50 minutes	Every 2 wk 1RM was retested and increased by 5% as tolerated in each group Supervision: an experienced therapist. 3 arms (n=34 per arm): high intensity, low intensity, control (no exercise)
Saragiotto 2016	20 days to 12 wk (median 8 wk (IQR 2.0))	1 to 5/wk (median 12 sessions (IQR 6.0))	Not reported	20 to 90 minutes (median 45 (IQR 30) minutes)	MCE is usually delivered in 1:1 supervised treatment sessions, and sometimes involves ultrasound imaging, the use of pressure biofeedback units or palpation to provide feedback on the activation of trunk muscles
Silva 2010	≥ 6 wk	2/wk	Balance training only	≥ 30 minutes	No studies found.
van der Heijden 2015	3 to 16 wk	2/wk to daily	Not reported	Not reported	No minimum requirement for inclusion. Assessed by duration (< or > 3 months), frequency (several times, or once a week), medium (land or water), etc
Yamato 2015	10 to 90 days (mostly 8 wk)	2/wk (mean session number 15.3, range 6 to 30)	Not reported	1 hour	No minimum requirement for inclusion. Must be supervised (for the Pilates technique).

1RM: one repetition maximum; ACSM: American College of Sport Medicine; HRmax: maximum heart rate; HRR: heart rate reserve, IQR: interquartile range; MCE: motor control exercise; MVC: maximum voluntary contraction; RPE: rating of perceived exertion; wk: week.

Table 6. Methodological quality of included reviews using the AMSTAR tool

Review	Criteria											Total "Y"	Total "N"	Total "n/a"
	1	2	3	4	5	6	7	8	9	10	11			
Bartels 2007	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Bidonde 2014	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N	8	3	-
Boldt 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Brown 2010	Y	Y	Y	N	Y	Y	Y	Y	n/a	N	N	7	3	1
Busch 2007	Y	Y	Y	N	Y	Y	Y	Y	Y	N	N	8	3	-
Busch 2013	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Cramp 2013	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Fransen 2014	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Fransen 2015	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	8	3	-
Gross 2015a	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Han 2004	Y	Y	Y	Y	Y	N	Y	Y	N	N	N	7	4	-
Hayden 2005	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	10	2	-
Hurkmans 2009	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9	2	-

Table 6. Methodological quality of included reviews using the AMSTAR tool (*Continued*)

Koopman 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	10	1	-
Lane 2014	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	9	2	-
Lauret 2014	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Regnaux 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10	1	-
Saragiotto 2016	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Silva 2010	Y	Y	Y	Y	Y	n/a	n/a	n/a	n/a	n/a	Y	6	0	5
van der Heijden 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9	2	-
Yamato 2015	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	9	2	-
Total "Y"	**20**	**21**	**21**	**19**	**21**	**10**	**20**	**20**	**17**	**10**	**3**	-	-	-
Total "N"	**1**	**-**	**-**	**2**	**-**	**10**	**-**	**-**	**2**	**10**	**18**	-	-	-
Total "n/a"	**-**	**-**	**-**	**-**	**-**	**1**	**1**	**1**	**2**	**1**	**-**	-	-	-

N: no; n/a: not applicable; Y: yes; out of maximum summative score of 11.
Following arbitration, the authors removed the response "cannot answer" due to no responses as such.

Table 7. Risk of bias - studies assessed as low risk of bias

Review	Number of studies in assessment	Selection bias		Performance bias	Detection bias	Attrition bias	Reporting bias	Other bias	
		Random sequence generation (studies)	Allocation concealment (studies)	Blinding of participants and personnel (studies)	Blinding of outcome assessment (studies)	Incomplete outcome data (studies)	Selective reporting (studies)	Sample size	Other biases (studies)
Bartels 2007	6	Not reported	3	Not reported	2	3	Not reported	2, n > 100 per arm	-
Bidonde 2014	9	5	3	2	8	8	5	1, n > 50 per arm	7
Boldt 2014	3	1	1	0	1	2	3	0	1
Brown 2010	1	0	0	0	0	1	1	1, n > 50 per arm	-
Busch 2007	34	17	10	8	20	Unclear	32	5, n > 50 per arm	-
Busch 2013	5	4	2	1	2	5	3	0, n > 50 per arm	-
Cramp 2013	7	5	2	0	Not reported	6	4		1
Fransen 2014	10	8	7	0	0	7	4	1, n > 50 per arm	7
Fransen 2015	54	40	22	3	4	29	10	5, total n > 200	
Gross 2015a	16	8	8	1	0	11	0	0	11
Han 2004	4	2	0	0	0	0	Not reported	0	
Hayden 2005	43	27	22	Not reported	12	29	Not reported	10, total n > 100 + 5, total n > 200	-

Table 7. Risk of bias - studies assessed as low risk of bias (*Continued*)

Hurkmans 2009	8	8	1	-	4	5	-	1, total n > 200	1
Koopman 2015	2	1	0	0	0	0	0	0	1
Lane 2014	30	16	14	30	7	19	29	3, total n > 100	
Lauret 2014	5	4	2	5	3	4	5	1, total n > 100	4
Regnaux 2015	1	1	0	0	1	0	0	1, total n > 100	1
Saragiotto 2016	7	5	4	1	1	2	7	1, total n > 100 + 1, total n > 200	7
Silva 2010	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
van der Heijden 2015	10	8	6	0	0	6	9	2, total n > 100	10
Yamato 2015	9	5	5	2	7	7	9	0	9
Studies with low risk of bias (number)	**264**	**165**	**112**	**53**	**72**	**144**	**121**	**total n > 100: 26 total n > 200: 15 total n > 400: 0**	**71**
Studies with low risk of bias (percentage)	**-**	**63%**	**42%**	**20%**	**27%**	**55%**	**46%**	**total n > 100: 10% total n > 200: 6% total n > 400: 0%**	**27%**

n: number of participants, n/a: not applicable.

Table 8. Interpretation of results by original review authors

Review	Review authors' conclusions	Overview authors' assessment of conclusions
Bartels 2007	"Aquatic exercise has some short-term beneficial effects on the condition of OA patients with hip or knee OA or both. The controlled and randomised studies in this area are still too few to give further recommendations on how to use this therapy... No long-term effects have been found."	Appropriate conclusions based on available data. No mention of quality/risk of bias in conclusions, though found to be high quality in results section
Bidonde 2014	"Low to moderate quality evidence relative to control suggests that aquatic training is beneficial for improving wellness, symptoms, and fitness in adults with fibromyalgia. Very low to low quality evidence suggests that there are benefits of aquatic and land-based exercise, except in muscle strength (very low quality evidence favoring land). No serious adverse effects were reported."	Appropriate conclusions based on available data.
Boldt 2014	"Evidence is insufficient to suggest that non-pharmacological treatments are effective in reducing chronic pain in people living with SCI. The benefits and harms of commonly used non-pharmacological pain treatments should be investigated in randomised controlled trials with adequate sample size and study methodology"	Appropriate conclusions based on available data.
Brown 2010	"There is a lack of available evidence to support the use of exercise in the alleviation of symptoms associated with dysmenorrhoea. The limited evidence implies that there are no adverse effects associated with exercise."	Review authors should not have commented on lack of adverse events as this was not reported in the included study. The comment on lack of adverse events contravened present Cochrane guidance
Busch 2007	"There is moderate quality evidence that short-term aerobic training (at the intensity recommended for increases in cardiorespiratory fitness) produces important benefits in people with FM in global outcome measures, physical function, and possibly pain and tender points. There is limited evidence that strength training improves a number of outcomes including pain, global wellbeing, physical function, tender points and depression. There is insufficient evidence regarding the effects of flexibility exercise. Adherence to many of the aerobic exercise interventions described in the included studies was poor."	Appropriate conclusions based on available data.
Busch 2013	"We have found evidence in outcomes representing wellness, symptoms, and physical fitness favoring resistance training over usual treatment and over flexibility exercise, and favoring aerobic training over re-	Appropriate conclusions based on available data.

Table 8. Interpretation of results by original review authors (*Continued*)

	sistance training. Despite large effect sizes for many outcomes, the evidence has been decreased to low quality based on small sample sizes, small number of randomized clinical trials (RCTs), and the problems with description of study methods in some of the included studies."	
Cramp 2013	"There is some evidence that physical activity interventions ... may help to reduce fatigue in RA. However, the optimal parameters and components of these interventions are not yet established."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion despite low/unclear quality score in results and discussion sections No conclusions about effect on pain (insufficient data).
Fransen 2014	"There is currently high-level evidence that land-based exercise will reduce hip pain, and improve physical function, among people with symptomatic hip osteoarthritis."	Evidence was good quality though sample sizes were often small (i.e. it is debatable if this was high level evidence as claimed by authors). Agree that results demonstrate small but significant benefit from intervention
Fransen 2015	"High-quality evidence suggests that land-based therapeutic exercise provides benefit in terms of reduced knee pain and quality of life and moderate-quality evidence of improved physical function among people with knee OA… Despite the lack of blinding we did not downgrade the quality of evidence for risk of performance or detection bias."	Appropriate conclusions based on available data. May have been generous with quality assessment but this was stated in conclusions for transparency
Gross 2015a	"…there is still no high quality evidence and uncertainty about the effectiveness of exercise for neck pain… Moderate quality evidence supports the use specific strengthening exercises as a part of routine practice … Moderate quality evidence supports the use of strengthening exercises, combined with endurance or stretching exercises may also yield similar beneficial results. However, low quality evidence notes when only stretching or only endurance type exercises … there may be minimal beneficial effects for both neck pain and function."	Appropriate conclusions based on available data.
Han 2004	"Tai chi appears to have no detrimental effects on the disease activity of RA in terms of swollen/tender joints and activities of daily living…tai chi appears to be safe, since only 1 participant out of 121 withdrew due to adverse effects and withdrawals were greater in the control groups than the tai chi groups."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias in conclusion despite very low quality score in results section
Hayden 2005	"Evidence from randomized controlled trials demonstrates that exercise therapy is effective at reducing	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of stud-

Table 8. Interpretation of results by original review authors (*Continued*)

	pain and functional limitations in the treatment of chronic low-back pain, though cautious interpretation is required due to limitations in this literature."	ies in conclusion despite low quality score in results and discussion sections
Hurkmans 2009	"Short-term, land-based dynamic exercise programs have a positive effect on aerobic capacity (aerobic capacity training whether or not combined with muscle strength training) and muscle strength (aerobic capacity training combined with muscle strength training) immediately after the intervention, but not after a follow-up period. Short-term, water-based dynamic exercise programs have a positive effect on functional ability and aerobic capacity directly after the intervention but it is unknown whether these effects are maintained after follow-up. Long-term, land-based dynamic exercise programs (aerobic capacity and muscle strength training) have a positive effect on functional ability, aerobic capacity, and muscle strength immediately after the intervention but it is unknown whether these effects are maintained after follow-up... Based on the evidence, aerobic capacity training combined with muscle strength training is recommended for routine practice in patients with RA."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion No conclusions regarding pain severity.
Koopman 2015	"Data from two single trials suggested that muscle strengthening of thumb muscles (very low-quality evidence) ... are safe and beneficial for improving muscle strength ... with unknown effects on activity limitations." "We found evidence varying from very low quality to high quality that ... rehabilitation in a warm or cold climate are not beneficial in PPS." "Due to a lack of good-quality data and randomised studies, it was impossible to draw definitive conclusions about the effectiveness of interventions in people with PPS."	Appropriate conclusions based on available data.
Lane 2014	"… Exercise therapy should play an important part in the care of selected patients with intermittent claudication, to improve walking times and distances. Effects were demonstrated following three months of supervised exercise although some programmes lasted over one year."	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion No conclusions regarding pain severity.
Lauret 2014	"There was no clear evidence of differences between supervised walking exercise and alternative exercise modes in improving the maximum and pain-free walking distance of patients with intermittent claudication…. The results indicate that alternative ex-	Appropriate conclusions based on available data. However, no mention of quality/risk of bias of studies in conclusion (in discussion)

Table 8. Interpretation of results by original review authors (*Continued*)

	ercise modes may be useful when supervised walking exercise is not an option for the patient."	
Regnaux 2015	"We found very low- to low-quality evidence for no important clinical benefit of high-intensity compared to low-intensity exercise programs in improving pain and physical function in the short term.... The included studies did not provide any justification for the levels of intensity of exercise programs. No authors reported evidence for the minimal and maximal intensity that could be delivered."	Appropriate conclusions based on available data. This overview has only used one study of the six included as it alone included a control group, for which we could not extract data as the control comparison was not used in the analysis by the review authors
Saragiotto 2016	"There is very low to moderate quality evidence that MCE has a clinically important effect compared with a minimal intervention for chronic low back pain... As MCE appears to be a safe form of exercise and none of the other types of exercise stands out, the choice of exercise for chronic low back pain should depend on patient or therapist preferences, therapist training, costs and safety."	Appropriate conclusions based on available data.
Silva 2010	"We were not able to provide any evidence to support the application of balance exercises (proprioceptive training) alone in patients with RA."	Appropriate conclusions based on available data (no included studies)
van der Heijden 2015	"This review has found very low quality but consistent evidence that exercise therapy for patellofemoral pain syndrome (PFPS) may result in clinically important reduction in pain and improvement in functional ability."	No subgroup analysis to differentiate between acute, subacute, and chronic pain made it difficult to extract appropriate data for this review
Yamato 2015	"No definite conclusions or recommendations can be made as we did not find any high quality evidence for any of the treatment comparisons, outcomes or follow-up periods investigated. However, there is low to moderate quality evidence that Pilates is more effective than minimal intervention in the short and intermediate term as the benefits were consistent for pain intensity and disability, with most of the effect sizes being considered medium."	Appropriate conclusions based on available data. There was no subgroup analysis to differentiate between acute, subacute, and chronic pain made it difficult to extract appropriate data for this review (one included study had subacute back pain (> 6 weeks), all others were chronic back pain (> 12 weeks)) but results are presented altogether as chronic pain

FM: fibromyalgia; MCE: motor control exercise; OA: osteoarthritis; PPS: postpolio syndrome; RA: rheumatoid arthritis; SCI: spinal cord injury.

Table 9. Pain severity

Review	Number of trials (and participants) assessing 'pain severity'	Baseline pain score	Post-intervention reported result or change data (or if only one data point reported in review)	Follow-up	Overall comment/ statement
Bartels 2007 (osteoarthritis)	Hip + knee OA: Post-intervention: 4 (638) Follow-up: 1 (310) Hip only: follow-up: 1 (17) Knee only: post-intervention: 1 (46)	Control baseline: Hip + knee OA WOMAC 0 to 20 (2 studies): 9.10 (SD 3.14) VAS 0 to 100 (1 study): 55.3 (SD 24.6) HAQ 0 to 3 (1 study): 1.05 (SD 0.61) Hip only VAS 0 to 100 (1 study): 56 (SD 21.89) Knee only VAS 0 to 10 (1 study): 5.6 (SD 1.4)	Hip + knee OA A minor effect of a 3% absolute reduction (0.6 fewer points on WOMAC 0 to 20 scale) and 6.6% relative reduction SMD 0.19 (95% CI 0.04 to 0.35) (P = 0.02) Knee only SMD 0.86 (95% CI 0.25 to 1.47) (P = 0.005) Absolute difference 12% (1.2 fewer points on a 0 to 10 scale) Relative change 22% improvement	Hip + knee OA Follow-up at 6 months: SMD 0.11 (95% CI -0.12 to 0.33) (ns) No difference Hip only SMD 1.00 (95% CI -0.04 to 2.04) (P = 0.06, ns)	Statistically significant post-intervention in hip + knee OA group, but not clinically significant Knee-only OA had moderate to large effect size (statistically significant) immediately post-intervention
Bidonde 2014 (fibromyalgia)	Post-intervention: 7 (382)	Weighted mean score at baseline (all participants): 69.59 median value for pain was 70.9 in studies comparing aquatic training to control	On 100-point scale: MD -6.59 (95% CI -10.71 to -2.48) SMD -0.53 (95% CI -0.76 to -0.31) Absolute difference -7% (95% CI -11 to -3) NNTB 5 (95% CI 3 to 8)	3 studies at 12, 48, or 52 weeks' post-intervention could not be combined. 2 studies showed SMD favouring intervention at follow-up.	"We found a moderate effect favouring the aquatic exercise training for pain" …"similar improvements in pain in the low pain groups (SMD -0.60, 95% CI -0.98 to -0.23) and in the high pain groups (SMD -0.57, 95% CI -1.11 to -0.03)." Among the major wellness outcomes, none of the outcomes met the threshold for clinically relevant differences (15%)

Table 9. Pain severity (*Continued*)

Boldt 2014 (spinal cord injury)	Post-intervention: 3 (149)	WUSPI score 22.6 (exercise group) to 11.05 (control group) in 1 group at baseline Not reported for 2 studies	WUSPI change score: Exercise group: -7.7 (SD 19.01) Control group: 12.8 (SD 12.74) SF-36 (pain experience): -1.9 (95% CI -3.4 to -0.4) favoured exercise (P = 0.01) VAS (0 to 10): MD -2.8 (95% CI -3.77 to -1.83) favoured exercise (P < 0.00001)	1 study at 4 weeks: VAS (0 to 10): -2.50 (95% CI -3.48 to -1.52) (P < 0.00001) WUSPI: -26.40 (95% CI -37.62 to -15.18favoured exercise (P < 0.00001)	"All three studies were fraught with high overall risk of bias. In particular, the comparison with 'no treatment' or waiting lists as control interventions likely leads to an overestimation of the effectiveness of the exercise programmes provided in these studies. Consequently, no conclusion on their effectiveness can be drawn."
Busch 2007 (fibromyalgia)	Strength training: 1 (21) Aerobic training: 3 (183)	Control baseline: Aerobic: 6.1/10 (VAS) (SD 1.97) Strength: 35/100 (VAS) (SD 19)	Aerobic training: SMD 0.65 (95% CI -0.09 to 1.39) (ns) Weighted absolute change 13% (1.3 cm lower on 10-cm scale) Relative change 21% Strength training: SMD 3.00 (95% CI 1.68 to 4.32) (ns) Weighted absolute change 49% (49 points lower on 100-point scale) Relative change 140%, NNTB 2	n/a	">30% improvement was seen in the strength training group as compared to an untreated control group in pain." Aerobic training led to an improvement of 1.3/10.
Busch 2013 (fibromyalgia)	Post-intervention: 2 (81) Follow-up at 8 weeks, 16 weeks, 28 weeks: 1 (60)	Not reported - change data only	Change score on VAS (in cm): MD -3.30 (95% CI -6.35 to -0.26) (P = 0.03) SMD -1.89 (95% CI -3.86 to 0.07) Relative % change	8 weeks: MD -0.68 (95% CI -1.62 to 0.26) (ns) 16 weeks: MD -1.79 (95% CI -2.70 to -0.88) (P < 0.001) 28 weeks: MD -0.85 (95% CI -1.77 to 0.	> 30% improvement post-intervention.

Table 9. Pain severity (*Continued*)

			44.6% (95% CI 3.5 to 85.9) favoured exercise NNTB 2 (95% CI 1 to 34)	07) (P = 0.07, ns) Overall (n = 180): MD -1.12 (95% CI -1.65 to -0.58) (P < 0.0001)	
Cramp 2013 (rheumatoid arthritis)	4 (not reported)	Not reported	In narrative only - Harkcom 1985: statistics not reported separately for pain data, but reported as improvement over time; Hakkinen 2003: "stat significant improvement in 24 months"; Evans 2012 and Wang 2008: no statistically significant effects	Not reported	"Improvement over time" with "significant improvement in 24 months." No actual data available.
Fransen 2014 (OA)	End of treatment: 9 (549) 3 to 6 months: 5 (391)	Not reported; land based exercise vs no exercise: mean pain in control group ~ 29/100 (based on 9 studies' control values)	End of treatment: SMD -0.38 (95% CI -0.55 to -0.20) "small to moderate" favoured exercise (P < 0.0001)	3 to 6 months: SMD -0.38 (95% CI -0.58 to -0.18) "small to moderate" favoured exercise (P = 0.0002)	"Small to moderate" statistically significant improvement, but only mild pain at baseline
Fransen 2015 (OA)	End of treatment: 44 (3537) Follow-up (2 to 6 months): 12 (1468) Follow-up (> 6 months): 8 (1272)	Not reported; land-based exercise vs no exercise: mean pain in control group 44/100 (based on 1 study control values)	Land-based exercise vs no exercise: Mean pain in intervention groups was 0.49 SDs lower (95% CI 0.39 to 0.59 lower). This translates to an absolute mean reduction of 12 points (95% CI 10 to 15) compared with control group on a 0 to 100 scale SMD -0.49 (95% CI -0.39 to -0.59) (P < 0.00001) Absolute reduction	2 to 6 months: SMD -0.24 (95% CI -0.35 to -0.14) favoured exercise (P < 0.00001) > 6 months: SMD -0.52 (95% CI -1.01 to -0.03) favoured exercise (P = 0.04)	Absolute improvement of 12/100 post-intervention (statistically significant)

Table 9. Pain severity (*Continued*)

			12% (95% CI 10% to 15%) Relative change 27% (95% CI 21% to 32%) NNTB 4 (95% CI 3 to 5)		
Gross 2015a (mechanical neck disorders)	12-week treatment: 2 (147) 24 week (or 12-week treatment + 12-week follow-up) : 2 (140)	Not reported, but control scores at end of treatment 40 to 60/100 (moderate pain)	12 weeks: pooled MD -14.90 (95% CI -22.40 to -7.39) favoured exercise (P = 0.0001)	24 weeks: pooled MD -10.94 (95% CI -18.81 to -3.08) favoured exercise (P = 0.0064)	2 trials showed a moderate (statistically significant) reduction in pain post-intervention (14.9/100)
Hayden 2005 (low back pain)	Earliest follow-up: 8 (370) Follow-up (time since randomisation) Short term (6 weeks): 6 (268) Intermediate term (6 months): 5 (249) Long term (12 months): 2 (126)	"Chronic group" at baseline: mean 46/100 (95% CI 41 to 50) (moderate pain)	Earliest: MD -10.20 (95% CI -19.09 to -1.31) (P = 0.02)	Short term: MD -8.58 (95% CI -18.46 to 1.29) (P = 0.09, ns) Intermediate term: MD -12.48 (95% CI -22.69 to -2.27) (P = 0.02) Long term: MD -3.93 (95% CI -9.89 to 2.02) (P = 0.2, ns)	Reduction of ~ 10/100 at earliest measurement point.
Hurkmans 2009 (rheumatoid arthritis)	4 studies (total 188 participants) in different categories (results not combined)	Not reported	Short-term (12 weeks): Short-term land-based (aerobic and strength training) SMD -0.53 (95% CI -1.09 to 0.04) Short-term land-based (aerobic only) SMD -0.27 (95% CI -0.79 to 0.26) Short-term water-based SMD 0.06 (95% CI -0.43 to 0.54)	Long-term (24 months) land-based (aerobic and strength training) SMD 0.35 (95% CI -0.46 to 1.16)	No significant difference between control and intervention.
Koopman 2015 (postpolio syndrome)	1 (55)	Not reported, but control scores at end of treatment mean 44 (SD 24) on a 0 to 100 scale (moderate	3 months post-intervention: VAS (0 to 100): MD 11.00 (95% CI -0.98 to 22.98) (P = 0.	n/a	No significant effect/no difference between groups.

Table 9. Pain severity (*Continued*)

		pain)	072)		
Regnaux 2015 (OA)	Only 1 study that had a no-exercise control: 1 (68) - excluded data for control (no exercise) from analysis (n = 34)	Not reported	Post-intervention: WOMAC (0 to 20) Change data presented for high- vs low-intensity groups only, not compared to control	n/a	Actual individual study data was extracted (where possible) instead of pooled MD or SMD due to comparison this overview wishes to make (exercise vs no-exercise only) Could not extract exercise vs control data.
Saragiotto 2016 (low back pain)	Short term (< 3 months): 4 (291) Intermediate term (3 to 12 months): 4 (348) Long term (> 12 months): 3 (279)	Not reported, but control scores at follow-up range 25 to 56/100 (mild-moderate pain)	Short term: MD -10.01 (95% CI -15.67 to -4.35) favoured exercise (P < 0.001)	Intermediate term: MD -12.61 (95% CI -20.53 to -4.69) favoured exercise (P = 0.002) Long term: MD -12.97 (95% CI -18.51 to -7.42) favoured exercise (P < 0.001)	Medium effect size favouring exercise at all follow-up assessments (moderate quality evidence at short- and long-term, low quality evidence at intermediate term) Clinically important effect.
van der Heijden 2015 (patellofemoral pain syndrome)	3 studies with pain > 3 months (135 participants), 2 studies used in analysis (41 participants) Long-term follow-up: 1 (94)	Not reported, but control scores at follow-up range 2.1 to 6.0/10 (mild-moderate pain)	Short-term (4 to 8 weeks): MD for usual pain in the exercise group was 0.93 (95% CI 1.60 to 0.25) SDs lower SMD -0.93 (95% CI -1.60 to -0.25) (P = 0.008)	"Long term" (16 weeks) VAS (0 to 10): MD -4.42 (95% CI -7.75 to -0.89) favoured exercise (P = 0.01)	Reduction in pain of 4/10 at 16 weeks' follow-up.
Yamato 2015 (low back pain)	Short term: 6 (265) Intermediate term: 2 (148)	Not reported, but control scores at earliest follow-up range 18 to 52/100 (mild-moderate pain)	Short-term follow-up (< 3 months): MD -14.05 (95% CI -18.91 to -9.19) (P < 0.001)	Intermediate term (3 to 12 months): MD -10.54, (95% CI -18.54 to -2.62) (P = 0.009)	"Low quality evidence (downgraded due to imprecision and risk of bias) that Pilates reduces pain compared with minimal intervention at short-term follow-up, with a medium effect size..

Table 9. Pain severity (*Continued*)

					intermediate-term follow-up, two trials, provided moderate quality evidence (downgraded due to imprecision) that Pilates reduces pain compared with minimal intervention, with a medium effect size"

CI: confidence interval; HAQ: Health Assessment Questionnaire; MD: mean difference; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; ns: not significant; OA: osteoarthritis; SD: standard deviation; SF-36: 36-item Short Form; SMD: standardised mean difference; VAS: visual analogue score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; WUSPI; Wheelchair User Shoulder Pain Index.

Table 10. Physical function

Review	Outcome measure	Number of trials (and participants) used in analysis	Post-intervention result (or if only 1 result reported)	Short-term follow-up (or if only 1 follow-up point reported)	Intermediate-term follow-up	Long-term follow-up	Overall comment/ statement
Bartels 2007 (OA)	Self-reported function (WOMAC and HAQ) and walking ability, and DRI	Post-intervention Hip + knee function: 4 (648) walking ability: 2 (355) Hip only function: 1 (28) Follow-up function hip + knee: 1 (306) hip only: 1 (17)	Function (hip + knee): SMD 0.26 (95% CI 0.11 to 0.42) favoured exercise (P < 0.001) Walking (hip + knee): SMD 0.18 (95% CI -0.03 to 0.39) favoured exercise (P = 0.08, ns) Function (hip only): SMD 0.76 (95% CI -0.02 to 1.53) favours exercise (P = 0.06, ns)	Hip only Disability, SMD 1.00 (95% CI -0.04 to 2.04) favoured exercise (P = 0.06, ns)	Hip + knee (6 months) Function, SMD 0.10 (95% CI -0.12 to 0.33) (ns)	n/a	Function was significantly improved in people with hip + knee OA immediately post-intervention only - small effect size only

Table 10. Physical function (*Continued*)

Bidonde 2014 (fibromyalgia)	Self-reported physical function (0 to 100 scale)	5 (285)	MD -4.35 (95% CI -7.77 to -0.94) SMD -0.44 (95% CI -0.76 to -0.11) Absolute difference -4 (95% CI -8 to -1) NNTB 6 (95% CI 3 to 22)	n/a	n/a	n/a	Small difference (improvement) in aquatic exercise group. Among the major wellness outcomes, none of the outcomes met the threshold for clinically relevant differences (15%)
Busch 2007 (fibromyalgia)	Physical function	Aerobic: 4 (253) Strength: 2 (47)	Aerobic: SMD 0.66 (95% CI 0.41 to 0.92) favoured exercise (P < 0.0001) Strength: SMD 0.52 (95% CI -0.07 to 1.10) favoured exercise (P = 0.08, ns)	n/a	n/a	n/a	Function was significantly improved from aerobic exercise training, strength training neared significance Moderate effect size.
Busch 2013 (fibromyalgia)	HAQ and SF-36 for function	3 (107)	Change score MD -6.29 (95% CI -10.45 to -2.13) favoured exercise (P < 0.01)	n/a	n/a	n/a	Significantly favourable effect of exercise.
Cramp 2013 (rheumatoid arthritis)	Disability	4 (not reported)	n/a	n/a	n/a	n/a	"Studies investigating hydrotherapy and tai chi demonstrated statistically significant improvements in the intervention arm compared to the con-

Table 10. Physical function (*Continued*)

							trol arm between baseline and follow-up. The studies investigating strength training and Ivengar yoga did not demonstrate a statistically significant difference between study arms."
Fransen 2014 (OA)	Physical function	Post-intervention: 9 (521) Follow-up (3 to 6 months): 5 (365)	SMD -0.30 (95% CI -0.54 to -0.05) "significant benefit" favoured exercise (P = 0.02) The demonstrated effect size for exercise was equivalent to an improvement of physical function of 7 points (95% CI 1 to 12) on a 0 to 100 scale compared with a control group	SMD -0.37 (95% CI -0.57 to -0.16) favoured exercise (P < 0.001)	n/a	n/a	Statistically significant, but small effect size only.
Fransen 2015 (OA)	Physical function	Post-intervention: 44 (3913) Follow-up (2 to 6 months): 10 (1279) Follow-up (> 6 months): 8 (1266)	SMD -0.52 (95% CI -0.64 to -0.39) favoured exercise (P < 0.0001); an improvement of 10 points (95% CI 8 to 13) on a 0- to 100-point scale	SMD -0.15 (95% CI -0.26 to -0.04) favoured exercise (P = 0.008)	SMD -0.57 (95% CI -1.05 to -0.10) favoured exercise (P = 0.02)	n/a	Significant effect from exercise at every follow-up point. Moderate effect size at short- and long-term follow-up, but only small ef-

Table 10. Physical function (*Continued*)

							fect at intermediate-term follow-up
Gross 2015a (mechanical neck disorders)	Physical function	12 wk: 2 (147) 24 wk: 2 (140)	12 wk treatment: pooled SMD -0.50 (95% CI -1.04 to 0.03) favoured exercise (P = 0.07, ns)	24 wk treatment (or 12 wk' treatment + 12 wk follow-up) : pooled SMD -0.40 (95% CI -0.74 to -0.06) favoured exercise (P = 0.02)	n/a	n/a	2 trials showed a moderate (statistical) improvement in function
Han 2004 (rheumatoid arthritis)	Functional assessment and 50-feet walk test	Function: 2 (52) Walk test: 2 (48)	Function: MD 0.01 (95% CI -2.94 to 2.97) (ns) Walk test: MD 0.35 seconds (95% CI -1.14 to 1.84) (ns)	n/a	n/a	n/a	No significant effect.
Hayden 2005 (low back pain)	Function	Earliest: 7 (337) Short term: 6 (268) Intermediate term: 4 (216) Long term: 2 (126)	Earliest: MD -2.98 (95% CI -6.48 to 0.53) favoured exercise (P = 0.09, ns)	Short term: MD -3.03 (95% CI -6.35 to 0.53) favoured exercise (P = 0.07, ns)	Intermediate term: MD -3.84 (95% CI -7.06 to -0.61) favoured exercise (P = 0.02)	Long term: MD -4.22 (95% CI -7.99 to -0.46) favoured exercise (P = 0.03)	Favoured exercise from the earliest measure, but only reached statistical significance at intermediate and long term after randomisation
Hurkmans 2009 (rheumatoid arthritis)	Functional ability	Land-based aerobic: 2 (66) Land-based aerobic + strength: 2 (74)	n/a	Short-term training (12 wk) Land-based aerobic only training SMD 0.03 (95% CI -0.46 to 0.51) (ns) Land-based aerobic and	n/a	n/a	No significant difference between control and intervention groups

Table 10. Physical function (*Continued*)

				strength training SMD -0.4 (95% CI -0.86 to 0.06) (ns)			
Koopman 2015 (postpolio syndrome)	Muscle strength; and activity limitation (Sunnaas ADL-index range 0 to 36; Rivermead Mobility Index (RMI) range 0 to 15)	Strength: 1 (10) Activity limitation: 1 (53)	Isometric muscle strength (postintervention): MD 39.00% (95% CI 6.12 to 71.88) Activity limitation: 3 months' postintervention: ADL-index: MD -2.70 (95% CI -4.53 to -0.87) Rivermead Mobility Index (RMI): MD -1.50 (95% CI -2.93 to -0.07)	Activity limitation: 6-months postintervention: ADL-index: MD -2.90 (95% CI -4.73 to -1.07) RMI: MD -1.80 (95% CI -3.19 to -0.41)	n/a	n/a	Activity limitation: favoured intervention at both assessment points "The baseline imbalance in favour of the usual care group probably biased these results."
Lane 2014 (intermittent claudication)	Maximal walking time and maximal walking distance	Postintervention Walking time: 12 (577) Walking distance: 9 (480) 3-month follow-up Walking time: 5 (174) Walking distance: 3 (116) 6-month follow-up Walking time: 4 (295) Walking distance: 3 (156)	Time: MD 4.51 minutes (95% CI 3.11 to 5.92) favoured exercise (P < 0.00001) Distance: 108.99 m (95% CI 38.20 to 179.78) favoured exercise (P = 0.003)	Time: MD 6.05 minutes (95% CI 5.47 to 6.62) favoured exercise (P < 0.00001) Distance: MD 104.46 m (95% CI -64.33 to 273.24) favoured exercise (ns)	Time: MD 3.20 minutes (2.04 to 4.36) favoured exercise (P < 0.0001) Distance: MD 138.36 m (95% CI 22.39 to 254.34) favoured exercise (P = 0.02)	n/a	Objectively measured walking time and distance showed significant improvement
Lauret 2014 (intermittent claudication)	Maximal walking time	No relevant studies	n/a	n/a	n/a	n/a	No relevant studies.

Table 10. Physical function (*Continued*)

	(mins) and maximal walking distance (metres)						
Regnaux 2015 (OA)	WOMAC (0 to 68) disability scale, and muscle strength	1 (68) - excluded control (no-exercise data: n = 34)	n/a	n/a	n/a	n/a	Could not extract exercise vs control data - data presented for high vs low intensity groups only, not compared to control
Saragiotto 2016 (low back pain)	Disability (Oswestry Disability Index, Roland Morris Disability Questionnaire)	Short-term follow-up (< 3 months): 5 (332) Intermediate term (3 to 12 months): 4 (348) Long term (> 12 months): 3 (279)	-	MD -8.63 (95% CI -14.78 to -2.47) (P < 0.01)	MD -5.47 (95% CI -9.17 to -1.77) (P = 0.004)	MD -5.96 (95% CI -9.81 to -2.11) (P = 0.002)	Small effect sizes, favoured exercise. Short term: CI included a clinically important effect.
Silva 2010 (rheumatoid arthritis)	HAQ function	No studies found	n/a	n/a	n/a	n/a	No studies found.
van der Heijden 2015 (patellofemoral pain syndrome)	Functional ability	Short-term follow-up: 7 (483) Long-term follow-up: 3 (274)	n/a	Short-term (4 to 8 wk): SMD 1.10 (95% CI 0.58 to 1.63) favoured exercise (P < 0.0001)	n/a	SMD 1.62 (95% CI 0.31 to 2.94) favoured exercise (P = 0.02)	Significant effect of exercise. Very large effect size at short- and long-term follow-up.
Yamato 2015 (low back pain)	Disability (all measures converted to 0 to 100 scale)	Short-term (< 3 months) follow-up: 5 (248) -Intermediate-term (3 to 12 months) follow-up: 2	n/a	MD -7.95 (95% CI -13.23 to -2.67) (P = 0.003)	MD -11.17 (95% CI -18.41 to -3.92) (P = 0.0025)	n/a	"Low quality evidence (downgraded due to imprecision and inconsistency) that Pilates improves

Table 10. Physical function (*Continued*)

		(146)					disability at short-term follow-up compared with minimal intervention, with a small effect size ... intermediate-term follow-up, two trials provided moderate quality evidence (downgraded due to imprecision) of a significant effect in favour of Pilates, with a medium effect size"

ADL: activities of daily living; CI: confidence interval; DRI: Disability Rating Index; HAQ: Health Assessment Questionnaire; MD: mean difference; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; ns: not significant; OA: osteoarthritis; SF-36: 36-item Short Form; SMD: standardised mean difference; wk: week; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index,

Table 11. Psychological function

Review	Outcome measure	Number of trials (and participants) reporting psychological function	Outcome result (postintervention or if only one measurement point)	Follow-up	Additional statement/comment
Mental health					
Bartels 2007	-	4 studies	SMD 0.16 (95% CI 0.01 to 0.032) favoured aquatic exercise	No significant difference at 6 months, 1 study	Very small effect size postintervention.
Busch 2013	SF-36 - Mental health scale	1 study	-	n/a	No group differences.

Table 11. Psychological function (*Continued*)

Bidonde 2014	SF-36 - mental Health scale SF-12 - Mental Health scale	4 studies, n = 243	MD -3.03 (95% CI -8.06 to 2.01)	n/a	No effect.
Anxiety					
Cramp 2013	Brief Symptom Inventory	1 study	"No significant effect"	n/a	-
Depression					
Boldt 2014	CES-D	1 study, n = 34	MD -6.0 (95% CI -15.87 to 3.87) (P = 0.23)	n/a	No effect.
Busch 2013	HADS - Depression Beck Depression Index	1 study, n = 21	MD -3.70 (95% CI -6.37 to -1.03) Relative difference 57%	n/a	Significant effect, favoured resistance training.
Cramp 2013	CES-D	Not reported	"Variable effect" reported in text only	n/a	-

CES-D: Centre for Epidemiological Studies-Depression; CI: confidence interval; HADS: Hospital Anxiety and Depression Scale; MD: mean difference; n: number of participants; n/a: not applicable; SF-12: 12-item Short Form; SF-36: 36-item Short Form; SMD: standardised mean difference.

Table 12. Quality of life

Review	Outcome measure	Number of trials (and participants) reporting Quality of Life (QoL)	Outcome result	Additional statement/ comment
(Health-related) Quality of Life				
Bartels 2007	QoL: SF-12 (Physical), PQoL, EuroQoL	Hip + knee OA (post-intervention): 3 studies, n = 599 Hip only OA (post-intervention): 1 study, n = 28 Hip only OA (follow-up): 1 study, n = 17	Hip + knee (post-intervention): SMD 0.32 (95% CI 0.03 to 0.61) (P = 0.028) Hip only (post-intervention): SMD 0.76 (95% CI -0.02 to 1.53) (ns) Hip only (follow-up): SMD 1.00 (95% CI -0.04 to 2.04) (ns)	Significantly favoured aquatic exercise post-intervention in hip + knee OA Small effect size only (when statistically significant).

Table 12. Quality of life (*Continued*)

Boldt 2014	PQoL (perceived quality of life) SQoL (subjective quality of life)	Post-intervention: 1 study, n = 34, PQoL; 1 study, n = 80, SQoL Follow-up (intermediate term): 1 study, n = 80, SQoL	Post-intervention: PQoL MD 10.8 (95% CI -4.2 to 25.8) (P = 0.16) SQoL MD 0.3 (95% CI -0.22 to 0.82) (P = 0.25) Follow-up: SQoL MD 0.5 (95% CI -0.03 to 1.03) (P = 0.07)	No difference between groups.
Fransen 2014	QoL	Post-intervention: 3 studies, n = 183	SMD 0.07 (95% CI -0.23 to 0.36) (ns)	No difference between groups.
Fransen 2015	QoL: self-report questionnaire, scale 0 to 100 (100 is maximum QoL)	Post-intervention: 13 studies, n = 1073	SMD 0.28 (95% CI 0.15 to 0.40) (P < 0.0001) Absolute difference 4% (95% CI 2% to 5%) relative difference 9% (95% CI 5% to 13%)	Statistically significant, but equates to an absolute improvement of 4 points (95% CI 2 to 5) on a 0 to 100 scale Small effect size only.
Gross 2015a	QoL: SF-36 (Physical Function subscale)	Post-intervention: 2 studies, n = 143	12-wk intervention: MD -2.22 (95% CI -5.17 to 0.72) (ns) 24-wk intervention: MD 0.06 (95% CI -4.06 to 4.17) (ns)	No significant difference between groups.
Lauret 2014	HRQoL	No relevant studies	n/a	n/a
Global assessment				
Busch 2007	Global wellbeing	Strength: 2 studies, n = 47 Aerobic: 4 studies, n = 269	Strength: SMD 1.43 (95% CI 0.76 to 2.10) Aerobic: SMD 0.49 (95% CI 0.23 to 0.75)	Favoured exercise - higher score showed better QoL, Strength: very large effect size. Aerobic: small-to-moderate effect size only.
Bidonde 2014	Participant-rated global (10-cm VAS)	1 study, n = 46	MD -0.87 (95% CI -1.74 to 0.00)	No effect.
Gross 2015a	Global perceived effect	1 study, n = 70	"No significant difference"	No significant difference.
Hayden 2005	Global assessment	7 studies, n = 16	Not reported	n/a
Saragiotto 2016	Global impression of recovery	1 study, n = 154	Short term, MD 1.30 (95% CI 0.30 to 2.30) (P = 0.01) Intermediate term, MD 1.	Medium effect size.

Table 12. Quality of life (*Continued*)

			20 (95% CI 0.31 to 2.09) (P = 0.008) Long term, MD 1.50 (95% CI 0.61 to 2.39) (P < 0.001)	
Yamato 2015	Global impression of recovery	1 study, n = 86	Short term (< 3 months): MD 1.50 (95% CI 0.70 to 2.30) Intermediate term (3 to 12 months): MD 0.70 (95% CI -0.11 to 1.51)	"Low quality evidence (downgraded due to imprecision and inconsistency), we found a significant short-term effect, with a small effect size, but not for intermediate/mid-term follow up."
Other method of assessment				
Bidonde 2014	Multi-dimensional function- FIQ	7 studies, n = 367	MD -5.97 (95% CI -9.06 to -2.88) SMD -0.55 (95% CI -0.83 to -0.27) Absolute difference -6 (95% CI -9 to -3) NNTB 5 (95% CI 3 to 9)	Favoured aquatic exercise - lower score showed reduced impact of pain on life "Moderate difference."
Busch 2013	Multi-dimensional function - FIQ	1 study, n = 60	SMD -1.27 (95% CI -1.83 to -0.72) Absolute difference -16.75 FIQ units (95% CI -23.31 to -10.19)	Favoured exercise - lower score showed reduced impact of pain on life Very large effect size.
Hayden 2005	Work status	9 studies, n = 21	Not reported	n/a
Silva 2010	Health Assessment Questionnaire (HAQ)	No included studies	n/a	n/a

FIQ: Fibromyalgia Impact Questionnaire; HRQoL: health-related quality of life; MD: mean difference; n: number of participants; n/a: not applicable; NNTB: number needed to treat for an additional beneficial outcome; OA: osteoarthritis; PQoL: perceived quality of life; QoL: quality of life; SF-36: 36-item Short Form; SMD: standardised mean difference; SQoL: subjective quality of life; VAS: visual analogue scale.

Table 13. Adherence/withdrawals

Review	Number of trials (and participants) reporting withdrawals	Number withdrawn (per 1000) - intervention group	Number withdrawn (per 1000) - control group	RR or OR

Table 13. Adherence/withdrawals (*Continued*)

Bidonde 2014 (fibromyalgia)	8 studies, n = 472	151 (imputed from reported 38/252)	129 (imputed from reported 30/232)	RR 1.13 (95% CI 0.73 to 1.77) (P = 0.45)
Busch 2013 (fibromyalgia)	3 studies, n = 107	134 (95% CI 30 to 439)	39	RR 3.50 (95% CI 0.79 to 15.49)
Fransen 2014 (osteoarthritis)	7 studies, n = 715	59 (95% CI 30 to 114)	34	OR 1.77 (95% CI 0.86 to 3.65)
Han 2004 (rheumatoid arthritis)	4 studies, n = 189	109 (imputed from reported 11/101)	284 (imputed from reported 25/88)	RR 0.37 (95% CI 0.19 to 0.72)
Regnaux 2015 (osteoarthritis)	1 study, n = 102	44 (imputed from reported 3/68 (4%) ; all from high-intensity group)	0	Calculated RR 3.55 (95% CI 0.19 to 66.8)
Saragiotto 2016 (low back pain)	7 studies, n = 671	0	0	-
Silva 2010 (rheumatoid arthritis)	No included studies	n/a	n/a	n/a
Total	**30 studies, n = 2256**	**82.8/1000**	**81/1000**	**Calculated RR 1.02 (95% CI 0.94 to 1.12) Calculated OR 1.05 (95% CI 0.88 to 1.25)**

CI: confidence interval; n: number of participants; n/a: not applicable; OR: odds ratio; RR: risk ratio.

Table 14. Adverse events (not death)

Review	Total number of trials (and participants) in review reporting exercise vs control in chronic pain population	Number of trials (and participants) reporting adverse events	Number of adverse events	Overall statement
Bartels 2007	4 (674)	2 (148)	0	Adverse events were recorded (and reported), but none occurred
Bidonde 2014	9 (519)	0	0	Review stated that no included studies actively reported on adverse events (some reported withdrawal)

Table 14. Adverse events (not death) (*Continued*)

Boldt 2014	3 (149)	2 (115)	5 events over 2 studies	"Neck, shoulder and elbow injuries in five participants in the intervention group."
Busch 2007	34 (2276)	6 (strength training: 115, aerobic: 1264)	Strength training: 3 Aerobic training: 5	-
Busch 2013	3 (81)	2 (86 exercising participants)	0	Adverse events were recorded (and reported), but none occurred
Cramp 2013	6 (371)	3	0	Adverse events were recorded (and reported), but none occurred
Fransen 2014	10 (> 549)	5	7 events over 3 studies	-
Fransen 2015	54 (5362)	11	42 events over 8 studies	-
Gross 2015a	16 (2485)	11	41 events over 6 studies	-
Han 2004	3 (206)	2	1 event in 1 study	In narrative: "approximately one-third of the patients complained of soreness in the knee, shoulder or lower back during the first 3 weeks… pain eventually subsided for all patients… only exception was one patient, who complained of knee pain."
Hayden 2005	43 (3907)	10	23 events over 10 studies	"Negative reported: 16 events over 7 trials."
Hurkmans 2009	5 (575)	2	0	Adverse events were recorded (and reported), but none occurred
Koopman 2015	2 (68)	1 (10)	0	Adverse events were recorded (and reported), but none occurred "The study investigated deleterious effects of this training on motor unit survival through

Table 14. Adverse events (not death) (*Continued*)

				motor unit number estimates (MUNE). Results showed that the MUNE did not change at the end of the training."
Lane 2014	30 (1822)	1 (88 exercising participants)	2 events in control group in 1 study	RR 0.20 (95% CI 0.01 to 4.15) in favour of exercise group.
Regnaux 2015	1 (102)	1 (68 exercising participants over 2 groups: low and high resistance)	3 events in 1 study	"3 participants in high resistance group discontinued the exercise intervention due to severe knee pain."
Saragiotto 2016	7 (671)	1 (154)	5 events in 1 study	"Five patients (three from the MCE [motor control exercise] group and two from the minimal intervention group) had mild adverse effects during the study (all temporary exacerbations of pain)."
van der Heijden 2015	10 (1690)	0	0	Of the relevant studies, none actively reported on adverse events
Yamato 2015	6 (265)	1 (86)	0	Adverse events were recorded (and reported), but none occurred
Total	**246 studies (> 21,772)**	**61 studies (> 2134 participants)**	**137 events over 39 studies**	**61/246 (25%) of studies have reported on adverse events; of which 39/61 (64%) did have adverse events occur as a result of the intervention or control.**

n: number of participants; RR: risk ratio.

WHAT'S NEW

Date	Event	Description
18 April 2017	New citation required but conclusions have not changed	Conclusions not changed; retrospective open access.
10 April 2017	Amended	See Published notes.

CONTRIBUTIONS OF AUTHORS

LG conceived the idea of the overview, and wrote the protocol and full overview.

LG, BHS, LC, and RAM developed the concept and details of the overview (participants, intervention, comparison, outcomes).

LG and CC carried out searches and selected reviews for inclusion (RAM and DM acted as arbitrators).

LG and CC carried out assessment of methodological quality using the AMSTAR tool (DM acted as arbitrator).

LG and RAM extracted data and interpreted initial findings.

LG, RAM, LC, and BHS formulated the focus of the discussion and made suggestions for future study and review authors.

All authors were involved in the interpretation of results, and in approving the final review.

LG and BHS will be responsible for future updates.

DECLARATIONS OF INTEREST

LG: none known.

RAM has received grant support from Grünenthal relating to individual patient level analyses of trial data regarding tapentadol in osteoarthritis and back pain (2015). He has received honoraria for attending boards with Menarini concerning methods of analgesic trial design (2014), with Novartis (2014) about the design of network meta-analyses, and RB on understanding pharmacokinetics of drug uptake (2015). He has received honoraria from Omega Pharma (2016) and Futura Pharma (2016) for providing advice on trial and data analysis methods.

CC: none known.

DM: none known.

LC: none known. For transparency, LC has received honoraria for speaking at educational meetings to healthcare professionals on a range of chronic pain topics (Pfizer (October 2015), Astellas (June 2014, March 2015)); editor on the *British Journal of Anaesthesia* (receives an honorarium plus a contribution toward related departmental expenses (October 2010 - to date)). LC is a medical clinician attending patients in the NHS Lothian Pain Service.

BHS: none known. BHS is a medical clinician attending patients in the NHS Tayside Pain Service.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- MRC, Lifelong Health and Wellbeing Grant, ID 91029, UK.
- National Institute of Academic Anaesthesia (BJA RCoA NIAA), UK.

NOTES

This overview review was re-published in April 2017 with retrospective open access.

INDEX TERMS

Medical Subject Headings (MeSH)

Chronic Pain [mortality; psychology; *therapy]; Exercise Therapy [adverse effects; *methods]; Health Services Needs and Demand; Myalgia [etiology]; Pain Measurement; Patient Compliance; Quality of Life; Randomized Controlled Trials as Topic; Review Literature as Topic

MeSH check words

Adult; Humans

Cochrane Library

Cochrane Database of Systematic Reviews

Traction for low-back pain with or without sciatica (Review)

Wegner I, Widyahening IS, van Tulder MW, Blomberg SEI, de Vet HCW, Brønfort G, Bouter LM, van der Heijden GJ

Wegner I, Widyahening IS, van Tulder MW, Blomberg SEI, de Vet HCW, Brønfort G, Bouter LM, van der Heijden GJ.
Traction for low-back pain with or without sciatica.
Cochrane Database of Systematic Reviews 2013, Issue 8. Art. No.: CD003010.
DOI: 10.1002/14651858.CD003010.pub5.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Traction for low-back pain with or without sciatica

Inge Wegner[1], Indah S Widyahening[2], Maurits W van Tulder[3], Stefan EI Blomberg[4], Henrica CW de Vet[5], Gert Brønfort[6], Lex M Bouter[7], Geert J van der Heijden[8]

[1]Department of Otorhinolaryngology - Head and Neck Surgery; G05.129, University Medical Center Utrecht, Utrecht, Netherlands. [2]Centre for Clinical Epidemiology & Evidence-Based Medicine, Faculty of Medicine Universitas Indonesia - Cipto Mangunkusumo Hospital, Jakarta, Indonesia. [3]Department of Health Sciences, Faculty of Earth and Life Sciences, VU University, Amsterdam, Netherlands. [4]Family Medicine Section, Department of Public Health and Caring Sciences, Uppsala, Sweden. [5]Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, Netherlands. [6]Wolfe-Harris Center for Clinical Studies, Northwestern Health Sciences University, Bloomington, MN, USA. [7]VU University, Amsterdam, Netherlands. [8]Department of Social Dentistry, Academic Center for Dentistry Amsterdam (ACTA), Amsterdam, Netherlands

Contact address: Inge Wegner, Department of Otorhinolaryngology - Head and Neck Surgery; G05.129, University Medical Center Utrecht, Heidelberglaan 100, Utrecht, 3584 CX, Netherlands. i.wegner@umcutrecht.nl.

Editorial group: Cochrane Back and Neck Group.
Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 8, 2013.

Citation: Wegner I, Widyahening IS, van Tulder MW, Blomberg SEI, de Vet HCW, Brønfort G, Bouter LM, van der Heijden GJ. Traction for low-back pain with or without sciatica. *Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No.: CD003010. DOI: 10.1002/14651858.CD003010.pub5.

ABSTRACT

Background

Traction has been used to treat low-back pain (LBP), often in combination with other treatments. We included both manual and machine-delivered traction in this review. This is an update of a Cochrane review first published in 1995, and previously updated in 2006.

Objectives

To assess the effects of traction compared to placebo, sham traction, reference treatments and no treatment in people with LBP.

Search methods

We searched the Cochrane Back Review Group Specialized Register, the Cochrane Central Register of Controlled Trials (2012, Issue 8), MEDLINE (January 2006 to August 2012), EMBASE (January 2006 to August 2012), CINAHL (January 2006 to August 2012), and reference lists of articles and personal files. The review authors are not aware of any important new randomized controlled trial (RCTs) on this topic since the date of the last search.

Selection criteria

RCTs involving traction to treat acute (less than four weeks' duration), subacute (four to 12 weeks' duration) or chronic (more than 12 weeks' duration) non-specific LBP with or without sciatica.

Data collection and analysis

Two review authors independently performed study selection, risk of bias assessment and data extraction. As there were insufficient data for statistical pooling, we performed a descriptive analysis. We did not find any case series that identified adverse effects, therefore we evaluated adverse effects that were reported in the included studies.

Main results

We included 32 RCTs involving 2762 participants in this review. We considered 16 trials, representing 57% of all participants, to have a low risk of bias based on the Cochrane Back Review Group's 'Risk of bias' tool.

For people with mixed symptom patterns (acute, subacute and chronic LBP with and without sciatica), there was low- to moderate-quality evidence that traction may make little or no difference in pain intensity, functional status, global improvement or return to work when compared to placebo, sham traction or no treatment. Similarly, when comparing the combination of physiotherapy plus traction with physiotherapy alone or when comparing traction with other treatments, there was very-low- to moderate-quality evidence that traction may make little or no difference in pain intensity, functional status or global improvement.

For people with LBP with sciatica and acute, subacute or chronic pain, there was low- to moderate-quality evidence that traction probably has no impact on pain intensity, functional status or global improvement. This was true when traction was compared with controls and other treatments, as well as when the combination of traction plus physiotherapy was compared with physiotherapy alone. No studies reported the effect of traction on return to work.

For chronic LBP without sciatica, there was moderate-quality evidence that traction probably makes little or no difference in pain intensity when compared with sham treatment. No studies reported on the effect of traction on functional status, global improvement or return to work.

Adverse effects were reported in seven of the 32 studies. These included increased pain, aggravation of neurological signs and subsequent surgery. Four studies reported that there were no adverse effects. The remaining studies did not mention adverse effects.

Authors' conclusions

These findings indicate that traction, either alone or in combination with other treatments, has little or no impact on pain intensity, functional status, global improvement and return to work among people with LBP. There is only limited-quality evidence from studies with small sample sizes and moderate to high risk of bias. The effects shown by these studies are small and are not clinically relevant.

Implications for practice

To date, the use of traction as treatment for non-specific LBP cannot be motivated by the best available evidence. These conclusions are applicable to both manual and mechanical traction.

Implications for research

Only new, large, high-quality studies may change the point estimate and its accuracy, but it should be noted that such change may not necessarily favour traction. Therefore, little priority should be given to new studies on the effect of traction treatment alone or as part of a package.

PLAIN LANGUAGE SUMMARY

Traction for low-back pain

We reviewed the evidence on the effect of traction on pain intensity, ability to perform normal daily activities, overall improvement and return to work among people with low back pain (LBP) in the acute (less than four weeks' duration), subacute (from four to 12 weeks' duration) or chronic (more than 12 weeks' duration) phase. Some patients also had sciatica. We examined the effects of traction immediately after the traction session, in the short-term (up to three months after traction) and in the long-term (around one year after traction).

LBP is a major health problem around the world and is a major cause of medical expenses, absenteeism and disability. One treatment option for LBP that has been used for thousands of years is traction, the application of a force that draws two adjacent bones apart from each other in order to increase their shared joint space. Various types of traction are used, often in combination with other treatments. The most commonly used traction techniques are mechanical or motorized traction (where the traction is exerted by a motorized pulley) and manual traction (in which the traction is exerted by the therapist, using his or her body weight to alter the force and direction of the pull).

The evidence is current to August 2012. The review included 32 studies and 2762 people with LBP. Most studies included a similar population of people with LBP with and without sciatica. The majority of studies included people with acute, subacute and chronic

LBP. Most studies reported follow-up of one to 16 weeks, and a limited number of studies reported long-term follow-up of six months to one year.

The included studies show that traction as a single treatment or in combination with physiotherapy is no more effective in treating LBP than sham (pretend) treatment, physiotherapy without traction or other treatment methods including exercise, laser, ultrasound and corsets. These conclusions are valid for people with and without sciatica. There was no difference regarding the type of traction (manual or mechanical).

Side effects were reported in seven of the 32 studies and included increased pain, aggravation of neurological signs and subsequent surgery. Four studies reported that there were no side effects. The remaining studies did not mention side effects.

The quality of the evidence ranged from very low to moderate. There was a scarcity of high-quality studies, especially those that distinguished between people with different symptom patterns (with and without sciatica, with pain of different duration).

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Traction compared with placebo, sham or no treatment for people with low-back pain with and without sciatica

Patient or population: people with low-back pain with and without sciatica
Settings: diverse
Intervention: traction
Comparison: placebo, sham or no treatment

Outcomes	Effects	No of Participants (studies)	Quality of the evidence (GRADE)
Pain intensity VAS (0-100 mm). Follow-up 12-16 weeks.	1 trial showed that there was no difference in pain intensity between the 2 groups (MD -4, 95% CI -17.7 to 9.7)	60 (1)	⊕⊕⊕◯ **moderate** Imprecision (< 400 participants)
Functional status Oswestry Disability Index or Roland Morris Disability Questionnaire Follow-up 12-16 weeks.	Not measured.		
Global improvement Follow-up 12-16 weeks.	1 trial showed that there was no difference in global improvement between the 2 groups (RD 0.06, 95% CI -0.16 to 0.28)	81 (1)	⊕⊕⊕◯ **moderate** Imprecision (< 300 participants)
Return to work Follow-up 12-16 weeks.	Not measured.		
Adverse effects	1 trial reported aggravation of neurological signs in 28% of the traction group, 20% of the light traction group and 20% of the placebo group		

CI: confidence interval; **MD:** mean difference; **RD:** risk difference; **VAS:** visual analogue scale.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Note. Each 'Summary of findings' table presents evidence for a specific comparison and a set of prespecified outcomes. Therefore, the information presented in the tables is limited by the comparisons and outcomes reported in the included studies.

BACKGROUND

Description of the condition

Low-back pain (LBP) is a major health problem around the world and a major cause of medical expenses, absenteeism and disability (Dagenais 2008; Lambeek 2011; Vos 2012). Although LBP is usually a self limiting and benign condition that tends to improve spontaneously over time, a large variety of therapeutic interventions is available for treatment (Chou 2007). Sciatica can result when the nerve roots in the lower spine are irritated or compressed. Most often, sciatica is caused when the L5 or S1 nerve root in the lower spine is irritated by a herniated disc. Degenerative disc disease may irritate the nerve root and cause sciatica, as can mechanical compression of the sciatic nerve, such as from spondylolisthesis, spinal stenosis or arthritis in the spine. For the purposes of this review, we define sciatica as pain radiating down the leg(s) along the distribution of the sciatic nerve (which is usually related to mechanical pressure, inflammation of lumbosacral nerve roots or both) (Bigos 1994).

Description of the intervention

One treatment for LBP and sciatica is traction, which has been used for thousands of years. It is used relatively frequently in North America (e.g. up to 30% of people with acute LBP and sciatica in Ontario, Canada) (Li 2001), and to a lesser extent in the UK, Ireland and the Netherlands (Harte 2005). Traction is often provided in combination with other treatment modalities (Harte 2005). The most commonly used traction techniques are mechanical or motorized traction (where the traction is exerted by a motorized pulley), manual traction (in which the traction is exerted by the therapist, using his or her body weight to alter the force and direction of the pull), and auto-traction (where the person controls the traction forces by grasping and pulling bars at the head of the traction table). There are also less common forms, such as underwater (where the person is fixed perpendicularly in a deep pool, a bar is grasped under the arms and traction is applied), and gravitational traction (e.g. bed rest traction, in which the person is fixed to a tilted table or bed, and inverted traction, where the participant is held in an inverted position by the ankles and another part of the lower extremities and gravity provides the force).

Lumbar traction uses a harness (with Velcro strapping) that is fitted around the lower rib cage and around the iliac crest. Duration and level of force exerted through this harness can be varied in a continuous or intermittent mode. The force can be standardized only in motorized traction or in methods using computer technology. With other techniques, total body weight and the strength of the person or therapist determine the forces exerted. In the application of traction force, consideration must be given to counter forces such as lumbar muscle tension, lumbar skin stretch and abdominal pressure, which depend on the participant's physical constitution. If the person is lying on the traction table, the friction of the body on the table or bed provides the main counter force during traction.

How the intervention might work

The exact mechanism through which traction might be effective is unclear. It has been suggested that spinal elongation, by decreasing lordosis and increasing intervertebral space, inhibits nociceptive impulses, improves mobility, decreases mechanical stress, reduces muscle spasm or spinal nerve root compression (due to osteophytes), releases luxation of a disc or capsule from the zygo-apophysial joint, and releases adhesions around the zygo-apophysial joint and the annulus fibrosus.

A more recent rationale, adapted to available neurophysiological research, suggests that stimulation of proprioceptive receptors in the vertebral ligaments and in the mono segmental muscles may modify and halt what is being conceptualized as a 'dysfunction'. Dysfunction is a relatively generalized disturbance involving higher cerebral centres as well as peripheral structures for postural control. The dysfunction involves self maintaining pain-provoking neuromuscular reflex patterns. In relation to benefits of traction, this rationale involves the 'shocking' of dysfunctional higher centres by means of relaying 'unphysiological' proprioceptive information centrally, and thus 'resetting' the dysfunction (Blomberg 2005). So far, none of the proposed mechanisms has been supported by sufficient empirical information.

Little is known about the adverse effects of traction. Only a few case reports are available, which suggest that there is some danger for nerve impingement in heavy traction (i.e. lumbar traction forces exceeding 50% of the total body weight). Other risks described for lumbar traction are respiratory constraints due to the traction harness or increased blood pressure during inverted positional traction. There is some debate about the effect of low traction forces. Beurskens 1997 says that a certain amount of force is required to achieve separation of the vertebra and widening of the intervertebral foramina. Forces below 20% of the participants' body weight do not achieve this goal and, therefore, can be considered to constitute a placebo or sham traction. Other reports say that these forces can still be expected to produce positive results, as even low traction forces can produce intervertebral separation due to flattening of lumbar lordosis, and relaxation of spinal muscles (Harte 2003; Krause 2000).

Why it is important to do this review

This systematic review updates our previous Cochrane review (Clarke 2006a). The 2006 review included 25 randomized controlled trials (RCTs) and was an update of a previous review of the

effectiveness of traction for back and neck pain (Van der Heijden 1995). The previous review stated that traction was not likely to be effective for people with and without sciatica, due to inconsistent results and methodological problems in most studies. This update integrated new literature on the subject and was performed using the latest methods.

OBJECTIVES

The objective of this systematic review was to determine if traction was more effective than reference treatments, placebo, sham traction or no treatment for LBP with or without sciatica, with a focus on pain intensity, functional status, global improvement and return to work.

METHODS

Criteria for considering studies for this review

Types of studies

We included only RCTs.

Types of participants

We included RCTs involving the following types of participants: male or female; aged 18 years or older; treated for LBP; in the acute, subacute or chronic phases, with or without sciatica. We excluded studies involving people with LBP due to specific causes (e.g. tumour, metastasis, fracture, inflammation, osteoporosis, rheumatoid arthritis).

Types of interventions

We included RCTs using any type of traction, such as mechanical traction, manual traction (unspecific or segmental traction), computerized traction, auto-traction, underwater traction, bed rest traction, inverted traction, continuous traction and intermittent traction. Additional treatment was allowed, provided that traction was the main contrast between the intervention and control groups. We included studies with any type of control group (i.e. those that used placebo, sham, no treatment or other treatments).

Types of outcome measures

The four primary outcome measures that we considered to be the most important were pain intensity (e.g. measured by a visual analogue scale (VAS) or a numerical rating scale (NRS)), back-pain-specific functional status (e.g. measured by the Roland Morris Disability Questionnaire or Oswestry Disability Index (ODI)), a global measure of improvement (e.g. overall improvement, proportion of participants recovered, subjective improvement of symptoms) and return to work (e.g. measured by return to work status or days off work). We also considered reported adverse effects. These outcomes could be measured immediately after the end of one traction session, immediately after a course of traction sessions, in the short-term after the end of the traction sessions (up to three months), or in the long-term (around one year).

Search methods for identification of studies

Electronic searches

We used the results of the literature search listed in Appendix 1, updating the three previous versions of this review (Clarke 2006a; Clarke 2006b; Van der Heijden 1995a). This included a computer-aided search the Cochrane Back Review Group Specialized Register (August 2012), the Cochrane Central Register of Controlled Trials (2012 Issue 8), MEDLINE (January 2006 to August 2012), EMBASE (January 2006 to August 2012) and CINAHL (January 2006 to August 2012).

Searching other resources

Furthermore, we screened reference lists of relevant reviews and identified RCTs, as well as references in personal files of the review authors.

Data collection and analysis

In this review, we followed the guidelines of the Cochrane Back Review Group (Furlan 2009), and the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Selection of studies

Two review authors independently selected the trials to be included in the systematic review using title, abstract and keywords. The same two review authors independently applied the selection criteria to the studies that were retrieved by our literature search. We used consensus to resolve disagreements concerning selection and inclusion of RCTs. There was the option to consult a third review author if disagreement had persisted, although this was not necessary. We only evaluated full papers and excluded papers written in languages other than English, Dutch, German, French and Swedish.

Data extraction and management

Two review authors (IW and ISW) independently extracted the data (using a standardized form) considering the study population (e.g. number of participants, age, gender, type and duration of back pain), the interventions (type, intensity, and frequency of index and reference interventions) and the primary outcomes (type and duration of follow-up). We used consensus to resolve disagreements and we would have consulted a third review author (GH) if disagreement persisted, although this was not necessary. We summarized key findings in a narrative format. We did not blind data extraction.

Assessment of risk of bias in included studies

We used the Cochrane Back Review Group's 'Risk of bias' tool to assess the risk of bias of the included RCTs (Furlan 2009). The 12 criteria are listed in Appendix 2. Studies included in the previous version of the review had not been assessed using this tool. Therefore, we re-assessed these studies according to the updated methods. We could not obtain two articles (Lind 1974; Reust 1988) and two articles were written in a language that the review authors did not master (Bihaug 1978; Walker 1982). We transformed the previous risk of bias assessments of these four trials to the new format without re-assessing them. As a result, supporting statements for the risk of bias assessments are missing for these studies. Two review authors (IW and ISW) independently assessed the methodological quality. Review authors resolved their initial discrepancies during discussion; the presented results are based on their full consensus. We did not blind quality assessment with regard to the authors, institution and journal. We did not contact study authors for additional information, because half the trials were published in the late 1990s. If the article did not contain the required information for the scoring of a specific item, we scored the item as 'unclear'.

We scored the criteria as 'low risk', 'high risk' or 'unclear risk', and reported them in the 'Risk of bias' table. We defined a study with a low risk of bias as one fulfilling six or more of the criteria and having no fatal flaws. In the previous review, a sensitivity analysis was performed in which six was considered the cut-off point for low risk of bias. A second sensitivity analysis was performed in which half of items that had been scored 'unclear' in each trial were included as 'positive'. The same cut-off point of six for low risk of bias is supported by empirical evidence (Van Tulder 2009). Blinding of participants and care providers to treatment allocation is nearly impossible in trials of traction therapy. Given that some of the primary outcomes assessed in this review are subjective measures (i.e. pain and functional status), any attempt to blind the outcome assessor regarding these outcomes can be considered irrelevant. However, most studies also assessed objective outcome measures. If the care provider assessing those outcomes was blinded, the item was scored as 'low risk'.

Measures of treatment effect

We analyzed dichotomous outcomes by calculating the risk difference. We analyzed continuous outcomes by calculating the mean difference (MD) when the same instrument was used to measure outcomes, or the standardized mean difference (SMD) when different instruments were used to measure the outcomes. We converted VAS or NRS scales to a 100-point scale. We expressed uncertainty using with 95% confidence intervals (CI).

We grouped outcomes by timing when they were measured: immediately after, short term and long term.

Unit of analysis issues

In several studies, we compared more than two intervention groups. We included these studies by making pair-wise comparisons between all possible pairs of intervention groups with traction being one of the intervention groups. The same group of participants was included more than once in these examples (e.g. underwater traction versus underwater massage and underwater traction versus balneotherapy in the study performed by Konrad 1992). These participants were not counted twice in the meta-analysis.

Dealing with missing data

In cases where data were reported as a median with an interquartile range (IQR), we assumed that the median was equivalent to the mean and the width of the IQR equivalent to 1.35 times the standard deviation in accordance with *Cochrane Handbook for Systematic Reviews of Interventions*, section 7.7.3.5 (Higgins 2011). If standard deviations were not given, we calculated them from the 95% CIs, P values based on a two-sided t-test or standard errors. We did not include data reported in graphs in this review.

Assessment of heterogeneity

We tested heterogeneity using the Chi^2 test and I^2 statistic; however, the decision regarding heterogeneity was dependent upon the I^2 statistic (Higgins 2011). We defined substantial heterogeneity as an I^2 greater than 50%, and where necessary, the effect of the interventions were synthesised narratively when the I^2 statistic was greater than 50%.

Assessment of reporting biases

We searched ClinicalTrials.org and ISRCTN.org for the protocols of included studies. When protocols were available, we checked studies for selective outcome reporting.

Data synthesis

A quantitative analysis had been planned, but most of the studies did not provide sufficient data to enable statistical pooling (e.g. some trials reported the mean score but not the standard deviation,

other trials reported median and IQR; some trials reported only post-intervention means and other trials reported mean change scores; some trials did not report any numerical data. Therefore, we used a descriptive analysis to summarize the data. In this analysis, we used a rating system of levels of evidence to summarize the results of the studies in terms of the strength of the scientific evidence. To accomplish this, we used the GRADE approach, as recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), and adapted in the updated Cochrane Back Review Group method guidelines (Furlan 2009). The system consists of five levels of evidence, based on performance against five principal domains or factors:

- **high-quality evidence** - consistent findings among at least 75% of RCTs with low risk of bias, consistent, direct and precise data and no known or suspected publication biases. Further research is unlikely to change either the estimate or our confidence of the results;
- **moderate-quality evidence** - one of the domains is not met. Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate;
- **low-quality evidence** - two of the domains are not met. Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate;
- **very-low-quality evidence** - three of the domains are not met. We are very uncertain about the results;
- **no evidence** - no RCTs were identified that addressed this outcome.

Factors that may decrease the quality of the evidence are: study design and risk of bias (downgraded when > 25% of the participants were from studies with a high risk of bias), inconsistency of results, indirectness (downgraded when > 50% of the participants were outside the target group), imprecision (downgraded when the total number of participants was less than 400 for continuous outcomes and 300 for dichotomous outcomes) and other factors (e.g. reporting bias).

Because the majority of studies contained a mix of participants with acute, subacute and chronic LBP, we did not separate out these groups in our analyses, other than in several trials involving only people with chronic LBP. We categorized studies as including people 'with sciatica' if more than66% of the participants were described as having sciatica (this may or may not have included those with nerve root symptoms) or if there was a separate analysis of outcomes in those with sciatica.

Clinical relevance

Two review authors independently carried out an analysis of the clinical relevance of each study. Without using an arbitrary predefined threshold, studies were judged as to whether: participants were described in enough detail to allow practitioners to decide whether they were similar to those in their practices; interventions and treatment settings were described well enough to allow practitioners to provide the same treatment for their participants; clinically relevant outcomes were measured and reported; the size of the effect; and the treatment benefits were worth the potential harms (see Table 1).

Subgroup analysis and investigation of heterogeneity

Predefined subgroup analyses included:

- different types of comparison (traction versus placebo, sham or no treatment; physiotherapy with traction versus physiotherapy without traction; different types of traction and traction versus other treatments);
- different symptom patterns in subjects (mixed population of people with LBP with and without sciatica; people with LBP with sciatica and people with LBP without sciatica).

However, we were not able to conduct these analyses, because of reasons stated above. Instead, the results were synthesized narratively. 'Summary of findings' tables were generated for all analyses of different types of comparison. Primary outcome measures at a follow-up duration of 12 to 16 weeks were included in the 'Summary of findings' tables.

Sensitivity analysis

In the previous review, sensitivity analyses were carried out to determine the cut-off for high-quality studies. The cut-off point was set at six criteria for risk of bias, which is supported by empirical evidence (Van Tulder 2009). We considered that studies that met six or more of the criteria for risk of bias carried low risk of bias, whereas studies that met fewer than six of the criteria carried high risk of bias. We did not plan or carry out any new sensitivity analyses.

RESULTS

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies; Characteristics of studies awaiting classification; Characteristics of ongoing studies.

We identified 32 trials that fulfilled the inclusion criteria. Seven new trials were published since the publication of the previous review (Fritz 2007; Gudavalli 2006; Harte 2007; Ozturk 2006; Schimmel 2009; Simmerman 2011; Unlu 2008). We included all 25 trials discussed in the previous review in this review. The total number of studies retrieved by all search methods over time was not available. In this review, we included 32 studies, involving 2762 participants. Two of these studies were reported in one publication

(Weber 1984); in four of the studies, there was more than one pertinent publication (Beurskens 1997; Gudavalli 2006; Mathews 1988; Van der Heijden 1995).

Presence of sciatica

Twenty-three of the studies included a relatively homogeneous population of people with LBP and sciatica (Bihaug 1978; Coxhead 1981; Fritz 2007; Güvenol 2000; Harte 2007; Larsson 1980; Lidström 1970; Lind 1974; Ljunggren 1984; Ljunggren 1992; Mathews 1975; Mathews 1988; Ozturk 2006; Pal 1986; Reust 1988; Sherry 2001; Simmerman 2011; Sweetman 1993; Unlu 2008; Walker 1982; Weber 1973; two trials in Weber 1984). Eight studies included a greater mix of participants with and without sciatica (Beurskens 1997; Borman 2003; Gudavalli 2006; Konrad 1992; Letchuman 1993; Tesio 1993; Van der Heijden 1995; Werners 1999). There was only one study that exclusively involved people who did not have sciatica (Schimmel 2009).

Duration of low-back pain

Ten studies included solely or primarily people with chronic LBP of more than 12 weeks (Borman 2003; Gudavalli 2006; Güvenol 2000; Ljunggren 1984; Schimmel 2009; Sherry 2001; Tesio 1993; Van der Heijden 1995; two in Weber 1984); in one study, participants were all in the subacute range (four to 12 weeks) (Konrad 1992); in 17 studies, the duration of LBP was a mixture of acute, subacute and chronic (Beurskens 1997; Bihaug 1978; Coxhead 1981; Fritz 2007; Harte 2007; Larsson 1980; Lidström 1970; Lind 1974; Ljunggren 1992; Mathews 1975; Mathews 1988; Ozturk 2006; Pal 1986; Simmerman 2011; Sweetman 1993; Unlu 2008; Walker 1982); in five studies duration was not specified (Letchuman 1993; Reust 1988; Weber 1973; and two in Weber 1984).

Comparisons

Thirteen studies compared traction with sham traction (Beurskens 1997; Letchuman 1993; Mathews 1975; Pal 1986; Reust 1988; Schimmel 2009; Van der Heijden 1995; Walker 1982; Weber 1973; and two in Weber 1984), with some kind of placebo (sham shortwave diathermy, Sweetman 1993; sham shortwave Lind 1974); or with no treatment (Konrad 1992). Fifteen studies compared traction with other treatments (Bihaug 1978; Coxhead 1981; Gudavalli 2006; Konrad 1992; Larsson 1980; Lidström 1970; Lind 1974; Ljunggren 1992; Mathews 1988; Sherry 2001; Simmerman 2011; Sweetman 1993; Unlu 2008; Werners 1999; Weber 1984). In one of these (Lind 1974), auto-traction was compared with physiotherapy, in which Tru-Trac traction was one of the range of treatments included. Five studies compared different types of traction (e.g. auto-traction versus manual traction or passive traction, continuous versus intermittent traction, inversion traction versus conventional traction) (Güvenol 2000; Letchuman 1993; Ljunggren 1984; Reust 1988; Tesio 1993). Four studies compared a standard physiotherapy programme (not including traction) with the same treatment with traction (Borman 2003; Fritz 2007; Harte 2007; Ozturk 2006). One study compared different types of underwater therapy, underwater traction being one of them (Konrad 1992).

Length of follow-up

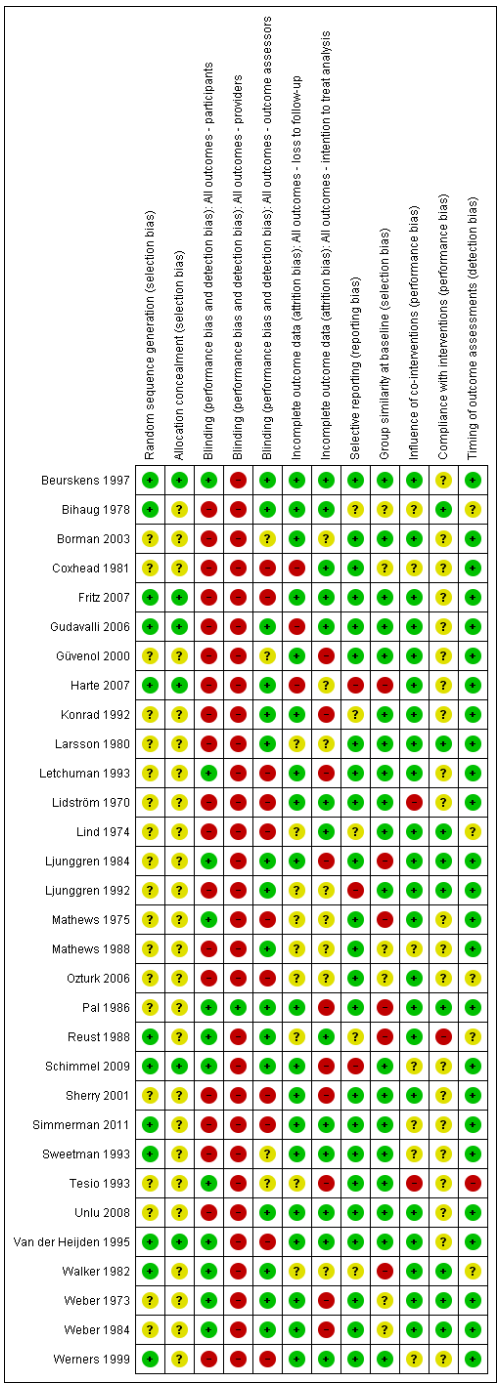
Fourteen studies reported short-term follow-up (one week) (Fritz 2007; Gudavalli 2006; Harte 2007; Larsson 1980; Ljunggren 1984; Ljunggren 1992; Ozturk 2006; Pal 1986; Simmerman 2011; Sweetman 1993; Unlu 2008; Weber 1973; two in Weber 1984). Fifteen studies reported follow-up at three to five weeks (Beurskens 1997; Bihaug 1978; Coxhead 1981; Fritz 2007; Konrad 1992; Lidström 1970; Lind 1974; Ljunggren 1984; Mathews 1975; Mathews 1988; Pal 1986, Reust 1988; Sherry 2001; Unlu 2008; Van der Heijden 1995). Fourteen studies reported follow-up at nine to 16 weeks (Beurskens 1997; Bihaug 1978; Borman 2003; Coxhead 1981; Gudavalli 2006; Güvenol 2000; Harte 2007; Larsson 1980; Ljunggren 1984; Schimmel 2009; Tesio 1993; Unlu 2008; Van der Heijden 1995; Werners 1999). Five studies reported follow-up at six months (Beurskens 1997; Gudavalli 2006; Harte 2007; Mathews 1988), or one year (Gudavalli 2006; Konrad 1992; Mathews 1988). One study did not report the timing at which the outcomes were measured (Walker 1982).

Risk of bias in included studies

See: Characteristics of included studies.

The results of the risk of bias analysis for the individual studies are summarized in Figure 1. Sixteen studies were considered to have a low risk of bias (Beurskens 1997; Fritz 2007; Gudavalli 2006; Larsson 1980; Letchuman 1993; Ljunggren 1984; Pal 1986; Schimmel 2009; Simmerman 2011; Sweetman 1993; Unlu 2008; Van der Heijden 1995; Weber 1973; both trials in Weber 1984; Werners 1999), representing 1568 (57%) participants. Overall, risk of bias scores ranged from two to 10 (maximum possible risk of bias score was 12). Some of the studies that were considered to have a low risk of bias based on the The Cochrane Collaboration's 'Risk of bias' tool were considered to have a high risk of bias in the previous review (Larsson 1980; Letchuman 1993; Ljunggren 1984; Pal 1986; Sweetman 1993; Weber 1973; Weber 1984). Overall completeness of data was assessed in this review, whereas previously, dropout during intervention and dropout during follow-up were scored. Selective reporting and timing of outcome assessments were not assessed previously.

Figure 1. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Allocation

The majority of the included studies did not properly report on their random and concealed allocation of treatment. In 20 of the included articles, there was no mention of the randomization procedure used and, in 26 of the included studies, it was unclear how concealment of treatment allocation was achieved. In six studies, both sequence generation and allocation procedure were conducted properly (Beurskens 1997; Fritz 2007; Gudavalli 2006; Harte 2007; Schimmel 2009; Van der Heijden 1991). In an additional six studies, the sequence generation was conducted properly, but the concealment of allocation was inadequately described (Bihaug 1978; Reust 1988; Simmerman 2011; Sweetman 1993; Walker 1982; Werners 1999). In the remaining studies, both randomization and allocation procedure were inadequately described or not mentioned at all. The authors claimed these studies were RCTs in the description of their methods and, therefore, these studies were included nevertheless.

Blinding

Blinding of outcomes was not achieved in the majority of the included studies. Blinding of the outcome assessor was achieved in 17 studies (Beurskens 1997; Bihaug 1978; Gudavalli 2006; Harte 2007; Konrad 1992; Larsson 1980; Ljunggren 1984; Ljunggren 1992; Mathews 1988; Pal 1986; Reust 1988; Schimmel 2009; Unlu 2008; Walker 1982; Weber 1973; both trials in Weber 1984), blinding of participants in 12 studies (Beurskens 1997; Letchuman 1993; Ljunggren 1984; Mathews 1975; Pal 1986; Reust 1988; Schimmel 2009; Tesio 1993; Van der Heijden 1995; Walker 1982; Weber 1973; Weber 1984), and blinding of care providers only in one study (Pal 1986). All of the studies that attempted to blind the participants to the assigned intervention did so by providing a sham treatment, with the exception of Tesio 1993. None of the studies evaluated the success of blinding post-treatment. It should be noted that blinding of care providers of traction is impossible in most cases. It is disputable whether the outcome is likely to be influenced by a lack of blinding of care providers when it comes to assessing subjective measures such as pain intensity and functional status, as mentioned earlier. However, in the case of objective outcome measures, blinding is of importance.

Incomplete outcome data

In three studies, loss to follow-up exceeded 20% of the study population (Coxhead 1981; Harte 2007), or significantly more subjects were lost to follow-up in one treatment group compared the number of subjects that were lost to follow-up in the other group (Gudavalli 2006). Loss to follow-up never exceeded 23%. In nine of the included trials, it was not clear how many subjects were lost to follow-up (Larsson 1980; Lind 1974; Ljunggren 1992; Mathews 1975; Mathews 1988; Ozturk 2006; Reust 1988; Tesio 1993; Walker 1982).

Selective reporting

None of the included RCTs had a published protocol in any of the protocol databases that were searched. The study's prespecified (primary and secondary) outcomes as reported in the article itself were compared with the reported outcomes. One study indicated that VAS scores, overall improvement and improvement in the straight leg raising test had been recorded at three and six months but did not report this (Harte 2007), while in another study, improvement in mobility, activities of daily living and the straight leg raising test were measured but not reported (Ljunggren 1992), and similarly for all outcome assessments at two and six weeks in another study (Schimmel 2009).

Other potential sources of bias

We identified no other potential sources of bias.

Effects of interventions

See: **Summary of findings for the main comparison** Traction compared with placebo, sham or no treatment for people with low-back pain with and without sciatica; **Summary of findings 2** Physiotherapy with traction compared with physiotherapy without traction for people with low-back pain with and without sciatica; **Summary of findings 3** Traction compared with another type of traction for people with low-back pain with and without sciatica; **Summary of findings 4** Traction compared with any other treatment for people with low-back pain with and without sciatica

Section (1) of the results describes those studies in which a mixed group of people with LBP is involved, i.e., some with and some without sciatica. In section (2), the participant populations include only people with LBP with sciatica. Section (3) describes the studies that included only people with LBP without sciatica. Studies that included more than 66% of participants with sciatica were categorized as studies that included people with sciatica.

(1) Traction for a mixed group of people with low-back pain, some with and some without sciatica

(1a) Traction versus placebo, sham or no treatment

There was low-quality evidence that decrease in pain intensity was greater in participants treated with traction at three to five weeks'

follow-up (MD 18.49 points on the VAS, 95% CI -24.12 to -12.87) (Beurskens 1997; Konrad 1992). However, the difference in pain intensity at one year' follow-up had an MD of only 9 points on the VAS (95% CI -19.32 to 1.12), favouring traction (Konrad 1992). Moderate-quality evidence indicated there was a small positive effect on functional status favouring the sham group at three to five weeks' follow-up (1.3 points on the Roland Morris Disability Questionnaire (RMDQ), 95% CI -2.90 to 0.30) (Beurskens 1997). There was no difference in global improvement at three to five weeks (RD -0.03, 95% CI -0.17 to 0.12) (Beurskens 1997; Van der Heijden 1995), or at six to 12 weeks (RD 0.03, 95% CI -0.12 to 0.18) (Beurskens 1997; Van der Heijden 1995). Moderate-quality evidence showed mean time to return to work in the traction group was two days earlier (Beurskens 1997).

(1b) Physiotherapy with traction versus physiotherapy without traction

There was low-quality evidence that there was no difference in pain intensity at one to two weeks' follow-up between the two groups (Borman 2003). There was a small mean difference of 5 points on the VAS (95% CI -5.67 to 15.67) in favour of physiotherapy at 12 to 16 weeks' follow-up (Borman 2003). Short-term and long-term functional status as measured by the ODI was better in the traction group than the physiotherapy group (short term mean points: 4, 95% CI -1.91 to 9.71; long term: 95% CI -2.78 to 10.78) (Borman 2003). There was low-quality evidence that global improvement at one to two weeks' follow-up was the same for both groups, whereas at 12 to 16 weeks' follow-up, global improvement was higher in the traction group (RD 0.53, 95% CI 0.28 to 0.79) (Borman 2003).

(1c) Different types of traction

One study with very-low-quality evidence showed that there was no difference in global improvement between participants undergoing static traction and participants undergoing intermittent traction (Letchuman 1993). Global improvement was higher in participants undergoing auto-traction than in participants undergoing mechanical traction (RD 0.53, 95% CI 0.32 to 0.73) (Tesio 1993). Outcomes on pain intensity and functional status were reported only for those participants responding to treatment.

(1d) Traction versus other treatments

Six studies compared traction with another treatment (Bihaug 1978; Gudavalli 2006; Konrad 1992; Lind 1974; Sweetman 1993; Werners 1999). Traction was compared with varying other treatments: physiotherapy, exercise, short-wave diathermy, interferential therapy, bed rest and analgesics.

There was low- to moderate-quality evidence that pain intensity was slightly lower in participants treated with traction in the short-term and the long-term (Gudavalli 2006; Konrad 1992; Sweetman 1993; Werners 1999). MDs varied from 1 to 8 points on the VAS with a follow-up duration varying from one week to one year. Moderate-quality evidence showed that functional status as measured by the ODI or RMDQ was the same for both groups at one to two weeks, 12 to 16 weeks and one year' follow-up (Gudavalli 2006; Werners 1999). There was a small difference in favour of the control group at three to five weeks (MD 0.2, 95% CI -0.05 to 0.46) and at six months (0.15 points, 95% CI -0.16 to 0.45) (Gudavalli 2006). There was a very small difference in global improvement favouring traction at 12 to 16 weeks (Bihaug 1978) (RD 0.05, 95% CI -0.11 to 0.20), for which there was high-quality evidence. The difference in global improvement at three to five weeks was much higher with an RD of 0.14 (95% CI -0.08 to 0.36) (Bihaug 1978) and 0.87 (95% CI 0.67 to 1.07) favouring traction (Lind 1974). However, the quality of evidence supporting this difference was very low.

(2) Traction for people with low-back pain and sciatica

(2a) Traction versus placebo, sham or no treatment for people with a mix of acute, subacute and chronic low back pain with sciatica

Low-quality evidence suggested that there was a small effect on pain intensity in favour of the sham group (MD 2.93 points on the VAS scale, 95% CI -14.73 to 20.59) at one to two weeks' follow-up (Pal 1986; Reust 1988), and at three to five weeks' follow-up (Pal 1986). There was low- to moderate-quality evidence that global improvement rates were higher in participants receiving traction at one to two weeks' follow-up (RD 0.13, 95% CI 0.04 to 0.22) (Larsson 1980; Sweetman 1993; Weber 1973; Weber 1984), and three to five weeks' follow-up (RD 0.27, 95% CI 0.12 to 0.43) (Larsson 1980; Lidström 1970). However, at 12 to 16 weeks' follow-up, there was no significant difference in global improvement between the two groups (RD 0.06, 95% CI -0.16 to 0.28) (Larsson 1980). Moderate-quality evidence suggested that more participants receiving traction returned to work compared with participants receiving sham treatment (RD 0.15, 95% CI -0.15 to 0.45) (Pal 1986).

(2b) Physiotherapy with traction versus physiotherapy without traction

Although moderate-quality evidence showed a lower mean pain intensity in the traction group (a difference of 7.96 points on the VAS, 95% CI -16.53 to 0.61) at one to two weeks' follow-up (Fritz 2007; Ozturk 2006), the difference in mean pain intensity between the two groups was 2.00 points (95% CI -10.02 to 14.02) in favour of the physiotherapy group at six weeks' follow-up (Fritz 2007). Functional status was measured by both the ODI and the RMDI. There was low- to moderate-quality evidence that there was no difference in functional outcome at one to two weeks',

six to 12 weeks', 12 to 16 weeks and six months' follow-up (Fritz 2007; Harte 2007). Low- to moderate-quality evidence showed no difference in global improvement at one to two weeks' (Ozturk 2006), three to five weeks' (Coxhead 1981), six weeks' (Fritz 2007) and 12 to 16 weeks' (Coxhead 1981) follow-up.

(2c) Different types of traction

We found three RCTs that compared two types of traction and reported on pain intensity (Ljunggren 1984; Reust 1988; Simmerman 2011). Reust 1988 compared auto-traction with mechanical traction. There was a small effect in favour of auto-traction (2.9 points on the VAS, 95% CI -14.73 to 20.59). Simmerman 2011 compared aquatic traction to a land-based supine position at one to two weeks' follow-up. There was a small effect in favour of auto-traction at one to two weeks' follow-up (8 points on the VAS, 95% CI -3.02 to 19.02). One RCT was identified that compared two types of traction, auto-traction versus manual traction, and reported on global improvement (Ljunggren 1984). There was a small effect in favour of manual traction at one to two weeks' follow-up (RD -0.16, 95% CI -0.40 to 0.09). Although one more RCT compared two types of traction (Güvenol 2000), this study only reported P values.

(2d) Traction versus other treatments

Three RCTs compared traction with other treatments and reported varying outcome measures (Lidström 1970; Ljunggren 1992; Unlu 2008). Traction was compared with physiotherapy, exercise, laser, ultrasound, manipulation and corset treatment.

There was moderate-quality evidence that mean pain intensity in the traction group was slightly lower at one to two weeks' follow-up (Ljunggren 1992; Unlu 2008), and three to five weeks' follow-up (Unlu 2008). The maximum MD in pain intensity was 4.9 points (95% CI -15.87 to 6.07) (Unlu 2008). However, at 12 to 16 weeks' follow-up the mean pain intensity in the traction group was higher (maximum MD 4.4 points, 95% CI -5.40 to 14.20) (Unlu 2008). There was no difference in functional status measured by the ODI or RMDI between the two groups at one to two weeks', three to five weeks' and 12 to 16 weeks' follow-up (Ljunggren 1992; Unlu 2008). There was low- to moderate-quality evidence that there is only a very small difference in global improvement between the two groups at one to two weeks' follow-up (RD 0.03, 95% CI -0.24 to 0.30) (Ljunggren 1992), and three to five weeks' follow-up (RD 0.42, 95% CI 0.17 to 0.67) (Lidström 1970).

(3) Traction for people with low-back pain and without sciatica

(3a) Traction versus sham treatment

There was moderate-quality evidence that there is a very small difference in pain intensity between the two groups, favouring the traction group by 4 points on the VAS (95% CI -17.65 to 9.65) (Schimmel 2009).

Adverse effects

Of the 32 studies, four stated that there were no adverse effects (Gudavalli 2006; Konrad 1992; Schimmel 2009; Walker 1982); seven studies reported some adverse effects, for example, increased pain in 11 of 14 inversion traction participants versus 2 of 13 conventional traction participants, and anxiety during treatment with "almost all of the inversion traction patients" (Güvenol 2000); increased pain in 31% of static traction group and 15% of intermittent traction group (Letchuman 1993); temporary deterioration in 4 of 24 of traction and 4 of 26 of exercise group (Ljunggren 1992); subsequent surgery in 7 of 83 in lumbar traction group versus none in control group (Mathews 1988); aggravation of neurological signs in 5 of 18 of traction group, 4 of 20 of light traction group and 4 of 20 of placebo group (Reust 1988); aggravation of symptoms in 5 of 43 of traction and 1 of 43 of sham (Weber 1973). Borman 2003 reported that 25% of the group receiving traction as part of standard physiotherapy and 37% of the physiotherapy without traction group felt "probably or definitely worse" at three-month' follow-up. The remaining 21 studies did not report adverse effects.

ADDITIONAL SUMMARY OF FINDINGS *[Explanation]*

Physiotherapy with traction compared with physiotherapy without traction for people with low-back pain with and without sciatica

Patient or population: people with low-back pain with and without sciatica
Settings: physical medicine and rehabilitation outpatient clinic of a larger hospital
Intervention: physiotherapy with traction
Comparison: physiotherapy without traction

Outcomes	Effects	No of Participants (studies)	Quality of the evidence (GRADE)
Pain intensity VAS (0-100 mm). Follow-up 12-16 weeks.	1 trial showed that there was no difference in pain intensity between the 2 groups (MD 5, 95% CI -5.7 to 15.7) in favour of the control group	39 (1)	⊕⊕◯◯ **low** Study design (high risk of bias) Imprecision (< 400 participants)
Functional status Oswestry Disability Index or Roland Morris Disability Questionnaire Follow-up 12-16 weeks.	2 trials showed that there was no difference in functional status between the 2 groups (SMD from 0.36 (95% CI -0.27 to 1.00) to 0.43 (95% CI -0.30 to 1.16))	69 (2)	⊕⊕◯◯ **low** Study design (high risk of bias) Imprecision (< 400 participants)
Global improvement Follow-up 12-16 weeks.	1 trial showed no difference in global improvement, another trial did show a clinically significant difference in global improvement (RD 0.53, 95% CI 0.28 to 0.79)	220 (2)	⊕⊕◯◯ **low** Study design (high risk of bias) Imprecision (< 300 participants)
Return to work Follow-up 12-16 weeks.	Not measured.		
Adverse effects	1 study reported that 25% of the physiotherapy with traction group and 37% of the physiotherapy without traction group felt worse at 3 months' follow-up		

CI: confidence interval; **MD:** mean difference; **RD:** risk difference.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Traction compared with another type of traction for people with low-back pain with and without sciatica

Patient or population: people with low-back pain with and without sciatica
Settings: diverse
Intervention: traction
Comparison: another type of traction

Outcomes	Effects	No of Participants (studies)	Quality of the evidence (GRADE)
Pain intensity VAS (0-100 mm). Follow-up 12-16 weeks.	Not measured.		
Functional status Oswestry Disability Index or Roland Morris Disability Questionnaire Follow-up 12-16 weeks.	Not measured.		
Global improvement Follow-up 12-16 weeks.	Not measured.		
Return to work Follow-up 12-16 weeks.	Not measured.		
Adverse effects	1 trial reported increased pain in 31% of the static traction group and 15% of the intermittent traction group		

VAS: visual analogue scale.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Traction compared with any other treatment for people with low-back pain with and without sciatica

Patient or population: people with low-back pain with and without sciatica
Settings: diverse
Intervention: traction
Comparison: other treatment

Outcomes	Effects	No of participants (studies)	Quality of the evidence (GRADE)
Pain intensity VAS (0-100 mm). Follow-up 12-16 weeks.	3 trials, of which 1 compared traction with 2 other types of treatment, showed no difference greater than 5 points on the VAS scale between the 2 groups (MD -2.90 (95% CI -8.53 to 2.93) to 4.50 (95% CI -0.45 to 9.45)	304 (3)	⊕⊕⊕○ **moderate** Imprecision (< 400 participants)
Functional status Oswestry Disability Index or Roland Morris Disability Questionnaire Follow-up 12-16 weeks.	3 trials, of which 1 compared traction to 2 other types of treatment and used 2 types of questionnaires to assess functional status, showed no difference between the 2 groups (SMD -0.08 (95% CI -0.39 to 0.23) to 0.51 (95% CI -0.12 to 1.14))	350 (3)	⊕⊕⊕○ **moderate** Imprecision (< 400 participants)
Global improvement Follow-up 12-16 weeks.	1 trial showed no difference in global improvement (RD 0.05, 95% CI -0.1 to 0.2)	42 (1)	⊕⊕○○ **low** Study design (high risk of bias) Imprecision (< 300 participants)
Return to work Follow-up 12-16 weeks.	Not measured.		
Adverse effects	1 trial reported temporary deterioration of low-back pain in 17% of the traction group and 15% of the exercise group		

MD: mean difference; **RD:** risk difference; **SMD:** standardized mean difference.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

DISCUSSION

Summary of main results

Many studies were identified on the effect of traction on pain intensity, functional status, global improvement and return to work in people with LBP. However, most evidence was imprecise and inconsistent and numerous studies carried substantial risk of bias. Many of the studies seemed to have sample sizes that were too small to detect a clinically significant difference. Furthermore, the heterogeneity in comparisons, outcomes and follow-up durations prohibited us, among other reasons, from pooling the data and, therefore, we used a descriptive analysis in this review. The sample sizes per comparison mostly did not reach the threshold of 400 for continuous outcomes and 300 for dichotomous outcomes (Furlan 2009; Higgins 2011). Therefore, we put little trust in positive effects that emerged.

The included studies largely differed in their population, outcome measures, and scales and duration of follow-up. Some studies included hospitalized participants with demonstrated herniated discs, neurological findings and sciatica, while other studies included people recruited from primary care or workers recruited through internal company newspapers. Some studies used the ODI, while others used the RMDI. Some studies reported on all four primary outcomes (pain intensity, functional status, global improvement and return to work), whereas others only reported on one or two, which might suggest publication bias.

The studies showed small differences in effects between traction and other treatment options on pain intensity, functional status, global improvement and return to work at short term. The effect was even smaller at longer-term follow-up. Mostly the MD between the two groups favours the traction group, but not always. For most of the outcomes, no effects of traction were shown and when they were, the effects were too small to be clinically relevant. The minimum important difference (between groups) in changes (within groups) for pain intensity and functional status established by Ostelo 2008 were used to judge clinical relevancy. A clinically relevant effect was achieved in pain intensity at three to five weeks' follow-up in people with and without sciatica undergoing traction when compared with sham treatment (Konrad 1992). A clinically relevant difference in changes in global improvement was seen in people with and without sciatica undergoing physiotherapy with traction at 12 to 16 weeks' follow-up (RD 0.53) (Borman 2003), and in global improvement in people with and without sciatica undergoing traction when compared to other treatments at 12 to 16 weeks (RD 0.57) (Bihaug 1978; Lind 1974). However, in all of these cases, the effects did not reach statistical significance and they were based on low- to very-low-quality evidence, which means that we are very uncertain about the findings. Studies with a high risk of bias typically overestimate the effect compared to studies with a low risk of bias (Van Tulder 2009).

Two articles examined the level of physical force applied in the treatment and concluded that even a low level of force may be effective (Harte 2003; Krause 2000). Beurskens 1997 maintained that traction at levels below 25% of body weight and using a split table can be regarded as sham (or low-dose) traction, and the sham traction group in their trial received treatment involving a force of 10% to 20% of the participant's body weight. In the other trials that classified their control groups as 'sham traction', the force applied varied (e.g. less than 25% of body weight in Van der Heijden 1995; 10 lb (4.5 kg) in Letchuman 1993; 1.8 kg in Pal 1986; 5 kg in Reust 1988; and a maximum of 20 lb (9 kg) in Mathews 1975). No differences between traction and sham traction were demonstrated in any of these trials.

Overall completeness and applicability of evidence

We minimized review bias by performing an extensive database search. Publication bias could be an issue. The many small RCTs are more likely to be published when positive. Authors possibly may refrain from publishing when results are negative. However, the review authors consider that it is unlikely that large trials on the subject were not published. Many of the published studies did not have a published protocol and, therefore, it is difficult to ascertain to what extent studies did not publish their findings because the results did not prove to be favourable.

Quality of the evidence

Sixteen of the 32 included studies demonstrated a low risk of bias. Items that were scored predominantly negatively or unclear were randomization, concealment and blinding. The majority of the included studies did not properly report on their random and concealed allocation of treatment. In 20 of the included articles, there was no mention of the randomization procedure used and, in 26 of the included studies, it was unclear how concealment of treatment allocation was achieved. Blinding of outcomes was not achieved in the majority of the included studies. Blinding of the outcome assessor was achieved in 17 studies and blinding of participants in 12 studies. The latter reflects the number of trials in which sham or simulated traction was used. Blinding of the care provider is virtually impossible given the nature of the intervention. As a result, only one study achieved blinding of the care provider.

Furthermore, relatively few participants were identified for any of the principal outcome measurements and, as a result, none of the findings should be considered robust.

Potential biases in the review process

Although content area experts may have inside knowledge, may be familiar with current interests in their field and may be aware of pressing questions in their field, they may also have personal prejudices and idiosyncrasies. Experts with strong opinions may make it difficult to prevent bias (Gotzsche 2012). To harness bias in this review, two non-experts (IW and ISW) in this area, trained in reviewing literature, were involved in writing this review. Data from previous reviews were verified, checked and changed where necessary by these two review authors.

Agreements and disagreements with other studies or reviews

In general, the results and conclusions of this updated review are consistent with the previous version of the review, namely that traction is no better than standard interventions for (acute, subacute and chronic) LBP. In this review, we discussed one high-quality study that included only people without sciatica that was not included in the previous review (Schimmel 2009). This study showed that traction in people without sciatica is no better than sham treatment. There was no significant difference between the traction and sham group in pain intensity or functional status.
Our findings were consistent with those reported in other systematic reviews on the subject (Chou 2007; Gay 2008). One review concluded there was insufficient data to draw firm conclusions on the clinical effect of traction (Van Middelkoop 2011). Only one RCT discussing the effect of traction was included in this review.

AUTHORS' CONCLUSIONS

Implications for practice

Effects of traction alone or as part of a package for people with low-back pain (LBP) with and without sciatica have not been shown. There are some randomized controlled trials (RCTs) showing benefit of traction, but the limited quality evidence from these small moderate to high risk of bias studies show very small effects that are not clinically relevant. In summary, to date the use of traction as treatment for non-specific LBP is not supported by the best available evidence.

Implications for research

New, large, high-quality studies may change the point estimate and its accuracy, but it should be noted that such change may not necessarily favour traction. Therefore, little priority should be given to new studies on the effect of traction treatment alone or as part of a package.

ACKNOWLEDGEMENTS

We acknowledge and thank Dr. J Clarke for her contribution to the previous review (Clarke 2006a), of which this is an update.

REFERENCES

References to studies included in this review

Beurskens 1997 *{published data only}*
* Beurskens AJ, Van der Heijden GJ, de Vet HCW, Koke AJ, Lindeman E, Regtop W, et al. The efficacy of traction for lumbar back pain: design of a randomized clinical trial. *Journal of Manipulative and Physiological Therapeutics* 1995; **18**:141–7.
Beurskens AJ, de Vet HC, Köke AJ, Lindeman E, Regtop W, Van der Heijden GJ, et al. Efficacy of traction for non-specific low back pain: a randomised clinical trial. *Lancet* 1995;**346**:1596–600.
Beurskens AJ, de Vet HC, Köke PT, Regtop W, Van der Heijden GJ, Lindeman E, et al. Efficacy of traction for nonspecific low back pain: 12-week and 6-month results of a randomized clinical trial. *Spine* 1997;**22**(23):2756–62.

Bihaug 1978 *{published data only}*
Bihaug O. [Autotraksjon for ischialgpasienter: en kontollert sammenlikning mellom effekten av Auto–traksjon–B og isometriske ovelser ad modum Hume endall og enkins]. *Fysioterapeuten* 1978;**45**:377–9.

Borman 2003 *{published data only}*
* Borman P, Keskin D, Bodur H. The efficacy of lumbar traction in the management of patients with low back pain. *Rheumatology International* 2003;**23**:82–6.

Coxhead 1981 *{published data only}*
Coxhead CE, Inskip H, Meade TW, North WR, Troup JD. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet* 1981;**1**(8229):1065–8.

Fritz 2007 *{published data only}*
Fritz JM, Lindsay W, Matheson JW, Brennan GP, Hunter SJ, Moffit SD, et al. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. *Spine* 2007;**32**(26):E793–800.

Gudavalli 2006 *{published data only}*
Cambron JA, Gudavalli MR, McGregor M, Jedlicka J, Keenum M, Ghanayem AJ, et al. Amount of health care and self-care following a randomized clinical trial comparing flexion-distraction with exercise program for chronic low back pain. *Chiropractic & Osteopathy* 2006;**14**:19.
* 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**:1070–82.

Güvenol 2000 ***{published data only}***
Güvenol K, Tüzün Ç, Peker O, Goktay Y. A comparison of inverted spinal traction and conventional traction in the treatment of lumbar disc herniations. *Physiotherapy Theory and Practice* 2000;**16**:151–160.

Harte 2007 ***{published data only}***
Harte AA, Baxter GD, Gracey JH. The effectiveness of motorised lumbar traction in the management of LBP with lumbo sacral nerve root involvement: a feasibility study. *BMC Musculoskeletal Disorders* 2007;**8**:118.

Konrad 1992 ***{published data only}***
Konrad K, Tatrai T, Hunka A, Vereckei E, Korondi I. Controlled trial of balneotherapy in treatment of low back pain. *Annals of the Rheumatic Diseases* 1992;**51**:820–2.

Larsson 1980 ***{published data only}***
Larsson U, Choler U, Lidström A, Lind G, Nachemson A, Nilsson B, et al. Auto-traction for treatment of lumbago-sciatica: a multicentre controlled investigation. *Acta Orthopaedica Scandinavica* 1980;**51**:791–8.

Letchuman 1993 ***{published data only}***
Letchuman R, Deusinger RH. Comparison of sacrospinalis myoelectric activity and pain levels in patients undergoing static and intermittent lumbar traction. *Spine* 1993;**18**(10): 1361–5.

Lidström 1970 ***{published data only}***
Lidström A, Zachrisson M. Physical therapy on low back pain and sciatica. *Scandinavian Journal of Rehabilitation Medicine* 1970;**2**:37–42.

Lind 1974 ***{published and unpublished data}***
* Lind GAM. *Auto-traction, Treatment of Low Back Pain and Sciatica. An Electromyographic, Radiographic and Clinical Study [thesis]*. Linköping: University of Linköping, 1974.

Ljunggren 1984 ***{published data only}***
Ljunggren AE, Weber H, Larsen S. Autotraction versus manual traction in patients with prolapsed lumber intervertebral discs. *Scandinavian Journal of Rehabilitation Medicine* 1984;**16**:117–24.

Ljunggren 1992 ***{published data only}***
Ljunggren AE, Walker L, Weber H, Amundsen T. Manual traction versus isometric exercises in patients with herniated intervertebral lumbar discs. *Physiotherapy Theory and Practice* 1992;**8**:207–13.

Mathews 1975 ***{published data only}***
Mathews JA, Hickling J. Lumbar traction: a double-blind controlled study for sciatica. *Rheumatology and Rehabilitation* 1975;**14**(4):222–5.

Mathews 1988 ***{published data only}***
Mathews JA, Mills SB, Jenkins VM, Grimes SM, Morkel MJ, Mathews W, et al. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *British Journal of Rheumatology* 1987;**26**(6): 416–23.
* Mathews W, Morkel M, Mathews J. Manipulation and traction for lumbago and sciatica: physiotherapeutic techniques used in two controlled trials. *Physiotherapy Practice* 1988;**4**:201–6.

Ozturk 2006 ***{published data only}***
Ozturk B, Gunduz OH, Ozoran K, Bostanoglu S. Effect of continuous lumbar traction on the size of herniated disc material in lumbar disc herniation. *Rheumatology International* 2006;**26**:622–6.

Pal 1986 ***{published data only}***
Pal B, Mangion P, Hossain MA, Diffey BL. A controlled trial of continuous lumbar traction in the treatment of back pain and sciatica. *British Journal of Rheumatology* 1986;**25** (2):181–3.

Reust 1988 ***{published data only}***
Reust P, Chantraine A, Vischer TL. Treatment of lumbar sciatica with or without neurological deficit using mechanical traction. A double-blind study. [Traitment par tractions mécaniques des lombosciatalgies avec ou sans déficit neurologique]. *Schweizerische Medizinische Wochenschrift* 1988;**118**(8):271–4.

Schimmel 2009 ***{published data only}***
Schimmel JJP, de Kleuver M, Horsting PP, Spruit M, Jacobs WCH, Van Limbeek J. No effect of traction in patients with low back pain: a single centre, single blind, randomized controlled trial of Intervertebral Differential Dynamics Therapy. *European Spine Journal* 2009;**18**:1843–50.

Sherry 2001 ***{published data only}***
Sherry E, Kitchener P, Smart R. A prospective randomized controlled study of VAX=D and TENS for the treatment of chronic low back pain. *Neurological Research* 2001;**23**: 780–4.

Simmerman 2011 ***{published data only}***
Simmerman SM, Sizer PS, Dedrick GS, Apte GG, Brismée JM. Immediate changes in spinal height and pain after aquatic vertical traction in patients with persistent low back symptoms: a crossover clinical trial. *PM & R: The Journal of Injury, Function, and Rehabilitation* 2011;**3**(5):447–57.

Sweetman 1993 ***{published data only}***
Sweetman BJ, Heinrich I, Anderson JAD. A randomized controlled trial of exercises, short wave diathermy, and traction for low back pain, with evidence of diagnosis-related response to treatment. *Journal of Orthopaedic Rheumatology* 1993;**6**:159–66.

Tesio 1993 ***{published data only}***
Tesio L, Merlo A. Autotraction versus passive traction: an open controlled study in lumbar disc herniation. *Archives of Physical Medicine and Rehabilitation* 1993;**74**:871–6.

Unlu 2008 ***{published data only}***
Unlu Z, Tasci S, Tarhan S, Pabuscu Y, Islak S. Comparison of 3 physical therapy modalities for acute pain in lumbar disc herniation measured by clinical evaluation and magnetic resonance imaging. *Journal of Manipulative and Physiological Therapeutics* 2008;**31**:191–8.

Van der Heijden 1995 *{published data only}*
Van der Heijden GJMG, Beurskens AJHM, Dirx MJM, Bouter LM, Lindeman E. Efficacy of lumbar traction: a randomised clinical trial. *Physiotherapy* 1995;**81**(1):29–35.

Walker 1982 *{published data only}*
Walker L, Svenkerud T, Weber H. [Traksjonbehandling ved lumbago–ischias: en kontrollert undersolske med Spina–trac]. *Fysioterapeuten* 1982;**49**:161-3, 177.

Weber 1973 *{published data only}*
Weber H. Traction therapy in sciatica due to disc prolapse. *Journal of the Oslo City Hospitals* 1973;**23**:167–76.

Weber 1984 *{published data only}*
Weber H, Ljunggren AE, Walker L. Traction therapy in patients with herniated lumbar intervertebral discs. *Journal of the Oslo City Hospitals* 1984;**34**:61–70.

Werners 1999 *{published data only}*
Werners R, Pynsent PB, Bulstrode CJK. Randomized trial comparing interferential therapy with motorized lumbar traction and massage in the management of low back pain in a primary care setting. *Spine* 1999;**24**(15):1579–84.

References to studies excluded from this review

Cevik 2007 *{published data only}*
Cevik R, Aslan B, Gur A, Sarac AJ, Yildiz H, Nas K, et al. Effect of new traction technique of prone position on distraction of lumbar vertebrae and its relation with different application of heating therapy in low back pain. *Journal of Back and Musculoskeletal Rehabilitation* 2007;**20**: 71–7.

Gose 1998 *{published data only}*
Gose E, Naguszewski K, Naguszewski K. Vertebral axial decompression therapy for pain associated with herniated or degenerated discs or facet syndrome: an outcome study. *Neurological Research* 1998;**20**:186–90.

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: a randomized, observer-blind trial. *Spine* 1993;**18**(1):98–107.

Moret 1998 *{published data only}*
Moret NC, Van der Stap M, Hagmeijer R, Molenaar A, Loes BW. Design and feasibility of a randomized clinical trial to evaluate the effect of vertical traction in patients with a lumber radicular syndrome. *Manual Therapy* 1998; **3**:203–11.

Olah 2008 *{published data only}*
Olah M, Molnar L, Dobai J, Olah C, Feher J, Bender T. The effects of weightbath traction hydrotherapy as a component of complex physical therapy in disorders of the cervical and lumbar spine: a controlled pilot study with follow-up. *Rheumatology International* 2008;**28**:749–56.

Ramos 1994 *{published data only}*
Ramos G, Martin W. Effects of vertebral axial decompression on intradiscal pressure. *Journal of Neurosurgery* 1994;**81**: 350–3.

Van der Heijden 1991 *{published data only}*
Van der Heijden GJMG, Bouter LM, Terpstra-Lindeman E, Essers AHM, Waltjè EMH, Köke AJA, et al. [De effectiviteit van tractie bij lage rugklachten: de resultaten van een garandomisserde en geblindeerde pilotstudy]. *Nederlands Tijdschrift voor Fysiotherapie* 1991;**101**(2): 37–41.

Additional references

Bigos 1994
Bigos SJ, Bowyer OR, Braen GR, Brown K, Deyo R, Haldeman S, et al. *Acute Low Back Problems in Adults*. Rockville, Maryland: U.S. Department of Health and Human Services, 1994. [AHCPR Publication No. 95–0642]

Blomberg 2005
Blomberg S. A new pragmatic management strategy for low-back pain - an integrated multimodal programme. In: Hutson M, Ellis R editor(s). *Textbook of Musculoskeletal Medicine*. Corby, UK: Oxford University Press, 2005:1–20. [ISBN 0–19–263050–4]

Boutron 2005
Boutron I, Moher D, Tugwell P, Giraudeau B, Poiraudeau S, Nizard R, et al. A checklist to evaluate a report of a non pharmacological trial (CLEAR NPT) was developed using consensus. *Journal of Clinical Epidemiology* 2005;**58**: 1233–40.

Chou 2007
Chou R, Huffman LH, American Pain Society, American College of Physicians. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine* 2007;**147**(7):492–504.

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. *The Spine Journal* 2008;**8**(1):8–20.

Furlan 2009
Furlan AD, Pennick V, Bombardier C, Van Tulder M, Editorial Board CBRG. 2009 Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009;**34**(18):1929–41.

Gay 2008
Gay RE, Brault JS. Evidence-informed management of chronic low back pain with traction therapy. *Spine* 2008;**8** (1):234–42.

Gotzsche 2012
Gotzsche P, Ioannidis J. Content area experts as authors: helpful or harmful for systematic reviews and meta-analyses? . *BMJ* 2012;**345**:e7031.

Harte 2003
Harte AA, Baxter GD, Gracey JH. The efficacy of traction for back pain: a systematic review of randomized controlled trials. *Archives of Physical Medicine and Rehabilitation* 2003; **84**:1542–53.

Harte 2005
Harte AA, Gracey JH, Baxter GD. Current use of lumbar traction in the management of low back pain: results of a survey of physiotherapists in the United Kingdom. *Archives of Physical Medicine and Rehabilitation* 2005;**86**:1164–9.

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 www.cochrane-handbook.org.

Krause 2000
Krause M, Refshauge KM, Dessen M, Boland R. Lumbar spine traction: evaluation of effects and recommended application for treatment. *Manual Therapy* 2000;**5**(2): 72–81.

Lambeek 2011
Lambeek LC, Van Tulder MW, Swinkels IC, Koppes LL, Anema JR, Van Mechelen W. The trend in total cost of back pain in the Netherlands in the period 2002 to 2007. *Spine* 2011;**36**(13):1050–8.

Li 2001
Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. *Physical Therapy* 2001;**81**(4):1018–28.

Ostelo 2008
Ostelo RW, Devo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 2008;**33**(1):90–4.

Sheffield 1996
Sheffield FJ. Adaptation of tilt table for lumbar traction. *Arch Phys Med Rehabil* 1964;**45**:469–72.

Van Middelkoop 2011
Van Middelkoop M, Rubinstein SM, Kuijpers T, Verhagen AP, Ostelo R, Koes BW, et al. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *European Spine Journal* 2011;**20**(1):19–39.

Van Tulder 2003
Van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in The Cochrane Collaboration Back Review Group. *Spine* 2003;**28**(12): 1290–9.

Van Tulder 2009
Van Tulder MW, Suttorp M, Morton S, Bouter LM, Shekelle P. Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine* 2009;**34**(16):111–6.

Vos 2012
Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;**380**(9859):2163–96.

References to other published versions of this review

Clarke 2006a
Clarke JA, Van Tulder MW, Blomberg SEI, de Vet HCW, Van der Heijden GJMG, Bronfort G. Traction for low-back pain with or without sciatica. *Cochrane Database of Systematic Reviews* 2005, Issue 4. DOI: 10.1002/14651858.CD003010.pub4

Clarke 2006b
Clarke J, Van Tulder M, Blomberg S, de Vet H, Van der Heijden G, Bronfort G. Traction for low back pain with or without sciatica: an updated systematic review within the framework of The Cochrane Collaboration. *Spine* 2006;**31** (14):1591–9.

Van der Heijden 1995a
Van der Heijden GJMG, Beurskens AJHM, Koes BW, Assendelft WJJ, de Vet HCW, et al. The efficacy of traction for back and neck pain: a systematic, blinded review of randomized clinical trials methods. *Physical Therapy* 1995; **75**(2):18–29.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

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

Beurskens 1997

Methods	RCT; participants randomly allocated by computer, sealed envelopes prepared by independent person, containing treatment code. Stratified on duration of complaints (< 6 or > 6 months), and according to PT practices
Participants	151 participants (85 male and 66 female, > 18 years old) recruited by physiotherapists and general practitioners in the Netherlands, with at least 6 wk of subacute and chronic nonspecific LBP, having never had any form of lumbar traction treatment. 150 completed 12-wk follow-up and 148 completed 6-month follow-up
Interventions	T) Traction: continuous mechanical traction with Eltrac, DIMEC Delft Instruments, the Netherlands. Traction force increased until participant indicated tolerance for pulling was reached, with minimum force of 35% and maximum of 50% of body weight. C) Comparison intervention: sham traction. Same as above except traction force was slowly increased until participant indicated feeling little pulling with maximum force of 20% body weight. Special brace worn around iliac crest, which became tighter in the back during treatment. Both groups treated 12 times in 5 wk for 20 min per session.
Outcomes	At 5 wk: global perceived effect (number and %): T) 34 (44%), C) 37 (51%); first main complain (mean): T) 28.5, C) 28.4; second main complaint (mean): T) 27, C) 24.6; RMDQ (mean): T) 3.5, C) 4.8; pain at the moment (mean): T) 21.2, C) 22.5; pain last wk (mean): T) 20.6, C) 23.7; severity of LBP (mean): T) 1.6, C)1.8; ROM (mean): T) -2.1, C) 0.1; ADL disability (mean): T) 26.7, C) 33.8; work absence (days) (mean): T) 21, C) 22.8. No significant differences on any outcome measures. At 12 wk: global perceived effect-recovery (number and %): T) 38 (50%), C) 35 (48%) ; first main complaint (mean): T) 33.7, C) 31.5; second main complaint (mean): T) 35. 4, C) 30.7; RMDQ (mean): T) 4.4, C) 4.3; pain at the moment (mean): T) 28.5, C) 22.8; severity of LBP (mean): T) 2.3, C) 2.2; ROM: T) -1.1, C) 1.2; ADL disability (mean): T) 27.1, C) 29.4; work absence (days) (mean): T) 23.5, C) 27.8. At 6 months: global perceived effect (number and %): T) 35 (47%), C) 32 (44%); first main complain (mean): T) 36.7, C) 36.0; second main complaint (mean): T) 35.8, C) 32.8; RMDQ (mean): T) 4.7, C) 4.0; pain at the moment (mean): T) 23.8, C) 20.1; ADL disability (mean): T) 25.7, C) 25.8; work absence (days) (mean): T) 35.7, C) 43.7 No significant differences on any outcome measures.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Allocation with the help of a random numbered list generated by computer

Beurskens 1997 (*Continued*)

Allocation concealment (selection bias)	Low risk	Sealed envelopes prepared by an independent person containing the treatment code
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Participants were blinded to treatment allocation.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	After admission of a participant into the trial, the treating physiotherapist received a sealed envelope that contained the treatment code. The envelope was opened at the first treatment session and, therefore, the care provider was not blinded for the assigned treatment
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessors were blinded to treatment allocation.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	Of the 151 participants, only 1 was lost to follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	The 2 treatment groups had similar demographic and clinical baseline characteristics
Influence of co-interventions (performance bias)	Low risk	Co-interventions, other than pain medication, were not allowed during the treatment period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Bihaug 1978

Methods	RCT; method of randomization not described.
Participants	42 participants (23 male, 19 female, aged 19-71 years (mean 44.1 years) referred from secondary care setting. All had radicular pain; in 32 radiating pain was below the knee. Pain duration was 3-52 wk (mean 9.7 wk). 25 participants were on sick leave at baseline (1-24 wk, mean 5.1 wk). 18 had severe pain, the remainder had moderate pain. 27 had neurological deficits (figures not given for the 2 different groups)
Interventions	T) Traction: auto-traction, using a combination of Lind's method and Myrin's method. Instead of pulling with the arms (as in Lind), participants pushed with 1 or both arms (according to Myrin/Spina-Trac). 4-12 sessions (mean 8.2), with interval of 3.1 days between sessions. (Force 70 kiloponds according to Lind.) All participants also received education in LBP/ biomechanics). C) Comparison intervention: exercise. Isometric exercises of the abdominal and pelvic floor muscles, to increase abdominal pressure (and, in turn, to increase intrinsic lumbar support) (Hume, Kendall and Jenkins; Fysioterapeuten number 3, Norway). 4-12 sessions (mean 10.6) with interval of 4.1 days between sessions)
Outcomes	Global improvement (symptom-free; mild symptoms with ability to work; some or no improvement; deterioration) (n). At end of treatment series: T) 5, 12, 3, 1; C) 2, 9, 10, 0. At 1 month AT: T) 12, 7, 2, 0; C) 5, 11, 5, 0. At 3 months AT: T) 16, 4, 1, 0; C) 12, 7, 2, 0.
Notes	Outcomes inappropriately dichotomized by authors, leading to P value < 0.05 at end of treatment series (ns at other follow-up points). Without this dichotomization, group differences are not statistically significant at any follow-up point

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Transformed from old format to new format.
Allocation concealment (selection bias)	Unclear risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Transformed from old format to new format.

Bihaug 1978 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	Transformed from old format to new format.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Transformed from old format to new format.
Selective reporting (reporting bias)	Unclear risk	Transformed from old format to new format.
Group similarity at baseline (selection bias)	Unclear risk	Transformed from old format to new format.
Influence of co-interventions (performance bias)	Unclear risk	Transformed from old format to new format.
Compliance with interventions (performance bias)	Low risk	Transformed from old format to new format.
Timing of outcome assessments (detection bias)	Unclear risk	Transformed from old format to new format.

Borman 2003

Methods	RCT; method of randomization not described.
Participants	42 participants (14 male, 28 female; age: T) 38.5 ± 8.4 years, C) 42.8 ± 10.5 years) with persistent (> 6 months) or recurring, non-specific LBP, or both; outpatients in physical medicine and rehabilitation department of large hospital. Duration of pain (months): T) 27 ± 19.5, C) 34.09 ± 14.1. Ratio of participants with/without radiation: T) 14:7, C) 13:8. Excluded those with neurological deficits.
Interventions	T) Traction and standard PT. Motorized traction (Eltrac 439, Enraf, the Netherlands) , 10 x 20-min sessions, participants lying on traction table in semi-fowler position. Canvas braces attached around iliac crest and lower thoracic region, with force increased to maximum of 50% body weight. Traction applied between ultrasound therapy and exercise sessions in standard PT programme (as below). C) Comparison intervention: standard PT. Included hot packs (10 min), ultrasound (10 min), exercise (20 min)
Outcomes	Pain (VAS) (mean, SD (range)): before: T) 5.7, 1.1 (3-8); C) 5.6, 1.7 (2-9); immediately after: T) 3.8, 1.1 (1-6); C) 3.8, 1.4 (1-7). Within-group difference P value < 0.01; between-group difference ns. 3 months. Follow-up: T) 4.1, 1.7 (0-7); C) 3.6, 1.7 (0-6). ODI: (mean, SD (range)): before: T) 32.3, 9.6 (12-44); C) 25.2, 10.4 (3-41); immediately after: T) 26.8, 9.1 (4-41); C) 22.9, 10.1 (3-43). Within-group differences P value < 0.01. 3 months. Follow-up: T) 23.7, 10.8 (6-38); C) 19.7, 10.8 (0-32). Within-group differ-

	ence P value < 0.05; between-group difference ns. Global improvement (complete/mild improvement, no change, no improvement and worse) (n): immediately after: T) 11, 6, 5; C) 10, 6, 5. 3 months follow-up: T) 8, 7, 5; C) 7, 5, 7. Between-group difference ns. Global satisfaction (n (%) of participants completely/somewhat satisfied; not satisfied) : immediately after: T) 17 (80.9%), 4 (19%); C) 15 (71.4%), 6 (28.6%); 3 months' follow-up: T) 12 (60%), 8 (40%); C) 11 (57.8%), 8 (42.1%). No differences were observed in outcomes for participants with and without radiation (P value > 0.05)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Unclear risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	4 participants were lost to follow-up (9.5%) : 2 in each group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	It is not clear whether an intention-to-treat analysis was used or not
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	There were no differences between groups in terms of age, sex, duration of pain, VAS and ODI scores at entry
Influence of co-interventions (performance bias)	Low risk	No co-interventions were allowed during the treatment period

Borman 2003 (*Continued*)

Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Coxhead 1981

Methods	RCT; randomly allocated treatment (method of randomization not described). The design was factorial - there were 16 treatment groups, enabling a comparison of combinations of methods as well as of individual methods
Participants	334 participants (185 men, 149 women, mean age 41.9 years) referred to the outpatient department with sciatic pain at least as far as the buttock crease, with/without back pain. Pain not due to malignant or infective disease, gynaecological disorders, sacroiliac disease, vertebral collapse or gross structural abnormality. Mean duration of symptoms 14.3 wk
Interventions	T) Traction: Tru-Trac apparatus, giving intermittent traction at pre-set forces and time intervals. Duration and intensity at the discretion of the physiotherapist. Comparison interventions: C1) Exercises based on a catalogue of exercises that brought in all ROM and muscle groups; C2) Manipulation by Maitland technique; C3) Corset - a ready-made fabric lumbar support available in 3 sizes. All participants received short-wave diathermy and a standardized 30-min "back school" lecture. For all interventions, participants treated daily for first wk, with decreasing frequency in the following 3 wk
Outcomes	Participant assessments at 4 wk, 16 wk (better): T) 82%, 72%; C1) 82%, 75%; C2) 80%, 69%; C3) 81%, 71%. Pain (-100 to +100 VAS) at 4 wk: T) 50.1 (37.9); C1) 52.6 (36.9); C2) 49.0 (40.0); C3) 49.8 (37.9). Statistical significance in C1 only. ROW at 4 wk: T) 36%; C1) 36%; C2) 33%; C3) 33%.
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.

Coxhead 1981 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Outcome assessors were not blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	High risk	At 4 months follow-up only 78% of the included participants were assessed
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Unclear risk	No information provided on demographic characteristics at baseline
Influence of co-interventions (performance bias)	Unclear risk	Unclear whether co-interventions were allowed during the treatment period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Fritz 2007

Methods	RCT; computer-generated random number lists and concealment of allocation by means of randomization envelopes
Participants	64 participants (33 in the extension group, 31 in the traction plus extension group) with symptoms of pain or numbness (or both) extending distal to the buttocks and signs of nerve root compression in the past 24 hours. All had LBP, 76.5% sciatica. Exclusion criteria included non-mechanical LBP and previous spinal fusion or spine surgery in the past 6 months. Mean age T) 41.7 years, C) 40.7 years. Duration of complaints: 47.5 days

Fritz 2007 (*Continued*)

Interventions	T) Traction: extension-oriented treatment and mechanical traction using an adjustable table. Traction during first 2 wk of treatment, 4 sessions per wk, 12 min per session, with a traction force of 40-60% of body weight. Extension-oriented treatment included 9 sessions of exercise, mobilization and education during a 6-wk treatment period C) Comparison intervention: extension-oriented treatment.
Outcomes	Assessment at 2 and 6 wk' post-treatment. ODI (all measurements: MD): 2 wk 7.2 (95% CI 0.13 to 14.3), 6 wk 1.8 (95% CI -6.4 to 10.1). Pain rating: 2 wk 0.23 (95% CI -1.4 to 1.9), 6 wk -0.17 (95% CI -1.4 to 1.1). FABQ - physical activity subscale: 2 wk 2.7 (95% CI 0.66 to 4.6), 6 wk 0.50 (95% CI -2.4 to 3.4). FABQ - work subscale: 2 wk -1.1 (95% CI -4.2 to 1.9), 6 wk -3.1 (95% CI -6.5 to 0.36)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random number list.
Allocation concealment (selection bias)	Low risk	Randomization envelopes.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Outcome assessors did not participate in the subject's treatment and were blinded to the treatment allocation. However, blinding was lost for 15 subjects (20%)
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	8 participants were lost to follow-up (12.5%).
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	There were no between-group differences at baseline, other than a higher percentage of participants using prescription pain

Fritz 2007 (*Continued*)

		medication in the TRACT group
Influence of co-interventions (performance bias)	Low risk	No co-interventions, other than analgesics, were allowed during the treatment period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Gudavalli 2006

Methods	RCT; random number tables and concealment of allocation by means of randomization envelopes
Participants	235 participants (123 in the flexion-distraction group, 112 in the active trunk exercise programme) with LBP with a duration of at least 3 months. All had LBP, 22.8% sciatica. Mean age: T) 42.2 years, C) 40.9 years
Interventions	T) Traction: flexion-distraction technique during 4 wk, 2-4 sessions per wk, 9-18 min of traction per session C) Comparison intervention: active trunk exercise programme. Treatment duration of 4 wk, 2-4 sessions per wk, 30-45 min per session
Outcomes	Assessment at 4 wk, 3 months and 12 months from baseline. VAS (mean change from baseline to time period indicated in MD (SE)): 4 wk: T) 20.57 (2.00), C) 12.34 (1.80); 3 months: T) 16.52 (2.95), C) 12.04 (2.53); 6 month: T) 18.26 (2.64), C) 8.92 (2.89) ; 12 months: T) 17.10 (2.55), C) 12.36 (2.43) RMDI: 4 wk: T) 2.81 (0.38), C) 2.30 (0.33); 3 months: T) 3.50 (0.50), C) 3.75 (0.51) ; 6 months: T) 3.89 (0.46), C) 3.42 (0.50); 12 months: T) 3.90 (0.53), C) 3.77 (0.44)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Random number tables.
Allocation concealment (selection bias)	Low risk	Sequentially numbered, sealed, manila envelopes.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.

Gudavalli 2006 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessors were blinded and all remained blinded for the entire study period. No incidents of unblinding were reported
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	High risk	Although total loss to follow-up was only 16.6%, significantly more subjects in the active trunk exercise programme group dropped out of the study (T) 13, C) 25)
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	No significant differences were found at baseline.
Influence of co-interventions (performance bias)	Low risk	Co-interventions were not allowed during the treatment period. Analgesics were not allowed 24 hours prior to measurements
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Güvenol 2000

Methods	RCT; method of randomization not described.
Participants	29 participants (mean age: T1) 33.8 years, T2) 39.6 years) with LBP and lower extremity pain of not less than 1 month, and lumbar disc herniation diagnosed by CT. Mean duration of pain (months): T1) 28.5 ± 26.5 months, T2) 39.3 ± 39.2 months). None had history of spinal surgery. Pain not due to disease such as malignant, inflammatory, infectious, metabolic, congenital or developmental disorders. Disc pathology at 2 levels was present in 10 subjects, 5 from each treatment group
Interventions	Traction: T1) Inversion spinal traction. Traction used a modified tilt table (Sheffield 1996). With participant lying supine, ankle straps mounted to the foot of the table; lumbar strap allowed vertical slide only. Table rotated until participant was upside down (inverted).

Güvenol 2000 (*Continued*)

	Inverted for 5 min on 1st day, 8 min on 2nd, 10 min on 3rd and onwards through 7 days (10 days total). T2) Conventional static traction. Initial force 30 kg, gradually increased up to 45 kg with 3-kg increments daily, according to participant's tolerance. Both T1) and T2) also received 15 min of infrared radiation, with abdominal and gluteal isometric exercises. Participants were not allowed to take NSAIDS; bed rest was required of all participants
Outcomes	Clinical parameters examined before, immediately after and 3 months after last treatment session. Pain cluster 1 - combination of: morning pain; pain throughout the day; night pain; pain with Valsalva manoeuvre; radicular pain. Pain cluster 2 - combination of: straight leg raising test pain onset; finger-to-floor distance; deep tendon reflex, sensory impairment, and motor strength; CT investigation. Results presented as P values only.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Unclear risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	4 participants were lost to follow-up (14%) : 2 from each group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.

Güvenol 2000 (*Continued*)

Group similarity at baseline (selection bias)	Low risk	There was no significant difference between groups regarding any of the baseline characteristics
Influence of co-interventions (performance bias)	Low risk	No co-interventions, other than analgesics, were allowed during the study period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Harte 2007

Methods	RCT; predetermined randomization table, concealment of allocation through sealed, opaque and sequentially numbered envelopes
Participants	30 participants (16 in the traction group and 14 in the manual therapy group) with acute or subacute LBP accompanied with radiculopathy. Exclusion in case of previous spinal surgery, co-existing conditions interventions within the last 3 months. Mean age T) 45.25 years, C) 42.79 years. Duration of complaints: T) 6.5 wk, C) 6 wk
Interventions	T) Traction: manual therapy (techniques described by Maitland or Cyriax), exercises, advice and motorized lumbar traction for 4-6 wk, 2-3 times per wk, 10-20 min per session, traction force 5-60 kg C) Comparison intervention: manual therapy, exercises and advice
Outcomes	Assessment at discharge, 3 months and 6 months post-treatment (all measures median (IQR), T vs. C). RMDQ: at discharge: 4 (5.8) vs. 4 (10.3), 3 months: 4.5 (10.8) vs. 1 (10.5), 6 months: 4.5 (15.3) vs. 2.5 (14). MPQ-PRI: at discharge: 4 (15.3) vs. 12 (16.5), 3 months: 6 (16.5) vs. 6 (21), 6 months: 10 (20.5) vs. 6.5 (21). SF36 PCS: at discharge: 38.5 (16.2) vs. 41.1 (21.1), 3 months: 41.6 (18.6) vs. 43.2 (24), 6 months: 40 (15) vs. 46 (22). SF36 MCS: at discharge: 52 (26.1) vs. 48.3 (25.6), 3 months: 49.5 (25.8) vs. 47.3 (21.3), 6 months: 51.8 (23) vs. 49.8 (19.8)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Predetermined randomization table.
Allocation concealment (selection bias)	Low risk	Sealed, opaque and sequentially numbered envelopes.

Harte 2007 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessors were blinded to treatment group allocation
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	High risk	7 participants were lost to follow-up (23%) .
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	High risk	Published results did not include all prespecified outcomes: VAS score, improvement and straight leg raising test
Group similarity at baseline (selection bias)	High risk	Baseline characteristics varied between groups: off work due to LBP, history of episodes, participation in physical activity and presence of neurological signs
Influence of co-interventions (performance bias)	Low risk	Participants were not permitted to receive any other type of manual therapy or any additional interventions during the treatment period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Konrad 1992

Methods	RCT; participants were randomly allocated to 1 of 4 groups in each factory. Method of randomization not described
Participants	170 participants (95 female, 75 male, mean age of 41.5 years) from 3 factories in Budapest, with non-specific back pain localized to the lumbosacral region, with or without radiation to the thigh. Duration of pain at least 1 month, but no longer than 3 months. A pain-free year before onset of the current episode.

Konrad 1992 (*Continued*)

	Exclusion criteria: participants with pregnancy, back surgery, spondylolisthesis, infections, tumours, fractures, ankylosing spondylitis, osteoporosis and structural scoliosis. 12 participants dropped out (3 from the balneotherapy group and 9 from the underwater massage group) and were analyzed separately
Interventions	T) Traction: underwater traction. Participant fixed perpendicularly in special deep pool, bar grasped under the arms and traction applied. 1st treatment - participant's own weight used. Then, in addition to traction due to gravity, traction belt applied to the pelvis with 3-kg weight on both sides. Comparison interventions: C1) Balneotherapy. Participants immersed in thermal water with minerals. C2) Underwater massage. Same water, with massage and movement while a stream of hot water (37 °C, 1 atm, 10 cm) played on the affected part. C3) Control group (no treatment). All treatments done for 15 min, 3 times per wk, for 4 wk. All participants taught how to use their back correctly. Only NSAIDs were offered to participants in the control group
Outcomes	Number of analgesics taken on admission, at 4 wk, at 1 year: T) 5.1 (2.9), 2.2 (0.9), 2.1 (1.2); C1) 4.8 (3.2), 2.3 (1.3), 1.9 (1.8); C2) 4.9 (3.4), 1.8 (0.7), 2.3 (1.7); C3) 5.1 (2.8), 3.9 (2.7), 3.7 (1.9). At 1 month, statistically significant difference in all treatment groups compared to control (P value < 0.01). No significant difference in analgesic consumption between the treatment groups. Pain intensity (100 mm VAS) on admission, at 4 wk, at 1 year: T) 56.7 (28.2), 24.6 (11.9), 45.8 (26.2); C1) 63.4 (24.1), 31.7 (16.2), 49.5 (25.7); C2) 68.4 (31.8), 33.5 (19.1), 54.7 (33.7); C3) 61.5 (32.88), 53.7 (23.8), 54.9 (24.8). At 1 month, statistically significant pain reduction in all treatment groups (P value < 0.01). No significant difference in control group At 1 year, no difference between groups. Reduction in analgesic consumption well maintained in treatment groups
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded

Konrad 1992 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	The investigator assessing the outcome was not aware of the treatment given
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	12 participants were lost to follow-up (7%) .
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Unclear risk	Published results did not include all prespecified outcomes: spinal ROM and straight leg raising
Group similarity at baseline (selection bias)	Low risk	Groups were comparable at baseline regarding age, sex and medical history
Influence of co-interventions (performance bias)	Low risk	No co-interventions, other than analgesics, were allowed during the study period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Larsson 1980

Methods	RCT; method of randomization not described.
Participants	82 participants (51 males and 31 female, age 20-55 years) in 6 departments of orthopaedic surgery in Sweden, with lumbago-sciatica with or without symptoms of neurological deficit. Duration of current episode at least 2 wk and not more than 3.5 months, positive straight leg raise test
Interventions	T) Traction: auto-traction: up to 3 treatments within 1 wk as per Lind (1974). Pelvis fixed to the foot end of bench, participant grasps bars at end and performs traction himself by pulling his arms. Participant supplied with reinforced, high, fabric corset and special pillow. Sessions < 1 hour. Participants treated as outpatients were usually taken home by ambulance. Participants confined to bed for first few days, then mobilized gradually in corset. C) Comparison intervention: corset of same type as traction group and same instructions with respect to rest. Standard analgesics (paracetamol) prescribed when required for both groups

Outcomes	Complete recoveries 1 wk, 3 wk: T) 15%, 17% C) 0%, 7%. Partial recoveries 1 wk, 3 wk: T) 27%, 32% C) 4%, 12%. Statistically significant between group differences in participant's recovery at 1 wk. At 3 wk, ns for those "completely recovered" but significant for those "completely recovered or free from pain in the leg" and "completely recovered or free from pain in the leg or the back", with traction group having better results
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessors were blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	It is unclear how many participants were lost to follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	It is unclear whether intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	Clinical characteristics were evenly distributed between the 2 groups at baseline
Influence of co-interventions (performance bias)	Low risk	No co-interventions, other than analgesics, were allowed during the treatment period
Compliance with interventions (performance bias)	Low risk	Participants were hospitalized, therefore, compliance with the given treatment was

Larsson 1980 (*Continued*)

		high
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Letchuman 1993

Methods	RCT, cross-over. Subjects randomly assigned to 1 of the 2 experimental groups, with each subject serving as his/her own control in the control group (method of randomization not described)
Participants	26 subjects (16 male, 10 female, aged 26-65 years) referred from physicians. Participants with LBP with/without lower extremity pain and neurological signs. Cough, sneeze or deep breaths did not cause severe pain, x-rays, MRI or CT scan of lumbar spine taken within past 6 months
Interventions	Traction: T1) Static (mechanical traction), continuous traction force (after sham treatment) for a 6-min period at magnitude of 50% bodyweight. T2) Intermittent traction, for a 6-min period (after sham treatment), with a 10-sec hold period at a magnitude of 50% body weight, followed by a 10-second rest period. C) Comparison intervention: sham treatment. 6 min of 'sham traction', using only 10 lb (4.5 kg) for a 10-sec hold, and 0 lb for a 10-sec rest
Outcomes	Pain intensity (0-10 VAS). Decreased pain: T1) 53.9% (7 of 13 participants), T2) 61.5% (8 of 13 participants). Increased pain: T1) 30.8% (4 of 13 participants), T2) 15.4% (2 of 13 participants)
Notes	Major thrust of study was to look at myoelectric activity for static or intermittent traction. Pain measures were recorded immediately after traction. Just 1 session of traction appears to have been given. Small sample size, frequency data only reported for pain measures

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	No mention of attempts to blind the participants. It is unlikely that the participants were aware of group assignment

Letchuman 1993 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	4 participants were lost to follow-up (13%) : 2 in each group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	Groups were similar at baseline with respect to age, sex and symptoms
Influence of co-interventions (performance bias)	Low risk	No co-interventions were used.
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Lidström 1970

Methods	RCT, subjects were placed by a physiotherapist in 1 of 3 groups according to a randomization procedure decided before the experiment (method of randomization not described)
Participants	62 participants (29 male, 33 female, aged 21-61 years) selected from an orthopaedic outpatient clinic. Participants had LBP and sciatic pain radiating down 1 leg for more than 1 month' duration. 32 participants had a history of pain > 1 year. Participants strongly suspicious of the presence of a disc prolapse were not accepted
Interventions	T) Traction: intermittent pelvic traction with a Tru-Trac traction table for 20 min with 4-sec hold intervals and a 2-sec rest. Traction force was correlated to the participant's weight according to the given figures. Instruction on Fowler position, strengthening exercises, regimental dispositions, every day at home. Comparison interventions: C1) Conventional treatment, hot packs for 15 min, massage and mobilizing exercises. C2) Control, hot packs for a length of time corresponding with the mean for the other

Lidström 1970 (*Continued*)

	methods of treatment
Outcomes	Global measure - participants opinion of noticeable improvement: T) 90% (18 of 20 participants), C1) 48% (10 of 21 participants), C2) 67% (14 of 21 participants). Need for analgesics before, after the treatments (of the 30 that were taking pills before the treatment): T) 9, 0; C1) 12, 7; C2) 9, 4. Traction appears to have reduced the subjective symptoms of the participants to a higher degree than the other methods
Notes	Authors stress the need for sufficient pull and duration of traction in order to influence the mechanical conditions of the spine effectively. No apparent follow-up after the treatment had finished (i.e. other than post-treatment)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Both the care provider and a blinded outcome assessor took part in the assessment of the outcome measures
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	4 participants (6.5%) did not complete follow-up evaluation.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	All 3 groups were similar at baseline.
Influence of co-interventions (performance bias)	High risk	The traction group received isometrical training in conjunction with traction. The comparison group was not treated with iso-

Lidström 1970 (*Continued*)

		metrical training
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Lind 1974

Methods	RCT, method of randomization not reported.
Participants	45 participants (29 male, 16 female; aged 30-50 years, mean 34.0 years) from waiting list of orthopaedic surgery department. All had several periods of attack, mean number 3.5. Participants with serious disorders (e.g. arteriosclerosis, hypertension) excluded. All had had some previous non-surgical therapy. Included participants with or without neurological signs
Interventions	T) Traction: auto-traction treatment followed initially by bed rest, correction of statico-dynamic disorders and advice on spinal hygiene. No PT or medicine. 1 participant given cotton corset. Mean number of treatments, approximately 1 hour long, over 1-3 wk: 3.7. Comparison interventions: C1) PT, with physiotherapist choosing individual treatment, including drugs. 12 of 15 participants received Tru-Trac traction; other treatments included isometric muscle training (n = 14), ergonomic instruction (n = 11), shortwave therapy (n = 7), heat (n = 7), cycle machine (n = 10), bath (n = 4) and manipulation (n = 1). C2) Bed rest and analgesics (Paraflex comp, 3-6 tablets/day), sham shortwave therapy
Outcomes	Disappearance of pain in lower back/legs without coughing/sneezing: T) 100%, C1) 53%, C2) 43%. Disappearance of pain in lower back/legs on coughing sneezing: T) 100%, C1) 50%, C2) 0%. Pain, mean distance radiated (initial radiation mean; at 3 wk; mean change score): T) 60 cm, 0 cm, 100%; C1) 66 cm; 23 cm, 65%; C2) 65 cm, 28 cm, 57%. Participant's own evaluation at 3 wk (1, 2, 3, 4, 0, -1 where 1 = highest improvement, 4 = unchanged, -1 = worse) T) 11, 2, 2, 0, 0, 0; C1) 0, 0, 6, 3, 5, 1; C2) 0, 2, 7, 3, 2, 0. (T vs. C1, P value < 0.000001; T vs. C2, P value < 0.0001) Recovery: T) 87%, C1) 0%, C2) 0%. P value < 0.00001 at 3 wk. Straight leg raising (% recovered) T) 100%, C1) 0%, C2) 0% (P value < 0.001). Regression of neurological deficits: auto-traction more effective in effecting a regression of neurological deficits
Notes	Although no final conclusions were made by the authors, we can assume it had a positive conclusion considering the P values reported. This is an underpowered study that would need replication

Lind 1974 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Transformed from old format to new format.
Allocation concealment (selection bias)	Unclear risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Transformed from old format to new format.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	Transformed from old format to new format.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Transformed from old format to new format.
Selective reporting (reporting bias)	Unclear risk	Transformed from old format to new format.
Group similarity at baseline (selection bias)	Low risk	Transformed from old format to new format.
Influence of co-interventions (performance bias)	Low risk	Transformed from old format to new format.
Compliance with interventions (performance bias)	Low risk	Transformed from old format to new format.
Timing of outcome assessments (detection bias)	Unclear risk	Transformed from old format to new format.

Ljunggren 1984

Methods	RCT (method of randomization not described)
Participants	52 hospitalized participants with lumbago-sciatica and prolapsed lumbar intervertebral discs, admitted to neurological department, and considered for operation. Inclusion criteria: radicular signs L5 or S1 (or both) nerve root; symptoms aggravated or unchanged in last 2-4 wk
Interventions	T1) Auto-traction and modified Gertrud Lind: traction force between 33% and 100% of participant's body weight; each pull for some seconds and sometimes up to 2 min. Every treatment lasted about 1 hour T2) Manual traction and modified manual therapy. Traction force scarcely reached 300 N. Static traction given twice, each pull lasting for 5 min
Outcomes	Immediately AT: overall assessment: no effect (number) T1) 21, T2) 15. Moderate effect (number): T1) 2, T2) 4. Good effect (number) T1) 3, T2) 4. At 2 wk: overall assessment: no effect (number) T1) 21, T2) 16. Moderate effect (number): T1) 1, T2) 4. Good effect (number) T1) 4, T2) 3. At 3 months: identical to results at 2 wk. Pain intensity (VAS) median (SD): BT: T1) 1.3 (0.3-3.5), T2) 3.5 (0.9-6.0). AT: T1) 0.8 (0-1.8), T2) 1.6 (0.2-3.0)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Participants were not informed about their participation in a randomized investigation with 2 treatment modalities
Blinding (performance bias and detection bias) All outcomes - providers	High risk	There is no mention of blinding of the care providers, but it is unlikely that they were
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	The outcome assessor was blinded to the treatment allocation
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	3 participants (5.8%) were lost to follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.

Ljunggren 1984 (*Continued*)

Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	High risk	Groups were not similar at baseline with regards to level of herniation, duration since first symptoms of sciatica and pain intensity in the lower back
Influence of co-interventions (performance bias)	Low risk	Participants were deprived of long-term working analgesics later than hours prior to the traction session
Compliance with interventions (performance bias)	Low risk	All participants were hospitalized, therefore, the compliance with the given treatment was high
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Ljunggren 1992

Methods	RCT (method of randomization not described)
Participants	50 participants (27 males, 23 females, aged 16-62 years) admitted to the department of neurology were included. Inclusion criteria: radiating pain, neurological symptoms and signs confirmed by a myelogram. Participants with previous spinal surgery, spondylolisthesis and root entrapment were excluded. The males had a mean duration of symptoms for 4.8 months, and the females for 5.3 months
Interventions	T) Traction: continuous manual (static) traction. The therapist exerted traction by gently leaning backwards against a belt placed around the back or hips, and attached below the knees of the participant. The traction force reached approximately 300 N. Repeated relief of pain was guiding factor; once per day for 10 min (in a few cases twice per day for 5 min). C) Comparison intervention: isometric exercises for the abdominal, back, hip and thigh muscles. Education about importance of these muscles was given. Contractions 6-8 sec, repeated 5-10 times, daily session approximately 20 min. Following treatment, all participants were instructed to lie in the most comfortable positions for 2 hours. Treatment for all participants lasted 5-7 days
Outcomes	Pain alleviation (1-10 VAS): pain-free or improved: T) 10 of 24 participants (41.6%), C) 10 of 26 participants (38.5%). Pain unchanged or worse: T) 14 of 24 participants (58.3%), C) 16 of 26 participants (61.5%). No significant difference between the 2 treatment groups found. 4 participants of each group deteriorated temporarily in connection with the treatment given

Ljunggren 1992 (*Continued*)

Notes		
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessors were blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	It is not clear how many participants were lost to follow-up
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	It is not clear whether an intention-to-treat analysis was used
Selective reporting (reporting bias)	High risk	Published results did not include all pre-specified outcomes: straight leg raising, mobility and ADL
Group similarity at baseline (selection bias)	Low risk	Groups were similar at baseline with respect to age, sex, habits of physical therapy and symptoms
Influence of co-interventions (performance bias)	Low risk	No co-interventions were used, except for analgesics.
Compliance with interventions (performance bias)	Low risk	All participants were hospitalized.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Mathews 1975

Methods	RCT, participants were allocated at random to either control or treatment groups (method of randomization not described)
Participants	27 participants (9 female and 18 male, aged 20-60 years). Participants had sciatica or cruralgia of at least 3 wk' duration with or without back pain. Back movement was required to be limited in at least 1 direction and either the sciatic or femoral nerve stretch test positive. All had root pain. Exclusion criteria: a recently acquired neurological deficit, psychological disturbance, were pregnant, a radiological evidence of sacro-iliitis or osteoporosis, previous traction
Interventions	T) Traction: traction on a plain couch using a force of at least 36.3 kg applied through a pelvic harness, the trunk being restrained by a thoracic harness; 30 min per day, 5 days per wk, 3 wk. C) Comparison intervention: sham traction; same routine as above except the traction did not exceed 9.1 kg
Outcomes	Mean improvement in pain (VAS): T) 28.8%, C)18.9%. Not statistically significant
Notes	Control group was low force traction. Small sample. Authors cited an improvement but it was not statistically significant. Questioned whether larger trial would have shown significance

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Participants were blinded. A sham condition was used.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	No mention of attempts to blind the outcome assessors. It is unlikely that the outcome assessors were blinded
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	It is not clear how many participants were lost to follow-up

Mathews 1975 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	It is not clear whether an intention-to-treat analysis was used
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	High risk	Groups were not similar at baseline with regards to age and heavy work
Influence of co-interventions (performance bias)	Low risk	No co-interventions were used.
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Mathews 1988

Methods	RCT, participants were allocated to treatment or control by the study methodologist, using a predetermined randomization system
Participants	143 participants (63 females, 80 males, aged 20-60 years), referred from a rheumatology clinic or general practitioner were included. Participants had low backache or pain, local tenderness, asymmetrical restriction of movement, limited straight leg raise and root pain with in the past 3 months
Interventions	T) Traction: continuous auto-traction at level required to relieve pain (usually approximately 45 kg), for 30 min, 5 days per wk, until pain was relieved, but for a maximum of 3 wk. C) Comparison intervention: 3 times per wk infrared heat treatment to the low back area at 60 cm for 15 min
Outcomes	Participant's assessment of pain (6-point scale). Number recovered (10-18 days, 1 year) : T) 40/77 (52%), 30/83 (36%); C) 27/54 (50%), 11/60 (18%). The 10-18 day and 1 year outcomes are based on different numbers of participants in each group. On 8th day, more than twice the number treated people as controls were recovered (statistically significant)
Notes	Data inconsistent between text and graph.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**

Mathews 1988 (*Continued*)

Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessors were blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	It is not clear how many participants were lost to follow-up
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	It is not clear whether an intention-to-treat analysis was used or not
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Unclear risk	No description of baseline characteristics given. No baseline table was added to the article
Influence of co-interventions (performance bias)	Unclear risk	It is not clear whether co-interventions were part of treatment protocol or whether co-interventions were allowed besides the treatment that was part of the protocol
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Ozturk 2006

Methods	RCT; method unknown.
Participants	46 participants (24 in the traction group, 22 in the control group) hospitalized with the diagnosis of lumbar disc herniation. Participants had LBP or sciatica, pain duration < 6 months and lumbar disc herniation verified by CT scan. People with LBP due to neoplastic, inflammatory, infectious or metabolic causes were excluded. Mean age: T) 40.2 years, C) 52.7 years
Interventions	T) Traction: physiotherapy programme, including hot pack, ultrasound and diadynamic current, and traction: continuous lumbar traction with Enraf Nonius Traction Eltrac 439. In total, 15 sessions, 5 sessions per wk, 15 min per session, traction force 255-0% of body weight C) Comparison intervention: physiotherapy programme without traction
Outcomes	Assessment before and immediately AT. VAS for pain (mean (SD)) AT: T) 2.4 (1.7), C) 3.6 (2.7)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	It is not clear how many participants were lost to follow-up
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	It is not clear whether an intention-to-treat analysis was used
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Unclear risk	No mention of group characteristics at baseline.

Ozturk 2006 (*Continued*)

Influence of co-interventions (performance bias)	Low risk	No co-interventions were used during the treatment period.
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Unclear risk	Unclear at what time outcome assessments (for all intervention groups) were measured

Pal 1986

Methods	RCT, participants were randomly allocated to groups A and B (method of randomization not described)
Participants	39 participants (23 male (mean age 38 years) and 16 female (mean age 39 years) were admitted to hospital for back pain and sciatica. Mean duration of pain: T) 42 days, C) 56 days. Neurological deficits at baseline: T) 50% of participants, 73% of participants
Interventions	T) Traction: continuous mechanical traction of 5.5-8.2 kg according to body weight, 2-6 wk (n = 25). C) Comparison intervention: sham traction (continuous mechanical) of 1.4-1.8 kg, 2-6 wk (n = 14). Both methods were applied with the participant supine on a tilted bed by means of a pelvic harness pulled by metal weights over a pulley
Outcomes	Pain score (0-100 VAS) baseline, 1 wk, 2 wk, 3 wk: T) 50, 25, 6, 5; C) 50, 15, 9, 3. No significant differences between groups. Number of participants returned to work, < 3 months, 3-6 months, > 6 months: T) 7, 6, 5; C) 3, 4, 2
Notes	Used median scores. Timing or RTW measures not clear. Conclusion is that all recovered, may be due to enforces immobilization. Suggest that "minimal wt traction at home as compliment to complete bed rest may have important place". Data inconsistent between text and graph.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.

Pal 1986 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - participants	Low risk	The participants were not aware of the amount of traction and, therefore, were blinded
Blinding (performance bias and detection bias) All outcomes - providers	Low risk	The ward sister was responsible for allocation. All other care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	The outcome assessors were not aware of the amount of traction and, therefore, were blinded
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	2 participants (4.9%) did not complete the trial: 1 participant in each group withdrew after a few days because of home circumstances
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	High risk	Groups were not similar at baseline. 24 participants were allocated to T and 15 participants were allocated to C
Influence of co-interventions (performance bias)	Low risk	No co-interventions were used.
Compliance with interventions (performance bias)	Low risk	Treatment was well tolerated by both groups. Participants were hospitalized
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Reust 1988

Methods	RCT, participants were randomized to 1 of 3 groups by a table of randomization
Participants	60 participants (35 male, 25 female, mean age 50.8 years) hospitalized for back pain, with or without neurological deficits, were included. Exclusion criteria: previous traction, fast progressing neurological deficit, behavioural problems, or bone aliments that may have caused the back pain. Duration of back pain unknown

Reust 1988 (*Continued*)

Interventions	Traction: T1) Continuous mechanical traction on an Eltrac 439. 5-kg force on day 1, 10 kg on day 2, 15 kg on day 3, increasing 5 kg each day up to a maximum of 50 kg. 10 min per day, 12 sessions, 12 days. Participants also received medication, 20 min lumbar 'parafango' per day, 20 min massage per day and strict bed rest. T2) Same as above, except traction force of up to maximum of 15 kg. C) Comparison intervention: same as above, except traction force to maximum of 5 kg
Outcomes	Pain (100-mm VAS): T1) 33.61 (29.55), T2) 30.68 (26.83), C) 30.25 (26.23). No significant difference between groups.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Transformed from old format to new format.
Allocation concealment (selection bias)	Unclear risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Transformed from old format to new format.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	Transformed from old format to new format.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Transformed from old format to new format.
Selective reporting (reporting bias)	Unclear risk	Transformed from old format to new format.
Group similarity at baseline (selection bias)	High risk	Transformed from old format to new format.

Reust 1988 (*Continued*)

Influence of co-interventions (performance bias)	Low risk	Transformed from old format to new format.
Compliance with interventions (performance bias)	High risk	Transformed from old format to new format.
Timing of outcome assessments (detection bias)	Unclear risk	Transformed from old format to new format.

Schimmel 2009

Methods	RCT; computer-generated random block lists and adequate allocation procedure
Participants	60 participants randomly allocated to 2 treatment groups (31 to the traction group, 29 to the sham group). All participants had LBP for > 3 months. Exclusion criteria were previous surgical treatment and radicular leg pain. Mean age: T) 42 years, C) 46 years
Interventions	T) Traction: intervertebral differential dynamics therapy: 20 sessions during 6 wk, 25-30 min per session, traction force 50% of body weight. After 2 wk a standard graded activity programme was added to the traction sessions, which consisted of 1-hour training for 2 days per wk during a total of 12 wk C) Comparison intervention: same as traction group, except for traction force of < 10% of body weight
Outcomes	Assessment at 2, 6 and 14 wk. VAS LBP (mean change (SD)) at 14 wk: 32 (26.8) in the intervertebral differential dynamics group vs. 36 (27.1) in the sham group. Significant improvement during the treatment period in both intervertebral differential dynamics and sham group for the ODI, SF-36 and VAS leg pain
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomization through computer-generated random block lists.
Allocation concealment (selection bias)	Low risk	Numbered, sealed envelopes.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	The participant was not informed about the intervention received until after the 14 wk' follow-up
Blinding (performance bias and detection bias) All outcomes - providers	High risk	The care provider was not blinded for the assigned treatment

Schimmel 2009 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Follow-up evaluation was carried out by an independent assessor, who was blinded to the treatment
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	4 participants were lost to follow-up (7%) : 1 from the T group, 3 from the C group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	High risk	Published results did not include all pre-specified outcomes: outcome assessments at 2 and 6 wk were not included or could not be extracted from the graphs
Group similarity at baseline (selection bias)	Low risk	No significant between-group differences at baseline.
Influence of co-interventions (performance bias)	Unclear risk	It is not clear whether co-interventions were allowed during the treatment period or whether co-interventions were part of treatment protocol
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Sherry 2001

Methods	RCT; participants randomized in sequential order and treatments determined by predefined central randomization list
Participants	44 participants recruited through advertisements in local newspapers. Inclusion criteria: pain of > 3 months' duration, associated leg pain and confirmed disc protrusion or herniation on CT scan or MRI. (T) 11 male, 11 female; C) 12 male, 10 female; age (mean/range) T) 41/27-57, C) 43/27-55; chronicity (mean/range years) T) 8.4/0.25-30, C) 6.2/0.5-28
Interventions	T) Traction: VAX-D: participant grasps handgrips with arms extended above head; pelvic harness connected to tensionometer, which provides feedback to programmed logic control and operating system; tension applied from baseline tension to therapeutic range of 50-95 lbs, with sessions 30 min long, comprising 15 cycles of decompression and relaxation. 5 sessions/wk over 4 wk, then once/week for 4 wk. C) Comparison intervention: transcutaneous electrical nerve stimulation treatment 30 min per day for 20 days, then once per wk for 4 wk

Sherry 2001 (*Continued*)

Outcomes	Post-treatment (8 wk): pain (10-cm VAS: pre/post): T) 5.99/1.85, C) 5.44/5.97. Disability (4-point self rating scale where 1 = cannot to, 4 = can do without limitation) (pre/post): T) 2.2/2.9, C) 2.2/2.2
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded to treatment allocation.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	2 participants (4.5%) did not complete the study: 1 participant from each group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	Groups were similar at baseline.
Influence of co-interventions (performance bias)	Low risk	Neither group received any physiotherapy modalities, epidural steroid injections or other treatments during the trial. Both groups were allowed to take non-narcotic analgesics and anti-inflammatory medication if necessary
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.

Sherry 2001 (*Continued*)

Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for both groups were measured at the same time

Simmerman 2011

Methods	RCT; flip of a coin followed by an inadequate allocation procedure
Participants	61 participants randomly allocated to 2 treatment groups (31 in the land-based supine flexion first group, 30 to the aquatic vertical traction first group). All participants had LBP and sciatica. Participants with neurological disorders or vertebral fractures were excluded. Mean age: T) 59.9 years, C) 59.3 years. Mean duration of pain complaints: T) 1.7 years, C) 8.9 years
Interventions	T) Traction: 1 session of aquatic vertical traction for 15 min with the use of 2 x 2-3 kg ankle weights, followed by 1 session of land-based supine flexion C) Comparison intervention: flexion group; 1 session of land-based supine flexion, followed by 1 session of aquatic vertical traction
Outcomes	Assessment at 2-7 days following treatment. Decrease in pain (mean (SD)) on a numerical rating scale (0-10 cm) after the first intervention: T) 2.7 (2.1), C) 1.7 (1.7)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Flip of a coin.
Allocation concealment (selection bias)	Unclear risk	Flip of a coin for the first subject, followed by assignment of all uneven-numbered subjects to the land-based supine flexion position as their first intervention and all even-numbered subjects to the aquatic vertical traction position
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.

Simmerman 2011 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Outcome assessors were not blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	No participants were lost to follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	There were no statistical differences between groups in terms of age, sex, body mass index, clinical signs and symptoms
Influence of co-interventions (performance bias)	Unclear risk	It is not clear whether co-interventions were allowed during the treatment period or whether they were part of the treatment protocol
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Sweetman 1993

Methods	RCT, randomization was organized by placing the sequentially numbered treatment folders in a random order according to Documenta Geigy random number tables
Participants	400 participants (200 males and 200 females, aged 14-78 years) referred from general practice. Inclusion criteria: LBP of sufficient severity to warrant PT, pain for > 1 wk. Exclusion criteria: serious causes for back pain including fractures, infection and malignancy, pregnancy, inflammatory arthritis, bone diseases, where physician suspected that treatments may precipitate or exacerbate spinal cord or nerve root compromise, when other therapy was specifically indicated, recent steroid injections, intercurrent treatment other than routine oral medication
Interventions	T) Traction: continuous mechanical traction, constant pull (10 min), 1st wk 33% body weight, 2nd wk 50% body weight, 3 times per wk. Comparison interventions: C1) Shortwave diathermy: 20 min, 3 times per wk, 2 wk. C2) Sham shortwave diathermy: once participant felt heat, output was turned down to minimum, 20 min, 3 times per wk, 2 wk.

	C3) Extension exercises: hump and hollow, alternate leg raise, alternate arm raise, opposite leg and arm raise (prone kneeling). Bridging (crouch lying), alternate leg raise, clasp hands behind head and shoulder, and both leg raise, head and shoulder raise (prone lying), 3 times per wk, 2 wk
Outcomes	Participant opinion of overall effect (better) at 2 wk: T) 49, C1) 39, C2) 37, C3) 45. Not statistically significant
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number tables.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Unclear risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	51 participants (12.8%) failed to attend for follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	Groups were similar at baseline.
Influence of co-interventions (performance bias)	Unclear risk	It is not clear whether co-interventions were allowed during the treatment period or whether they were part of the treatment protocol

Sweetman 1993 (*Continued*)

Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Tesio 1993

Methods	RCT, participants allocated at random (method of randomization not described)
Participants	44 participants (25 males, 19 females, aged 23-63 years), referred from an outpatient service of a rehabilitation unit in a large teaching hospital. Inclusion criteria: LBP with or without radiation, duration > 1 month, herniation or protrusion, failure of 1 or more conservative approaches. Exclusion criteria: neoplastic, inflammatory or metabolic causes of back pain, or indication for urgent surgery
Interventions	Traction: T1) Intermittent auto-traction, participant provides traction force by pulling vigorously on the bar at the head of the table for a period of 3-6 sec, 1 min rest, 30-60 min session, every 2nd or 3rd day, total 3-10 sessions. If the participant reported benefit, the treatment was continued for 3-6 more sessions until no further improvement. T2) Passive traction. Traction force was adjusted approximately every 10 min, 35% of body weight, 45 min, daily bases for 5-10 sessions
Outcomes	Immediate outcomes (improved): T1) 17 of 22 participants, T2) 4 of 22 participants (statistically significant) Cross-over: non-responders to either treatment were crossed over to the other modality after a delay of 4-5 days
Notes	Most results given for only auto-traction responses (they openly favoured the treatment of the researchers)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	No mention of attempts to blind the participants. It is likely that the participants were blinded

Tesio 1993 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Unclear risk	No mention of attempts to blind the outcome assessors.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	It is not clear how many participants were lost to follow-up
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	No significant differences were found between groups with respect to sex, age, pain duration and score, presence of positive straight leg raise test or neural deficits, presence of more than 1 disc affected, presence of spinal stenosis, history of previous episodes and possible psychological bias
Influence of co-interventions (performance bias)	High risk	Co-interventions were allowed.
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	High risk	All important outcome assessments for all intervention groups were not measured at the same time. The auto-traction group was evaluated after 3 sessions, whereas the passive traction group was assessed after 5 treatment sessions

Unlu 2008

Methods	RCT; method of randomization unclear.
Participants	60 participants (20 in the traction group, 20 in the ultrasound group and 20 in the low power laser group) with acute LBP and leg pain that was definitely being caused by lumbar disc herniation. All participants had complaints of sciatica. Mean age: T) 42.5 years, C1) 48.2 years, C2) 42.8 years. Symptom duration: T) 47.9 days, C1) 36.8 days, C2) 49 days

Interventions	T) Traction: standard motorized traction therapy system (Tru-Trac 401) for 15 min per session, traction force 35-50% of total body weight Comparison interventions: C1) Ultrasound treatment, using 1 MHz at an intensity of 1.5 W/cm^2, at the right and left sides of the lumbar region. The ultrasound head was moved using small, continuous, circular movements for 8 min C2) Laser: a Gal-Al-As diode laser device (Endolaser 476) at power input of 50 mV and wavelength of 830 nm. Diameter of the laser beam was 1 mm. Stimulation time of 4 min at each point (both sides of the herniated disc)
Outcomes	Assessment BT, AT and at 1 and 3 months. VAS for LBP (mean (SD)): T) BT 58.2 (18.1), AT 29.5 (16.4), 1 month 25.5 (13.3) , 3 months 31.3 (16.4); C1) BT 51.7 (18.7), AT 29.7 (17.9), 1 month 27.2 (18.6), 3 months 26.9 (15.2); C2) BT 54.0 (17.0), AT 34.4 (18.9), 1 month 30.7 (19.1), 3 months 30.0 (16.9) VAS for radicular pain (mean (SD)): T) BT 59.6 (15.4), AT 27.7 (15.4), 1 month 21.8 (15.4), 3 months 29.5 (16.7); C1) BT 56.0 (15.3), AT 29.1 (14.4), 1 month 26.8 (18. 6), 3 months 25.2 (13.9); C2) BT 53.1 (25.9), AT 32.9 (23.6), 1 month 25.6 (21.1), 3 months 23.6 (17.7) RMDQ (mean (SD)): T) BT 14.2 (4.3), AT 9.8 (3.9), 1 month 8.5 (3.5), 3 month 8.9 (4.0); C1) BT 13.4 (4.5), AT 9.3 (5.7), 1 month 8.2 (6.0), 3 month 8.6 (6.0); C2) BT 12.5 (5.0), AT 9.9 (4.1), 1 month 7.3 (4.3), 3 months 6.7 (4.5) MODQ (mean (SD)): T) BT 19.3 (5.3), AT 14.6 (4.7), 1 month 13.5 (5.0), 3 months 14.9 (4.9); C1) BT 19.6 (6.4), AT 14.4 (5.0), 1 month 14.3 (5.5), 3 months 14.4 (5.9) ; C2) BT 18.4 (7.1), AT 14.7 (6.0), 1 months 13.5 (5.9), 3 months 13.6 (6.2)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	No mention of attempts to blind the participants. It is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Outcome assessor was blinded to treatment allocation during the assessments

Unlu 2008 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	No loss to follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	No statistically significant differences between groups.
Influence of co-interventions (performance bias)	Low risk	Co-interventions were not allowed during the treatment period. After the treatment period, participants were asked to restrict further treatment as much as possible
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Van der Heijden 1995

Methods	RCT (using sealed envelopes allocated from a list of random numbers)
Participants	25 participants (13 men, 12 women) recruited from hospital setting. Mean (SD) age: T) 46(8); C) 47(8). At baseline: mean duration: T) 18% < 6 months, 82% > 24 months; C) 17% < 6 months, 83% > 24 months. Severity: mean (SD) on pain VAS: T) 47 (27), C) 37 (23). Radiation: T) 73%, C) 58%
Interventions	T) Continuous traction: force slowly increased from 30% of body weight until participant indicated a distinct but tolerable pulling; maximum force 30-50% of body weight. C) Comparison intervention: force slowly increased from zero until participant indicated a little pulling. Maximal force 25% of body weight. For both groups: 10-12 sessions during 4 consecutive wk; also received leaflet about LBP and ADL
Outcomes	VAS at 5 wk (median improvement): T) 14, C) 16. Difference (95% CI): 2 (-29 to 14). VAS at 9 wk (median improvement): T) 14, C) 4. Difference (95% CI): -10 (-31 to 17) . Global improvement/recovery at 5 wk (% recovered): T) 54, C) 34. Difference (95% CI): 20% (-18% to 58%). Global improvement/recovery at 9 wk (% recovered): T) 38, C) 25. Difference (95% CI): 13% (-25% to 51%)

Notes		
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Random number list.
Allocation concealment (selection bias)	Low risk	Treatment allocation with sealed envelopes with a code for either treatment group. Envelopes were prepared by an independent person
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Participants were blinded to treatment allocation.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Outcome assessors were not blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	4 participants (16%) were lost to follow-up: 3 from the traction group and 1 from the comparison group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	Both groups were comparable with respect to age, sex and back pain history
Influence of co-interventions (performance bias)	Low risk	No co-interventions were allowed for the duration of the treatment period
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Walker 1982

Methods	RCT, methods of randomization judged adequate.
Participants	29 participants (18 male, 11 female, mean age: T) 37.8 years, C) 37.3) chosen by a specialist in neurology at the department of neurology in a hospital in Oslo, Norway. Non-specific LBP and radiating pain, of mixed duration (18 subjects with pain > 12 wk; 11 with < 12 wk)
Interventions	T) Traction: Spina-Trac according to Myrin; 20 min daily with 2 hours rest afterwards, for 4-8 days. 40-70 kiloponds force. Other: "traditional regimen for sciatica: 1 wk of strict bed-rest, back school, unspecified analgesics when needed (but never in morning BT sessions). C) Comparison intervention: sham traction. Same as (T) except that forces greater than 10 kiloponds not possible
Outcomes	Pain (number improved, unchanged or worse). T) 4, 13. C) 2, 10 (not statistically significant). Lasègue (number improved, unchanged or worse). T) 7, 10. C) 2, 10 (not statistically significant). Mobility (number improved, unchanged or worse). T) 4, 13. C) 2, 10 (not statistically significant)
Notes	Underpowered study with invalid pain outcome measure.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Transformed from old format to new format.
Allocation concealment (selection bias)	Unclear risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Transformed from old format to new format.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Transformed from old format to new format.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Unclear risk	Transformed from old format to new format.

Walker 1982 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Unclear risk	Transformed from old format to new format.
Selective reporting (reporting bias)	Unclear risk	Transformed from old format to new format.
Group similarity at baseline (selection bias)	High risk	Transformed from old format to new format.
Influence of co-interventions (performance bias)	Low risk	Transformed from old format to new format.
Compliance with interventions (performance bias)	Low risk	Transformed from old format to new format.
Timing of outcome assessments (detection bias)	Unclear risk	Transformed from old format to new format.

Weber 1973

Methods	No randomization methods mentioned.
Participants	72 participants (42 men, 32 women, 85% aged 30-60 years), admitted to neurology department. All had radiating pains and neurological signs corresponding to a lesion in the L5 or S1 root (or both), positive radiculogram. Exclusion criteria: people with bladder paresis, strong persistent pains, acutely occurring pareses or considerable constraint of the spinal column (or both). Duration unknown
Interventions	T) Traction: intermittent mechanical traction, 33% of body weight, Tru-Trac motor, 5-sec pauses, 20 min once per day for 5-7 days. C) Comparison intervention: sham traction with a force of up to 7 kg, 20 min once per day for 5-7 days
Outcomes	Back pain (improved): T) 14 of 37 participants, C) 15 of 35 participants. Leg pain (improved): T) 19 of 37, C) 16 of 35. No difference between the groups.
Notes	Did not test for statistical significance.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.

Weber 1973 (*Continued*)

Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Participants were not informed as to the amount of traction applied, therefore, they were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers. It is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	The investigator was not informed as to which participant belonged to which group, therefore, the outcome assessor was blinded
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	14 participants were lost to follow-up: 6 in the traction group and 8 in the comparison group
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Unclear risk	No description of baseline characteristics, no baseline table included
Influence of co-interventions (performance bias)	Low risk	No co-interventions were allowed/administered during the treatment period
Compliance with interventions (performance bias)	Low risk	Participants were hospitalized during the course of treatment
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Weber 1984

Methods	RCT, allocation to the treatment groups was done by randomization (method of randomization not described)
Participants	94 participants (54 males, 40 females). All had sciatica, radiating pain, neurological symptoms and signs corresponding to a lesion of the L5 or S1 root and positive radiculogram. Exclusion criteria: spondylolisthesis or previous operations of the spine, root entrapment caused mainly by hypertrophic facet joints or a narrow bony canal in the last 3 studies. Duration unknown

Weber 1984 (*Continued*)

Interventions	Traction: T1) Spina-Trac, intermittent manual traction, force 40-70 Kp for 10-12 sec followed by rest. 20 min once per day. T2) Continuous manual traction, therapist exerted traction by gently leaning back against a belt placed below the knees of participant, force < 30 Kp. Comparison intervention: C1) Simulated traction (for comparison against Spina-Trac). C2) Isometric exercises (for comparison against continuous manual traction) Duration of treatment unknown.
Outcomes	Improved (overall assessment): T1) 5 of 21 participants, T2) 10 of 24 participants, C1) 5 of 23 participants, C2) 10 of 26 participants. No significant difference between T1 and C1. No significant difference between T2 and C2. Temporary, immediate relief of pain obtained in the manual traction group, but not in the exercise group
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No mention of randomization procedure.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	Low risk	Study 1: participants were blinded for treatment allocation. Study 2: no mention of attempts to blind the participants, but it is unlikely that the participants were blinded
Blinding (performance bias and detection bias) All outcomes - providers	High risk	No mention of attempts to blind the care providers, but it is unlikely that the care providers were blinded
Blinding (performance bias and detection bias) All outcomes - outcome assessors	Low risk	Study 1: without knowledge of the method used, a neurologist recorded the results Study 2: without knowledge of the method used, a physiotherapist recorded the results
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	Study 1: 4 participants (9.1%) were lost to follow-up: 6 from the treatment group and 8 from the control group Study 2: 1 participant (2%) was lost to follow-up.

Weber 1984 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	High risk	No intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Unclear risk	No description of baseline characteristics, no baseline table provided
Influence of co-interventions (performance bias)	Low risk	Except for analgesics, no co-interventions were allowed during the treatment period
Compliance with interventions (performance bias)	Low risk	Participants were hospitalized.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

Werners 1999

Methods	RCT, randomization was done by the orthopaedic practitioner using a minimization computer program
Participants	147 participants (79 males, 68 females, mean age 38.75 years). Entry criterion was LBP severe enough to warrant seeking the help of an orthopaedic general practitioner. Participants with sciatica not excluded. No participant had objective neurology Exclusion: age < 20, > 60 years, previous surgery, significant medical condition and spinal disorder demonstrable on plain x-ray
Interventions	T) Traction: motorized, intermittent lumbar traction, with simultaneous massage applied by 2 motorized, mechanical wheels moving up and down the spine while the participant is lying on their back, 10-20 kg, 6 sessions, 2-3 wk. C) Comparison intervention: interferential therapy, standard Galva electrotherapy system, 6 sessions, 2-3 wk
Outcomes	ODI 1st, 2nd, 3rd visit: T) 29.5 (14.8), 24.5 (15.0), 21.7 (14.7); C) 29.7 (15.1), 25.4 (14.0), 21.1 (14.6). Pain (VAS 1-100) 1st, 2nd, 3rd visit: T) 50.6 (15.1), 44.3 (14.7), 39.2 (13.5); C) 49.7 (13.3), 45.5 (13.7), 42.0 (12.8). No differences between groups.
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**

Werners 1999 (*Continued*)

Random sequence generation (selection bias)	Low risk	Minimization computer program with stratification.
Allocation concealment (selection bias)	Unclear risk	No information provided on allocation concealment.
Blinding (performance bias and detection bias) All outcomes - participants	High risk	Participants were not blinded.
Blinding (performance bias and detection bias) All outcomes - providers	High risk	Care providers were not blinded.
Blinding (performance bias and detection bias) All outcomes - outcome assessors	High risk	Outcome assessors were not blinded.
Incomplete outcome data (attrition bias) All outcomes - loss to follow-up	Low risk	24 participants (16%) were lost to follow-up.
Incomplete outcome data (attrition bias) All outcomes - intention to treat analysis	Low risk	Intention-to-treat analysis was used.
Selective reporting (reporting bias)	Low risk	Published results included all prespecified outcomes.
Group similarity at baseline (selection bias)	Low risk	The demographics of the participants entering were similar for both groups with respect to age, sex, type of work, sick leave, weight, height and previous treatment for back pain
Influence of co-interventions (performance bias)	Unclear risk	It is not clear whether co-interventions were allowed during the treatment period or whether they were part of the treatment protocol
Compliance with interventions (performance bias)	Unclear risk	Not mentioned.
Timing of outcome assessments (detection bias)	Low risk	All important outcome assessments for all intervention groups were measured at the same time

ADL: activities of daily living; AT: after treatment; BT: before treatment: C: comparison; CI: confidence interval; CT: computed tomography; FABQ: Fear-Avoidance Beliefs Questionnaire; IQR: interquartile range; LBP: low-back pain; MD: mean difference; in: minute; MODQ: Modified Oswestry Disability Questionnaire; MPQ-PRI: McGill Pain Questionnaire - Pain Rating Index; MRI:

magnetic resonance imaging; ns: not significant; NSAID: non-steroidal anti-inflammatory drug; ODI: Oswestry Disability Index; PT: physiotherapy; RCT: randomized controlled trial; RMDQ: Roland Morris Disability Questionnaire; ROM: range of motion; RTW: return to work; SD: standard deviation; SE: standard error; sec: second; SF36 MCS: Short Form-36 Mental Component Summary; SF36 PCS: Short Form-36 Physical Component Summary; T: traction; VAS: visual analogue scale; wk: week.

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

Study	Reason for exclusion
Cevik 2007	Study was not randomized.
Gose 1998	Study was not an RCT.
Hansen 1993	Used low-force traction as a sham treatment and included regular traction as 1 component of a physiotherapy programme
Moret 1998	Article described a feasibility study, not a full trial.
Olah 2008	Study was not randomized.
Ramos 1994	Study is not an RCT; outcome is intradiscal pressure.
Van der Heijden 1991	Pilot study only, in preparation for Van der Heijden 1995.

RCT: randomized controlled trial.

DATA AND ANALYSES

Comparison 1. Low-back pain with/without radiation, traction versus placebo, sham or no treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.1 3-5 weeks	2	247	Mean Difference (IV, Fixed, 95% CI)	-18.49 [-24.12, -12.87]
1.2 6-12 weeks	1	150	Mean Difference (IV, Fixed, 95% CI)	0.30 [-9.91, 10.51]
1.3 6 months	1	150	Mean Difference (IV, Fixed, 95% CI)	-0.5 [-11.55, 10.55]
1.4 1 year	1	97	Mean Difference (IV, Fixed, 95% CI)	-9.10 [-19.32, 1.12]
2 Functional status	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.1 3-5 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.2 6-12 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.3 6 months	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Global improvement	2		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
3.1 3-5 weeks	2	175	Risk Difference (M-H, Fixed, 95% CI)	-0.03 [-0.17, 0.12]
3.2 6-12 weeks	2	175	Risk Difference (M-H, Fixed, 95% CI)	0.03 [-0.12, 0.18]
3.3 6 months	1	150	Risk Difference (M-H, Fixed, 95% CI)	0.02 [-0.14, 0.18]
4 Return to work (days)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
4.1 3-5 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
4.2 6-12 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
4.3 6 months	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 2. Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.1 1-2 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 12-16 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Functional status	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.1 1-2 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.2 12-16 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Global improvement	1		Risk Difference (M-H, Fixed, 95% CI)	Totals not selected
3.1 1-2 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.2 12-16 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 3. Low-back pain with/without radiation, two types of traction

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Global improvement	2		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
1.1 1-2 weeks	2	93	Risk Difference (M-H, Fixed, 95% CI)	0.35 [0.17, 0.54]

Comparison 4. Low-back pain with/without radiation, traction versus other treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	4		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.1 1-2 weeks	2		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 3-5 weeks	2		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.3 12-16 weeks	2		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.4 6 months	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.5 1 year	2		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Functional status	2		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.1 1-2 weeks	1	138	Std. Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.40, 0.27]
2.2 3-5 weeks	1	235	Std. Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.05, 0.46]
2.3 12-16 weeks	2	290	Std. Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.26, 0.21]
2.4 6 months	1	168	Std. Mean Difference (IV, Fixed, 95% CI)	0.15 [-0.16, 0.45]
2.5 1 year	1	173	Std. Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.25, 0.34]
3 Global improvement	3		Risk Difference (M-H, Fixed, 95% CI)	Totals not selected
3.1 1-2 weeks	2		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.2 3-5 weeks	2		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.3 12-16 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 5. Low-back pain with radiation, traction versus placebo, sham or no treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.1 1-2 weeks	2	79	Mean Difference (IV, Fixed, 95% CI)	2.93 [-14.73, 20.59]
1.2 3-5 weeks	1	39	Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Global improvement	5		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
2.1 1-2 weeks	4	398	Risk Difference (M-H, Fixed, 95% CI)	0.13 [0.04, 0.22]
2.2 3-5 weeks	2	123	Risk Difference (M-H, Fixed, 95% CI)	0.27 [0.12, 0.43]
2.3 12-16 weeks	1	81	Risk Difference (M-H, Fixed, 95% CI)	0.06 [-0.16, 0.28]
3 Return to work	1		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
3.1 2 years	1	39	Risk Difference (M-H, Fixed, 95% CI)	0.15 [-0.15, 0.45]

Comparison 6. Low-back with radiation, physiotherapy with traction versus physiotherapy without traction

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.1 1-2 weeks	2	110	Mean Difference (IV, Fixed, 95% CI)	-7.96 [-16.53, 0.61]
1.2 6 weeks	1	64	Mean Difference (IV, Fixed, 95% CI)	2.0 [-10.02, 14.02]
2 Functional status	2		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.1 1-2 weeks	2	94	Std. Mean Difference (IV, Fixed, 95% CI)	-0.08 [-0.49, 0.32]
2.2 6-12 weeks	1	64	Std. Mean Difference (IV, Fixed, 95% CI)	0.14 [-0.35, 0.63]
2.3 12-16 weeks	1	30	Std. Mean Difference (IV, Fixed, 95% CI)	0.43 [-0.30, 1.16]
2.4 6 months	1	30	Std. Mean Difference (IV, Fixed, 95% CI)	0.18 [-0.54, 0.90]
3 Global improvement	3		Risk Difference (M-H, Fixed, 95% CI)	Totals not selected
3.1 1-2 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.2 3-5 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.3 6 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.4 12-16 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
4 Return to work	1		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected
4.1 3-5 weeks	1		Odds Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 7. Low-back pain with radiation, traction versus other treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	2		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.1 1-2 weeks	2		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 3-5 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.3 12-16 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Functional status	2		Std. Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.1 1-2 weeks	2		Std. Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.2 3-5 weeks	1		Std. Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.3 12-16 weeks	1		Std. Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Global improvement	2		Risk Difference (M-H, Fixed, 95% CI)	Totals not selected
3.1 1-2 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
3.2 3-5 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 8. Low-back pain with radiation, two types of traction

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.1 1-2 weeks	3	149	Mean Difference (IV, Fixed, 95% CI)	6.58 [-2.77, 15.93]
2 Global improvement	1		Risk Difference (M-H, Fixed, 95% CI)	Totals not selected
2.1 1-2 weeks	1		Risk Difference (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 9. Low-back pain without radiation, traction versus sham

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain intensity	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.1 12-16 weeks	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]

Analysis 1.1. Comparison 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment

Outcome: 1 Pain intensity

Study or subgroup	Favours traction N	Mean(SD)	Placebo, sham or no tx N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
1 3-5 weeks							
Beurskens 1997 (1)	77	20.6 (27.5)	73	23.7 (27.5)		40.8 %	-3.10 [-11.90, 5.70]
Konrad 1992 (2)	44	24.6 (11.9)	53	53.7 (23.8)		59.2 %	-29.10 [-36.41, -21.79]
Subtotal (95% CI)	**121**		**126**			**100.0 %**	**-18.49 [-24.12, -12.87]**
Heterogeneity: Chi² = 19.83, df = 1 (P<0.00001); I² =95%							
Test for overall effect: Z = 6.45 (P < 0.00001)							
2 6-12 weeks							
Beurskens 1997	77	24.2 (31.9)	73	23.9 (31.9)		100.0 %	0.30 [-9.91, 10.51]
Subtotal (95% CI)	**77**		**73**			**100.0 %**	**0.30 [-9.91, 10.51]**

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Favours traction		Placebo, sham or no tx		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	IV,Fixed,95% CI		IV,Fixed,95% CI
Heterogeneity: not applicable							
Test for overall effect: Z = 0.06 (P = 0.95)							
3 6 months							
Beurskens 1997	77	25 (34.5)	73	25.5 (34.5)		100.0 %	-0.50 [-11.55, 10.55]
Subtotal (95% CI)	**77**		**73**			**100.0 %**	**-0.50 [-11.55, 10.55]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.09 (P = 0.93)							
4 1 year							
Konrad 1992	44	45.8 (26.2)	53	54.9 (24.8)		100.0 %	-9.10 [-19.32, 1.12]
Subtotal (95% CI)	**44**		**53**			**100.0 %**	**-9.10 [-19.32, 1.12]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.74 (P = 0.081)							

(1) Traction versus sham

(2) Underwater traction versus no treatment

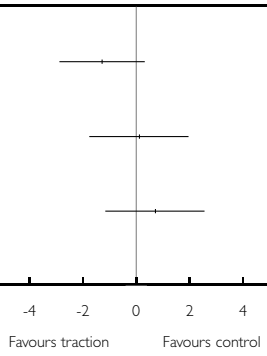
Analysis 1.2. Comparison 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment, Outcome 2 Functional status.

Review: Traction for low-back pain with or without sciatica

Comparison: 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment

Outcome: 2 Functional status

Study or subgroup	Traction N	Mean(SD)	Placebo, sham or no tx N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Mean Difference IV,Fixed,95% CI
1 3-5 weeks						
Beurskens 1997 (1)	77	3.5 (5)	73	4.8 (5)		-1.30 [-2.90, 0.30]
2 6-12 weeks						
Beurskens 1997	77	4.4 (5.8)	73	4.3 (5.8)		0.10 [-1.76, 1.96]
3 6 months						
Beurskens 1997	77	4.7 (5.8)	73	4 (5.8)		0.70 [-1.16, 2.56]



(1) Traction versus sham

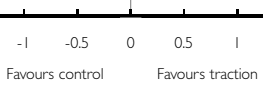
Analysis 1.3. Comparison 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment, Outcome 3 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment

Outcome: 3 Global improvement

Study or subgroup	Traction n/N	Placebo, sham or no tx n/N	Risk Difference M-H,Fixed,95% CI	Weight	Risk Difference M-H,Fixed,95% CI
1 3-5 weeks					
Beurskens 1997 (1)	34/77	37/73		85.7 %	-0.07 [-0.22, 0.09]
Van der Heijden 1995 (2)	7/13	4/12		14.3 %	0.21 [-0.18, 0.59]
Subtotal (95% CI)	**90**	**85**		**100.0 %**	**-0.03 [-0.17, 0.12]**
Total events: 41 (Traction), 41 (Placebo, sham or no tx)					
Heterogeneity: Chi^2 = 1.65, df = 1 (P = 0.20); I^2 =39%					
Test for overall effect: Z = 0.36 (P = 0.72)					
2 6-12 weeks					
Beurskens 1997	38/77	35/73		85.7 %	0.01 [-0.15, 0.17]
Van der Heijden 1995	5/13	3/12		14.3 %	0.13 [-0.23, 0.50]
Subtotal (95% CI)	**90**	**85**		**100.0 %**	**0.03 [-0.12, 0.18]**
Total events: 43 (Traction), 38 (Placebo, sham or no tx)					
Heterogeneity: Chi^2 = 0.36, df = 1 (P = 0.55); I^2 =0.0%					
Test for overall effect: Z = 0.42 (P = 0.68)					
3 6 months					
Beurskens 1997	35/77	32/73		100.0 %	0.02 [-0.14, 0.18]
Subtotal (95% CI)	**77**	**73**		**100.0 %**	**0.02 [-0.14, 0.18]**
Total events: 35 (Traction), 32 (Placebo, sham or no tx)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.20 (P = 0.84)					



(1) Traction versus sham

(2) Traction versus sham

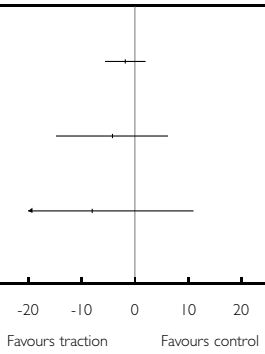
Analysis 1.4. Comparison 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment, Outcome 4 Return to work (days).

Review: Traction for low-back pain with or without sciatica

Comparison: 1 Low-back pain with/without radiation, traction versus placebo, sham or no treatment

Outcome: 4 Return to work (days)

Study or subgroup	Traction N	Mean(SD)	Placebo, sham or no tx N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Mean Difference IV,Fixed,95% CI
1 3-5 weeks						
Beurskens 1997 (1)	77	21 (11.6)	73	22.8 (11.6)		-1.80 [-5.51, 1.91]
2 6-12 weeks						
Beurskens 1997	77	23.5 (32.5)	73	27.8 (32.5)		-4.30 [-14.71, 6.11]
3 6 months						
Beurskens 1997	77	35.7 (59.3)	73	43.7 (59.3)		-8.00 [-26.99, 10.99]



(1) Traction versus sham

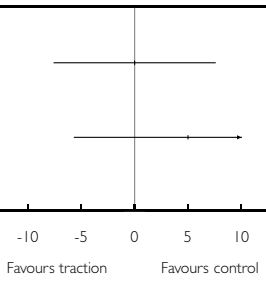
Analysis 2.1. Comparison 2 Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 2 Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 1 Pain intensity

Study or subgroup	Physio with traction		Physio without traction		Mean Difference	Mean Difference
	N	Mean(SD)	N	Mean(SD)	IV,Fixed,95% CI	IV,Fixed,95% CI
1 1-2 weeks						
Borman 2003	21	38 (11)	21	38 (14)		0.0 [-7.61, 7.61]
2 12-16 weeks						
Borman 2003	20	41 (17)	19	36 (17)		5.00 [-5.67, 15.67]



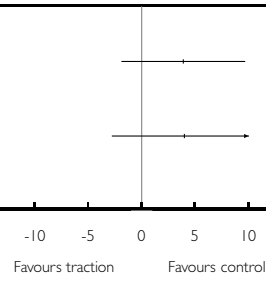
Analysis 2.2. Comparison 2 Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 2 Functional status.

Review: Traction for low-back pain with or without sciatica

Comparison: 2 Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 2 Functional status

Study or subgroup	Physio with tract		Physio without tract		Mean Difference	Mean Difference
	N	Mean(SD)	N	Mean(SD)	IV,Fixed,95% CI	IV,Fixed,95% CI
1 1-2 weeks						
Borman 2003	21	26.8 (9.1)	21	22.9 (10.1)		3.90 [-1.91, 9.71]
2 12-16 weeks						
Borman 2003	20	23.7 (10.8)	19	19.7 (10.8)		4.00 [-2.78, 10.78]

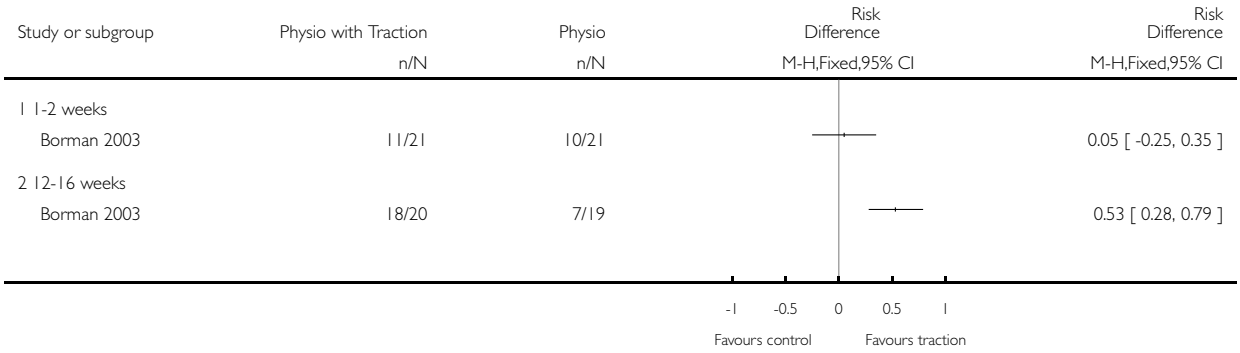


Analysis 2.3. Comparison 2 Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 3 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 2 Low-back pain with/without radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 3 Global improvement

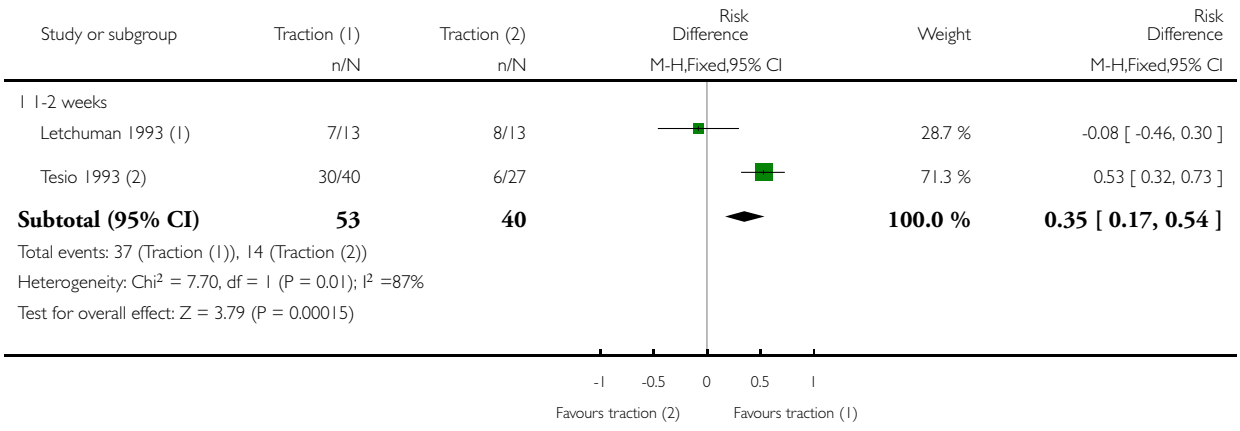


Analysis 3.1. Comparison 3 Low-back pain with/without radiation, two types of traction, Outcome 1 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 3 Low-back pain with/without radiation, two types of traction

Outcome: 1 Global improvement



(1) Static traction (1) versus intermittent traction (2)

(2) Auto-traction (1) versus mechanical traction (2)

Analysis 4.1. Comparison 4 Low-back pain with/without radiation, traction versus other treatment, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 4 Low-back pain with/without radiation, traction versus other treatment

Outcome: 1 Pain intensity

Study or subgroup	Traction N	Mean(SD)	Other treatment N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Mean Difference IV,Fixed,95% CI
1 1-2 weeks						
Sweetman 1993 (1)	100	0 (0)	100	0 (0)		Not estimable
Werners 1999 (2)	51	44.3 (14.7)	50	45.5 (13.7)		-1.20 [-6.74, 4.34]
2 3-5 weeks						
Gudavalli 2006 (3)	123	20.6 (13.8)	112	12.3 (13.8)		8.30 [4.77, 11.83]
Konrad 1992 (4)	44	24.6 (11.9)	26	33.5 (19.1)		-8.90 [-17.04, -0.76]
Konrad 1992 (5)	44	24.6 (11.9)	35	31.7 (16.2)		-7.10 [-13.52, -0.68]
3 12-16 weeks						
Gudavalli 2006 (6)	87	16.5 (16.1)	76	12 (16.1)		4.50 [-0.45, 9.45]
Werners 1999	43	39.2 (13.5)	38	42 (12.8)		-2.80 [-8.53, 2.93]
4 6 months						
Gudavalli 2006	90	18.2 (18.4)	74	8.9 (18.4)		9.30 [3.64, 14.96]
5 1 year						
Gudavalli 2006 (7)	96	17.1 (15.9)	78	12.4 (15.9)		4.70 [-0.05, 9.45]
Konrad 1992 (8)	44	45.8 (26.2)	35	49.5 (25.7)		-3.70 [-15.21, 7.81]
Konrad 1992 (9)	44	45.8 (26.2)	26	54.7 (33.7)		-8.90 [-23.99, 6.19]

(1) Traction versus exercise (no numbers given)

(2) Traction versus interferential therapy

(3) Traction versus exercise

(4) Underwater traction versus underwater massage

(5) Underwater traction versus balneotherapy

(6) Traction versus exercise

(7) Traction versus exercise

(8) Underwater traction versus balneotherapy

(9) Underwater traction versus underwater massage

Analysis 4.2. Comparison 4 Low-back pain with/without radiation, traction versus other treatment, Outcome 2 Functional status.

Review: Traction for low-back pain with or without sciatica

Comparison: 4 Low-back pain with/without radiation, traction versus other treatment

Outcome: 2 Functional status

Study or subgroup	Traction N	Mean(SD)	Other treatment N	Mean(SD)	Std. Mean Difference IV,Fixed,95% CI	Weight	Std. Mean Difference IV,Fixed,95% CI
1 1-2 weeks							
Werners 1999 (1)	70	24.5 (15)	68	25.4 (14)		100.0 %	-0.06 [-0.40, 0.27]
Subtotal (95% CI)	**70**		**68**			**100.0 %**	**-0.06 [-0.40, 0.27]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.36 (P = 0.72)							
2 3-5 weeks							
Gudavalli 2006 (2)	123	2.81 (2.5)	112	2.3 (2.5)		100.0 %	0.20 [-0.05, 0.46]
Subtotal (95% CI)	**123**		**112**			**100.0 %**	**0.20 [-0.05, 0.46]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.55 (P = 0.12)							
3 12-16 weeks							
Gudavalli 2006	86	3.5 (3.2)	76	3.75 (3.2)		55.8 %	-0.08 [-0.39, 0.23]
Werners 1999	67	21.7 (14.7)	61	21.1 (14.6)		44.2 %	0.04 [-0.31, 0.39]

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Traction N	Mean(SD)	Other treatment N	Mean(SD)	Std. Mean Difference IV,Fixed,95% CI	Weight	Std. Mean Difference IV,Fixed,95% CI
Subtotal (95% CI)	**153**		**137**			**100.0 %**	**-0.03 [-0.26, 0.21]**
Heterogeneity: Chi^2 = 0.25, df = 1 (P = 0.62); I^2 =0.0%							
Test for overall effect: Z = 0.22 (P = 0.83)							
4 6 months							
Gudavalli 2006	90	3.89 (3.2)	78	3.42 (3.2)		100.0 %	0.15 [-0.16, 0.45]
Subtotal (95% CI)	**90**		**78**			**100.0 %**	**0.15 [-0.16, 0.45]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.94 (P = 0.35)							
5 1 year							
Gudavalli 2006	95	3.9 (2.9)	78	3.77 (2.9)		100.0 %	0.04 [-0.25, 0.34]
Subtotal (95% CI)	**95**		**78**			**100.0 %**	**0.04 [-0.25, 0.34]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.29 (P = 0.77)							

(1) Traction versus interferential therapy (ODQ)

(2) Traction versus exercise (RMDQ)

Analysis 4.3. Comparison 4 Low-back pain with/without radiation, traction versus other treatment, Outcome 3 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 4 Low-back pain with/without radiation, traction versus other treatment

Outcome: 3 Global improvement

Study or subgroup	Favours control n/N	Other treatment n/N	Risk Difference M-H,Fixed,95% CI	Risk Difference M-H,Fixed,95% CI
1 1-2 weeks				
Bihaug 1978 (1)	17/21	11/21		0.29 [0.01, 0.56]
Sweetman 1993 (2)	49/100	45/100		0.04 [-0.10, 0.18]
Sweetman 1993 (3)	49/100	39/100		0.10 [-0.04, 0.24]
2 3-5 weeks				
Bihaug 1978 (4)	19/21	16/21		0.14 [-0.08, 0.36]
Lind 1974 (5)	13/15	0/15		0.87 [0.67, 1.06]
Lind 1974 (6)	13/15	0/14		0.87 [0.67, 1.07]
3 12-16 weeks				
Bihaug 1978 (7)	20/21	19/21		0.05 [-0.11, 0.20]

(1) Traction versus exercise

(2) Traction versus exercise

(3) Traction versus short wave diathermy

(4) Traction versus exercise

(5) Traction versus physiotherapy

(6) Traction versus bedrest and analgesics

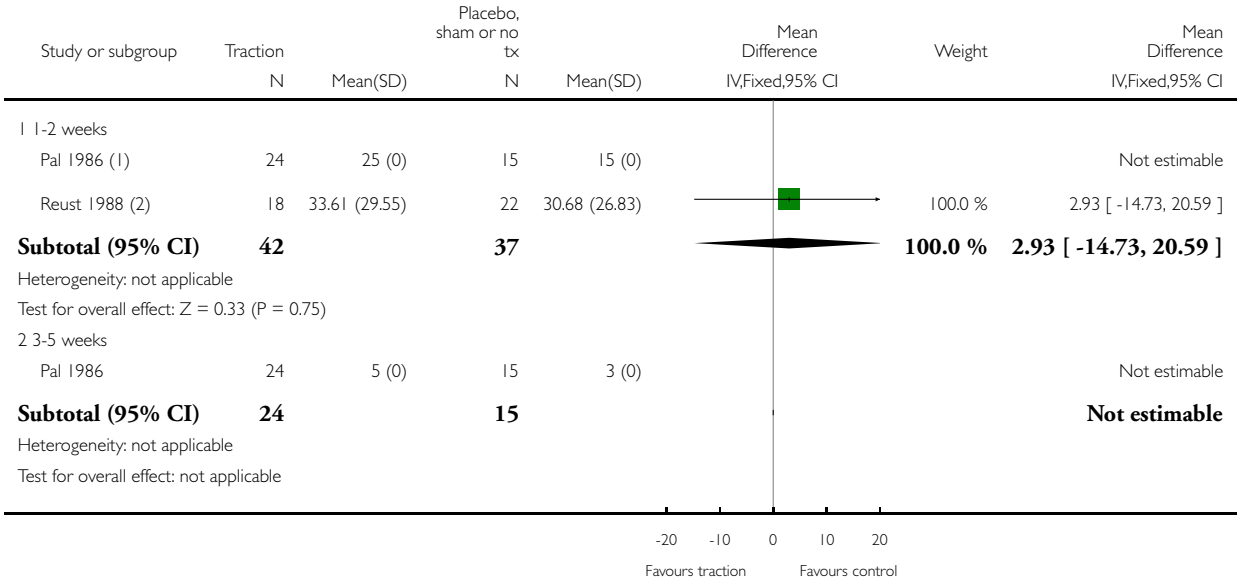
(7) Traction versus exercise

Analysis 5.1. Comparison 5 Low-back pain with radiation, traction versus placebo, sham or no treatment, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 5 Low-back pain with radiation, traction versus placebo, sham or no treatment

Outcome: 1 Pain intensity



(1) Traction versus sham

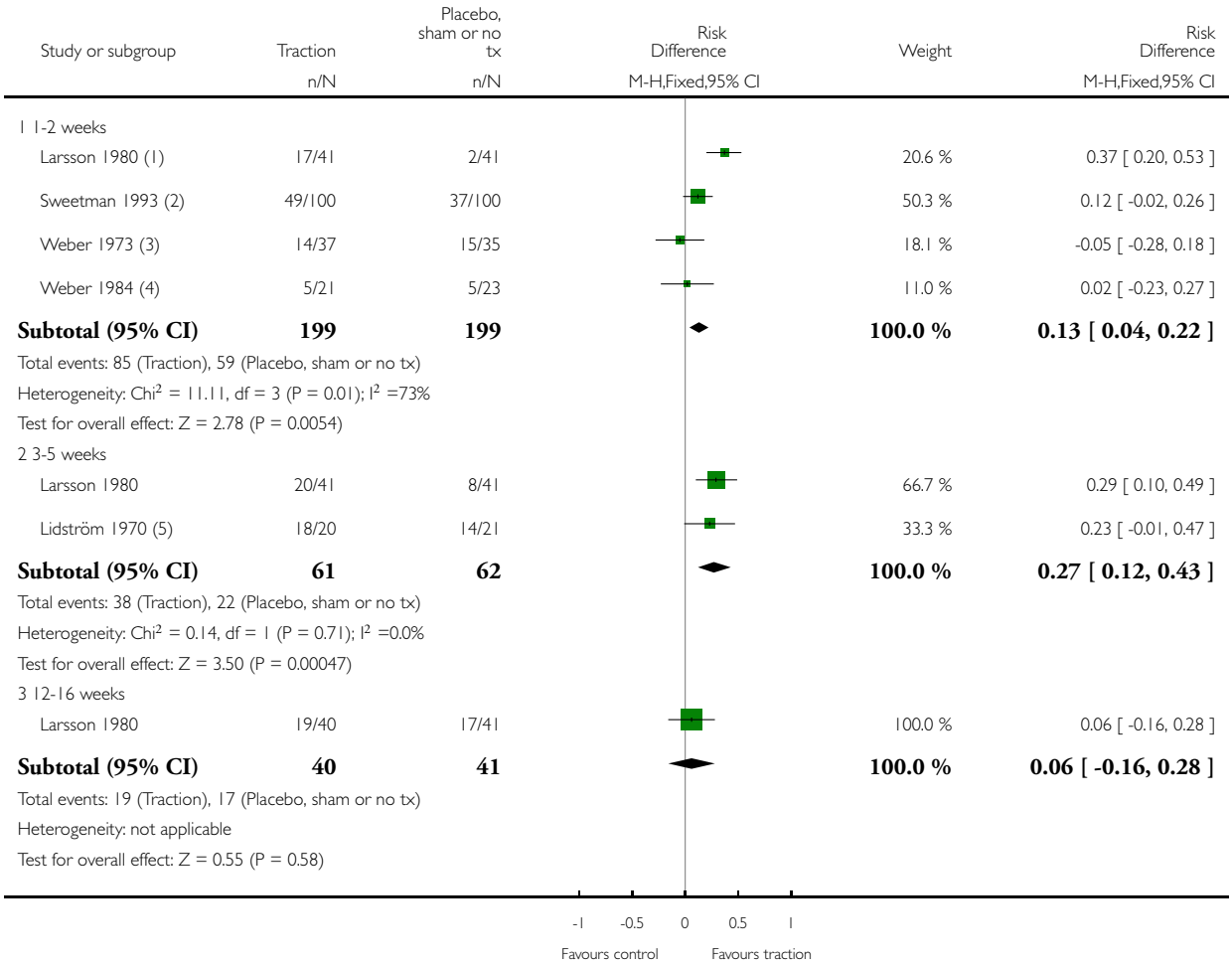
(2) Traction versus sham

Analysis 5.2. Comparison 5 Low-back pain with radiation, traction versus placebo, sham or no treatment, Outcome 2 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 5 Low-back pain with radiation, traction versus placebo, sham or no treatment

Outcome: 2 Global improvement



(1) Traction versus no treatment

(2) Traction versus sham

(3) Traction versus sham

(4) Traction versus sham

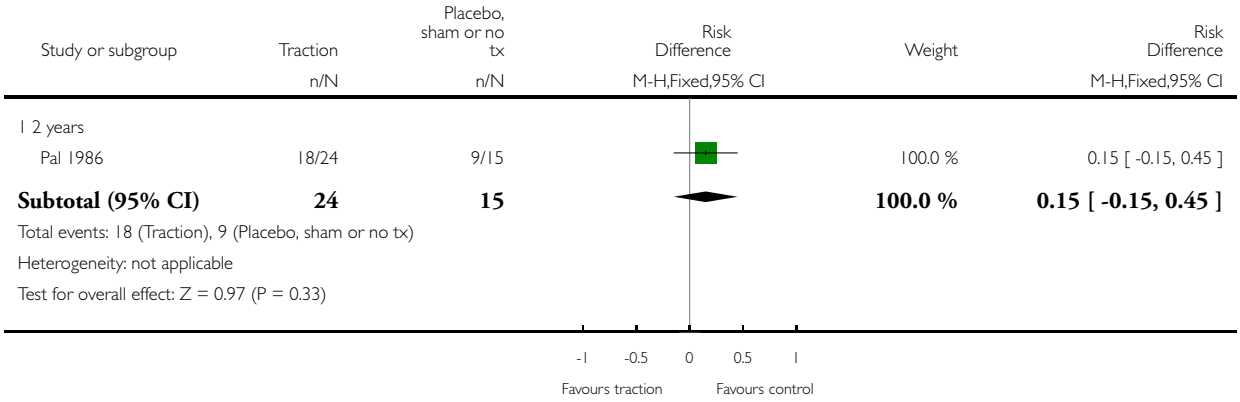
(5) Traction versus no treatment

Analysis 5.3. Comparison 5 Low-back pain with radiation, traction versus placebo, sham or no treatment, Outcome 3 Return to work.

Review: Traction for low-back pain with or without sciatica

Comparison: 5 Low-back pain with radiation, traction versus placebo, sham or no treatment

Outcome: 3 Return to work

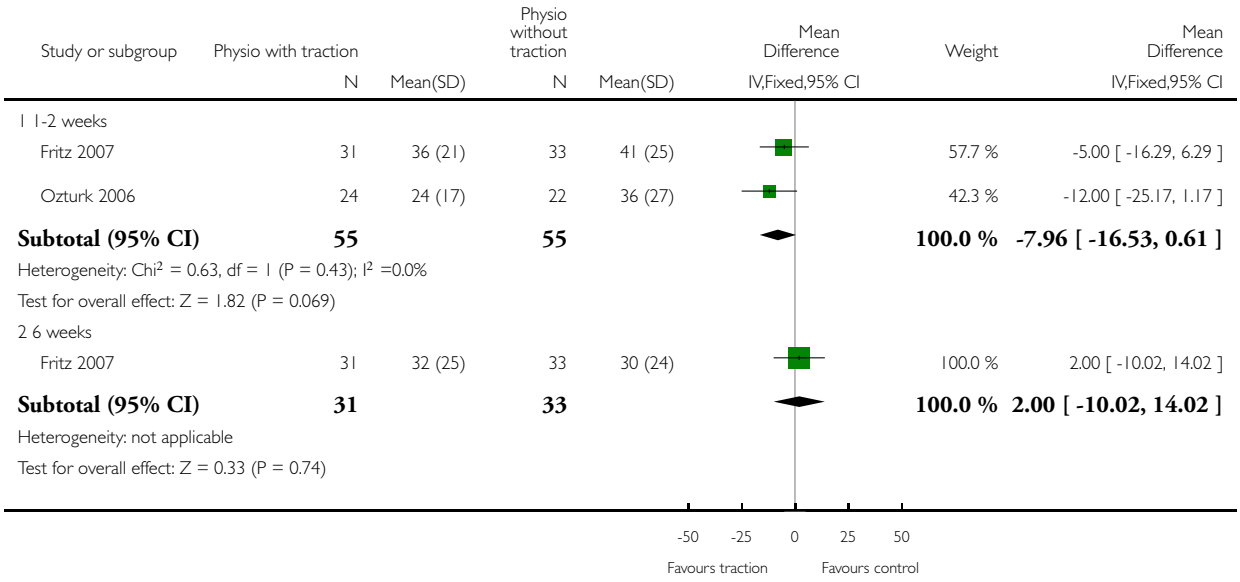


Analysis 6.1. Comparison 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 1 Pain intensity

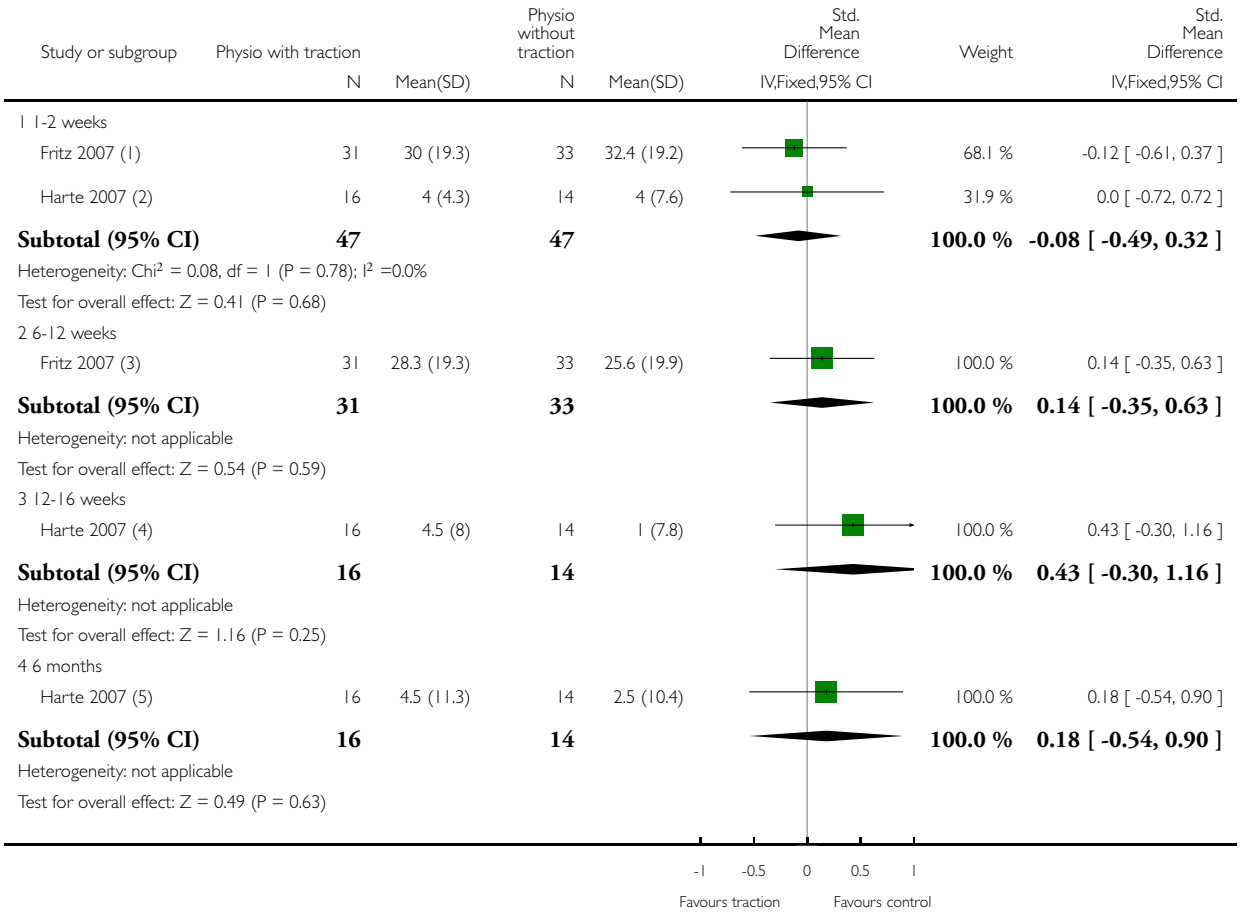


Analysis 6.2. Comparison 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 2 Functional status.

Review: Traction for low-back pain with or without sciatica

Comparison: 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 2 Functional status



(1) ODQ

(2) RMDQ

(3) ODQ

(4) RMDQ

(5) RMDQ

Analysis 6.3. Comparison 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 3 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 3 Global improvement

Study or subgroup	Physio with traction n/N	Physio without traction n/N	Risk Difference M-H,Fixed,95% CI	Risk Difference M-H,Fixed,95% CI
1 1-2 weeks				
Ozturk 2006	12/19	12/20		0.03 [-0.27, 0.34]
2 3-5 weeks				
Coxhead 1981	117/143	110/149		0.08 [-0.01, 0.17]
3 6 weeks				
Fritz 2007	21/31	21/33		0.04 [-0.19, 0.27]
4 12-16 weeks				
Coxhead 1981	89/123	92/127		0.00 [-0.11, 0.11]

Analysis 6.4. Comparison 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction, Outcome 4 Return to work.

Review: Traction for low-back pain with or without sciatica

Comparison: 6 Low-back with radiation, physiotherapy with traction versus physiotherapy without traction

Outcome: 4 Return to work

Study or subgroup	Physio with traction n/N	Physio without traction n/N	Odds Ratio M-H,Fixed,95% CI	Odds Ratio M-H,Fixed,95% CI
1 3-5 weeks				
Coxhead 1981	20/56	13/46		1.41 [0.61, 3.28]

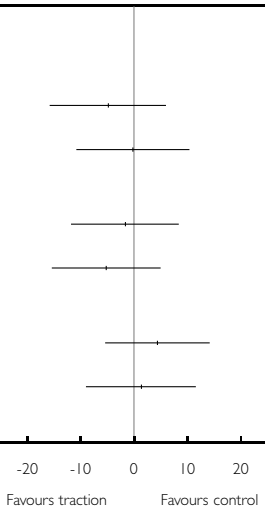
Analysis 7.1. Comparison 7 Low-back pain with radiation, traction versus other treatment, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 7 Low-back pain with radiation, traction versus other treatment

Outcome: 1 Pain intensity

Study or subgroup	Traction N	Mean(SD)	Other treatment N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Mean Difference IV,Fixed,95% CI
1 1-2 weeks						
Ljunggren 1992 (1)	24	0 (0)	26	0 (0)		Not estimable
Unlu 2008 (2)	20	29.5 (16.4)	20	34.4 (18.9)		-4.90 [-15.87, 6.07]
Unlu 2008 (3)	20	29.5 (16.4)	20	29.7 (17.9)		-0.20 [-10.84, 10.44]
2 3-5 weeks						
Unlu 2008 (4)	20	25.5 (13.3)	20	27.2 (18.6)		-1.70 [-11.72, 8.32]
Unlu 2008 (5)	20	25.5 (13.3)	20	30.7 (19.1)		-5.20 [-15.40, 5.00]
3 12-16 weeks						
Unlu 2008 (6)	20	31.3 (16.4)	20	26.9 (15.2)		4.40 [-5.40, 14.20]
Unlu 2008 (7)	20	31.3 (16.4)	20	30 (16.8)		1.30 [-8.99, 11.59]



(1) Traction versus exercise (VAS-scores only presented in graph)

(2) Traction versus laser

(3) Traction versus ultrasound

(4) Traction versus ultrasound

(5) Traction versus laser

(6) Traction versus ultrasound

(7) Traction versus laser

Analysis 7.2. Comparison 7 Low-back pain with radiation, traction versus other treatment, Outcome 2 Functional status.

Review: Traction for low-back pain with or without sciatica

Comparison: 7 Low-back pain with radiation, traction versus other treatment

Outcome: 2 Functional status

Study or subgroup	Traction N	Mean(SD)	Other treatment N	Mean(SD)	Std. Mean Difference IV,Fixed,95% CI	Std. Mean Difference IV,Fixed,95% CI
1 1-2 weeks						
Ljunggren 1992 (1)	24	0 (0)	26	0 (0)		Not estimable
Unlu 2008 (2)	20	14.6 (4.7)	20	14.4 (5)		0.04 [-0.58, 0.66]
Unlu 2008 (3)	20	14.6 (4.7)	20	14.7 (6)		-0.02 [-0.64, 0.60]
Unlu 2008 (4)	20	9.8 (3.9)	20	9.3 (5.7)		0.10 [-0.52, 0.72]
Unlu 2008 (5)	20	9.8 (3.9)	20	9.9 (4.1)		-0.02 [-0.64, 0.60]
2 3-5 weeks						
Unlu 2008 (6)	20	13.5 (5)	20	14.3 (5.5)		-0.15 [-0.77, 0.47]
Unlu 2008 (7)	20	8.5 (3.5)	20	7.3 (4.3)		0.30 [-0.32, 0.92]
Unlu 2008 (8)	20	8.5 (3.5)	20	8.2 (6)		0.06 [-0.56, 0.68]
Unlu 2008 (9)	20	13.5 (5)	20	13.5 (5.9)		0.0 [-0.62, 0.62]
3 12-16 weeks						
Unlu 2008 (10)	20	8.9 (4)	20	6.7 (4.5)		0.51 [-0.12, 1.14]
Unlu 2008 (11)	20	14.9 (4.9)	20	14.4 (5.9)		0.09 [-0.53, 0.71]
Unlu 2008 (12)	20	8.9 (4)	20	8.6 (6)		0.06 [-0.56, 0.68]
Unlu 2008 (13)	20	14.9 (4.9)	20	13.6 (6.2)		0.23 [-0.39, 0.85]

(1) Traction versus exercise (functional status was evaluated, but no results were reported)

(2) Traction versus ultrasound (ODQ)

(3) Traction versus laser (ODQ)

(4) Traction versus ultrasound (RMDQ)

(5) Traction versus laser (RMDQ)

(6) Traction versus laser (ODQ)

(7) Traction versus ultrasound (RMDQ)

(8) Traction versus laser (RMDQ)

(9) Traction versus ultrasound (ODQ)

(10) Traction versus laser (RMDQ)

(11) Traction versus laser (ODQ)

(12) Traction versus ultrasound (RMDQ)

(13) Traction versus ultrasound (ODQ)

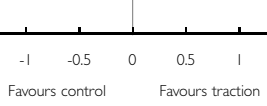
Analysis 7.3. Comparison 7 Low-back pain with radiation, traction versus other treatment, Outcome 3 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 7 Low-back pain with radiation, traction versus other treatment

Outcome: 3 Global improvement

Study or subgroup	Traction n/N	Other treatment n/N	Risk Difference M-H,Fixed,95% CI	Risk Difference M-H,Fixed,95% CI
1 1-2 weeks				
Ljunggren 1992 (1)	10/24	10/26		0.03 [-0.24, 0.30]
2 3-5 weeks				
Lidström 1970 (2)	18/20	10/21		0.42 [0.17, 0.67]



(1) Traction versus exercise

(2) Traction versus physiotherapy

Analysis 8.1. Comparison 8 Low-back pain with radiation, two types of traction, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

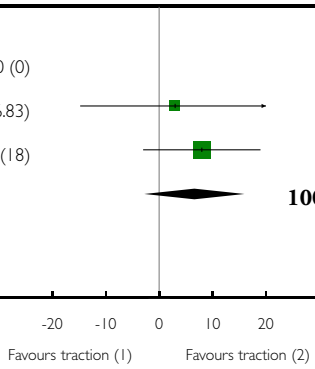
Comparison: 8 Low-back pain with radiation, two types of traction

Outcome: 1 Pain intensity

Study or subgroup	Traction (1) N	Mean(SD)	Traction (2) N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
1 1-2 weeks							
Ljunggren 1984 (1)	26	0 (0)	23	0 (0)			Not estimable
Reust 1988 (2)	18	33.61 (29.55)	22	30.68 (26.83)		28.0 %	2.93 [-14.73, 20.59]
Simmerman 2011 (3)	30	47 (25)	30	39 (18)		72.0 %	8.00 [-3.02, 19.02]
Subtotal (95% CI)	**74**		**75**			**100.0 %**	**6.58 [-2.77, 15.93]**

Heterogeneity: Chi² = 0.23, df = 1 (P = 0.63); I² =0.0%

Test for overall effect: Z = 1.38 (P = 0.17)



(1) Auto-traction (1) versus manual traction (2) (was scores only presented in a graph)

(2) Auto-traction (1) versus mechanical traction (2)

(3) Aquatic traction (1) versus land based supine position (2)

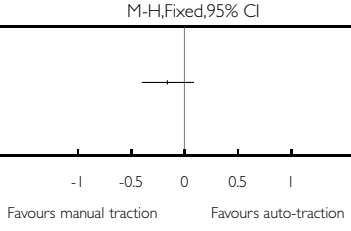
Analysis 8.2. Comparison 8 Low-back pain with radiation, two types of traction, Outcome 2 Global improvement.

Review: Traction for low-back pain with or without sciatica

Comparison: 8 Low-back pain with radiation, two types of traction

Outcome: 2 Global improvement

Study or subgroup	Traction (1) n/N	Traction (2) n/N	Risk Difference M-H,Fixed,95% CI	Risk Difference M-H,Fixed,95% CI
1 1-2 weeks				
Ljunggren 1984 (1)	5/26	8/23		-0.16 [-0.40, 0.09]



(1) Auto-traction (1) versus manual traction (2)

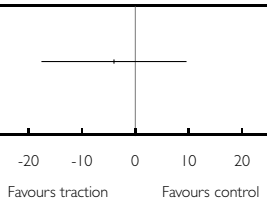
Analysis 9.1. Comparison 9 Low-back pain without radiation, traction versus sham, Outcome 1 Pain intensity.

Review: Traction for low-back pain with or without sciatica

Comparison: 9 Low-back pain without radiation, traction versus sham

Outcome: 1 Pain intensity

Study or subgroup	IDD therapy N	Mean(SD)	Sham N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Mean Difference IV,Fixed,95% CI
1 12-16 weeks						
Schimmel 2009 (1)	31	32 (26.8)	29	36 (27.1)		-4.00 [-17.65, 9.65]



(1) IDD therapy versus sham

ADDITIONAL TABLES

Table 1. Clinical relevance

Author	Participants	Interventions	Outcomes	Effect size	Benefits/harms
Beurskens 1997	+	+	+	-	-
Bihaug 1978	+	+	+	-	-
Borman 2003	+	+	+	-	-
Coxhead 1981	+	-	+	-	-
Fritz 2007	+	+	+	-	-
Gudavalli 2006	+	+	+	-	-
Güvenol 2000	+	+	+	?	-

Table 1. Clinical relevance (*Continued*)

Harte 2007	+	+	+	-	-
Konrad 1992	+	?	+	-	-
Larsson 1980	+	+	+	-	-
Letchuman 1993	-	+	+	-	-
Lidström 1970	+	+	+	?	-
Lind 1974	+	+	+	+	+
Ljunggren 1984	+	+	+	-	-
Ljunggren 1992	+	+	+	-	-
Mathews 1975	+	+	+	-	-
Mathews 1988	+	+	+	-	-
Ozturk 2006	+	+	+	-	-
Pal 1986	+	+	+	-	-
Reust 1988	-	+	+	-	-
Schimmel 2009	+	+	+	-	-
Sherry 2001	+	+	+	+	?
Simmerman 2011	+	+	+	-	-
Sweetman 1993	+	+	+	-	-
Tesio 1993	+	+	+	?	-
Unlu 2008	+	+	+	-	-
Van der Heijden 1995	+	+	+	-	-
Walker 1982	+	+	+	-	-
Weber 1973	-	+	+	-	-
Weber 1984 (1)	-	+	+	-	-
Weber 1984 (2)	-	+	+	-	-

Table 1. Clinical relevance (*Continued*)

Werners 1999	+	+	+	-	-

+: yes; -: no; ?: unknown.

APPENDICES

Appendix 1. Search strategy

MEDLINE (Ovid) (1966 to August 2013)

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab,ti.
5. drug therapy.fs.
6. randomly.ab,ti.
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.ti,ab.
13. exp Back Pain/
14. backache.ti,ab.
15. exp Low Back Pain/
16. (lumbar adj pain).ti,ab.
17. coccyx.ti,ab.
18. coccydynia.ti,ab.
19. sciatica.ti,ab.
20. sciatic neuropathy/
21. spondylosis.ti,ab.
22. lumbago.ti,ab.
23. or/12-22
24. exp Spine/
25. discitis.ti,ab.
26. exp Spinal Diseases/
27. (disc adj degeneration).ti,ab.
28. (disc adj prolapse).ti,ab.
29. (disc adj herniation).ti,ab.
30. spinal fusion.sh.
31. spinal neoplasms.sh.
32. (facet adj joints).ti,ab.
33. intervertebral disk.sh.
34. intervertebral disc.sh.
35. Intervertebral Disc Displacement.sh.
36. postlaminectomy.ti,ab.

37. arachnoiditis.ti,ab.
38. (failed adj back).ti,ab.
39. or/24-38
40. 23 or 39
41. 11 and 40
42. exp Traction/
43. exp "Physical Therapy (Specialty)"/
44. 42 or 43
45. exp Fractures, Bone/
46. 44 not 45
47. 11 and 41 and 46

EMBASE Ovid (1980 to August 2013)

1. Clinical Article/
2. exp Clinical Study/
3. Clinical Trial/
4. Controlled Study/
5. Randomized Controlled Trial/
6. Major Clinical Study/
7. Double Blind Procedure/
8. Multicenter Study/
9. Single Blind Procedure/
10. Phase 3 Clinical Trial/
11. Phase 4 Clinical Trial/
12. crossover procedure/
13. placebo/
14. or/1-13
15. allocat$.mp.
16. assign$.mp.
17. blind$.mp.
18. (clinic$ adj25 (study or trial)).mp.
19. compar$.mp.
20. control$.mp.
21. cross?over.mp.
22. factorial$.mp.
23. follow?up.mp.
24. placebo$.mp.
25. prospectiv$.mp.
26. random$.mp.
27. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).mp.
28. trial.mp.
29. (versus or vs).mp.
30. or/15-29
31. 14 and 30
32. human/
33. Nonhuman/
34. exp ANIMAL/
35. Animal Experiment/
36. 33 or 34 or 35
37. 32 not 36
38. 31 not 36
39. 37 and 38
40. 38 or 39
41. dorsalgia.mp.

42. back pain.mp.
43. exp LOW BACK PAIN/
44. exp BACKACHE/
45. (lumbar adj pain).mp.
46. coccyx.mp.
47. coccydynia.mp.
48. sciatica.mp.
49. exp ISCHIALGIA/
50. spondylosis.mp.
51. lumbago.mp.
52. or/41-50
53. exp SPINE/
54. discitis.mp.
55. exp Spine Disease/
56. (disc adj degeneration).mp.
57. (disc adj prolapse).mp.
58. (disc adj herniation).mp.
59. spinal fusion.mp.
60. spinal neoplasms.mp.
61. (facet adj joints).mp.
62. intervertebral disk.mp.
63. postlaminectomy.mp.
64. arachnoiditis.mp.
65. (failed adj back).mp.
66. or/53-65
67. 52 or 66
68. 40 and 67
69. exp traction therapy/
70. exp fracture/
71. 69 not 70
72. 68 and 71

CENTRAL (*The Cochrane Library*, 2012 Issue 8)

1. MeSH descriptor Back Pain explode all trees
2. dorsalgia
3. backache
4. MeSH descriptor Low Back Pain explode all trees
5. (lumbar next pain) or (coccyx) or (coccydynia) or (sciatica) or (spondylosis)
6. MeSH descriptor Sciatica explode all trees
7. MeSH descriptor Spine explode all trees
8. MeSH descriptor Spinal Diseases explode all trees
9. (lumbago) or (discitis) or (disc near degeneration) or (disc near prolapse) or (disc near herniation)
10. spinal fusion
11. facet near joints
12. MeSH descriptor Intervertebral Disk explode all trees
13. postlaminectomy
14. arachnoiditis
15. failed near back
16. MeSH descriptor Cauda Equina explode all trees
17. lumbar near vertebra*
18. spinal near stenosis
19. slipped near (disc* or disk*)
20. degenerat* near (disc* or disk*)
21. stenosis near (spine or root or spinal)

22. displace* near (disc* or disk*)
23. prolap* near (disc* or disk*)
24. (#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)
25. MeSH descriptor Traction explode all trees
26. MeSH descriptor Physical Therapy (Specialty) explode all trees
27. (#25 OR #26)
28. MeSH descriptor Fractures, Bone explode all trees
29. (#27 AND NOT #28)
30. (#24 AND #29)

CINAHL (Ebsco) (January 2006 to August 2013)
S53 S49 and S52
S52 S50 NOT S51
S51 (MH "Fractures+")
S50 (MH "Traction") OR "traction"
S49 S47 or S48
S48 S35 or S43 or S47
S47 S44 or S45 or S46
S46 "lumbago"
S45 (MH "Spondylolisthesis") OR (MH "Spondylolysis")
S44 (MH "Thoracic Vertebrae")
S43 S36 or S37 or S38 or S39 or S40 or S41 or S42
S42 lumbar N2 vertebra
S41 (MH "Lumbar Vertebrae")
S40 "coccydynia"
S39 "coccyx"
S38 "sciatica"
S37 (MH "Sciatica")
S36 (MH "Coccyx")
S35 S29 or S30 or S31 or S32 or S33 or S34
S34 lumbar N5 pain
S33 lumbar W1 pain
S32 "backache"
S31 (MH "Low Back Pain")
S30 (MH "Back Pain+")
S29 "dorsalgia"
S28 S26 NOT S27
S27 (MH "Animals")
S26 S7 or S12 or S19 or S25
S25 S20 or S21 or S22 or S23 or S24
S24 volunteer*
S23 prospectiv*
S22 control*
S21 followup stud*
S20 follow-up stud*
S19 S13 or S14 or S15 or S16 or S17 or S18
S18 (MH "Prospective Studies+")
S17 (MH "Evaluation Research+")
S16 (MH "Comparative Studies")
S15 latin square
S14 (MH "Study Design+")
S13 (MH "Random Sample")
S12 S8 or S9 or S10 or S11

S11 random*
S10 placebo*
S9 (MH "Placebos")
S8 (MH "Placebo Effect")
S7 S1 or S2 or S3 or S4 or S5 or S6
S6 triple-blind
S5 single-blind
S4 double-blind
S3 clinical W3 trial
S2 "randomi?ed controlled trial*"
S1 (MH "Clinical Trials+")

Appendix 2. Criteria for assessing risk of bias for internal validity

Random sequence generation (selection bias)

Selection bias (biased allocation to interventions) due to inadequate generation of a randomized sequence

There is a low risk of selection bias if the investigators describe a random component in the sequence generation process such as: referring to a random number table, using a computer random number generator, coin tossing, shuffling cards or envelopes, throwing dice, drawing of lots, minimization (minimization may be implemented without a random element, and this is considered to be equivalent to being random).

There is a high risk of selection bia*s* if the investigators describe a non-random component in the sequence generation process, such as: sequence generated by odd or even date of birth, date (or day) of admission, hospital or clinic record number; or allocation by judgement of the clinician, preference of the participant, results of a laboratory test or a series of tests, or availability of the intervention.

Allocation concealment (selection bias)

Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment

There is a low risk of selection bias if the participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: central allocation (including telephone, web-based and pharmacy-controlled randomization); sequentially numbered drug containers of identical appearance; or sequentially numbered, opaque, sealed envelopes.

There is a high risk of bias if participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: using an open random allocation schedule (e.g. a list of random numbers); assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or non-opaque or not sequentially numbered); alternation or rotation; date of birth; case record number; or other explicitly unconcealed procedures.

Blinding of participants

Performance bias due to knowledge of the allocated interventions by participants during the study

There is a low risk of performance bias if blinding of participants was ensured and it was unlikely that the blinding could have been broken; or if there was no blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding.

Blinding of personnel/care providers (performance bias)

Performance bias due to knowledge of the allocated interventions by personnel/care providers during the study

There is a low risk of performance bias if blinding of personnel was ensured and it was unlikely that the blinding could have been broken; or if there was no blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding.

Blinding of outcome assessor (detection bias)

Detection bias due to knowledge of the allocated interventions by outcome assessors

There is low risk of detection bias if the blinding of the outcome assessment was ensured and it was unlikely that the blinding could have been broken; or if there was no blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding, or:

- for participant-reported outcomes in which the participant was the outcome assessor (e.g. pain, disability): there is a low risk of bias for outcome assessors if there is a low risk of bias for participant blinding (Boutron 2005);
- for outcome criteria that are clinical or therapeutic events that will be determined by the interaction between participants and care providers (e.g. co-interventions, length of hospitalization, treatment failure), in which the care provider is the outcome assessor: there is a low risk of bias for outcome assessors if there is a low risk of bias for care providers (Boutron 2005);
- for outcome criteria that are assessed from data from medical forms: there is a low risk of bias if the treatment or adverse effects of the treatment could not be noticed in the extracted data (Boutron 2005).

Incomplete outcome data (attrition bias)

Attrition bias due to amount, nature or handling of incomplete outcome data

There is a low risk of attrition bias if there were no missing outcome data; reasons for missing outcome data were unlikely to be related to the true outcome (for survival data, censoring unlikely to be introducing bias); missing outcome data were balanced in numbers, with similar reasons for missing data across groups; for dichotomous outcome data, the proportion of missing outcomes compared with the observed event risk was not enough to have a clinically relevant impact on the intervention effect estimate; for continuous outcome data, the plausible effect size (difference in means or standardized difference in means) among missing outcomes was not enough to have a clinically relevant impact on observed effect size, or missing data were imputed using appropriate methods (if dropouts are very large, imputation using even 'acceptable' methods may still suggest a high risk of bias) (Van Tulder 2003). The percentage of withdrawals and dropouts should not exceed 20% for short-term follow-up and 30% for long-term follow-up and should not lead to substantial bias (these percentages are commonly used but arbitrary, not supported by literature) (Van Tulder 2003).

Selective Reporting (reporting bias)

Reporting bias due to selective outcome reporting

There is low risk of reporting bias if the study protocol is available and all of the study's prespecified (primary and secondary) outcomes that are of interest in the review have been reported in the prespecified way, or if the study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were prespecified (convincing text of this nature may be uncommon).

There is a high risk of reporting bias if not all of the study's prespecified primary outcomes have been reported; one or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not prespecified; one or more reported primary outcomes were not prespecified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); one or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; the study report fails to include results for a key outcome that would be expected to have been reported for such a study.

Group similarity at baseline (selection bias)

Bias due to dissimilarity at baseline for the most important prognostic indicators.
There is low risk of bias if groups are similar at baseline for demographic factors, value of main outcome measure(s), and important prognostic factors (examples in the field of back and neck pain are duration and severity of complaints, vocational status, percentage of participants with neurological symptoms) (Van Tulder 2003).

Co-interventions (performance bias)

Bias because co-interventions were different across groups

There is low risk of bias if there were no co-interventions or they were similar between the index and control groups (Van Tulder 2003).

Compliance (performance bias)

Bias due to inappropriate compliance with interventions across groups

There is low risk of bias if compliance with the interventions was acceptable, based on the reported intensity/dosage, duration, number and frequency for both the index and control intervention(s). For single-session interventions (e.g. surgery), this item is irrelevant (Van Tulder 2003).

Intention-to-treat-analysis

There is low risk of bias if all randomized participants were reported/analysed in the group to which they were allocated by randomization.

Timing of outcome assessments (detection bias)

Bias because important outcomes were not measured at the same time across groups

There is low risk of bias if all important outcome assessments for all intervention groups were measured at the same time (Van Tulder 2003).

Other bias

Bias due to problems not covered elsewhere in the table

There is a low risk of bias if the study appears to be free of other sources of bias not addressed elsewhere (e.g. study funding).

FEEDBACK

Personal experience with traction, 2 January 2010

Summary

Individual shared personal experience with traction as a positive alternative to surgery for his back pain.
Personal correspondence between Managing Editor and contributor. Not appropriate to include.

Reply

Contributor responded appreciatively to correspondence.

Contributors

Victoria Pennick, Managing Editor, Cochrane Back Review Group

WHAT'S NEW

Last assessed as up-to-date: 13 May 2013.

Date	Event	Description
27 May 2013	New citation required but conclusions have not changed	Conclusions not changed.
13 May 2013	New search has been performed	Review updated. Seven new RCTs were incorporated. The review was performed using the latest methods concerning risk of bias assessment and reporting as stated in the Handbook

HISTORY

Protocol first published: Issue 2, 2001

Review first published: Issue 4, 2005

Date	Event	Description
14 January 2010	Amended	Feedback added
27 June 2008	Amended	Converted to new review format.

(*Continued*)

25 January 2007	New citation required but conclusions have not changed	Conclusions were not changed by the addition of the newly identified trial. Author by-line changed
31 October 2006	New search has been performed	There was only one additional trial identified for this update. It did not change the conclusions

CONTRIBUTIONS OF AUTHORS

All authors were involved in writing the protocol and the final manuscript. I Wegner, IS Widyahening and GJMG van der Heijden were involved in the quality assessment, data extraction processes and the data analysis.

DECLARATIONS OF INTEREST

Two review authors (GJMG van der Heijden, HCW de Vet) were also authors of two included studies. They did not participate in the quality assessment or data extraction processes in these studies.

SOURCES OF SUPPORT

Internal sources

- Institute for Work & Health, Canada.
- EMGO+ Institute for Health and Care Research, VU University Medical Centre, Netherlands.
- Department of Public Health and Caring Sciences, Family Medicine, Uppsala Science Park, Sweden.
- Northwestern Health Sciences University, Wolff-Harris Center for Clinical Studies, USA.
- Julius Centre for Health Sciences and Primary Care, University Medical Center Utrecht, Netherlands.

External sources

- No sources of support supplied

INDEX TERMS

Medical Subject Headings (MeSH)

*Traction [adverse effects]; Acute Pain [therapy]; Chronic Pain [therapy]; Low Back Pain [complications; *therapy]; Pain Measurement; Randomized Controlled Trials as Topic; Sciatica [complications; *therapy]

MeSH check words

Humans