ABSTRACT

Background

Exercise or physical activity is recommended for improving pain and functional status in people with knee or hip osteoarthritis. These are complex interventions whose effectiveness depends on one or more components that are often poorly identified. It has been suggested that health benefits may be greater with high-intensity rather than low-intensity exercise or physical activity.

Objectives

To determine the benefits and harms of high- versus low-intensity physical activity or exercise programs in people with hip or knee osteoarthritis.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL; issue 06, 2014), MEDLINE (1948 to June 2014), EMBASE (1980 to June 2014), CINAHL (1982 to June 2014), PEDro (1929 to June 2014), SCOPUS (to June 2014) and the World Health Organization (WHO) International Clinical Registry Platform (to June 2014) for articles, without a language restriction. We also handsearched relevant conference proceedings, trials, and reference lists and contacted researchers and experts in the field to identify additional studies.

Selection criteria

We included randomized controlled trials of people with knee or hip osteoarthritis that compared high- versus low-intensity physical activity or exercise programs between the experimental and control group.

High-intensity physical activity or exercise programs training had to refer to an increase in the overall amount of training time (frequency, duration, number of sessions) or the amount of work (strength, number of repetitions) or effort/energy expenditure (exertion, heart rate, effort).
Data collection and analysis

Two review authors independently assessed study eligibility and extracted data on trial details. We contacted authors for additional information if necessary. We assessed the quality of the body of evidence for these outcomes using the GRADE approach.

Main results

We included reports for six studies of 656 participants that compared high- and low-intensity exercise programs; five studies exclusively recruited people with symptomatic knee osteoarthritis (620 participants), and one study exclusively recruited people with hip or knee osteoarthritis (36 participants). The majority of the participants were females (70%). No studies evaluated physical activity programs. We found the overall quality of evidence to be low to very low due to concerns about study limitations and imprecision (small number of studies, large confidence intervals) for the major outcomes using the GRADE approach. Most of the studies had an unclear or high risk of bias for several domains, and we judged five of the six studies to be at high risk for performance, detection, and attrition bias.

Low-quality evidence indicated reduced pain on a 20-point Western Ontario and McMaster Universities Arthritis Index (WOMAC) pain scale (mean difference (MD) -0.84, 95% confidence interval (CI) -1.63 to -0.04; 4% absolute reduction, 95% CI -8% to 0%; number needed to treat for an additional beneficial outcome (NNTB) 11, 95% CI 14 to 22) and improved physical function on the 68-point WOMAC disability subscale (MD -2.65, 95% CI -5.29 to -0.01; 4% absolute reduction; NNTB 10, 95% CI 8 to 13) immediately at the end of the exercise programs (from 8 to 24 weeks). However, these results are unlikely to be of clinical importance. These small improvements did not continue at longer-term follow-up (up to 40 weeks after the end of the intervention). We are uncertain of the effect on quality of life, as only one study reported this outcome (0 to 200 scale; MD 4.3, 95% CI -6.5 to 15.2; 2% absolute reduction; very low level of evidence).

Our subgroup analyses provided uncertain evidence as to whether increased exercise time (duration, number of sessions) and level of resistance (strength or effort) have an impact on the exercise program effects.

Three studies reported withdrawals due to adverse events. The number of dropouts was small. Only one study systematically monitored adverse effects, but four studies reported some adverse effects related to knee pain associated with an exercise program. We are uncertain as to whether high intensity increases the number of adverse effects (Peto odds ratio 1.72, 95% CI 0.51 to 5.81; -2% absolute risk reduction; very low level of evidence). None of the included studies reported serious adverse events.

Authors’ conclusions

We found very low-quality 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. There was insufficient evidence to determine the effect of different types of intensity of exercise programs.

We are uncertain as to whether higher-intensity exercise programs may induce more harmful effects than those of lower intensity; this must be evaluated by further studies. Withdrawals due to adverse events were poorly monitored and not reported systematically in each group. We downgraded the evidence to low or very low because of the risk of bias, inconsistency, and imprecision.

The small number of studies comparing high- and low-intensity exercise programs in osteoarthritis underscores the need for more studies investigating the dose-response relationship in exercise programs. In particular, further studies are needed to establish the minimal intensity of exercise programs needed for clinical effect and the highest intensity patients can tolerate. Larger studies should comply with the Consolidated Standards of Reporting Trials (CONSORT) checklist and systematically report harms data to evaluate the potential impact of highest intensities of exercise programs in people with joint damage.

PLAIN LANGUAGE SUMMARY

Benefits and harms of high- versus low-intensity exercise programs for hip or knee osteoarthritis

Review question

We searched the literature until June 2014 for studies on the benefits and harms of high- versus low-intensity exercise programs for people with hip or knee osteoarthritis.

Background
Osteoarthritis is a chronic condition that affects the joints (commonly hips, knees, spine, and hands). Over time, cartilage wears down in joints. People with osteoarthritis generally feel pain and can have difficulties performing daily activities such as walking. Exercise or physical activity programs are non-drug treatments usually recommended for people with hip or knee osteoarthritis. Many types of exercises are prescribed, but it may be unclear whether or not they are effective. Several different components can play a role in the effectiveness of an exercise regimen, such as exercise duration, frequency, or level of resistance. High intensity can be defined as an extra amount of time (duration or frequency) or resistance (strength or effort) required in the exercise programs.

**Study characteristics**

We identified six randomized controlled trials with 656 participants. Five studies (620 participants) enrolled people with knee osteoarthritis, and one study (36 participants) enrolled people with knee or hip osteoarthritis. The studies included more women (70%) than men.

**Key results**

On a scale of 0 to 20 points (lower scores mean reduced pain), people who completed a high-intensity exercise program rated their pain 0.84 points lower (4% absolute improvement) than people who completed a low-intensity exercise program. People who performed a low-intensity exercise program rated their pain at 6.6 points.

On a scale of 0 to 68 points (lower scores mean better function), people who completed a high-intensity exercise program rated their physical function 2.65 points lower (4% absolute improvement) than people who completed a low-intensity exercise program. People who performed a low-intensity exercise program rated their pain at 20.4 points.

On a scale of 0 to 200 mm visual analog scale (higher score means better function), people who completed a high-intensity exercise program rated their quality of life 4.3 mm higher (6.5 mm lower to 15.2 mm higher) (2% absolute improvement) than people who performed a low-intensity exercise program. People who performed a low-intensity exercise program rated their quality of life at 66.7 mm.

Two per cent more people had adverse effects with high-intensity exercise, or 17 more people out of 1000.

• 39 out of 1000 people reported an adverse effect related to high-intensity exercise program

• 22 out of 1000 people reported an adverse effect related to low-intensity exercise program

Adverse events were not systematically monitored and were incompletely reported by group. None of the included studies reported serious adverse events.

Based on the evidence, people with knee osteoarthritis who perform high-intensity exercise may experience slight improvements in knee pain and function at the end of the exercise program (8 to 24 weeks) when compared with a low-intensity exercise program. We are uncertain as to whether high-intensity exercise improves quality of life or increases the number of people who experience adverse events.

**Quality of evidence**

We graded the quality of evidence as low for pain and function and very low for quality of life. The small number of studies and participants included in some analyses reduced the robustness and precision of these findings.

Adverse effects were poorly recorded. Very low quality evidence shows we are uncertain whether higher-intensity exercise programs may result in more side effects than lower-intensity exercise programs. Further research may change the result.
### Summary of Findings for the Main Comparison

**Physical activity and exercise programs in osteoarthritis**

**Patient or population:** People with hip or knee osteoarthritis  
**Settings:** Hospital or primary care  
**Intervention:** High- versus low-intensity exercise programs

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed risk</td>
<td>Corresponding risk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>High- versus low-intensity exercise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain (at study completion)</td>
<td>The mean pain (at study completion) in the control groups was <strong>6.6 points</strong></td>
<td>The mean pain (at study completion) in the intervention groups was <strong>0.84 lower</strong> (1.63 to 0.04 lower)</td>
<td>-</td>
<td>313 (4 studies)</td>
<td>⊕⊕⊕⊕ low¹ ²</td>
</tr>
<tr>
<td>Physical function (at study completion)</td>
<td>The mean physical function (at study completion) in the control groups was <strong>20.4 points</strong></td>
<td>The mean physical function (at study completion) in the intervention groups was <strong>2.65 lower</strong> (5.29 to 0.01 lower)</td>
<td>-</td>
<td>310 (4 studies)</td>
<td>⊕⊕⊕⊕ low¹ ²</td>
</tr>
</tbody>
</table>

MD -0.84 (95% CI -1.63 to -0.04). Absolute mean reduction 4% with high intensity exercise programs (95% CI -8% more to 0% fewer). Relative reduction -13% (95% CI -25% more to 0% fewer)  
NNTB = 11 (95% CI 14 to 22)³ ⁴

MD -2.65 (95% CI -5.29 to -0.01). Absolute mean reduction 4% with high intensity exercise programs (95% CI -8% more to 0% fewer). Relative mean reduction 13% (95% CI -26% more to 0% fewer)  
NNTB = 10 (95% CI 8 to 13)³ ⁴ ⁵
### Quality of life (at study completion)

| EuroQol (VAS) from 0 to 200 mm; higher score means better function | The mean quality of life (at study completion) in the control groups was 66.7 mm | The mean quality of life (at study completion) in the intervention groups was 4.3 higher (6.5 to 15.2 higher) | 214 (1 study) | MD 4.3 (95% CI -6.5 to 15.2). Absolute mean improvement 2% (95% CI -3% fewer to 8% more). Relative improvement 6% (95% CI -10% fewer to 23% more) NNTB = NA |

### Adverse effects (related to the exercise programs)

| Follow-up: 8 to 24 weeks | 22 per 1000 (11 to 131 higher) | 39 per 1000 (11 to 131 higher) | Peto OR 1.72 (0.51 to 5.81) | 364 (4 studies) | Absolute risk reduction 2% fewer events with low intensity exercise programs (95% CI 11% fewer to -1% more). Relative risk reduction: 69% fewer with low intensity exercise programs (95% CI 425% more to -48% fewer) NNTH = 65 (95% CI NNTB 92 to NNTH 11) |

### Severe adverse events or withdrawals (due to adverse events) - not reported

| See comment | See comment | Not estimable | - | See comment | 3 studies reported drop-outs or withdrawals (due to adverse events). No severe adverse events were observed |

---

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: confidence interval; MD: mean difference; NA: not applicable; NNTB: number needed to treat for an additional beneficial outcome; NNTH: number needed to treat for an additional harmful outcome; NRS: numeric rating scale; OR: odds ratio; SD: standard deviation; VAS: visual analog scale; WOMAC: Western Ontario and McMaster Universities Arthritis Index

**GRADE Working Group grades of evidence**

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

1. Downgraded because of lack of blinding. No intention-to-treat analysis. Incomplete outcome data.
2. Downgraded because of imprecision.
3. Calculations based on the control group baseline mean (SD) WOMAC pain: -1.54 (3.84) points on 0-20 scale (from McCarthy 2004) and an assumed minimal clinically important difference of 4 points (Tubach 2012).
4. NNT for continuous outcomes calculated using the Wells calculator (from the CMSG Editorial office; http://musculoskeletal.cochrane.org), and for dichotomous outcomes using the Cates NNT calculator (www.nntonline.net/visualrx).
5. Calculations based on the control group baseline mean (SD) WOMAC function: -4.5 (14.7) points on 0-68 scale (from McCarthy 2004) and an assumed minimal clinically important difference of 14 points (Tubach 2012).
6. Only one study reported EuroQol data.
7. Calculations based on the control group baseline mean (SD) EuroQol: 66.7 (18.2) points on 200 mm scale (from McCarthy 2004) and an assumed minimal clinically important difference of 15% of mean baseline.
8. Downgraded because of inconsistency (only 1 of 6 studies systematically monitored adverse effects. Unbalanced withdrawals across exercise groups and are > 10%).
9. Some studies did not report whether or not adverse events occurred in either group.
BACKGROUND

Description of the condition

Osteoarthritis is the most common type of arthritis and related diseases in the world (Murray 2012). The prevalence increases with age, and lifestyle factors such as obesity and lack of physical activity are risk factors (Woolf 2003). Approximately 10% of the world’s population aged 60 or older have symptomatic osteoarthritis (Zhang 2010).

Osteoarthritis is a chronic condition that affects the joints and occurs when cartilage in joints wears down over time. The disease process can affect almost any joint, but occurs mostly in the knees, hips, spine, and hands. The population impact is greatest for osteoarthritis of the hips and knees (Vos 2012).

People with osteoarthritis generally experience pain, reduced joint motion, and muscle weakness and are unable to perform a variety of daily living activities (Moskowitz 2009). Although osteoarthritis is a degenerative disease and therefore has no cure, a number of treatments can control symptoms and improve quality of life.

Description of the intervention

People with osteoarthritis experiencing pain have reduced activities (Moskowitz 2009). Similarly, reduced muscle strength is associated with pain and functional disability (Jan 2008). Current international guidelines, in Hochberg 2012 and Brosseau 2014, recommend managing osteoarthritis by promoting activity and participation in regular physical activities and exercise therapy (Vignon 2006; McAlindon 2014). Several clinical studies have shown that aerobic physical activity and muscle-strengthening exercise may help reduce symptoms of osteoarthritis and improve function (Latham 2010). Prescribed physical activity or exercise therapies usually target aerobic capacity, muscular strength, and flexibility. The World Health Organization defines physical activity as all forms of activity (for example occupational, recreational, sports related) involving skeletal muscles that require energy expenditure (World Health Organization 2010). Exercise refers to a form of physical activity that is planned and structured and is often developed by a fitness or rehabilitation specialist for the client or patient (Bouchard 2007). The delivery of exercise programs varies by amount and magnitude of work (level of resistance, frequency, duration, and progression), supervision (type, mode of delivery), and setting (home, community/gym, healthcare setting).

How the intervention might work

Physical activity or exercise may be effective for people with osteoarthritis (Bijlsma 2011). Exercise prescription includes different components: intensity, frequency, duration, and mode. Intensity is a feature of exercise programs that may be high, vigorous, moderate, or low depending on the treatment goal (for example muscle weakness) or the subject population. Studies exploring the impact of intensity level of exercise on physical performance, Kraus 2002, Heiwe 2011, and Robbins 2012, have suggested that a more intense program of physical activity or exercise may be more effective for stroke patients, in Hunter 2011, or older adults, in Galvao 2005. Health benefits may be greater with high-versus low-intensity exercise programs.

Why it is important to do this review

Several systematic reviews have highlighted a minimum intensity of exercise programs that is necessary for health benefits (Fransen 2008a; Fransen 2008b), but evidence for the effect of high-intensity physical activity or exercise programs on pain and physical function in people with hip or knee osteoarthritis is lacking. In a Cochrane review including only a single trial (39 participants), the benefits of an exercise program did not differ by high or low intensity for people with knee osteoarthritis (Brosseau 2003). Moreover, the interaction between the effect of intensity and type of treatment (physical activity, exercise) or the joint involved could not be explored. More recently, several new clinical studies have reported that more intensive exercise programs increasing the strength of muscles and overall activity level may be beneficial for adults with osteoarthritis (Jan 2008). Several trials have been published since 2003, and a reappraisal of the available evidence regarding the effect of intensity on both physical activity (for example walking or cycling) and exercise programs is warranted.

OBJECTIVES

To determine the benefits and harms of high-versus low-intensity physical activity or exercise programs on pain and physical function in people with hip or knee osteoarthritis.

METHODS

Criteria for considering studies for this review

Types of studies

We only included randomized controlled trials (RCTs).
Types of participants
We included studies if they recruited at least 75% of participants with clinically or radiographically confirmed primary osteoarthritis of the knee or hip. We excluded studies of people with inflammatory arthritis, such as rheumatoid arthritis. We did not consider studies of children.

Types of interventions
Studies were eligible if they compared high- and low-intensity physical activity or exercise programs between the experimental and control groups.

The World Health Organization 2010 study defines physical activity as "any movement produced by skeletal muscles that requires energy expenditure." "Exercise program" is defined as a form of physical activity that is planned, structured, and repeated over a period of time (Bouchard 2007), with "the intention of improving or maintaining physical fitness or health" (Kwakkel 2004; Umpierre 2011).

We considered high-intensity programs in the experimental group as the following:
- the additional amount of time spent in an activity or exercise compared to the control group (session duration or number of sessions, or both); and
- the amount of resistance work (strength, rates) or effort (magnitude) required to perform an activity or exercise compared to the control group (resistance exercise).

More specifically, high-intensity physical activity or exercise program training could refer to an increase in (1) the overall amount of training time: the program length (week, months), frequency (days/week), duration of sessions (minutes), number of sessions, or both; and (2) the amount of work (torque, repetitions, strength) or amount of effort/energy expenditure (exertion, heart rate, effort).

The review included studies involving the following types of comparisons:
1. The same exercise or physical activity programs performed in the experimental and control groups but with different intensity.
2. Any exercise or physical activity performed in the experimental and control groups but with additional exercise or physical activity program(s) in the experimental group.

We excluded studies that compared exercise programs with no exercise training (for example passive stretch, educational advice, placebo or sham).

Types of outcome measures
We used data from the outcomes assessment conducted immediately on completion of the intervention program. When data were available, we analyzed the effects at mid-term (6 to 12 months) and long-term (after 12 months) follow-up.

Major outcomes
The major outcomes were pain, function, and quality of life, as currently recommended for osteoarthritis trials (Altman 1996; Pham 2004). For safety, the major outcomes were the number of participants who withdrew because of adverse events and number of participants experiencing any serious adverse events.

Pain
If a trial provided data on more than one pain scale, we extracted data on the pain scale that was highest on the following list according to a previously described hierarchy of pain-related outcomes (Juni 2006; Reichenbach 2007).
- Pain overall
- Pain on walking
- Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale
- Pain during activities other than walking
- WOMAC global scale
- Lequesne osteoarthritis index global score
- Other algofunctional scale
- Patient global assessment
- Physician global assessment
- Other outcome
- No continuous outcome reported

Physical function
If a trial provided data on more than one physical function scale, we extracted data according to the following hierarchy.
- Global disability score
- Walking disability
- WOMAC disability subscore
- Composite disability scores other than WOMAC
- Disability other than walking
- WOMAC global scale
- Lequesne osteoarthritis index global score
- Other algofunctional scale

Quality of life
We extracted quality-of-life data collected by the Medical Outcomes Survey Short Forms 12 and 36, EuroQoL, Sickness Impact Profile, or Nottingham Health Profile.

Safety
- Withdrawals due to adverse events
- Severe adverse events outcomes: inpatient hospitalization, life-threatening events or death
- Adverse effects associated with the exercise intervention including joint or muscle contractures, fatigue, pain, falls, function limitations
Minor outcomes

Physical global performance
- Walking ability, including gait speed and walking endurance
- Muscle strength, using quantitative strength sensors
- Aerobic capacity, including peak VO₂ or peak work rate
- Range of motion
- Physical activity levels
- Number of steps

Joint imaging
- Joint space narrowing measurement on radiography
- The Kellgren-Lawrence classification

Search methods for identification of studies

Electronic searches
We first searched the following databases the Cochrane Central Register of Controlled Trials (CENTRAL; 2012, Issue 10), MEDLINE through OVID (1948 to October 2012), EMBASE through Elsevier (1980 to October 2012), CINAHL (1982 to October 2012) and PEDro (from 1929 to October 2012). The MEDLINE search involved the Cochrane highly sensitive search strategy, sensitivity-maximizing version (2008 revision). The EMBASE search involved the UK Cochrane Centre search filter to identify reports of RCTs. We used the Google Scholar search engine to find additional references.

We performed an updated search in all the databases on June 2014. The Trials Search Co-ordinator for the Cochrane Musculoskeletal Review Group helped develop search equations. The queries combined free text words and controlled vocabulary. The search strategy was based on synonyms of (“physical activity” OR “exercise”) AND “osteoarthritis.” We used an adapted search strategy to search MEDLINE (Appendix 1), EMBASE (Appendix 2), the Cochrane Central Register of Controlled Trials (CENTRAL; Cochrane Library) (Appendix 3), CINAHL (Appendix 4), and PEDro (Appendix 5). We did not restrict the search by language of publication or publication status.

Searching other resources
we searched for additional relevant systematic reviews in the Cochrane Database of Systematic Reviews and the Database of Abstracts of Reviews of Effects (DARE) (to June 2014). We handsearched the reference lists of selected trials and systematic reviews identified from electronic searches.

We also searched conference proceedings available online for the American College of Rheumatology, European League Against Rheumatism, and Osteoarthritis Research Society International (up to the two latest editions). We contacted authors and field experts for any additional published or unpublished data.

To identify trials in progress, we used the WHO International Clinical Trials Registry Platform (www.who.int/trialsearch); ClinicalTrials.gov (www.clinicaltrials.gov); and the metaRegister of Controlled Trials (mRCT) (www.controlled-trials.com).

We contacted authors of active or completed trials for provisional results if they had not yet been published. We supplemented database searching and hand-searching and for the RCTs identified by the first step by tracking citations in SCOPUS (Appendix 6).

Data collection and analysis

Selection of studies
We removed duplicate records from retrieved references. Using the inclusion and exclusion criteria, two review authors (JPR, CN) independently screened the titles and abstracts identified by the search strategy in order to identify potentially relevant studies. These review authors obtained and screened the full-length articles for selected titles and abstracts to check for eligibility and decide on their inclusion. Disagreements were resolved by discussion and with the assistance of a third review author (MMLC) if needed.

If results of eligible trials were available in an abstract only, we contacted the trial authors to ask for a report of the trial results. We linked multiple reports relating to the same trial or trials with potentially overlapping populations. If we could not exclude the possibility of overlapping populations, we selected the more recent trial.

To confirm the eligibility of studies, we assessed the intensity of exercise programs. Interventions in rehabilitation are in fact complex (Boutron 2008), involving several components that may interact separately or together, for difficulties in classification. For each included study, the two review authors (JPR, CN) used the following steps:

1. identified the experimental and the control group; and
2. rated the intensity of physical activity or exercise programs in the experimental and control group as high or low on the basis of their description.

We used a consensus method to resolve disagreements and consulted a third review author (IB) if necessary. If the article did not contain information on the exercise program intensity, we contacted the trial authors for additional information. In the absence of sufficient information, we excluded the study. Review authors were blinded to all characteristics of the trial except for the content of the interventions.

Data extraction and management
Two review authors (JPR, MMLC) independently extracted results of individual trials by using a standardized piloted extraction method.
We entered data into Review Manager and checked it for accuracy (RevMan 2011).

Assessment of risk of bias in included studies
We evaluated the risk of bias in each included study according to the 'Risk of bias' tool recommended by The Cochrane Collaboration. Two review authors (JPR, MMLC) independently examined seven specific domains: sequence generation, allocation concealment, blinding of participants or personnel, blinding of outcome assessors, incomplete outcome data, selective outcome reporting, and other potential source of bias (that is design specific, baseline imbalance). We scored each criterion as “high risk of bias,” “low risk of bias,” or “unclear risk of bias,” depending on the information supplied in the report.

We classified studies as at low risk of bias if all key domains had low risk of bias and no serious flaws; high risk of bias if one or more domains had high risk of bias; and unclear risk of bias if one or more domains had unclear risk of bias (Higgins 2011). We resolved any disagreements by involving a third review author (IB).

Measures of treatment effect
For dichotomous data, we expressed the results of each RCT as risk ratios (RRs) with corresponding 95% confidence intervals (CIs). We used the Peto method to calculate a Peto Odds ratio from the number of adverse effects reported by each study before combining them (Deeks 2011).

For continuous outcomes, we summarized results as mean difference (MD) if the same tool was used to measure the same outcome across separate studies. Alternatively, we calculated the standardized mean difference (SMD) when studies measured the same outcome but used different tools. The SMD expresses the size of the intervention effect in each study relative to the variability observed in that study. We calculated the SMD by dividing the MD by the standard deviation (SD) for the outcome among participants. An SMD greater than 0 indicates a beneficial effect in favor of high-intensity exercise or physical activity. We computed a 95% CI for the SMD. We interpreted the SMD as described in Cohen 1988: SMD = 0.2 is considered a small beneficial effect; 0.5 a medium effect; and 0.8 a large effect.

If the meta-analysis resulted in statistically significant overall estimates, we transformed the treatment effect measures to relative improvement on the original units to express the final results of the review. We back-translated the results by multiplying the SMD by the SD for a representative study (Akl 2011).

Unit of analysis issues
For cross-over trials, we planned to extract data from the first period only, but we included none in this review. Whenever possible, we used results from an intention-to-treat analysis. For studies containing more than two intervention groups, allowing for multiple pair-wise comparisons between all possible pairs of intervention groups, we included the same group of participants only once in the meta-analysis following the procedure recommended by The Cochrane Collaboration (Deeks 2011).

Dealing with missing data
In case of missing outcome data, we contacted the original investigators to request data. We performed sensitivity analyses to
assess how sensitive the results were to changes and addressed the potential impact of missing data on the review findings in the Discussion section.

Assessment of heterogeneity

We evaluated clinical heterogeneity by determining if different clinical factors (characteristics of participants, interventions, outcome measures) varied between trials and could have an influence on the treatment effects. We assessed statistical heterogeneity by a visual inspection of graphs and by using the I² statistic, which describes the proportion of variability in effect estimates due to heterogeneity rather than sampling error (Higgins 2002).

We interpreted the value of the I² statistic according to the following thresholds (Higgins 2011): 0% to 40%, might not be important; 30% to 60%, may represent moderate heterogeneity; 50% to 90%, may represent substantial heterogeneity; and 75% to 100%, considerable heterogeneity. We also computed the 95% CI for the I² statistic (Ioannidis 2007a), as well as the between-study variance Tau², estimated from the random-effects model (Rucker 2008). In all cases, we considered the results from both the fixed-effect and random-effects models and chose the most appropriate.

Assessment of reporting biases

We planned to draw contour-enhanced funnel plots for each meta-analysis to assess the presence of small-study effects (Peters 2008). As the required statistical conditions were not met (10 or more studies, no statistical significant heterogeneity, and ratio of the maximal to minimal variance across studies greater than 4), we did not perform these analyses.

Data synthesis

We performed a meta-analysis if the data of the studies were clinically and statistically sufficiently homogeneous. If not sufficiently homogeneous, we explored heterogeneity in stratified analyses. The starting point for all meta-analyses of studies of effectiveness involved identifying the data type for the outcome measurements. We performed separate meta-analyses for each outcome of interest. We conducted fixed-effect meta-analyses using inverse weighting by variances of treatment contrasts. We conducted random-effects analyses using the approach in the Hardy 1996 study with CIs of profile likelihood form. We considered the results from both the fixed-effect and random-effects models and chose a model based on the assessment of heterogeneity, the size of trials, and the risk of bias within trials.

We planned to perform a bivariate random-effects meta-analysis to address issues of correlated outcome and missingness, but, since no data were missing, we did not perform the analysis.

Subgroup analysis and investigation of heterogeneity

To explore heterogeneity in estimating the effect of intensity, we performed subgroup analyses according to the type of exercise intensity. We compared exercise program effect for pain and function whether the intensity varied in time (duration) and in resistance (strength or effort).

Sensitivity analysis

We performed a sensitivity analysis to assess how the results of meta analysis might be affected by a selection bias (study recruitment participants with hip and knee osteoarthritis) on immediate post-treatment pain and physical function outcomes. We planned to perform a sensitivity analysis to assess this effect on the meta-analysis results, excluding studies at high or unclear risk of bias, but since all of the identified studies had high or unclear risk of bias, we were unable to perform this analysis.

'Summary of findings' tables

We presented the primary outcomes of the review in 'Summary of findings' tables (pain, physical function, quality of life, and adverse effects associated with the exercise program). We included:

1. 'Summary of findings' tables that provided key information concerning the quality of evidence, the magnitude of effect of the interventions examined, and the sum of available data on the main outcomes, as recommended in the Cochrane Handbook for Systematic Reviews of Interventions (Schünemann 2011a); and
2. an overall grading of the evidence related to each main outcome, using the GRADE approach (Schünemann 2011b).

Overall outcome data presented in the 'Summary of findings' tables are based on the time of measurement immediately after the end of the exercise program. For dichotomous outcomes, we calculated the absolute risk difference by using the risk difference statistic in Review Manager (RevMan 2011), with results expressed as a percentage. The relative percentage change (RPC) was calculated using the risk ratio computed from the Peto Odds Ratio (Schünemann 2011b). RPC was obtained as the RR -1 and reported as a percentage. The number needed to treat for an additional harmful outcome (NNTH) from the control group event rate was calculated using the Visual Rx NNT calculator (Cates 2008).

For continuous outcomes, the absolute risk difference was calculated as the MD between high- and low-intensity physical activity or exercise program groups in the original measurement units (divided by the scale), expressed as a percentage; the relative difference was calculated as the absolute change (or MD) divided by the pooled baseline mean obtained from the low-intensity exercise program group in Review Manager. We used the Wells calculator to obtain the number needed to treat for an additional beneficial outcome for continuous measures (available at the Cochrane Musculoskeletal Group editorial office; http://
We determined the minimal clinically important difference for the WOMAC score. We assumed that for each subscale, a change of at least 4 points on the 15-point WOMAC pain scale (15% for absolute improvement) and 10 points on the 68-point WOMAC function scale (15% for absolute improvement) was needed to be considered a clinically meaningful difference (Tubach 2012).

**RESULTS**

**Description of studies**

**Results of the search**

We included six studies (nine reports) (Mangione 1999; McCarthy 2004; Jan 2008; Ng 2010; Foroughi 2011; Singh 2011). The search retrieved 6493 citations, and 3374 citations after duplicates were removed. We excluded 2969 studies on citation screening and 385 studies on abstract screening (see Figure 1). After selecting 20 full-text reports, we excluded 11 reports because they did not meet the selection criteria: no randomization (one study) and no comparison of high- versus low-intensity interventions (10 studies). The remaining nine eligible full-text reports corresponded to six studies of interventions that were exclusively high-intensity exercise programs compared with low-intensity exercise programs. McCarthy 2004 reported on additional variables in two different reports, which were counted as one study for analysis. Likewise, Foroughi 2011 reported on additional variables in two reports that were counted as one study. We found one report and a thesis of one study that were counted as one study (Singh 2011). We had insufficient information to determine inclusion eligibility for one trial (Steinhilber 2012), and we could not contact the authors, so we listed the study in the Characteristics of studies awaiting classification section. In addition, we identified four ongoing trials (see Characteristics of ongoing studies). We requested additional information from eight authors and received responses from six (Ng 2010; Teixeira 2011; Foroughi 2011; Messier 2011; Pua 2012; Østerås 2012) (see Appendix 7). The last searches were performed in June 2014.
Figure 1. Study flow diagram.

6386 records identified through database searching (up to June 2014):
- search (up to 2012): 5242 records found
- update search (2012 to June 2014): 1250 records found

1 additional record identified through other sources (list of references, expert consultation, web search)

3374 titles after duplicates removed

405 abstracts screened

365 abstracts excluded

20 full-text reports assessed for eligibility

11 full-text reports excluded, with reasons
not RCT (n = 1)
not high vs low intensity (n = 10)

6 studies included in qualitative synthesis (from 9 reports)

5 studies included in quantitative synthesis (meta-analysis)
Included studies
We have provided a full description of the six included studies in Characteristics of included studies.

Methods
All included studies were RCTs with a parallel-group design.

Sample size
The six studies included 656 participants. The median sample size was 78 (lower quartile = 43; upper quartile = 176).

Participants
Two studies were conducted in Australia and one study each in the United Kingdom, United States, Taiwan, and India. Most participants (70%) were female, and one study included females only (Foroughi 2011). The mean age of participants was 61 years (range 56 to 71 years). All studies recruited participants in a single center.

The location of osteoarthritis was reported in all studies. The predominant location was the knee. Only one study included participants with hip or knee osteoarthritis (Ng 2010). Two studies reported the mean duration of osteoarthritis: 14 and 12 years (Mangione 1999; Foroughi 2011).

Interventions
All six included studies examined an exercise program and compared outcomes of high and low intensity of exercise. No study examined a physical activity program with different intensity.

Three studies assessed the effect of exercise programs with the two levels of intensity by amount of time spent in the program (McCarthy 2004; Ng 2010; Singh 2011), and two compared the two levels by resistance (strength or effort) (Mangione 1999; Jan 2008; Foroughi 2011).

The duration of programs ranged from eight to 24 weeks. The mean frequency was three sessions per week (range two to five per week).

See Characteristics of included studies for a description of the components of the exercise programs.

Three studies examined exercise programs with a single component: walking (Ng 2010), set of muscle repetitions (Foroughi 2011), or cycling (Mangione 1999), and three studies examined an exercise program with multiple components (Jan 2008; McCarthy 2004; Singh 2011). Most exercise programs were supervised. Two were partially or completely unsupervised, with an exercise program executed at home (McCarthy 2004; Ng 2010). The intervention was delivered by an experienced or trained exercise therapist in four studies and not clearly reported in two studies (Mangione 1999; Singh 2011).

Adherence to training interventions
Adherence to the interventions was defined in terms of (1) attendance at an appointment and (2) compliance with the training advice or the content of the sessions provided by the healthcare professional (Brazzelli 2011).

We were not able to perform an analysis on attendance. In most of the studies, attendance or compliance was not clearly reported. The included studies did not systematically report compliance. Ng 2010 reported greater compliance with walking programs performed for three versus five days (100% versus 58% to 100%).

Outcomes
Only a limited number of studies reported the same outcomes prespecified in the protocols (see Additional tables). The six studies considered the end of the intervention as the final data collection point (range eight to 24 weeks).

Primary outcomes (Table 1): all six studies reported pain outcomes. Four studies used the WOMAC pain scale (McCarthy 2004; Jan 2008; Ng 2010; Foroughi 2011), one used a visual analog scale (VAS) (Singh 2011), and one used the Arthritis Impact Measurement Scale 2 (AIM2) subscale to assess pain intensity (Mangione 1999). Five studies assessed physical function, all used the WOMAC disability subscores (McCarthy 2004; Jan 2008; Ng 2010; Foroughi 2011; Singh 2011). Only one study reported quality of life, by the Short Form 36 and the EuroQol consisting of a 200 mm vertical VAS (McCarthy 2004).

Secondary outcomes (Table 2): included studies reported multiple secondary outcomes but only few studies used the same outcome measures. Two studies reported gait speed, for a fast pace, in Mangione 1999 and Jan 2008, and normal pace, in Mangione 1999 and Foroughi 2011. Three studies reported muscle strength for knee extensor muscles (McCarthy 2004; Jan 2008; Foroughi 2011), two studies knee flexor muscles (Jan 2008; Foroughi 2011), and two studies global strength (Foroughi 2011; Singh 2011). One study examined aerobic capacity (Mangione 1999). Two studies examined range of motion (McCarthy 2004; Singh 2011). One study examined physical activity and number of steps (Ng 2010), and data could not be extracted.

Follow-up assessment: two studies provided follow-up assessments at mid-term (six to 16 weeks after the end of the interventions) (McCarthy 2004; Ng 2010), and one study at long-term (40 weeks after the end of the interventions) (McCarthy 2004).
Safety

- Severe adverse events or withdrawal due to adverse events: three studies reported dropouts or adverse events. Foroughi 2011 reported dropout events in the high- and low-intensity exercise program groups. Mangione 1999 reported adverse events but did not specify in which group they occurred. Ng 2010 reported dropout events in both groups for health reasons. McCarthy 2004, Jan 2008, and Singh 2011 did not report the presence or absence of adverse events.

- Adverse effects: only one study systematically reported adverse effects related to the exercise programs (Foroughi 2011), and three additional studies made specific reference to the presence of adverse effects (McCarthy 2004; Jan 2008; Ng 2010) (see Characteristics of included studies).

Excluded studies

After screening of citations and abstracts, we excluded 12 reports on examining the full text. We based exclusions on unmet criteria related to (1) no randomization assignment (1 study) and (2) no clear difference in intensity in exercise programs between groups (11 studies).

Ongoing studies

See Characteristics of ongoing studies

We identified four ongoing studies registered in WHO ICTRP as potentially eligible for inclusion, but no findings were available yet. Three studies are comparing high versus low strength training on knee osteoarthritis: in the United States (Messier 2011), in Sweden (Ång 2013) and in Singapore (Pua 2012). One study is examining the effects of high versus low exercise program on hip osteoarthritis in Norway (Østerås 2012).

Risk of bias in included studies

Results of the 'Risk of bias' assessment are in Characteristics of included studies and Figure 2. Figure 2 provides a summary of the judgments of each methodological quality item for each study.
Figure 2. Risk of bias summary: review authors’ judgements about each risk of bias item for each included study.

<table>
<thead>
<tr>
<th></th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias): Subjective</th>
<th>Blinding of outcome assessment (detection bias): Objective</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foroughi 2011</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>?</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Jan 2008</td>
<td>+</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>
Allocation
We considered allocation sequence generation adequate in five studies (Mangione 1999; McCarthy 2004; Jan 2008; Ng 2010; Foroughi 2011), and unclear in one study (Singh 2011). Four of these studies used a computer-generated list for sequence generation and one study a random table (Jan 2008). Although one study used a central allocation to conceal allocation from the investigator enrolling participants and was rated as low risk (McCarthy 2004), four studies were rated as at unclear risk of bias because they provided insufficient information to determine allocation methods. We rated one study as at high risk of bias because the treatment allocation was not concealed (Ng 2010).

Blinding
In exercise studies, participants and care providers are rarely blinded from treatment allocation. In five studies, we rated performance bias at high risk (Mangione 1999; McCarthy 2004; Jan 2008; Ng 2010; Singh 2011). We rated one study as at low risk of bias because the participants were blinded to the investigators’ hypothesis (Foroughi 2011). For detection bias, we separated the assessment of blinded assessors for subjective and objective outcomes. As five studies reported that participants were not blinded, we rated these as at high risk of bias; we rated one study as at low risk of bias (Foroughi 2011). For the objective outcomes, we rated the studies as at low (McCarthy 2004; Jan 2008; Foroughi 2011), high (Ng 2010; Singh 2011), and unclear risk of bias (Mangione 1999).

Incomplete outcome data
We rated five studies as at high risk of bias. Five studies featured a median of 16% drop-out after inclusion (range 4% to 28%). Two studies reported an intention-to-treat (ITT) analysis (McCarthy 2004; Jan 2008). Only one study mentioned a method of imputation (last observation carried forward) to replace the missing data (McCarthy 2004). We considered this method inappropriate and might introduce a bias in estimating the treatment effect. The three other studies did not report an ITT analysis and showed a statistically significant rate (greater than 10%) of dropouts (Mangione 1999; Ng 2010; Foroughi 2011). The last study reported no information on dropouts (Singh 2011).

The attrition rate at the end of exercise programs ranged from 0% to 32%. Overall, the proportion was larger with high- versus low-intensity programs, except in one trial in which the attrition was lower in low- versus high-intensity programs (32% versus 12%) (Ng 2010).

Selective reporting
We found the research protocol for two studies (Ng 2010; Foroughi 2011). We assessed four trials as at unclear risk of bias because they did not distinguish between primary and secondary outcomes or outcomes data was not reported in a valid format, or both (Mangione 1999; Jan 2008; Foroughi 2011; Singh 2011). We assessed the remaining two studies as at high risk of bias because additional outcomes data were reported in separate reports or were not reported in the final publication, or both (McCarthy 2004; Ng 2010). No studies were assessed as at low risk of bias.

Other potential sources of bias
The studies appeared to be free of other serious potential sources of bias. Only one of the included studies reported differences in compliance between the high- and low-intensity interventions (Ng 2010). The compliance was lower for high-intensity programs, which could have biased nal results. We noted no difference in baseline participant characteristics. Of the six included studies, three did not report a source of funding (Jan 2008; Foroughi 2011; Singh 2011). Ng 2010 declared that Sanofi-Aventis Consumer Health Care, a pharmaceutical company, supplied the study glucosamine intakes but did not report if it has another role in the trial. McCarthy 2004 and Mangione 1999 reported that their trials were funded by the National Institute for Health Research and the Arthritis Foundation, respectively.

Effects of interventions
See: Summary of findings for the main comparison Summary of findings table: Physical activity and exercise programs in osteoarthritis; Summary of findings 2 Subgroup analysis: Exercise duration versus resistance in people with knee or hip osteoarthritis All studies reported means and standard deviation data for the outcomes at baseline and at the end of the intervention.

Primary outcomes

Pain
Four studies assessed pain on the WOMAC pain subscale at the end of the exercise program (McCarthy 2004; Jan 2008; Ng 2010; Foroughi 2011). Pooled analysis (Figure 3; Analysis 1.1) revealed a statistically significant reduction (mean difference (MD) (fixed-effect model): -0.84, 95% confidence interval (CI) -1.63 to -0.04, I² = 0%, P = 0.04, low-quality evidence) on the 20-point WOMAC pain scale for the high- versus low-intensity exercise programs, an absolute reduction in pain of 4% (8% better to 0% worse), and
relative reduction of 13% (25% better to 0% worse) (Summary of findings for the main comparison). Between-study heterogeneity was negligible ($I^2 = 0\%$). One study with 200 participants found a statistically significant reduction (MD -1.7 cm, 95% CI -1.9 to -1.4) on a 10-cm VAS equivalent to a MD of -0.67 (95% CI -0.8 to -0.6) on a Likert scale (Singh 2011). One study of 39 participants found no statistically significant difference on the AIM2 pain subscale (MD -0.11, 95% CI -1.3 to 1.1) (Mangione 1999).

Figure 3. Forest plot of comparison: 1 High- versus low-intensity exercise, outcome: 1.1 Pain (WOMAC).

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>High Intensity</th>
<th>Low Intensity</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>1.1.1 at study completion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>7.5</td>
<td>3.9</td>
<td>93</td>
<td>9.84</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>3.8</td>
<td>2.7</td>
<td>20</td>
<td>4.4</td>
</tr>
<tr>
<td>Ng 2013</td>
<td>5.36</td>
<td>4.24</td>
<td>15</td>
<td>5.60</td>
</tr>
<tr>
<td>Jun 2009</td>
<td>4.3</td>
<td>3.5</td>
<td>34</td>
<td>4.8</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>162</td>
<td>151</td>
<td>100.0%</td>
<td>-0.34 [1.63, -0.04]</td>
</tr>
</tbody>
</table>

Two studies including 199 participants assessed pain with a high-versus low-intensity exercise program at the mid-term (six to 16 weeks) (McCarthy 2004; Ng 2010), and one study included 139 participants at the long-term (40 weeks) after the end of the intervention (McCarthy 2004): there was difference at the mid-term (MD -0.82, 95% CI -1.90 to 0.26, $I^2 = 41\%$, $P = 0.14$) and an effect at the long-term (MD -1.30, 95% CI -2.56 to -0.04, $P = 0.03$). For these two studies, the programs varied in the amount of time spent in exercise (total program duration or session) between the groups.

Physical function

Five studies evaluated the effect of high- versus low-intensity exercise programs on physical function. At the end of the exercise program, four studies including 313 participants reported a statistically significant mean reduction on the 68-point WOMAC disability subscale (fixed-effect model) of -2.65, 95% CI -5.29 to -0.01, $I^2 = 0\%$, $P = 0.05$, low-quality evidence, for high-versus low-intensity exercise (McCarthy 2004; Jan 2008; Ng 2010; Foroughi 2011) (Figure 4; Analysis 1.2), an absolute reduction in function of 4% (8% better to 0% worse) and relative reduction of 13% (26% better to 0% worse) (Summary of findings for the main comparison). Between-study heterogeneity was negligible ($I^2 = 0\%$). We did not include one study of 200 participants in the pooled analysis because we identified that the results were discordant and inconsistent (score ranged between 0 and 4) with the findings of the other four studies (Singh 2011). We requested of study authors how they scored the WOMAC function test. We received no additional information despite two requests (Appendix 7). Singh 2011 found a statistically significant difference between high- and low-intensity programs for physical function (MD -0.51, 95% CI -0.6 to -0.4; lower score favoring high-intensity exercise).
exercise), which is unlikely to be of clinical significance.

Figure 4. Forest plot of comparison: 1 High- versus low-intensity exercise, outcome: 1.2 Physical Function (WOMAC).

Based upon data from two studies, we observed no statistical difference in physical function at mid-term, McCarthy 2004 and Ng 2010, and long-term, McCarthy 2004, between high- and low-intensity exercise programs (Analysis 1.2).

Quality of life
One study (McCarthy 2004) of 214 participants found no statistically significant difference between high- and low-intensity exercise programs on quality of life measured by the EuroQol at the end of the intervention (MD 4.3, 95% CI -6.5 to 15.2, very low-quality evidence), an absolute reduction in quality of life of 2% (8% better to -3% worse) and relative reduction of 6% (10% better to -23% worse) and in the mid-term (MD -2.9%, CI -16.3 to 12.9) and long-term (MD 0.9, 95% CI -13.0 to 14.7) (Summary of findings for the main comparison).

Safety

Serious adverse events or withdrawals due to adverse events
Three studies reported adverse events (Mangione 1999; Ng 2010; Foroughi 2011). Due to the limited number of events and lack of information, we decided not to pool the data. Ng 2010 and Foroughi 2011 reported the same number of dropouts or withdrawals between high- (n = 3) and low-intensity exercise groups (n = 3). Dropouts were due to medical reasons. Mangione 1999 reported two adverse events (fall during the warm-up session, hit the shin with the cycling pedal) without specifying the group in which events occurred. No severe adverse events were reported.

Adverse effects
Four studies including 364 participants provided information on adverse effects related to exercise programs (Analysis 1.3). Only one study with 54 participants declared systematic monitoring (Foroughi 2011). Three studies reported some adverse effects in the high intense groups: Jan 2008 (three), Foroughi 2011 (two), Ng 2010 (two). Adverse effects were primarily related to knee pain associated with an exercise program.

In our analysis of all reported data, the number of effects was not statistically different between the high- versus low-intensity exercise programs (Peto odds ratio 1.72, 95% CI 0.51 to 5.81, I² = 22%, P = 0.39, very low-quality evidence) (Analysis 1.3).

Secondary outcomes
Gait speed

Three studies reported gait speed (Mangione 1999; Jan 2008; Foroughi 2011). In two studies, no statistically significant difference was found between the high- and low-intensity exercise programs in free walking speed (MD (random-effects model): 0.04, 95% CI -0.05 to 0.13, I² = 14%, P = 0.37) or fast walking speed in the short-term (one trial, MD 0.08, 95% CI -0.13 to 0.29, P = 0.45) (Analysis 1.4). Jan 2008 reported data that we could not transform for the analysis. Between-study heterogeneity was negligible (I² = 14%).

Muscle strength

Many different muscle groups were tested, with heterogeneity in number of methods used to evaluate muscle strength. We then calculated the effect size using standardized mean difference (SMD) to allow for pooling data resulting from different units of measurement. Three studies of 285 participants reported muscle strength on the knee extensor (McCarthy 2004; Jan 2008; Foroughi 2011), and showed no statistically significant difference between high- and low-intensity exercise programs immediately after treatment (SMD (random-effects model): 0.38, 95% CI 0.04 to 0.72, I² = 42%, P = 0.03) (Analysis 1.5), although with substantial statistical heterogeneity (I² = 42%). Two studies of 113 participants found no statistically significant difference (SMD (random-effects model): 0.18, 95% CI -0.64 to 1.00) on strength of knee flexor muscles immediately after treatment (Jan 2008; Foroughi 2011) (Analysis 1.5), with large statistical heterogeneity (I² = 78%). Two studies of 245 participants measured global strength (Foroughi 2011; Singh 2011), and showed an improvement with high-versus low-intensity exercise programs after treatment (SMD 1.01, 95% CI 0.74 to 1.27, I² = 0%, P = 0.001). Between-study heterogeneity was negligible (I² = 0%).

Aerobic capacity

One study of 39 participants found no statistically significant difference on aerobic capacity between the high- and low-intensity exercise programs at the end of treatment (MD -1.40, 95% CI -4.2 to 1.4) (Mangione 1999) (Analysis 1.6). We considered this study as high risk of bias because of the lack of blinding and high attrition (30% in the two groups).

Range of motion

Two studies, of 190 and 200 participants, reported statistically significant effects of high-versus low-intensity exercise programs on range of motion (McCarthy 2004; Singh 2011). As the data showed signs of heterogeneity (I² = 92%), we did not pool the data (Analysis 1.7).

Physical activity level and number of steps

Only one study reported the two outcomes (Ng 2010). Data were incomplete. We performed no analysis for these two outcomes.

Subgroup analysis

We conducted subgroup analyses (Figure 5, Figure 6) to explore possible differences in pooled MDs for immediate post-treatment on pain and physical function outcomes according to the type of intensity of exercise programs (time and level of resistance).

Figure 5. Forest plot of comparison: 2 Subgroup analysis: Exercise duration versus resistance, outcome: 2.1 WOMAC Pain.
Figure 6. Forest plot of comparison: 2 Subgroup analysis: Exercise duration versus resistance, outcome: 2.2 Physical Function.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Fixed, 95% CI</td>
<td>IV, Fixed, 95% CI</td>
<td>IV, Fixed, 95% CI</td>
<td>IV, Fixed, 95% CI</td>
</tr>
<tr>
<td>2.2.1 Duration exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>-1.37 (-2.47 to -0.28)</td>
<td>-1.37 (-2.47 to -0.28)</td>
<td>-1.37 (-2.47 to -0.28)</td>
<td>-1.37 (-2.47 to -0.28)</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>-1.23 (-2.23 to -0.23)</td>
<td>-1.23 (-2.23 to -0.23)</td>
<td>-1.23 (-2.23 to -0.23)</td>
<td>-1.23 (-2.23 to -0.23)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>-1.37 (-2.47 to -0.28)</td>
<td>-1.37 (-2.47 to -0.28)</td>
<td>-1.37 (-2.47 to -0.28)</td>
<td>-1.37 (-2.47 to -0.28)</td>
</tr>
<tr>
<td>Heterogeneity (Chi²)</td>
<td>0.00 (df = 1)</td>
<td>0.00 (df = 1)</td>
<td>0.00 (df = 1)</td>
<td>0.00 (df = 1)</td>
</tr>
<tr>
<td>Test for overall effect Z</td>
<td>0.65 (P = 0.52)</td>
<td>0.65 (P = 0.52)</td>
<td>0.65 (P = 0.52)</td>
<td>0.65 (P = 0.52)</td>
</tr>
</tbody>
</table>

2.2.2 Resistance exercise

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Fixed, 95% CI</td>
<td>IV, Fixed, 95% CI</td>
<td>IV, Fixed, 95% CI</td>
<td>IV, Fixed, 95% CI</td>
</tr>
<tr>
<td>Jan 2000</td>
<td>-1.57 (-5.06 to 1.93)</td>
<td>-1.57 (-5.06 to 1.93)</td>
<td>-1.57 (-5.06 to 1.93)</td>
<td>-1.57 (-5.06 to 1.93)</td>
</tr>
<tr>
<td>Farnight 2011</td>
<td>-1.89 (-5.06 to 1.32)</td>
<td>-1.89 (-5.06 to 1.32)</td>
<td>-1.89 (-5.06 to 1.32)</td>
<td>-1.89 (-5.06 to 1.32)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>-1.57 (-5.06 to 1.93)</td>
<td>-1.57 (-5.06 to 1.93)</td>
<td>-1.57 (-5.06 to 1.93)</td>
<td>-1.57 (-5.06 to 1.93)</td>
</tr>
<tr>
<td>Heterogeneity (Chi²)</td>
<td>0.22 (df = 1)</td>
<td>0.22 (df = 1)</td>
<td>0.22 (df = 1)</td>
<td>0.22 (df = 1)</td>
</tr>
<tr>
<td>Test for overall effect Z</td>
<td>0.38 (P = 0.71)</td>
<td>0.38 (P = 0.71)</td>
<td>0.38 (P = 0.71)</td>
<td>0.38 (P = 0.71)</td>
</tr>
</tbody>
</table>

Pain

Subgroup analysis (Analysis 2.1) showed a statistical significant difference (MD -1.37, 95% CI -2.47 to -0.28, I² = 0%, P = 0.01, absolute mean reduction 7%, very low-quality evidence). With different amounts of time spent in exercise programs, pain on the 20-point WOMAC pain scale was lower with higher than with lower intensity exercise. This finding did not seem to be of clinical significance. No statistical significant differences was found for amount of resistance (MD -0.23, 95% CI -1.40 to 0.93, I² = 0%, P = 0.71, absolute mean reduction 1%, very low-quality evidence) (Summary of findings 2).

Physical Function

Subgroup analysis (Analysis 2.2) showed a statistically significant difference (MD -1.0.10, 95% CI -8.12 to -0.07, I² = 0%, P = 0.05, absolute mean reduction 6%, very low-quality evidence) on the 68-point WOMAC disability subscale.

This finding did not seem to be of clinical significance. High- and low-intensity exercise programs did not differ statistically in the effect of resistance (MD -1.57, 95% CI -5.06 to 1.93, I² = 33%, P = 0.38, absolute mean reduction 2%, very low-quality evidence).

Adverse events

No statistical difference was found between the subgroup exercise programs (Analysis 2.3).

We did not conduct other subgroup analyses as described in the protocol, as we found that data were insufficient.

Sensitivity analysis

The results were robust on excluding one trial, Ng 2010, that included participants with hip and knee osteoarthritis (results not shown).
### ADDITIONAL SUMMARY OF FINDINGS

#### Explanation

Subgroup analysis: Exercise duration versus resistance in people with knee or hip osteoarthritis

**Patient or population:** People with knee or hip osteoarthritis  
**Settings:** Hospital or primary care  
**Intervention:** Subgroup analysis: exercise type of intensity (duration and resistance)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed risk</td>
<td>Corresponding risk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain - duration exercise</td>
<td>The mean pain - duration exercise in the control groups was <strong>8.6 points</strong></td>
<td>The mean pain - duration exercise in the intervention groups was <strong>1.37 lower</strong> (2.47 to 0.28 lower)</td>
<td>200 (2 studies)</td>
<td>⊕⊕⊕⊕ very low 1,2</td>
<td>MD -1.37 (95% CI -2.47 to -0.29). Absolute mean reduction 7% with high intensity exercise programs (95% CI -12% more to 1% fewer). Relative mean reduction 16% (95% CI -29% more to 3% fewer). NNTB = 11 (95% CI 9 to 14) 3,4</td>
</tr>
<tr>
<td>WOMAC (VAS or Likert) from 0 to 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up: 8 to 12 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lower scores mean reduced pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain - resistance exercise</td>
<td>The mean pain - resistance exercise in the control groups was <strong>4.6 points</strong></td>
<td>The mean pain - resistance exercise in the intervention groups was <strong>0.23 lower</strong> (1.4 lower to 0.93 lower)</td>
<td>113 (2 studies)</td>
<td>⊕⊕⊕⊕ very low 1,2</td>
<td>MD -0.23 (-1.40 to 0.93). Absolute mean reduction 1% with high intensity exercise programs (95% CI -7% more to 5% fewer). Relative mean reduction 5% (95% CI -30% more to 20% fewer). NNTB = 17 (95% CI 13 to 22) 5,6</td>
</tr>
<tr>
<td>WOMAC (VAS or Likert) from 0 to 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up: 8 to 24 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lower scores mean reduced pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Function - duration exercise</td>
<td>WOMAC from 0 to 68</td>
<td>Follow-up: 8 to 12 weeks</td>
<td>lower scores mean better function</td>
<td>The mean function - duration exercise in the control groups was <strong>27 points</strong></td>
<td>The mean function - duration exercise in the intervention groups was <strong>4.10 lower</strong> (8.12 to 0.07 lower)</td>
</tr>
<tr>
<td>Function - resistance exercise</td>
<td>WOMAC from 0 to 68</td>
<td>Follow-up: 8 to 24 weeks</td>
<td>lower scores mean better function</td>
<td>The mean function - resistance exercise in the control groups was <strong>16.3 points</strong></td>
<td>The mean function - resistance exercise in the intervention groups was <strong>1.57 lower</strong> (5.06 to 1.93 lower)</td>
</tr>
</tbody>
</table>

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; MD: mean difference; NA: not applicable; NNTB: number needed to treat for an additional beneficial outcome; NNTH: number needed to treat for an additional harmful outcome; NRS: numeric rating scale; SD: standard deviation; VAS: visual analog scale; WOMAC: Western Ontario and McMaster Universities Arthritis Index

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

¹ Downgraded because of lack of blinding. No intention-to-treat analysis. Incomplete outcome data.
² Downgraded because of imprecision (large confidence interval with small sample size).
³ Calculations based on the control group baseline mean (SD) pain: 10.0 (3.7) points on 0-20 scale (from McCarthy 2004).
⁴ NNT for continuous outcomes calculated using the Wells calculator (from the CMSG Editorial office; http://musculoskeletal.cochrane.org/), and for dichotomous outcomes using the Cates NNT calculator (www.nntonline.net/visualrx/).
⁵ Calculations based on the control group baseline mean (SD) pain: 7.8.0 (3.3) points on 0-20 scale (from Jan 2008).
Calculations based on the control group baseline mean (SD) function: 30.8 (14.4) points on 0-68 scale (from McCarthy 2004).

Calculations based on the control group baseline mean (SD) function: 26.1 (8.1) points on 0-68 scale (from Jan 2008).
DISCUSSION

Summary of main results

The main purpose of this review was to evaluate the effect of high-versus low-intensity physical activity or exercise programs on reducing pain and improving physical function and quality of life for people with hip or knee osteoarthritis. We characterized the intensity of an activity or an exercise program either by the overall amount of training time (duration, number of sessions) or the amount of resistance (strength or effort). We included six studies involving 656 participants. Based upon low-quality evidence, we found a small improvement in pain and function of high-intensity exercise programs in people with knee osteoarthritis, but this is unlikely to be of clinical importance. We are uncertain of the effect on quality of life. The pooled analysis showed small statistical effect sizes for pain (0.84 points) on a 20-point WOMAC scale and physical function (2.65 points) on a 68-point WOMAC scale. The minimal clinically important difference values we used for the WOMAC measures were 15% for absolute improvement and 20% for relative improvement (Tubach 2012). The statistically significant differences we observed between high- and low-intensity exercise were much smaller, and therefore unlikely to be of clinical significance. In addition, the effects of high-intensity exercise were found only in the short-term after treatment. Based upon very low-quality evidence, we found a statistically significant difference in subgroup analyses depending on the type of intensity of exercise programs (time and level of resistance). However, these findings were unlikely to be of clinical importance.

In most of the included trials adverse events were poorly monitored and poorly reported. We are uncertain as to whether higher-intensity exercise programs may induce more harmful effects than lower-intensity programs. We downgraded the evidence from high to low or very low because of the risk of bias, imprecision (small number of studies or participants, large confidence interval, small effects), and inconsistency in the reporting of adverse effects.

Overall completeness and applicability of evidence

Our evidence is limited to exercise programs for participants with knee osteoarthritis. We did not perform a meta-analysis of physical activity interventions or participants with hip osteoarthritis because of the lack of data. Generalization of our findings to other populations should be limited. Most participants in the included trials were women (greater than 70%), confirming the gender differences noted in the prevalence of knee osteoarthritis (O’Connor 2007). However, the extracted data did not allow us to consider gender-related differences in the evaluation of the effectiveness of exercise programs. We were also unable to determine whether knee osteoarthritis severity plays a role. More research is needed into the potential impact of disease severity on the effects of exercise programs.

All included studies assessed a variety of interventions, with different levels of intensity. The type of exercise programs differed among the six studies. None of the studies compared the same exercise interventions. Programs included walking (Ng 2010), cycling (Mangione 1999), global muscle strengthening (McCarthy 2004), dynamic resistance training (Jan 2008; Foroughi 2011), and isometric resistance training (Singh 2011). We could not determine the optimal type of exercise. The studies assessed a variety of outcome measures relating to pain, physical function, quality of life, and physical performance. The outcomes measured often differed between studies. For pain (4 studies) and physical function (5 studies), the WOMAC scales were most commonly used (see Table 1). The same outcome (that is WOMAC pain or disability subscore) was sometimes available in different units (Likert, visual analog scale, or numeric rating scale) or different scales, and we had to rescale the data to pool the measures in the analysis.

Apart from pain and physical function, quality of life and adverse events were the main outcomes assessed in the studies. However, only one study specified quality of life as an outcome, and few studies (n = 3) reported information on any adverse events. In addition, only one study reported the secondary outcomes we examined. Finally, the results were limited to short-term effects because a limited number of studies reported follow-up assessments for longer times.

Quality of the evidence

We found limitations in the included studies inherent to study design (lack of blinding, incomplete data reporting, no intention-to-treat (ITT) analysis) and imprecision (inadequate and small sample sizes, and small number of studies). We found the overall quality of evidence to be low for pain and physical function to very low for adverse effects for the primary outcomes using the GRADE approach. Most of the studies had an unclear or high risk of bias for several domains. We downgraded most studies (n = 5) to high risk of bias for performance, detection, and attrition bias because of lack of blinding and incomplete outcome data, which can affect the quality of the randomization.

Given that the primary outcomes of this review were participant self-reported pain and physical function, the treatment effect sizes may be inflated. We assessed potential bias due to incomplete outcome data as at high risk of bias because of the lack of ITT analysis and high level of attrition (see Characteristics of included studies). Although sample size does not contribute to the assessment of study risk of bias, most of the included six studies were underpowered to provide clear answers, and we downgraded evidence because of imprecision. Three studies had small sample sizes (less than 100 participants), and two did not report a prospective sample size calculation. Due to the small sample sizes, we are unclear
whether the recruited participants represent all people with knee osteoarthritis. The intervention may benefit only a subset of people (that is selection bias).

Despite performing an extensive search, we included a limited number of studies (n = 6). We were unable to estimate the impact of publication bias on our results.

**Potential biases in the review process**

The studies included in this review form the best available evidence for the review question. We conducted an extensive search of the literature in all relevant databases and included six studies of interventions for exercise programs in participants with knee osteoarthritis. We also identified two ongoing trials for potential additional data. We made all attempts to reduce the bias involved with the review process. For missing data, we made attempts to extract data that were graphically displayed using software tools (that is http://arohatgi.info/WebPlotDigitizer/index.html) or to gather the information from authors of the included studies.

**Agreements and disagreements with other studies or reviews**

Systematic reviews, Wang 2012, Uthman 2013, and Juhl 2014, and recommendations, Fernandez 2013, investigating the effects of exercise programs for knee osteoarthritis have been published, and have reported that exercise programs are effective for reducing pain and improving functional outcome in people with knee osteoarthritis. However, we lack information regarding the optimal intensity of exercise for people with knee or hip osteoarthritis. Only a few studies have directly compared different intensities of exercise programs, and the type of exercise varied widely, so specifying the required dose for optimal benefit for symptoms of lower-limb osteoarthritis is difficult.

Previous systematic reviews of exercise suggested a relationship between the time spent in an exercise program and the size of outcomes for various health conditions such as stroke, in Veerbeek 2014, or back pain, in Ferreira 2010. A similar (positive dose-response) relationship between the time spent in the exercise program (number of supervised exercise sessions and session duration) and the effect size for pain was reported by Juhl 2014 when considering only aerobic exercise in the meta-regression analysis. However, this evidence was obtained from indirect comparisons, with a global effect size computed by meta-regression (EUnetHTA 2013). All randomized controlled trials we included compared exercise programs for which participants were randomized to a high- or low-intensity group (direct comparisons) within each study. This situation is reinforced by the results found by indirect comparisons and adds confidence to the findings.

In a meta-analysis considering only the exercise programs based on resistance, Juhl 2014 found no statistically significant regression coefficients on meta-regression analysis of intensity resistance or effort, but equally by length of exercise program, number of supervised sessions, duration of individual supervised sessions, or number of sessions per week as covariates. This absence of the effect of high-intensity strength training of the lower limb was also reported in a study of older adults (Raymond 2013). In our subgroup analysis, we also found no statistically significant difference between level of intensity with resistance exercises with the same length of program. Resistance exercises seem to have a positive impact on muscle strength improvements in healthy elderly, in Fiararone Singh 2004 and Porter 2006, or rheumatoid patient population, in Baillet 2012. Furthermore, there is a large body of evidence for many years showing changes in strength and muscle composition in relation to different intensities (Beijersbergen 2013).

Strengthening may also be effective in people with knee osteoarthritis, but we lack data favoring a greater effect of exercise with high versus low resistance. The intensity of exercises for both groups may not have been separated enough to differentiate between the loadings achieved by the study groups. Future studies should investigate a specific exercise program designed for a sham-exercise group, whereby the joint receives low loading. Adverse events have been reported during exercising (Chilibeck 2011). Recently, some data have suggested that physical exercise may induce adverse effects in knee joints of animals (Siebels 2014). High-resistance exercise should perhaps be used with caution, particularly for people with depleted cartilage or severity stage of osteoarthritis. The lack of evidence of the beneficial effect of high resistance suggests a preference for moderate resistance in strengthening exercises for people with knee osteoarthritis, but with a long-duration program.

**Authors’ Conclusions**

**Implications for practice**

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. We did not find important clinical difference in subgroup analyses depending on the type of intensity of exercise programs (time and level of resistance).

We are unable to make a conclusion about the effects for quality of life between high- and low-intensity exercise programs. Adverse events related to exercise were minor, but they were poorly recorded. It is uncertain if higher-intensity exercise programs may induce more harmful effects than lower-intensity programs; this must be evaluated by further studies.

Our review highlights the need for better reporting of exercise programs in clinical trials with explicit descriptions that enable...
replication (Slade 2012). We could not investigate the effect of other exercise characteristics (for example delivery modalities such as supervision or mono- or multi-modal exercise). Studies directly comparing prescriptive elements (modality, intensity, duration of exercise) are critically needed to advance this field.

Implications for research
The small number of studies comparing high- and low-intensity exercise programs in osteoarthritis underscores the need for more studies to investigate the dose-response relationship. More research is required to further study the potential impact of clinical characteristics of participants in terms of gender and disease severity on the effects of an exercise program.

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. We could not ensure that exercise of sufficient intensity was performed because exercises performed at high resistance failed to increase muscle strength of knee extensor or knee flexor, in Jan 2008 and Foroughi 2011, or aerobic capacity, in Mangione 1999. The difference (small or large) between the high- and low-intensity programs varied across studies. The differences may be insufficient to induce any effects, particularly for resistance exercise programs. In particular, further studies are needed to establish the minimal effective intensity of programs required to produce a clinical effect and the highest intensity that patients can tolerate. Dose selection should be better justified, especially when the suboptimal dose is unknown.

Clinical studies in the field of rehabilitation should be carefully designed and follow the most recent Consolidated Standards of Reporting Trials (CONSORT) guidelines to report their data. Most studies included in this review did not use ITT analysis, concealed allocation, or blinding of assessors or did not report data on common outcome measures. In addition, specific designs can also be considered to address some limitations that traditional randomized controlled trials face, such as the small number of participants in clinical trials of rehabilitation (Graham 2012).

ACKNOWLEDGEMENTS
The review authors are grateful to Louise Falzon, the Trials Search Co-ordinator for the Cochrane Musculoskeletal Group, for her assistance in designing the search strategy, and to Pr Serge Poiraud and for his comments. We thank Dr Elizabeth Tanjong Ghogomu and Lara Maxwell, Cochrane Musculoskeletal Review Group, for their valuable comments and expert advice. We also thank Dr Ann Moseley for help searching studies.

REFERENCES

References to studies included in this review

Foroughi 2011 [published and unpublished data]

Jan 2008 [published data only]

Mangione 1999 [published data only]

McCarthy 2004 [published data only]

Ng 2010 [published and unpublished data]

Singh 2011 [published data only]
Singh J. Effects of Exercise Rehabilitation Programme on Osteoarthritic Knee With Special Reference to Biochemical Changes (Thesis submitted to the faculty of the Punjabi University). 2010.
Singh J, Singh P, Sohal MS. Effect of exercise rehabilitation...

References to studies excluded from this review

Chang 2012 *published data only*

Diracoglu 2005 *published data only*

Eyigor 2004 *published data only*

Fitzgerald 2011 *published and unpublished data*

Green 1993 *published data only*

Kruger 1997 *published data only*

Sevick 2000 *published data only*

Shakoor 2007 *published data only*

Teixeira 2011 *published data only*

Topp 2002 *published data only*

Veenhof 2006 *published data only*

References to studies awaiting assessment

Steinhilber 2012 *published data only*

References to ongoing studies

Ång 2013 *unpublished data only*

Messier 2011 *published data only*
Strength Training for ARthritis Trial (START). Ongoing study March 2012.

Pua 2012 *published data only*
Eccentric resistance ergometry in knee osteoarthritis rehabilitation: a randomized clinical trial (EUREKA). Ongoing study April 2012.

Österäs 2012 *unpublished data only*

Additional references

Ald 2011

Altman 1996

Baillet 2012
High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis (Review)

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Beijersbergen 2013

Bijlsma 2011

Bouchard 2007

Boutron 2008

Brazzelli 2011

Brosseau 2003

Brosseau 2004

Cates 2008

Chilibeck 2011

Cohen 1988

Deeks 2011

EUnetHTA 2013

Fernandes 2013

Ferreira 2010

Fiatarone Singh 2004

Fransen 2008a

Fransen 2008b

Galvao 2005

Graham 2012

Hardy 1996

Heiwe 2011

Higgins 2002

Higgins 2011
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0
High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis (Review)

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hochberg 2012

Hunter 2011

Ioannidis 2007a

Juhl 2014

Jüni 2006

Kraus 2002

Kwakkel 2004

Latham 2010

McAlindon 2014

Moskowitz 2009

Murray 2012

O’Connor 2007

Peters 2008

Pham 2004

Porter 2006

Raymond 2013

Reichenbach 2007

RevMan 2011

Robbins 2012

Rucker 2008

Schünemann 2011a
Schünemann 2011b

Siebelt 2014

Slade 2012

Tubach 2012

Umpierre 2011

Uthman 2013

Veerbeek 2014

Vignon 2006

Vos 2012

Wang 2012

Woolf 2003

World Health Organization 2010

Zhang 2010

* Indicates the major publication for the study.
## Characteristics of included studies [ordered by study ID]

### Foroughi 2011

<table>
<thead>
<tr>
<th>Methods</th>
<th>Randomized Controlled Trial (RCT) with 2 groups</th>
</tr>
</thead>
</table>
| Participants | Location: Australia  
Randomized: 54  
Age: mean 65.5 years  
Sample: 54 women with knee primary osteoarthritis  
Settings: academic, monocenter  
Inclusion criteria: women > 40 years old with primary osteoarthritis of at least 1 knee, following ACR criteria  
Exclusion criteria: secondary osteoarthritis, joint injury, injection or surgery within the past 6 months or joint replacement, already participated in structured exercise, contraindications to exercise and/or functional magnetic resonance imagery (fMRI), severe functional limitation or cognitive impairment |
| Interventions | Intensity: Exercise program with different levels of strength resistance  
High resistance (n = 26): frequency: 3 times/week * 24 weeks; session duration: 30 min; intensity: trained at 80% of their peak muscle strength with 8 repetitions * 3 sets  
Progression: 3% increments in resistance per session as tolerated  
Low resistance (n = 28): frequency: 3 times/week * 24 weeks; session duration: 30 min; intensity: minimal resistance was set and no progression with 8 repetitions * 2 sets  
Supervision: an experienced physiologist |
| Outcomes | At 24 weeks:  
WOMAC (pain, physical function, global), adverse events, walking speed, muscle strength (unilateral knee extension, bilateral knee flexion, leg press)  
A Likert scale (range 0-20 or 0-68) in WOMAC |
| Notes | • Dropouts: 5 (9%)  
• Adherence: “Average training session compliance was with the resistance training group attending an average of 97% sessions and the sham-exercise group attending 99% sessions among those who completed their 6-month assessments”. One participant in the high resistance group discontinued the exercise intervention due to severe knee pain. Distinction between attendance and compliance is unclear.  
• Adverse events: Some adverse events were reported. Adverse effects related to exercise programs: “2 dropouts related to intervention (knee and/or back pain), 1 in the resistance group and one in the low resistance group. One subject discontinued training in the resistance group.”  
• Attrition: high resistance, n = 6 (23%), low resistance, n = 3 (11%)  
• Exercise under supervision provided by an exercise physiologist.  
• We identified 2 published reports for the same trial. All data were extracted from Foroughi* except walking speed, which is from Foroughi.  
• We contacted the author, who provided additional information.  
• No information on funding. |
### Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: “participants were randomly allocated using a computer randomization program.”</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>Quote: “A co-investigator uninvolved in participant testing or training managed the randomization procedure.”&lt;br&gt;Comment: It is unclear if the sequence of randomization was concealed from the investigator enrolling participants</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias) All outcomes</td>
<td>Unclear risk</td>
<td>Quote: “All participants were blinded to the investigator’s hypothesis.”&lt;br&gt;Comment: Staff were not blinded. It unclear how this influenced the data</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Subjective</td>
<td>Low risk</td>
<td>Comment: Performance was measured by a 3D motion acquisition system. It is unlikely that this provided risk of bias, even though the outcome assessor was not blinded</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Objective</td>
<td>Low risk</td>
<td>Comment: Performance was measured through a 3D motion acquisition system. It is unlikely that this provided risk of bias, even though the outcome assessor was not blinded</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias) All outcomes</td>
<td>High risk</td>
<td>Comment: Unbalanced drop-out rate, &gt;10%. No imputation method used. No ITT analysis, which might introduce a bias</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear risk</td>
<td>Comment: Trial registered (ACTRv12605000116628). All outcomes specified in the protocol were separately reported in the 2 publications. It is unclear how this could influence the data</td>
</tr>
<tr>
<td>Other bias</td>
<td>Low risk</td>
<td>Comment: There is no indication that there are other important risks of bias. Power sample size calculation provided</td>
</tr>
</tbody>
</table>
Methods | RCT with 3 groups
| ITT analysis

Participants | Location: Taiwan
| Randomized: 102
| Age: mean 62.6 years
| Sample: 79 females and 19 males with knee osteoarthritis
| Settings: university hospital, monocenter
| Inclusion criteria: > 50 years old, knee pain and osteophytes confirmed by radiography, osteoarthritis grade 3 on the Kellgren-Lawrence classification, history of pain > 6 months, no NSAIDs during the study
| Exclusion criteria: knee physical therapy 3 months prior, other problems with knee joint, neuropathy, unstable medical conditions

Interventions | Intensity: Exercise program with different levels of strength resistance
| High resistance (n = 34): frequency: 3 times/week * 8 weeks; session duration: 30 min; intensity: trained at 60% of 1RM (about 45 to 50 kg) with 8 repetitions * 3 sets
| Low resistance (n = 34): frequency: 3 times/week * 8 weeks; session duration: 50 min; intensity: trained at 10% 1RM (about 7 to 10 kg) with 15 repetitions * 10 sets
| Progression: every 2 weeks 1RM was retested and increased by 5% as tolerated in each group
| Control group (n = 34): received no exercise
| Supervision: an experienced therapist

Outcomes | At 8 weeks: WOMAC (pain, physical function), walking time (level ground, stairs, figure-eight pattern, spongy surface), muscle strength (flexion and extension of the knee; torque tested at 60, 120, and 180°/s). A Likert scale (range 0-20 or 0-68) for the WOMAC

Notes | • Dropouts: 3 (4%) in total
| • Adherence: “Although there was 100% adherence to the exercise intervention in the low-resistance group, 3 participants in high-resistance group discontinued the exercise intervention due to severe knee pain.”
| • Adverse events: Adverse effects related to exercise: “3 participants in HR group discontinued the exercise intervention due to severe knee pain.”
| • Attrition: high resistance, n = 3 (9%); low resistance, n = 0 (0%), control group n = 0 (0%)
| • No information on funding

Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: &quot;using a random numbers table from a random integer generator&quot;</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>Comment: Unclear if the list of randomization is concealed</td>
</tr>
</tbody>
</table>
### Mangione 1999

**Method**
- RCT with 2 groups

**Participants**
- Location: USA
- Randomized: 39
- Age: mean 71 years
- Sample: 26 females and 13 males with knee osteoarthritis
- Settings: academic, monocenter
- Inclusion criteria: community dwelling, ≥ 50 years, knee pain and positive radiography or clinical signs of osteoarthritis
- Exclusion criteria: rheumatoid disease, history of unstable metabolic diseases, neurological disorders or cardiopulmonary conditions that precluded participation in aerobic exercise
### Interventions

- **High effort (n = 19):** frequency: 3 times/week * 10 weeks; session duration: 25 min; intensity: cycling fixed at 70% HRR
- **Low effort (n = 20):** frequency: 3 times/week * 10 weeks; session duration: 25 min; intensity: cycling fixed at 40% HRR

All groups: Warm-up, then cycling exercises for 25-min training period on a stationary cycle and cool-down exercise. Each session lasted about 1 hr. Intensity determined from the maximum HR.

The authors did not provide any information regarding the progression or the supervision of sessions.

Using a Likert scale (range 0-20 or 0-68) for WOMAC.

### Outcomes

At 10 weeks: subscale pain, Arthritis Impact Measurement Scale 2, walking speed, 6-min walk test, peak oxygen consumption.

### Notes

- Dropouts: 15 (28%)
- Adherence: “92.2% and was nearly identical between the groups.”
- Adverse events: “Two adverse events occurred during the testing and training” and did not seem related to the group allocation.
- Attrition: overall 15 participants. Number per group is unknown.
- Funded by the Arthritis Foundation.

### Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Comment: Computer generated random numbers used</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>Comment: No procedure was described</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Comment: Insufficient information, but it is unlikely that participants and care providers were blinded</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High risk</td>
<td>Comment: Insufficient information, but it is unlikely that the participants were blinded</td>
</tr>
<tr>
<td>Subjective</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>Unclear risk</td>
<td>Comment: No information on blinding of outcome assessors was provided. Unclear how this influenced the data</td>
</tr>
<tr>
<td>Objective</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>High risk</td>
<td>Comment: No ITT analysis. High rate of drop-out (30%), which might introduce bias</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear risk</td>
<td>Comment: Insufficient information. No protocol registration. Major and minor outcomes not declared; unclear what they are</td>
</tr>
</tbody>
</table>
Other bias | Unclear risk | Comment: No sample size calculation provided. Baseline performance seems similar between the 2 groups

McCarty 2004

Methods | RCT with 2 groups

Participants | Location: UK  
Randomized: 225  
Age: mean 64.7 years  
Sample: 125 females and 89 males with knee osteoarthritis  
Settings: academic, monocenter  
Inclusion criteria: met the ACR clinical criteria (knee pain, knee radiograph, clinical features) and radiological evidence of osteophytes  
Exclusion criteria: symptomatic back or hip disease, knee osteoarthritis secondary to inflammatory arthritis, symptoms affecting the ankles or feet, person unable or unwilling to attend a physiotherapy treatment, psychiatric or medical morbidity that precludes participation in exercise treatment, receiving an intra-articular steroid injection within 3 months

Interventions | Intensity: Exercise program with different levels of time spent in exercise  
High intensity (n = 111): frequency: 2 times/week * 8 weeks; session duration: 45 min; intensity: home exercise + class exercise = circuit of exercise (5-min warm-up, 5-min stretching, 5-min balance training, 10-min isotonic exercises, 5-min isometric quadriceps exercises, 5-min cool-down). Progression: the class exercise program was increased and decreased by the senior physiotherapist by clinical discretion and in discussion with the participant  
Low intensity (n = 103): frequency: no information; session duration: no information; intensity: home exercise = muscle strengthening exercises, muscular endurance exercise, range of motion, standing balance. The home exercise was progressed. The level of intensity was fixed at 60% of the initial assessment, then reassessed and increased to 70% after 4 weeks, then to 80%  
All groups: Home and class exercises were progressed or reduced in case of an exacerbation of symptoms  
Supervision: senior therapist in the class exercise

Outcomes | At 8 weeks: Pain VAS, WOMAC pain, WOMAC physical function, SF-36, EuroQol (ED-5D), muscle strength, range of motion  
3 months: EuroQol (ED-5D)  
6 months: SF-36, EuroQol (ED-5D), muscle strength, pain VAS, WOMAC, range of motion  
9 months: EuroQol (ED-5D)  
12 months: SF-36, EuroQol (ED-5D), muscle strength, pain VAS, WOMAC, range of motion  
Using a Likert scale (range 0-20 or 0-68) in WOMAC
Notes

- Dropouts: 35 participants (16%) at the post-treatment assessment
- Adherence: attendance and compliance are not clearly reported
- Adverse events: no severe adverse events reported. Adverse effects related to exercise programs: “Despite compliance and the wide range of abilities of the patients in the sample only one adverse event was recorded. One patient, while performing one of the home exercises, developed an inguinal hernia that required surgical repair.”
- Attrition: high resistance, n = 17 (17%); low resistance, n = 7 (6%)
- We identified 2 published reports for the same trial. Data were extracted from McCarthy 2004
- Funded by the Health Technology Assessment (HTA) Programme

<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: “Subject allocation was carried out using a computerised minimisation algorithm built into an access.” Comment: Adequate method was used</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: “The lead investigator provided participants’ details to the trial data manager, who used a computerised, password-secured randomisation system (Microsoft Access) to allocate patients, at a location separate from the trial investigator’s place of work.” Comment: Allocation seemed to have remained concealed and the risk of selection bias low</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Quote: “This design feature allowed the patients to know which of the two treatment programs they were receiving.” Comment: Participants and staff were not blinded</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High risk</td>
<td>Comment: Participants were not blinded, introducing a risk of bias for subjective outcomes</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>Low risk</td>
<td>Quote: “To protect against bias, outcome assessments were made blind to allocation.”</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>High risk</td>
<td>Quote: “An ITT analysis on the major outcome’s 12-month data was conducted using last value carried forward (LVCF) imputation to examine further the effect of missing data.” Comment: ITT conducted only on 12 months’ data. A dropout rate &gt; 10% is noted and unbalanced among the 2 groups immediately at the end of the intervention (83% and 94%)</td>
</tr>
</tbody>
</table>
### Selective reporting (reporting bias)
- **High risk**
  - **Comment:** No protocol was found. Major and minor outcomes were specified, but additional outcomes data was incompletely reported in different reports (i.e. the SF-36 score is reported in McCarthy a but not McCarthy b).

### Other bias
- **Low risk**
  - **Comment:** There is no indication of other important risks of bias. Baseline performances seem similar between the 2 groups. A sample size calculation was provided.

---

**Ng 2010**

**Methods**
- **RCT with 2 groups?**

**Participants**
- **Location:** Australia
- **Randomized:** 36
- **Age:** mean 60.4 years
- **Sample:** 17 females and 11 males with hip or knee osteoarthritis
- **Settings:** university
  - Inclusion criteria: age from 40 to 75 years, osteoarthritis diagnosed in one hip or knee and experiencing clinical symptoms within the previous month, ability to walk at least 15 min continuously and safely participate in an exercise program
  - Exclusion criteria: other forms of arthritis, corticosteroid or visco supplement injections, history of infection, living in a dependent environment, taking daily medication for osteoarthritis, planning surgery in the next 6 months, receiving psychiatric or psychological treatment, pregnant or planning to become pregnant, exercising more than 60 min per week, or participating in another research study

**Interventions**
- **Intensity:** Exercise program with different levels of time spent in exercise
  - High intensity (n = 17): frequency: 5 times/week * 12 weeks; duration: 30 to 60 minutes
  - Low intensity (n = 19): frequency: 3 times/week * 12 weeks; duration: 30 to 60 minutes
- **Intensity and progression:** all groups were asked to walk at least 3000 steps (30 min) during 6 weeks and 6000 steps (60 min) during the 6 other weeks
- **Supervision:** unsupervised

**Outcomes**
- **At 12 weeks:** WOMAC (pain, function, global) using a numerical rating scale (range 0-10); Active Australia Physical Activity Questions, number of steps
- **Follow-up at 18 weeks:** WOMAC (pain, function, global); Active Australia Physical Activity Questions, number of steps
  - We converted the final scores expressed in a VAS to a Likert scale for the WOMAC pain (range 0-20) and function (range 0-68) subscales

**Notes**
- **Dropouts:** 8 (22%) did not complete the study
- **Adherence:** “Most participants in both groups were compliant with walking the number of ‘walking days’ called for in the protocol (i.e., they walked the prescribed 3 or 5 days per week), but compliance was higher in the 3-day walking group than in the 5-day walking group. Participants in the 3-day walking group showed 100% compliance with walking 3 days per week at Week 18. Participants in the 5-day walking group showed compliance ranging from 93% (Week 7) to 58% (Week 16)”
during the 12-week walking program.”
  - Adverse events: 5 participants dropped out during the walking program. Adverse effects reported related to exercise programs: Reasons for drop-out from the 3-day walking group were a death in the family (n = 1; dropout in Week 8), pain in the knees (n = 1; Week 7), and a torn Achilles tendon (n = 1; Week 7); from the 5-day walking group were pain while walking due to leg length discrepancies (n = 1; Week 12) and development of Baker’s cyst causing pain while walking (n = 1; Week 9)
  - Attrition: 5-day group, n= 2 (12%); 3-day group, n= 6 (32%)
  - We contacted the author, who provided additional information
  - Funded by the University of Queensland and Sanofi-Aventis Consumer Health Care, who supplied glucosamine sulphate intakes

### Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Comment: A computer random-number generator was used</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>High risk</td>
<td>Quote: “For practical reasons, allocation to group was not concealed.” “The assessor and main analyst (NTMN) was not blinded to group allocation and conducted the randomisation process before baseline, which may have contributed to ascertainment or performance biases.” Comment: Confirmed by the author that the allocation was not concealed</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias) All outcomes</td>
<td>High risk</td>
<td>Comment: Insufficient information, but it is unlikely that participants and care providers were blinded</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Subjective</td>
<td>High risk</td>
<td>Quote: “assessor and main analyst (NTMN) were not blinded to group allocation”</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Objective</td>
<td>High risk</td>
<td>Quote: “assessor and main analyst (NTMN) were not blinded to group allocation” Comment: Participants were not likely blinded, introducing a risk of bias</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias) All outcomes</td>
<td>High risk</td>
<td>Quote: “data were analysed on a per protocol basis” Comment: No ITT. Rate of drop-out is unbalanced between the 2 groups (12% and 32%), and high in 1 group</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>High risk</td>
<td>Comment: Protocol registration. All the important outcomes data were not reported in the final publication. The authors kindly provided the data (WOMAC) on request</td>
</tr>
</tbody>
</table>
### Ng 2010  
(Continued)

<table>
<thead>
<tr>
<th>Other bias</th>
<th>High risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comment: Compliance in the 5-day group was lower and different than in the 3-day group, for a high risk of bias. No sample size calculation</td>
<td></td>
</tr>
</tbody>
</table>

### Singh 2011

<table>
<thead>
<tr>
<th>Methods</th>
<th>RCT with 2 groups</th>
</tr>
</thead>
</table>
| Participants | Location: India  
Randomized: 200  
Age: mean 54.1 years  
Sample: 138 females and 62 males with knee osteoarthritis  
Settings: unknown  
Inclusion criteria: age between 40 and 65 years. Knee osteoarthritis  
Exclusion criteria: conditions that precluded participation in exercise such as coronary heart disease, myocardial infarction, unstable angina, chronic bronchitis, emphysema, peripheral disease, thrombophlebitis, embolism, kidney failure, uncontrolled hypertension |
| Interventions | Intensity: Exercise program with different levels of time spent in exercise  
High intensity (n = 100): frequency: 5 times/week * 8 weeks; session duration: 40 min; intensity: aerobic warm-up (5-10 min), walking (5-10 min), cycling at 60% of maximum heart rate (15-20 min) in addition to the conventional exercise program  
Low intensity (n = 100): frequency: 5 times/week * 8 weeks; duration: unknown min; intensity: conventional exercise program = hot packs, isometric exercises to quadriceps and hamstrings, range of motion, stretching and joint mobilization exercises, progressive resisted exercises  
All groups: participated in a conventional exercise program  
The authors did not provide any information on the progression or the supervision of the sessions |
| Outcomes | At 8 weeks: pain (VAS), WOMAC (function), muscle strength (isometric and isotonic strength measures), range of motion. Unclear whether VAS or Likert was used in WOMAC |
| Notes | ● Dropouts: not reported  
● Adherence: not reported  
● Adverse events: presence/absence not reported  
● Attrition: no information  
● We contacted the authors for additional informations but received no response |
Singh 2011  *(Continued)*

- No information on funding

### Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors' judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Unclear risk</td>
<td>Comment: No information on the process of randomization</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>Comment: Insufficient information on allocation concealment</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias) All outcomes</td>
<td>High risk</td>
<td>Comment: Insufficient information, but it is unlikely that participants and care providers were blinded, for possible risk of bias</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Subjective</td>
<td>High risk</td>
<td>Comment: Insufficient information, but it is unlikely that participants and care providers were blinded, for possible risk of bias</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Objective</td>
<td>High risk</td>
<td>Comment: Insufficient information, but it is unlikely that participants and care providers were blinded</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias) All outcomes</td>
<td>Unclear risk</td>
<td>Comment: Insufficient information</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear risk</td>
<td>Comment: Insufficient information. WOMAC score could not be extracted because of insufficient data presentation</td>
</tr>
<tr>
<td>Other bias</td>
<td>Unclear risk</td>
<td>Comment: Baseline performance seems similar between the 2 groups. No sample size calculation</td>
</tr>
</tbody>
</table>

1RM: one repetition maximum  
ACR: American College of Rheumatology  
HR: heart rate  
HRR: heart rate reserve  
ITT: intention to treat  
NSAIDS: nonsteroidal anti-inflammatory drugs  
RCT: randomized controlled trial  
SF-36: Short Form 36  
VAS: visual analog scale  
WOMAC: Western Ontario and McMaster Universities Arthritis Index
### Characteristics of excluded studies [ordered by study ID]

<table>
<thead>
<tr>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang 2012</td>
<td>RCT but no variation of intensity between the 2 interventions. Compared exercise with elastic bands + conventional modality treatments (shortwave diathermy, hot packs, transcutaneous electrical nerve stimulation, interferential current versus conventional modality treatments)</td>
</tr>
<tr>
<td>Diracoglu 2005</td>
<td>RCT but no variation of intensity between the 2 interventions. Compared strengthening exercise + balance versus strengthening exercise alone</td>
</tr>
<tr>
<td>Eyigor 2004</td>
<td>Intervention not appropriate. Effect of intensity could not be evaluated between the 2 groups. Compared isokinetic exercise program versus progressive resisted exercises program (isotonic regimen)</td>
</tr>
<tr>
<td>Fitzgerald 2011</td>
<td>RCT but no variation of intensity between the 2 interventions. Same amount of time in training between the 2 groups</td>
</tr>
<tr>
<td>Green 1993</td>
<td>RCT but no variation of intensity between the 2 interventions. Study compared home exercise versus hydrotherapy</td>
</tr>
<tr>
<td>Kruger 1997</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Sevick 2000</td>
<td>RCT but no variation of intensity between the 2 interventions. Study compared an aerobic exercise program consisting of walking on a treadmill at 50% to 70% of their heart reserve versus a resistance exercise program consisting of strengthening major muscle groups of both the upper and lower extremities</td>
</tr>
<tr>
<td>Shakoor 2007</td>
<td>RCT but no variation in intensity between the 2 interventions. Compared exercise versus exercise + activity modification advice</td>
</tr>
<tr>
<td>Teixeira 2011</td>
<td>This study appears to be the same trial as Fitzgerald 2011 and was excluded</td>
</tr>
<tr>
<td>Topp 2002</td>
<td>RCT but no variation in intensity between the 2 interventions. Compared dynamic resistance versus isometric resistance versus no intervention</td>
</tr>
<tr>
<td>Veenhof 2006</td>
<td>RCT but no variation in intensity between the 2 interventions. Compared behavioral versus usual care</td>
</tr>
</tbody>
</table>

RCT: randomized controlled trial

### Characteristics of studies awaiting assessment [ordered by study ID]

<table>
<thead>
<tr>
<th>Study</th>
<th>Methods</th>
<th>Participants</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steinhilber 2012</td>
<td>RCT</td>
<td>26 participants with hip osteoarthritis and 10 with hip replacement</td>
<td>Institutional hip exercises versus home-based strengthening</td>
</tr>
</tbody>
</table>
Steinhilber 2012  *(Continued)*

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Adherence, strength, Short Form 36</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes</td>
<td>Contacted for additional information but received no response</td>
</tr>
</tbody>
</table>

RCT: randomized controlled trial

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

**Messier 2011**

<table>
<thead>
<tr>
<th>Trial name or title</th>
<th>Strength Training for ARthritis Trial (START)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>RCT</td>
</tr>
<tr>
<td>Participants</td>
<td>Knee osteoarthritis</td>
</tr>
<tr>
<td>Interventions</td>
<td>High-intensity strength training versus low-intensity strength training versus attention control</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Pain, knee joint compressive forces, function, mobility, inflammatory markers, thigh composition measured at baseline, 6, 12, 18 months</td>
</tr>
<tr>
<td>Starting date</td>
<td>March 2012</td>
</tr>
<tr>
<td>Contact information</td>
<td><a href="mailto:jollajk@wfu.edu">jollajk@wfu.edu</a>, United States</td>
</tr>
<tr>
<td>Notes</td>
<td>NCT01489462</td>
</tr>
</tbody>
</table>

**Pua 2012**

<table>
<thead>
<tr>
<th>Trial name or title</th>
<th>Eccentric resistance ergometry in knee osteoarthritis rehabilitation: a randomized clinical trial (EUREKA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>RCT</td>
</tr>
<tr>
<td>Participants</td>
<td>Knee osteoarthritis</td>
</tr>
<tr>
<td>Interventions</td>
<td>Standard physiotherapy treatment and exercise on an eccentric ergometer versus standard physiotherapy</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Isometric knee extensor torque, gait speed, bridging, knee pain, knee passive range of motion, Short Form-36, standing balance</td>
</tr>
<tr>
<td>Starting date</td>
<td>April 2012</td>
</tr>
<tr>
<td>Contact information</td>
<td><a href="mailto:pua.yong.hao@sgh.com.sg">pua.yong.hao@sgh.com.sg</a>, Singapore</td>
</tr>
</tbody>
</table>
Continued

Notes

ACTRN12612000411842

Ång 2013

<table>
<thead>
<tr>
<th>Trial name or title</th>
<th>Dose-response effects of medical exercise therapy in patients with osteoarthritis of the knee - a Nordic multicenter clinical trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>RCT</td>
</tr>
<tr>
<td>Participants</td>
<td>Knee osteoarthritis</td>
</tr>
<tr>
<td>Interventions</td>
<td>High-dosage exercise therapy versus low-dosage exercise therapy</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Knee Injury and Osteoarthritis Outcome Score, VAS pain scale, self-rated patient satisfaction scale, Hospital Anxiety and Depression scale, Catastrophizing Scale, Tampa Scale of Kinesiophobia, 20-meter walk test, 30s Maximal Repeated Unilateral Knee Bending test, Five Time Repeated Chair Stands</td>
</tr>
<tr>
<td>Starting date</td>
<td>December 2013</td>
</tr>
<tr>
<td>Contact information</td>
<td><a href="mailto:Tom.Torstensen@ki.se">Tom.Torstensen@ki.se</a>; <a href="mailto:Bjorn.Ang@ki.se">Bjorn.Ang@ki.se</a>, Sweden</td>
</tr>
<tr>
<td>Notes</td>
<td>NCT02024126</td>
</tr>
</tbody>
</table>

Østerås 2012

<table>
<thead>
<tr>
<th>Trial name or title</th>
<th>Dose-response: exercise therapy on hip osteoarthritis, a pilot study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>RCT</td>
</tr>
<tr>
<td>Participants</td>
<td>Hip osteoarthritis</td>
</tr>
<tr>
<td>Interventions</td>
<td>High-dosage exercise therapy versus low-dosage exercise therapy</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Stairs, squats, WOMAC Index of Osteoarthritis (questionnaire) measured 8 weeks and 6 months</td>
</tr>
<tr>
<td>Starting date</td>
<td>September 2012</td>
</tr>
<tr>
<td>Contact information</td>
<td><a href="mailto:freddy.p123@hotmail.com">freddy.p123@hotmail.com</a>, Norway</td>
</tr>
<tr>
<td>Notes</td>
<td>NCT01700933</td>
</tr>
</tbody>
</table>

RCT: randomized controlled trial
VAS: visual analog scale
WOMAC: Western Ontario and McMaster Universities Arthritis Index
### Comparison 1. High versus low intensity exercise

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Pain (WOMAC)</td>
<td>4</td>
<td>313</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>1.1 at study completion</td>
<td>4</td>
<td>313</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-0.84 [-1.63, -0.04]</td>
</tr>
<tr>
<td>1.2 at mid term</td>
<td>2</td>
<td>199</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-0.82 [-1.90, 0.26]</td>
</tr>
<tr>
<td>1.3 at long-term</td>
<td>1</td>
<td>138</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-1.33 [-2.56, -0.10]</td>
</tr>
<tr>
<td>2 Physical Function (WOMAC)</td>
<td>4</td>
<td>313</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>2.1 at study completion</td>
<td>4</td>
<td>313</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-2.65 [-5.29, -0.01]</td>
</tr>
<tr>
<td>2.2 at mid term</td>
<td>2</td>
<td>199</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-2.95 [-7.00, 1.10]</td>
</tr>
<tr>
<td>2.3 at long-term</td>
<td>1</td>
<td>138</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-4.20 [-9.32, 0.92]</td>
</tr>
<tr>
<td>3 Adverse effects</td>
<td>4</td>
<td>364</td>
<td>Peto Odds Ratio (Peto, Fixed, 95% CI)</td>
<td>1.72 [0.51, 5.81]</td>
</tr>
<tr>
<td>4 Gait speed</td>
<td>3</td>
<td>77</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>4.1 Free walking speed</td>
<td>2</td>
<td>77</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>0.04 [-0.05, 0.13]</td>
</tr>
<tr>
<td>4.2 Fast walking speed</td>
<td>2</td>
<td>107</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>0.08 [-0.13, 0.29]</td>
</tr>
<tr>
<td>5 Muscle strength</td>
<td>4</td>
<td>285</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>5.1 Knee extensor</td>
<td>3</td>
<td>285</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>0.38 [0.04, 0.72]</td>
</tr>
<tr>
<td>5.2 Knee Flexor</td>
<td>2</td>
<td>113</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>0.18 [-0.64, 1.00]</td>
</tr>
<tr>
<td>5.3 Global strength</td>
<td>2</td>
<td>245</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>1.01 [0.74, 1.27]</td>
</tr>
<tr>
<td>6 Aerobic capacity</td>
<td>1</td>
<td>200</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
<tr>
<td>7 Range of Motion</td>
<td>2</td>
<td>113</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>Totals not selected</td>
</tr>
</tbody>
</table>

### Comparison 2. Subgroup analysis: Exercise duration versus resistance

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Pain</td>
<td>4</td>
<td>200</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>1.1 Duration exercise</td>
<td>2</td>
<td>200</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-1.37 [-2.47, -0.28]</td>
</tr>
<tr>
<td>1.2 Resistance exercise</td>
<td>2</td>
<td>113</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-0.23 [-1.40, 0.93]</td>
</tr>
<tr>
<td>2 Function</td>
<td>4</td>
<td>200</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>2.1 Duration exercise</td>
<td>2</td>
<td>200</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-4.10 [-8.12, -0.07]</td>
</tr>
<tr>
<td>2.2 Resistance exercise</td>
<td>2</td>
<td>113</td>
<td>Mean Difference (IV, Fixed, 95% CI)</td>
<td>-1.57 [-5.06, 1.93]</td>
</tr>
<tr>
<td>3 Adverse effects</td>
<td>5</td>
<td>242</td>
<td>Odds Ratio (IV, Fixed, 95% CI)</td>
<td>Subtotals only</td>
</tr>
<tr>
<td>3.1 Duration exercise</td>
<td>2</td>
<td>242</td>
<td>Odds Ratio (IV, Fixed, 95% CI)</td>
<td>0.62 [0.11, 3.64]</td>
</tr>
<tr>
<td>3.2 Resistance Exercise</td>
<td>3</td>
<td>161</td>
<td>Odds Ratio (IV, Fixed, 95% CI)</td>
<td>3.68 [0.55, 24.73]</td>
</tr>
</tbody>
</table>
### Analysis 1.1. Comparison 1 High versus low intensity exercise, Outcome 1 Pain (WOMAC).

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 1 High versus low intensity exercise

Outcome: 1 Pain (WOMAC)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV,Fixed,95% CI</td>
</tr>
<tr>
<td>1 at study completion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>93</td>
<td>7.5 (3.95)</td>
<td>79</td>
<td>9.04 (3.84)</td>
<td>-1.54 [-2.71, -0.37]</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>20</td>
<td>3.8 (2.7)</td>
<td>25</td>
<td>4.4 (3.7)</td>
<td>-0.60 [-2.47, 1.27]</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15</td>
<td>5.39 (4.34)</td>
<td>13</td>
<td>5.48 (4.4)</td>
<td>-0.09 [-3.34, 3.16]</td>
</tr>
<tr>
<td>Jan 2008</td>
<td>34</td>
<td>4.8 (3.5)</td>
<td>34</td>
<td>4.8 (2.7)</td>
<td>0.00 [-1.49, 1.49]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>162</strong></td>
<td><strong>151</strong></td>
<td></td>
<td></td>
<td><strong>100.0 % -0.84 [-1.63, -0.04]</strong></td>
</tr>
<tr>
<td>Heterogeneity: Chi² = 2.88, df = 3 (P = 0.41); I² =0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.05 (P = 0.040)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 at mid term</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>95</td>
<td>8.04 (3.6)</td>
<td>76</td>
<td>9.13 (3.99)</td>
<td>-1.09 [-2.24, 0.06]</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15</td>
<td>5.01 (5.2)</td>
<td>13</td>
<td>3.9 (3)</td>
<td>1.11 [-1.99, 4.21]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>110</strong></td>
<td><strong>89</strong></td>
<td></td>
<td></td>
<td><strong>100.0 % -0.82 [-1.90, 0.26]</strong></td>
</tr>
<tr>
<td>Heterogeneity: Chi² = 1.70, df = 1 (P = 0.19); I² =41%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.49 (P = 0.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 at long-term</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>76</td>
<td>8.05 (3.81)</td>
<td>62</td>
<td>9.38 (3.53)</td>
<td>-1.33 [-2.56, -0.10]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>76</strong></td>
<td><strong>62</strong></td>
<td></td>
<td></td>
<td><strong>100.0 % -1.33 [-2.56, -0.10]</strong></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.12 (P = 0.034)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-10 -5 0 5 10
Favors High intensity Favors Low intensity
## Analysis 1.2. Comparison 1 High versus low intensity exercise, Outcome 2 Physical Function (WOMAC)

**Review:** High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

**Comparison:** 1 High versus low intensity exercise

**Outcome:** 2 Physical Function (WOMAC)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV,Fixed,95% CI</td>
</tr>
<tr>
<td>1 at study completion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>93</td>
<td>23.6 (13.9)</td>
<td>79</td>
<td>28.1 (4.7)</td>
<td>-4.50 [-8.80, -0.20]</td>
</tr>
<tr>
<td>Jan 2008</td>
<td>34</td>
<td>14.7 (8.5)</td>
<td>34</td>
<td>14.8 (9.2)</td>
<td>-0.10 [-4.31, 4.11]</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15</td>
<td>16.45 (12.61)</td>
<td>13</td>
<td>17.68 (17.46)</td>
<td>-1.23 [-12.67, 10.21]</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>20</td>
<td>13.3 (9.4)</td>
<td>25</td>
<td>18.1 (12)</td>
<td>-4.80 [-11.05, 1.45]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>162</td>
<td>151</td>
<td>100.0 %</td>
<td>-2.65 [-5.29, -0.01]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: \( \chi^2 = 2.63, \text{df} = 3 (P = 0.45); I^2 = 0.0\%

Test for overall effect: \( Z = 1.97 (P = 0.049) \)

2 at mid term

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV,Fixed,95% CI</td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>95</td>
<td>26.6 (14.2)</td>
<td>76</td>
<td>29.8 (14.5)</td>
<td>-3.20 [-7.53, 1.13]</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15</td>
<td>16.45 (12.61)</td>
<td>13</td>
<td>17.68 (17.4)</td>
<td>-1.23 [-12.64, 10.18]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>110</td>
<td>89</td>
<td>100.0 %</td>
<td>-2.95 [-7.00, 1.10]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: \( \chi^2 = 0.10, \text{df} = 1 (P = 0.75); I^2 = 0.0\%

Test for overall effect: \( Z = 1.93 (P = 0.15) \)

3 at long-term

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCarthy 2004</td>
<td>76</td>
<td>26.5 (13.6)</td>
<td>62</td>
<td>30.7 (16.5)</td>
<td>-4.20 [-9.32, 0.92]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>76</td>
<td>62</td>
<td>100.0 %</td>
<td>-4.20 [-9.32, 0.92]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable

Test for overall effect: \( Z = 1.61 (P = 0.11) \)

Test for subgroup differences: \( \chi^2 = 0.28, \text{df} = 2 (P = 0.87), I^2 = 0.0\% \)

---

High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis (Review)  
Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
**Analysis 1.3. Comparison 1 High versus low intensity exercise, Outcome 3 Adverse effects.**

*Review:* High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

*Comparison:* 1 High versus low intensity exercise

*Outcome:* 3 Adverse effects

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Experimental</th>
<th>Control</th>
<th>Peto Odds Ratio</th>
<th>Peto,Fixed 95% CI</th>
<th>Weight</th>
<th>Peto Odds Ratio</th>
<th>Peto,Fixed 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foroughi 2011</td>
<td>2/26</td>
<td>1/28</td>
<td></td>
<td></td>
<td>27.9%</td>
<td>2.16 [0.21, 21.75]</td>
<td></td>
</tr>
<tr>
<td>Jan 2008</td>
<td>3/34</td>
<td>0/34</td>
<td></td>
<td></td>
<td>28.2%</td>
<td>7.86 [0.79, 78.20]</td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>0/111</td>
<td>1/103</td>
<td></td>
<td></td>
<td>9.7%</td>
<td>0.13 [0.00, 6.33]</td>
<td></td>
</tr>
<tr>
<td>Ng 2010</td>
<td>2/15</td>
<td>2/13</td>
<td></td>
<td></td>
<td>34.2%</td>
<td>0.85 [0.11, 6.84]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>186</strong></td>
<td><strong>178</strong></td>
<td></td>
<td></td>
<td><strong>100.0%</strong></td>
<td><strong>1.72 [0.51, 5.81]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 7 (Experimental), 4 (Control)

Heterogeneity: Chi² = 3.87, df = 3 (P = 0.28); I² = 22%

Test for overall effect: Z = 0.87 (P = 0.39)

Test for subgroup differences: Not applicable

---

High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis (Review)

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Analysis 1.4. Comparison 1 High versus low intensity exercise, Outcome 4 Gait speed.

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 1 High versus low intensity exercise

Outcome: 4 Gait speed

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Free walking speed</td>
<td>1</td>
<td>19</td>
<td>1.2 (0.17)</td>
<td>19</td>
<td>1.2 (0.17)</td>
</tr>
<tr>
<td>Mangione 1999</td>
<td>19</td>
<td>1.1 (0.2)</td>
<td>20</td>
<td>1.01 (0.19)</td>
<td>44.6%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>38</td>
<td>39</td>
<td>0.0</td>
<td>100.0%</td>
<td>0.04 [-0.05, 0.13]</td>
</tr>
<tr>
<td>Fast walking speed</td>
<td>2</td>
<td>34</td>
<td>0 (0)</td>
<td>34</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Mangione 1999</td>
<td>19</td>
<td>1.71 (0.3)</td>
<td>20</td>
<td>1.63 (0.36)</td>
<td>100.0%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>53</td>
<td>54</td>
<td>0.2</td>
<td>100.0%</td>
<td>0.08 [-0.13, 0.29]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau^2 = 0.00; Chi^2 = 1.17, df = 1 (P = 0.28); I^2 =14%
Test for overall effect: Z = 0.90 (P = 0.37)

Heterogeneity: not applicable
Test for overall effect: Z = 0.76 (P = 0.45)
### Analysis 1.5. Comparison 1 High versus low intensity exercise, Outcome 5 Muscle strength.

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 1 High versus low intensity exercise

Outcome: 5 Muscle strength

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Std. Mean Difference</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
</tr>
<tr>
<td>1 Knee extensor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>93</td>
<td>71.1 (32.1)</td>
<td>79</td>
<td>59.9 (21.6)</td>
</tr>
<tr>
<td>Jan 2008</td>
<td>34</td>
<td>88.1 (21.6)</td>
<td>34</td>
<td>86.7 (24.2)</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>20</td>
<td>57.4 (25.8)</td>
<td>25</td>
<td>40.3 (16.9)</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>147</td>
<td>138</td>
<td>100.0 %</td>
<td>0.38 [ 0.04, 0.72 ]</td>
</tr>
<tr>
<td>Heterogeneity: Tau^2 = 0.04; Chi^2 = 3.47, df = 2 (P = 0.18); I^2 = 42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.21 (P = 0.027)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Knee Flexor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 2008</td>
<td>34</td>
<td>57.4 (18.9)</td>
<td>34</td>
<td>61.7 (20.1)</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>20</td>
<td>110.6 (24.1)</td>
<td>25</td>
<td>93.5 (29.1)</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>54</td>
<td>59</td>
<td>100.0 %</td>
<td>0.18 [-0.64, 1.00 ]</td>
</tr>
<tr>
<td>Heterogeneity: Tau^2 = 0.28; Chi^2 = 4.58, df = 1 (P = 0.03); I^2 = 78%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.43 (P = 0.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Global strength</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singh 2011</td>
<td>100</td>
<td>531.5 (47.1)</td>
<td>100</td>
<td>481.5 (46.9)</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>20</td>
<td>1461.8 (504.1)</td>
<td>25</td>
<td>1092.1 (438.2)</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>120</td>
<td>125</td>
<td>100.0 %</td>
<td>1.01 [ 0.74, 1.27 ]</td>
</tr>
<tr>
<td>Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.67, df = 1 (P = 0.41); I^2 = 0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 7.39 (P &lt; 0.00001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Analysis 1.6. Comparison 1 High versus low intensity exercise, Outcome 6 Aerobic capacity.**

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 1 High versus low intensity exercise

Outcome: 6 Aerobic capacity

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (Mean(SD))</td>
<td>N (Mean(SD))</td>
<td>IV,Fixed,95% CI</td>
<td>IV,Fixed,95% CI</td>
</tr>
<tr>
<td>Mangione 1999</td>
<td>19 (20.53(4.23)</td>
<td>20 (21.93(4.56)</td>
<td>-1.40 [-4.16, 1.36]</td>
<td></td>
</tr>
</tbody>
</table>

Favors High intensity Favors Low intensity

---

**Analysis 1.7. Comparison 1 High versus low intensity exercise, Outcome 7 Range of Motion.**

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 1 High versus low intensity exercise

Outcome: 7 Range of Motion

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (Mean(SD))</td>
<td>N (Mean(SD))</td>
<td>IV,Random,95% CI</td>
<td>IV,Random,95% CI</td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>104 (124.9(12.2))</td>
<td>86 (119 (13.9))</td>
<td>5.90 [2.14, 9.66]</td>
<td></td>
</tr>
<tr>
<td>Singh 2011</td>
<td>100 (121.3 (11.4))</td>
<td>100 (107.3 (13.6))</td>
<td>14.00 [10.52, 17.48]</td>
<td></td>
</tr>
</tbody>
</table>

Favors Low intensity Favors High intensity
### Analysis 2.1. Comparison 2 Subgroup analysis: Exercise duration versus resistance, Outcome 1 Pain.

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 2 Subgroup analysis: Exercise duration versus resistance

Outcome: 1 Pain

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV,Fixed,95% CI</td>
</tr>
<tr>
<td>Duration exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>93</td>
<td>7.5 (3.95)</td>
<td>79</td>
<td>9.04 (3.84)</td>
<td>88.6 %</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15</td>
<td>5.39 (4.34)</td>
<td>13</td>
<td>5.48 (4.4)</td>
<td>11.4 %</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>108</td>
<td></td>
<td>92</td>
<td></td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Heterogeneity: Chi² = 0.68, df = 1 (P = 0.41); I² = 0.0%
Test for overall effect: Z = 2.45 (P = 0.014)

<table>
<thead>
<tr>
<th>Resistance exercise</th>
<th>N</th>
<th>Mean(SD)</th>
<th>N</th>
<th>Mean(SD)</th>
<th>IV,Fixed,95% CI</th>
<th>IV,Fixed,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 2008</td>
<td>34</td>
<td>4.8 (3.5)</td>
<td>34</td>
<td>4.8 (2.7)</td>
<td>61.3 %</td>
<td>0.00 [-1.49, 1.49]</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>20</td>
<td>3.8 (2.7)</td>
<td>25</td>
<td>4.4 (3.7)</td>
<td>38.7 %</td>
<td>-0.60 [-2.47, 1.27]</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>54</td>
<td></td>
<td>59</td>
<td></td>
<td>100.0 %</td>
<td>-0.23 [-1.40, 0.93]</td>
</tr>
</tbody>
</table>

Heterogeneity: Chi² = 0.24, df = 1 (P = 0.62); I² = 0.0%
Test for overall effect: Z = 0.39 (P = 0.70)
Test for subgroup differences: Chi² = 1.96, df = 1 (P = 0.16); I² = 49%
### Analysis 2.2. Comparison 2 Subgroup analysis: Exercise duration versus resistance, Outcome 2 Function.

**Review:** High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

**Comparison:** 2 Subgroup analysis: Exercise duration versus resistance

**Outcome:** 2 Function

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Mean (SD)</th>
<th>Low intensity</th>
<th>Mean (SD)</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration exercise</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>93</td>
<td>23.6 (13.9)</td>
<td>79</td>
<td>28.1 (14.7)</td>
<td>-4.50 [-8.80, -0.20]</td>
<td>87.6%</td>
<td></td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15</td>
<td>16.45 (12.61)</td>
<td>13</td>
<td>17.68 (17.46)</td>
<td>-1.23 [-12.67, 10.21]</td>
<td>12.4%</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>108</td>
<td>92</td>
<td></td>
<td></td>
<td>-4.10 [-8.12, -0.07]</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Chi² = 0.28, df = 1 (P = 0.60); I² = 0.0%

Test for overall effect: Z = 1.99 (P = 0.046)

| **Resistance exercise** |               |           |               |           |               |        |                |
| Jan 2008              | 34             | 14.7 (8.5)  | 34            | 14.8 (9.2)  | -0.10 [-4.31, 4.11] | 68.8%  |               |
| Foroughi 2011         | 20             | 13.3 (9.4)  | 25            | 18.1 (12)   | -4.80 [-11.05, 1.45] | 31.2%  |               |
| **Subtotal (95% CI)** | 54             | 59        |               |           | -1.57 [-5.06, 1.93] | 100.0% |               |

Heterogeneity: Chi² = 1.49, df = 1 (P = 0.22); I² = 33%

Test for overall effect: Z = 0.88 (P = 0.38)

Test for subgroup differences: Chi² = 0.87, df = 1 (P = 0.35); I² = 0.0%

---

High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis (Review)  
Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Analysis 2.3. Comparison 2 Subgroup analysis: Exercise duration versus resistance, Outcome 3 Adverse effects.

Review: High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis

Comparison: 2 Subgroup analysis: Exercise duration versus resistance

Outcome: 3 Adverse effects

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>High intensity</th>
<th>Low intensity</th>
<th>Odds Ratio (IV/Fixed, 95% CI)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>0/11</td>
<td>1/103</td>
<td>0.31 [0.01, 7.61]</td>
<td>30.3%</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>2/15</td>
<td>2/13</td>
<td>0.85 [0.10, 7.04]</td>
<td>69.7%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>126</td>
<td>116</td>
<td>0.62 [0.11, 3.64]</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Total events: 2 (High intensity), 3 (Low intensity)
Heterogeneity: Chi² = 0.27, df = 1 (P = 0.60); I² = 0%
Test for overall effect: Z = 0.53 (P = 0.60)

<table>
<thead>
<tr>
<th>Resistance Exercise</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mangione 1999</td>
<td>0/19</td>
<td>0/20</td>
<td>Not estimable</td>
<td></td>
</tr>
<tr>
<td>Jan 2008</td>
<td>3/34</td>
<td>0/34</td>
<td>7.67 [0.38, 154.34]</td>
<td>40.2%</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>2/26</td>
<td>1/28</td>
<td>2.25 [0.19, 26.41]</td>
<td>59.8%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>79</td>
<td>82</td>
<td>3.68 [0.55, 24.73]</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Total events: 5 (High intensity), 1 (Low intensity)
Heterogeneity: Chi² = 0.38, df = 1 (P = 0.54); I² = 0%
Test for overall effect: Z = 1.34 (P = 0.18)

Additional tables

Table 1. Primary outcomes reported in included studies

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Primary outcome reported</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pain (WOMAC)</td>
</tr>
<tr>
<td>Mangione 1999</td>
<td>no</td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>yes</td>
</tr>
</tbody>
</table>
Table 1. Primary outcomes reported in included studies (Continued)

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Gait speed (fast)</th>
<th>Gait speed (normal)</th>
<th>Muscle strength (knee extensor)</th>
<th>Muscle strength (knee flexor)</th>
<th>Muscle strength (global)</th>
<th>Aerobic capacity</th>
<th>Range of motion</th>
<th>Physical activity</th>
<th>Number of steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mangione 1999</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Jan 2008</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Singh 2011</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

AIM2: Arthritis Impact Measurement Scale 2  
ALF: aggregated locomotor function  
QoL: quality of life  
SF-36: Short Form 36  
VAS: visual analog scale  
WOMAC: Western Ontario and McMaster Universities Arthritis Index

Table 2. Secondary outcomes reported in included studies

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Gait speed (fast)</th>
<th>Gait speed (normal)</th>
<th>Muscle strength (knee extensor)</th>
<th>Muscle strength (knee flexor)</th>
<th>Muscle strength (global)</th>
<th>Aerobic capacity</th>
<th>Range of motion</th>
<th>Physical activity</th>
<th>Number of steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mangione 1999</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>McCarthy 2004</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Jan 2008</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Singh 2011</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
APPENDICES

Appendix 1. MEDLINE search strategy

1. osteoarthritis/
2. (degenerative adj2 arthritis).tw.
3. (osteoarthr$ or arthrosis).tw.
4. or/1-3
5. Knee/
6. exp Knee Joint/
7. knee$.tw.
8. Hip/
9. Hip Joint/
10. (hip or hips).tw.
11. or/5-10
12. 4 and 11
13. Osteoarthritis, Knee/
14. Osteoarthritis, Hip/
15. or/12-14
16. exp Exercise/
17. exp Exercise Movement Techniques/
18. exp Exercise Therapy/
19. Rehabilitation/
20. exp Motor Activity/
21. exercis$.tw.
22. physical activit$.tw.
23. strength$.tw.
24. ((isometric$ or isokinetic$ or aerobic$ or endurance or weigh$ or resistance) adj3 (train$ or therap$ or rehab$ or program$)).tw.
25. ((high or low) adj (intens$ or impact or dose$ or amount$)).tw.
26. Physical Therapy Modalities/
27. (physical ther$ or physiother$).tw.
28. (run$ or jog$ or walk$ or treadmill$ or cycl$ or row$ or gait).tw.
29. or/16-28
30. randomized controlled trial.pt.
31. controlled clinical trial.pt.
32. randomized.ab.
33. placebo.ab.
34. drug therapy.fs.
35. randomly.ab.
36. trial.ab.
37. groups.ab.
38. or/30-37
39. (animals not (humans and animals)).sh.
40. 38 not 39
41. and/15,29,40
Appendix 2. EMBASE search strategy

1. osteoarthritis/
2. (degenerative adj2 arthritis).tw.
3. (osteoarthr$ or arthrosis).tw.
4. or/1-3
5. knee/
6. knee$.tw.
7. hip/
8. (hip or hips).tw.
9. or/5-8
10. 4 and 9
11. knee osteoarthritis/
12. hip osteoarthritis/
13. or/10-12
14. exp exercise/
15. exp kinesiotherapy/
16. rehabilitation/
17. exp motor activity/
18. exercis$.tw.
19. physical activit$.tw.
20. strength$.tw.
21. ((isometric$ or isokinetic$ or aerobic$ or endurance or weigh$ or resistance) adj3 (train$ or therap$ or rehab$ or program$)).tw.
22. ((high or low) adj (intens$ or impact or dose$ or amount$)).tw.
23. exp physiotherapy/
24. (physical ther$ or physiother$).tw.
25. (run$ or jog$ or walk$ or treadmill$ or cycl$ or row$ or gait).tw.
26. or/14-25
27. 13 and 26
28. (random$ or placebo$).ti,ab.
29. ((single$ or double$ or triple$ or treble$) and (blind$ or mask$)).ti,ab.
30. controlled clinical trial$.ti,ab.
31. RETRACTED ARTICLE/
32. or/28-31
33. (animal$ not human$).sh,hw.
34. 32 not 33
35. 27 and 34

Appendix 3. CENTRAL search strategy

#1 MeSH descriptor Osteoarthritis, this term only
#2 degenerative Near/2 arthritis:ti,ab
#3 (osteoarthr* or arthrosis):ti,ab
#4 (#1 OR #2 OR #3)
#5 MeSH descriptor Knee, this term only
#6 MeSH descriptor Knee Joint explode all trees
#7 knee*:ti,ab
#8 MeSH descriptor Hip, this term only
#9 MeSH descriptor Hip Joint, this term only
#10 (hip or hips):ti,ab
#11 (#5 OR #6 OR #7 OR #8 OR #9 OR #10)
#12 (#4 AND #11)
#13 MeSH descriptor Osteoarthritis, Knee, this term only
Appendix 4. CINAHL search strategy

S1 (MH "Osteoarthritis")
S2 TI degenerative N2 arthritis OR AB degenerative N2 arthritis
S3 TI osteoarthr* or AB arthrosis OR TI osteoarthr* OR AB arthrosis
S4 S1 or S2 or S3
S5 (MH "Knee")
S6 (MH "Knee Joint")
S7 TI knee* OR AB knee*
S8 (MH "Hip")
S9 (MH "Hip Joint")
S10 TI hip OR TI hips OR AB hip OR AB hips
S11 S5 or S6 or S7 or S8 or S9 or S10
S12 S4 and S11
S13 (MH "Osteoarthritis, Knee")
S14 (MH "Osteoarthritis, Hip")
S15 S12 or S13 or S14
S16 (MH "Exercise")
S17 (MH "Therapeutic Exercise")
S18 (MH "Rehabilitation")
S19 (MH "Motor Activity")
S20 TI exercis* OR AB exercis*
S21 TI physical activit* OR AB physical activit*
S22 TI strength* OR AB strength*
S23 TI Isometric* N3 train* OR AB Isometric* N3 train*
S24 TI Isokinetic* N3 train* OR AB Isokinetic* N3 train*
S25 TI Aerobic* N3 train* OR AB Aerobic* N3 train*
S26 TI endurance N3 train* OR AB endurance N3 train*
S27 TI weigh* N3 train* OR AB weigh* N3 train*
S28 TI resistance N3 train* OR AB resistance N3 train*
S29 TI Isometric* N3 therap* OR AB Isometric* N3 therap*
S30 TI Isokinetic* N3 therap* OR AB Isokinetic* N3 therap*
S31 TI Aerobic* N3 therap* OR AB Aerobic* N3 therap*
Appendix 5. PEDro search strategy

1. Osteoarthritis OR arthritis (title or abstract)
2. Hip OR Knee (body part)
3. Exercise OR Physical Activity (title or abstract)
4. Training OR hydrotherapy, balneotherapy OR skill training OR strength training (therapy)
5. 3 OR 4
6. Systematic Review OR Clinical Trial (method)
7. 1 AND 2 AND 5 AND 6
Appendix 6. SCOPUS search strategy

1 TITLE-ABS-KEY(osteoarthr* OR degenerative arthritis OR arthrosis)
2 TITLE-ABS-KEY(hip OR hips OR knee*)
3 TITLE-ABS-KEY(exercis* OR physical activit* OR strength* OR isometric* train* OR isokinetic* train* OR aerobic* train* OR endurance train* OR weigh* train* OR resistance train* OR isometric* therap* OR isokinetic* therap* OR aerobic* therap* OR endurance therap* OR weigh* therap* OR resistance therap* OR isometric* rehab* OR isokinetic rehab* OR aerobic* rehab* OR endurance rehab* OR weigh* rehab* OR resistance rehab* OR isometric program* OR isokinetic program* OR aerobic program* OR endurance program* OR weigh* program* OR resistance program*)
4 TITLE-ABS-KEY(high intens* OR low intens* OR high impact OR low impact OR high dose* OR low dose* OR high amount* OR low amount*)
5 TITLE-ABS-KEY(physical ther* OR physiother* OR run* OR jog* OR walk* OR treadmill* OR cycle* OR row* OR gait)
6 3 OR 4 OR 5
7 1 AND 2 AND 6
8 Limit 7 to Conference Paper

Appendix 7. Survey of authors’ reactions to provide information on trials

<table>
<thead>
<tr>
<th>[Study ID]</th>
<th>[Study author contacted]</th>
<th>[Study author replied]</th>
<th>Current status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teixeira 2011</td>
<td>26/06/2013</td>
<td>28/06/2013</td>
<td>provided more data</td>
</tr>
<tr>
<td>Ng 2010</td>
<td>15/07/2013</td>
<td>16/07/2013</td>
<td>provided more data</td>
</tr>
<tr>
<td>Foroughi 2011</td>
<td>04/10/2013</td>
<td>08/10/2013</td>
<td>provided more data</td>
</tr>
<tr>
<td>Singh 2011</td>
<td>10/03/2014</td>
<td>no contact</td>
<td>no data provided</td>
</tr>
<tr>
<td>Messier 2011</td>
<td>06/11/2013</td>
<td>07/11/2013</td>
<td>data not available</td>
</tr>
<tr>
<td>Pua 2012</td>
<td>06/11/2013</td>
<td>07/11/2013</td>
<td>data not available</td>
</tr>
<tr>
<td>Østerås 2012</td>
<td>06/11/2013</td>
<td>07/11/2013</td>
<td>data not available</td>
</tr>
<tr>
<td>Steinhilber 2012</td>
<td>13/11/2013</td>
<td>No contact</td>
<td>data not available</td>
</tr>
</tbody>
</table>

Contributions of Authors

JPR and MMLC wrote the review.
LB wrote the previous systematic review that is the foundation of the current work.
JPR, MMLC, IB, LB, PR, and LT conceived of and designed the review and interpreted the data.
JPR and CN independently screened retrieved clinical studies for inclusion.
JPR and MMLC extracted data from included studies and performed the methodological quality assessment.
JPR, MMLC, and LT analyzed the data.
JPR, MMLC, IB, LT, LB, and PR reviewed the final manuscript prior to submission.
DECLARATIONS OF INTEREST

None of the researchers have a commercial for-profit interest in this review. There was no financial support from the pharmaceutical industry for this review.

SOURCES OF SUPPORT

Internal sources
• EHESP, French school of public health, France.
  in-kind support
• Centre de recherche Epidémiologies et Biostatistique, INSERM U1153, France.
  in-kind support
• Hôpital Hôtel-Dieu, APHP, France.
  in-kind support

External sources
• No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. Measure of treatment effect in continuous outcomes: added mean difference calculation when the same tool was used to measure the same outcome across separate studies.
2. 'Summary of findings' tables: added minimum clinical important difference threshold and methods of calculation to express absolute and relative changes for dichotomous or continuous outcome measures. We followed the recommendations provided by the Cochrane Musculoskeletal Review Group.
3. We did not perform contour-enhanced funnel plots to assess the presence of small-study effects as the required statistical conditions were not met.
4. We stated in the protocol that we would attempt to fit a bivariate random-effects meta-analysis to address the issue between correlated outcomes and missing data. As no data were missing, we did not perform the analysis.
5. We did not perform subgroup analysis to explore whether a relationship exists between the type of treatments, type of intervention, type of joint, indication of exercise, duration of treatment as insufficient data was available. Meta-regression was also not possible due to the small number of included studies.
6. We stated in the protocol that would attempt to perform a sensitivity analysis to explore how the results of meta-analysis might be affected by including only studies at low risk of bias. However, as all the identified studies were at high risk of bias, we did not perform the analysis.