Effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review of randomized clinical trials
Van Baar M E, Assendelft W J, Dekker J, Oostendorp R A, Bijlsma J W

Authors' objectives
To determine the effectiveness of exercise therapy in patients with osteoarthritis (OA) of the hip or knee.

Searching
The following databases were searched: MEDLINE (January 1966 to September 1997); EMBASE (January 1988 to September 1997); and CINAHL (January 1982 to September 1997) using a highly sensitive search strategy for RCTs and systematic reviews. The Cochrane Controlled Trials Register was also searched and references in relevant review articles and trials were screened. No language restrictions were applied.

Study selection
Study designs of evaluations included in the review
Randomised controlled trials (RCTs) that evaluated exercise in patients with OA of the hip or knee were included if results had been published as a full report. Trials were excluded if intervention and control groups had not received identical exercise therapy. Reasons were given for exclusion of identified studies.

Specific interventions included in the review
Exercise therapy, defined as follows, was studied: a range of activities designed to improve strength, range of motion, endurance, balance, coordination, posture, motor function, or motor development. Exercise could be performed actively, passively, or against resistance and no restrictions were applied as to type of supervision or group size. Additional interventions were allowed. Interventions in the review included the following: individual and group therapy; supervised and home-based programmes; aerobic and aerobic hydrotherapy; resistance exercises; 'fitness walking'; strength training monitored on a dynameter; weight bearing and non weight bearing exercises; and non aerobic programmes directed to range of motion. Perioperative exercise therapy was excluded.

Participants included in the review
Patients with OA of the hip or knee (assessed clinically, radiographically or a combination) were studied. Subjects included those with mild to moderate OA and included those recruited from the community and through clinical settings.

Outcomes assessed in the review
The following outcomes were assessed: pain; self-reported disability; observed disability in walking; and patient's global assessment of effect.

How were decisions on the relevance of primary studies made?
The authors do not state how the papers were selected for the review, or how many of the authors performed the selection.

Assessment of study quality
Validity was assessed using the following criteria as defined in the Maastricht-Amsterdam consensus list (see Other Publications of Related Interest no.2): validity criteria (including adequacy of randomisation; randomisation independent of person determining eligibility; control for conterventions; reporting of conterventions; adherence to interventions; binding of carer, patient, and outcome assessor; use of relevant outcome measure; withdrawals and drop-outs with no substantial bias; <= 20% withdrawals for short-term follow up and <= 30% withdrawals for long term follow-up; identical timing of outcome assessment for all intervention groups; and intention-to-treat analysis); descriptive criteria (including specification of eligibility criteria; baseline similarity of groups; description of interventions; adverse effects described and attributed to allocated treatment; short term follow-up with outcome
assessed at end of intervention period; and long-term follow-up with outcome assessment at last 6 months after
randomisation); and statistical criteria (including sample size and presentation of point estimate and distribution
measures). Studies satisfying at least 50% of the validity criteria were classified as having ‘acceptable validity’ versus
other studies classified as having ‘low validity’. Studies with sufficient power of at least 0.8 (based on ES of 0.5) were
distinguished from studies with low power. Validity was assessed by two reviewers independently according to the
above criteria. Each internal validity criteria was scored as positive if bias was unlikely, negative if bias was likely or
inconclusive if there was insufficient or missing information. A total score for internal validity was calculated (range
of validity score from 0 to 12). Disagreements were resolved by consensus or with arbitration by a third reviewer. The
study that was written by one reviewer was assessed by the other reviewer and by another uninvolved assessor.

Data extraction

Quantitative data were extracted by one reviewer. Effect sizes (ES) and 95% confidence intervals were calculated using
Hedge’s ‘g’ statistic for continuous outcome measures and Cohen’s ‘h’ for differences in proportions. Calculation of
Hedge’s ‘g’ required calculation of the means and standard deviation for each group. In the absence of these data, ES
was calculated from Z score and sample sizes. If possible ES was based on change scores. Where these were lacking,
posttreatment scores were used. Power estimates for an ES of 0.2 and 0.5 were made.

Methods of synthesis

How were the studies combined?
The studies were combined in a narrative review.

How were differences between studies investigated?
Studies were weighted as to validity and their power levels and conclusions were based on studies that had both
acceptable validity and sufficient power.

Results of the review

Twelve studies concerning 10 trials were included.

Six RCTs satisfied at least 50% of the validity criteria. One trial did not satisfy any validity criteria. Methodological
flaws included: lack of control for cointerventions; lack of intention-to-treat analysis; and insufficient information on
methods used to allocate treatment, level of compliance, control for cointerventions, blinding of outcome assessment,
eligibility criteria and description of interventions. Sample size and power of studies varied widely. Two studies had
sufficient power (0.8 or more) to detect medium effects (ES = 0.5) and were of acceptable validity.

Most trials compared exercise therapy with either placebo therapy or no treatment.

Pain: 4 different outcome measures were used. Results in the two trials with acceptable validity and sufficient power
were ES (hip or knee OA) = 0.58 (95% CI: 0.54, 0.62) and ES for the other trial (knee OA only) = 0.47 (95% CI: 0.44,
0.50) for aerobic exercise and ES = 0.31 (95% CI: 0.28, 0.34) for resistance exercise. In both trials, subjects had
radiographic evidence of mild to moderate OA. Conflicting results were noted in the other 4 trials, with one favouring
exercise, one reporting ES of borderline significance, and 2 low validity, low power studies reporting no treatment
effect. Evidence indicates that exercise therapy has a small to moderate beneficial effect on pain in knee OA and to a
lesser extent hip OA.

Self-reported disability: Three different outcome measures were used. Results in the two trials with acceptable validity
and sufficient power were ES (hip or knee OA) = 0.26 (95% CI: 0.22, 0.30) and ES for the other trial (knee OA) = 0.41
(95% CI: 0.38, 0.44) for aerobic exercise and ES = 0.36 (95% CI: 0.33, 0.39) for resistance exercise. Results in the
other 3 trials (all low validity and low power) were conflicting with two favouring exercise and one favouring the
control therapy. Evidence indicates that exercise therapy has a small beneficial effect on self-reported disability in knee
OA and to a lesser extent hip OA.

Observed disability in walking: Four different outcome assessments were used. Results in the two trials with acceptable
validity and sufficient power were ES (hip or knee OA) = 0.28 (95% CI: 0.24,0.32) and ES for the other trial (knee
OA) = 0.89 (95% CI: 0.85, 0.93) for aerobic exercise and ES = 0.31 (95% CI: 0.28, 0.34) for resistance exercise. Results in the other two studies were conflicting with one (acceptable validity, low power) favouring the control intervention and the other (low validity, low power) trial favouring exercise. Exercise has a small beneficial effect on walking performance.

Patient’s global assessment of effect: Exercise has a medium to great beneficial effect according to one trial with acceptable validity and sufficient power reported ES (hip or knee OA) = 0.64 (95% CI: 0.60, 0.68). The other trial reporting this outcome (acceptable validity, low power) also favoured exercise.

Comparisons between different exercise therapy: None of the 4 trials comparing interventions had both acceptable validity and high power. No evidence was available in favour of one particular type of exercise programme.

**Authors' conclusions**

There is evidence of short-term beneficial effects of exercise therapy in patients with mild to moderate OA of the knee and, to a lesser extent, the hip. However, the small number of good studies restricts drawing firm conclusions.

**CRD commentary**

The aims and inclusion criteria were stated. No language restrictions were applied to primary studies. Methodology was rigorously assessed and results from this assessment presented. Methods used to assess validity and data extraction were described. Given the heterogeneity among studies, a narrative review was appropriate. The discussion included consideration of the following limitations of the review: conclusions based on small numbers; only two RCTs had an acceptable validity score as well as sufficient power; limited number of studies assessed observed disability and patient’s global assessment of effect; a number of different instruments were used to assess outcome measures; hardly any information was available on the long-term effects of exercise therapy; effectiveness of exercise therapy in patients with hip OA has hardly been studied; insufficient evidence was available to draw conclusions on the optimal content of an exercise therapy intervention; major threats to validity were present in the included studies such as lack of blinding of outcome assessors, absence of information on adherence to intervention, the quality of the outcome assessment instrument was not assessed, insufficient data was presented to calculate effect size, and intervention groups not comparable at baseline.

More comprehensive details of the literature search such as keywords employed would have been helpful as would fuller details of the characteristics of participants. Only published studies were included leaving the review open to publication bias. Only one reviewer extracted quantitative data.

The authors’ conclusions are supported by the evidence presented.

**Implications of the review for practice and research**

Practice: The authors consider that exercise therapy may be recommended for patient with OA of the knee and also for patients with OA of the hip with a mild to moderate stage of disease.

Research: The authors consider that clinical trials are needed to study the long-term effectiveness of exercise therapy and the effectiveness of exercise therapy in patient with hip OA. Future studies should attend to the following: sufficient sample size; adherence to exercise therapy; controls for cointerventions; blinded outcome assessment; an adequate data analysis including an intention to treat analysis; and the incorporation of a standard set of outcome measures (see Other Publications of Related Interest).

**Funding**

Dutch Fund of Investigative Medicine of the Dutch Health Insurance Council, grant number OG92.066.

**Bibliographic details**

PubMedID
10403263

DOI
10.1002/1529-0131(199907)42:7<1361::AID-ANR9>3.0.CO;2-9

Other publications of related interest

This additional published commentary may also be of interest. Reginster JY. Review: exercise therapy may reduce pain and disability in osteoarthritis of the hip or knee. Evid Based Med 2000;5:53.

Indexing Status
Subject indexing assigned by NLM

MeSH
Exercise Therapy; Humans; Osteoarthritis, Hip /therapy; Osteoarthritis, Knee /therapy; Randomized Controlled Trials as Topic; Reproducibility of Results

AccessionNumber
11999001430

Date bibliographic record published
30/09/2000

Date abstract record published
30/09/2000

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.