RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Effectiveness of exercise therapy for osteoarthritis: an overview of systematic reviews and randomised controlled trials

Tim Schleimer ,[1,2,3] Florian Teichert ,[1,3] Marius Henriksen,[2] Rebekka Doeding,[1] Tiziano Innocenti,[4,5] Helena Brisby,[6,7] Matthias C Klotz,[8] Marianne Korinth,[9] Patrick J Owen,[10,11] Dawid Pieper,[12,13] Daniel L Belavy[1]

To cite: Schleimer T, Teichert F, Henriksen M, *et al*. Effectiveness of exercise therapy for osteoarthritis: an overview of systematic reviews and randomised controlled trials. *RMD Open* 2026;**12**:e006275. doi:10.1136/rmdopen-2025-006275

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/rmdopen-2025-006275).

Received 20 August 2025
Accepted 15 December 2025

ABSTRACT

Objective To provide an overview of the effects of exercise for osteoarthritis.

Design Overview.

Data sources Medline, Embase, Epistemonikos, PEDro, Cochrane and registries from inception to 8 November 2025.

Eligibility criteria Reviews comparing exercise with placebo, no intervention or other interventions on pain and function for osteoarthritis. Supplementary trials were included to update inconclusive areas.

Data extraction and synthesis Two independent reviewers extracted data and assessed bias. Data were standardised to a 0–100 scale and reanalysed using random-effects meta-analysis. Certainty was rated using Grading of Recommendations Assessment, Development and Evaluation.

Results Five reviews (κ=100; n=8631) and 28 supplementary trials (knee/hip κ=23, hand κ=3, ankle κ=2; n=4360) were included. Evidence indicated small, short-term effects of exercise versus placebo (mean difference −10.8, 95% CI −19.1 to −2.6) and no-treatment (−12.4, 95% CI −15.6 to −9.2) for knee osteoarthritis pain, but certainty was very low and effects in larger or longer-term trials were smaller. Moderate evidence suggested negligible effects in hip (−6.7 95% CI −9.3 to −4.0) and small effects in hand (−10.0 95% CI −15.5 to −4.5) osteoarthritis. Varying certainty evidence indicated comparable outcomes to education, manual therapy, analgesics, injections and arthroscopy. Single trials in selected populations showed exercise was less effective than knee osteotomy (12.4 95% CI 4.7 to 20.2) and joint replacement (knee 17.1 95% CI 10.4 to 23.8; hip 24.2 95% CI 18.2 to 30.2) at longer term.

Conclusion and relevance Evidence on exercise for osteoarthritis remains largely inconclusive, suggesting negligible or short-lasting small effects comparable to, or less effective than, other treatments. These findings question its universal promotion and highlight the need to revisit research priorities and clinical discussions around its worthwhileness.

Registration CRD42023446888.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Exercise is recommended as first-line treatment for various types of osteoarthritis. However, the effectiveness of exercise for osteoarthritis is still controversial and emerging evidence suggests that the benefits of exercise may be overestimated. Despite a plethora of systematic reviews, an overview of the evidence on the effectiveness of exercise compared with a range of different comparators (ie, placebo, natural history/usual care, non-pharmacological and pharmacological treatments) is lacking.

WHAT THIS STUDY ADDS

⇒ Our study indicates that exercise therapy provides negligible or small effects in pain and physical function when compared with placebo or no treatment across knee, hip and hand osteoarthritis. The evidence is largely inconclusive and effects are clinically irrelevant in larger, longer-term trials. Scarce head-to-head evidence of varying certainty, mainly pertaining to knee osteoarthritis, suggests that exercise is comparable to pharmacological and non-pharmacological treatments, but less effective than osteotomy for medial compartment and total joint replacement for severe osteoarthritis.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The study highlights considerable uncertainties in the evidence on exercise for osteoarthritis, particularly concerning prevailing research practices (with ongoing publication and inclusion of small trials distorting effects), the long-term durability of effects, its effects compared with other active treatments and its effectiveness across less-studied osteoarthritis subtypes. Research should focus on priority questions with sound methodology to reliably inform policy decisions. Practice should consider discussing with patients the worthwhileness of exercise relative to other treatment options, incorporating preferences, harms and their own expertise.

Check for updates

For numbered affiliations see end of article.

Correspondence to
Tim Schleimer;
tim.schleimer@hs-bochum.de

INTRODUCTION

Despite the consistent recommendation[1 2] and widespread promotion[3 4] of exercise therapy as first-line treatment for patients with osteoarthritis, the underlying evidence is primarily based on studies comparing exercise therapy with no treatment.[5–7] While numerous systematic reviews with meta-analysis have synthesised this evidence,[8] uncertainty remains about the worthwhileness and durability of exercise effects, as reported improvements vary widely—from as little as five points[9 10] to as much as 20 points[11] when back-transformed to a standardised 0–100 scale—and most reviews report outcomes only in the short term.[11 12] Moreover, most of the evidence focuses on knee osteoarthritis, with considerably less available for other burdensome types such as hip, hand, ankle or shoulder osteoarthritis.[13 14] Furthermore, these studies have largely ignored how exercise therapy performs relative to other comparisons of interest. Doctors and physiotherapists often need to decide on when to provide one or another of different interventions in the clinic: general practitioners may consider exercise and/or analgesics; physiotherapists and orthopaedists may consider exercise, manual therapy or intra-articular injections, and orthopaedic surgeons have to make decisions on whether to offer surgery or continued conservative care.

The current evidence base lacks a comprehensive overview of the comparative effectiveness of exercise therapy for osteoarthritis, with existing overviews mainly summarising effect sizes for exercise versus no treatment across multiple overlapping reviews.[15 16] Without reliable comparative evidence, clinicians are unable to make optimal treatment decisions.[17] Thus, gathering knowledge about the effectiveness of exercise compared with various alternative treatments (ie, placebo, natural history/usual care, non-pharmacological and pharmacological treatments and surgery) is a critical next step to thoroughly inform decision-making of doctors and physiotherapists. Overviews of systematic reviews can address several questions within a topic, particularly when multiple reviews already exist and also allow for the critical reassessment of meta-analyses and updating with supplementary trials when the synthesised evidence is incomplete or outdated.[18] Thus, they allow for the integration of heterogeneous designs, making them a more suitable approach for broad research questions than network meta-analysis (NMA).[19] Consequently, overviews serve as a valuable tool for comprehensively informing decision making and highlighting uncertainties in the evidence through a single, accessible document.[20]

Thus, the aim of this overview was to provide the latest review, or where reviews were unavailable, randomised controlled trial (RCT) evidence on the comparative effects and durability of exercise on pain and physical function compared with a range of different comparators—including placebo, natural history/usual care, non-pharmacological and pharmacological treatments and surgery—across knee, hand, hip, shoulder and ankle osteoarthritis and to identify areas of remaining uncertainties.

METHODS

We prospectively registered the protocol for this overview in PROSPERO (CRD42023446888). Protocol deviations were minor (online supplemental data 3). We reported our results in accordance with the Preferred Reporting Items for Overviews of Reviews statement[18] (online supplemental data 2, table S2). Further, we followed methodological guidance by the Cochrane Collaboration.[21 22] The data and code of the underlying analyses can be found in an open science repository.

Patient and public involvement

Originally, we did not involve a patient research partner due to limited resources during the initial stages of the study. However, following suggestions during peer review, a patient research partner (MK) from the German Rheumatism League, with lived experience of osteoarthritis, contributed to the interpretation of the results and critically revised the manuscript.

Eligibility criteria

We included reviews and RCTs, published in peer-reviewed journals in the English language only, that examined the effects of exercise therapy in adult patients with either clinically or radiologically confirmed osteoarthritis of the knee, hip, hand, shoulder or ankle joint and that reported one of the core outcomes self-reported pain and/or physical function.[23–25] The broad inclusion criteria for osteoarthritis types were chosen to also provide evidence on joints that are under-researched yet in need of evidence synthesis to inform clinical practice and guide future research. Comparators of interest were placebo/sham/attention control (defined as such by the study authors[26]), no intervention (eg, wait-list) or other relevant interventions that are most likely to be prescribed as an alternative to exercise at various stages of care/disease (education, manual therapy, oral/topical non-steroidal anti-inflammatory drugs[NSAIDs], corticosteroids, viscosupplementation, platelet-rich plasma[PRP], prolotherapy, arthroscopic surgery, osteotomy, partial/total joint replacement[1]). We defined a systematic review as a review that reports a clearly stated research question, a systematic search in at least two databases, stated inclusion and exclusion criteria, selection methods conducted by at least two reviewers, critical appraisal of the included RCTs and presentation of at least one pairwise meta-analysis of aggregate data including at least two studies.[27 28] We excluded studies examining broader populations (eg, chronic musculoskeletal pain), presurgical and postsurgical rehabilitation interventions (unless outcomes for a timepoint prior to surgery were presented) and NMAs (as necessary information may be difficult to access from them).[18] Supplementary RCTs were included when: (1) the included reviews were deemed inconclusive (defined as a low or very-low certainty of evidence level after the initial re-analysis)[29] or (2) the included reviews provided incomplete coverage in relation to the scope of this overview (absence of a meta-analysis for certain type of OA or a relevant comparator intervention)[30 31] (online supplemental data 4, table S6).

Search methods

Information sources

To identify reviews, we searched MEDLINE, Epistemonikos, PEDro and the Cochrane Database of Systematic Reviews from inception to 31 October 2025.[32] Additionally, we conducted forward citation tracking of the included reviews[33] (online supplemental data 5). Supplementary RCTs were searched in MEDLINE, EMBASE and CENTRAL from inception to 8 November 2025.[21] In addition, we searched for registered studies in the WHO International Clinical Trials Registry Platform and clinicaltrials.gov (online supplemental data 6).

Search strategy

A sequential search strategy by searching first for reviews and second for RCTs was used. The search strategy for reviews combined the concepts (1) osteoarthritis, (2) exercise and (3) systematic review. We searched for RCTs from the reviews' last date of search if the included review was deemed inconclusive (new studies may change the certainty of evidence). If the review provided incomplete coverage regarding population or comparator (no meta-analysis), the terms for the population were modified (searched for the specific osteoarthritis type) or terms for the comparator intervention were added.

Data collection and analysis

Selection process

We used Covidence (Veritas Health Innovation, Melbourne, Australia) for automated deduplication and manual screening. Two authors (TS and FT) independently screened the titles and abstracts of the retrieved records against the inclusion criteria. The full texts of eligible studies were screened again to confirm eligibility for inclusion. Conflicts were resolved by consensus. If eligibility was confirmed, the prioritisation process was performed. The screening process was piloted on 20 articles.

Prioritisation and selection of reviews

We performed a prioritisation process to avoid conflicting results and double-counting among multiple reviews.[34 35] To prepare the prioritisation process, all eligible reviews were sorted by their population, intervention, comparison and outcome (PICOs) characteristics. The different exercise forms (eg, land-based, aquatic) were grouped together in the same category. Reviews that covered different types of osteoarthritis (eg, knee and hip) were classified according to the joint that was the primary focus of the majority of included trials. We had prospectively planned to select the most comprehensive review in each category. However, following new guidance on selecting the most appropriate reviews for overviews,[36] we post hoc expanded our selection criteria to include: (a) up-to-dateness (ie, reviews published within the last ten years) and (b) adequacy of the reported results. This was defined as (a) close alignment of the review with our research question (addressing the effectiveness of general exercise at different time points for pain and physical function), (b) data presented in a manner that allows for re-analysis and (c) no apparent inconsistencies in the data (online supplemental data 7, figure S1). The primary study overlap (corrected covered area, CCA) among the prioritised reviews was calculated in R statistical software V.4.2.1 (R Foundation for Statistical Computing, Vienna, Austria)[37] using the *ccaR* package[38] to rule out a moderate (6–10%) or high (≥11%) amount of overlapping study information.[39] We conducted a sensitivity analysis by describing the effect sizes among the non-prioritised reviews.

Data extraction and management

One author (TS) extracted the information on study characteristics using a customised Excel extraction sheet. Two authors (TS and FT) independently extracted outcome data for meta-analyses. The items extracted from the prioritised reviews were the main study characteristics, number of trials included, information on study designs and participant characteristics, judgements of risk of bias for each individual study, outcomes for pain and physical function for each time point, information on meta-analytic models, heterogeneity estimator and certainty of the evidence ratings. Items extracted from the supplementary RCTs were the main study characteristics, participant characteristics, description of the exercise intervention and comparator and outcomes for pain and physical function for each time point (online supplemental data 8). If available, we extracted the mean and SD at baseline and follow-up, the mean change scores and SDs from each group and the adjusted (based on analysis of covariance) or unadjusted mean between-group difference (MD). Data from intention-to-treat (ITT) analysis were extracted with prioritisation of data from analyses based on actual observations ('modified' ITT[40]). When available, we extracted outcomes for short-term (≥1 day but ≤3 months), medium-term (≥3 months but <12 months) and long-term (≥12 months). When the study assessed multiple time points that fell into a given time period, we extracted data closest to 3 months (post intervention/short-term), closest to 6 months (medium-term) and closest to 12 months (long-term). The equations published by Higgins *et al*[21] were used to convert the data if necessary (eg, SE to SD, 95% CI to SD). The conversion sheet used is available on the open science repository. If relevant data were presented in figures only, the web app *WebPlotDigitizer*[41] (available at https://automeris.io/WebPlotDigitizer/) was used to extract the data. When relevant information was not available within a study, the authors were contacted a minimum of three times over a 4-week period.

Assessment of methodological quality of included reviews

Quality of included reviews

The methodological quality of the included reviews was assessed with AMSTAR (A MeaSurement Tool to Assess systematic Reviews) version 2.[42] To arrive at an overall judgement, we interpreted no critical weakness as high confidence, one critical weakness as moderate confidence, two critical weaknesses as low confidence and three or more critical weaknesses as very-low confidence in the results of the review. Two authors (TS and RD) independently assessed the items and formed judgements of confidence by following an internal scheme (online supplemental data 9, table S27). The overall quality of the included reviews was classified as high, moderate, low or critically low.

Risk of bias of primary studies

The risk of bias assessments conducted by the included reviews was relied on, if they assessed bias using the original Cochrane Risk of Bias tool (ROB-1)[43] or the revised Risk of Bias tool (ROB-2).[44] As ROB-1 does not provide an algorithm for an overall rating, we defined studies as high risk of bias if more than one of the domains (selection bias, blinding, attrition bias, reporting bias, other bias) was rated as high risk. If a review used a different risk of bias tool, two authors (TS and FT) independently assessed bias of each included trial using ROB-2. The identified supplementary RCTs were also assessed using ROB-2. A combined rating was prepared for pain and physical function at the short-term follow-up time point. The effect of assignment to the intervention (ITT effect) was the specified effect of interest. The criteria used are outlined in online supplemental data 10.

Statistical analysis

A re-analysis of the original and updated (incorporating the supplementary RCTs) data of the included reviews was performed. To ensure adequate homogeneity, each osteoarthritis type, comparison and time point was analysed separately. For meta-analysis, a random effects (RE) model with the restricted maximum likelihood heterogeneity estimator was used to estimate the effect of exercise therapy on pain and physical function for each comparison.[45 46] The Hartung-Knapp Sidik-Jonkmann (HKSJ) adjustment was used to calculate the 95% CI around the pooled effect estimate.[47] As the HKSJ can be uninformative because the CI may be overly wide (when few studies are included in RE meta-analysis) or implausibly narrow (if true heterogeneity is equal to zero),[48] the suggestions by Schulz *et al*[49] were then followed to decide on the synthesis method (online supplemental data 12, figure S2).

The 95% prediction interval (95% PI) was computed to report statistical heterogeneity in meta-analyses including at least five studies. It represents the distribution of treatment effects that can be expected in a new trial that is similar to the ones included in the meta-analysis. Experts strongly recommend the use of 95% PIs to aid interpretation of results.[50 51] A bootstrapped approach[52] was used in the presence of <10 studies. If no 95% PI could be calculated, we visually examined the forest plot for apparent differences in trial-specific treatment effects ('qualitative heterogeneity'[53]), inspected whether the Q-test was significant[49] and whether τ was greater than six, which we considered indicative of moderate to high heterogeneity.

Data extracted from the included reviews were assessed for the common error of conflating SD with SE[54] leading to large outlying effect sizes (values ≥ three SDs from the mean). If an error was suspected, the primary study was re-examined and the error was corrected. When studies reported SDs that were unexpectedly small, the primary study authors were contacted for clarification.

To pool effect sizes for pain and physical function, all used scales were rescaled to a 0–100 scale by dividing the mean and SD by the range of the original scale and multiplying it by the range of the new scale.[55] The results of the conducted re-analyses and new analyses are presented in Forest plots. All analyses were performed in R statistical software using the packages *meta*,[56] *metafor*[57] and *dmetar*.[58] The data and code are available on an open data repository.

Sensitivity analyses

To assess the robustness of the findings, sensitivity analyses were conducted by removal of outlying studies[59] and studies with overall high risk of bias.[60] We decided post hoc to change the prespecified sensitivity analysis threshold for exclusion of small study samples, increasing it from 20 participants per arm to studies with <100 participants because many studies met the original threshold, and trials with at least 100 participants can be considered large in the conservative, non-pharmacological treatment field.[61] Further analysis included the use of the standard DerSimonean-Laird (DL) method to calculate 95% CI for the pooled effect.[49]

Subgroup analyses

Pending the number of studies (≥10 included in meta-analysis), prespecified subgroup analyses were performed. These included the treatment delivery mode (individually, class or home) and interventions that combined exercise with another conservative treatment.

Reporting bias

In the presence of ≥10 studies in meta-analysis, we conducted assessment for reporting bias as outlined in our certainty of evidence criteria (online supplemental data 13). In addition, extended funnel plots were created to assess the impact of a hypothetical new study on the overall meta-analytic effect estimate. This can inform future trial planning.[62] An extended funnel plot visualises which particular sample and effect size is needed to significantly influence the overall effect estimate.[63] This analysis was performed using the R package *extfunnel2*.[64]

Certainty assessment

Two authors (TS and FT) independently conducted new certainty assessments using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.[65] The certainty of evidence was rated as high, moderate, low or very low. The rating was initially set to a high level of certainty as only evidence from RCTs was included. The domain of indirectness was not judged as the eligibility criteria were strictly defined. We rated imprecision using a minimally contextualised approach.[66] For this, we prespecified the minimum clinically important difference (MCID), which is the smallest difference in outcome patients may perceive as clinically relevant,[67] as 8 points on a 0–100 scale to serve as the smallest effect threshold for pain and physical function. Given the lack of patient-centred, intervention-specific and context-dependent estimates of clinical relevance for interpreting treatment effects in osteoarthritis,[68] we used the conceptually more robust smallest worthwhile effect previously estimated in patients with chronic musculoskeletal conditions,[69] reasoning that this threshold is also likely to be considered worthwhile by patients with osteoarthritis for a low-risk, low-cost intervention such as exercise. To better capture the magnitude of effects, we classified effects as follows: <8 points as clinically irrelevant, 8–<12.5 points as small, 12.5–<20 points as moderate and ≥20 points as large. These thresholds, when back-transformed using the median pooled SD of 25 from large osteoarthritis trials,[70] align with Cohen's rule-of-thumb thresholds for moderate and large effect sizes.[71] We rated inconsistency in accordance with the Confidence in Network Meta-Analysis, which is also applicable to pairwise meta-analysis and overcomes limitations such as relying solely on I^2 to interpret heterogeneity.[72] The informative statements suggested by Santesso *et al*[73] were used to communicate the findings and facilitate interpretation. The detailed criteria are outlined in online supplemental data 13.

RESULTS

Review selection process and characteristics

The search identified 3028 unique records, from which 155 reviews were assessed at full text, 55 were classified as eligible and 14 of these were considered for the prioritisation process (figure 1; Reasons for exclusion in online supplemental data 14, table S30). After the prioritisation (online supplemental data 14, table S31), we included five reviews (two Cochrane and three non-Cochrane reviews),[74–78] involving 100 unique RCTs on three different treatment comparisons (exercise compared with placebo κ=1, no treatment/usual care κ=3, and education κ=1) with a total of 8631 patients across three different types of osteoarthritis (knee κ=2, hip κ=1, hand k=1, mixed κ=1) (table 1). The primary study overlap among the reviews

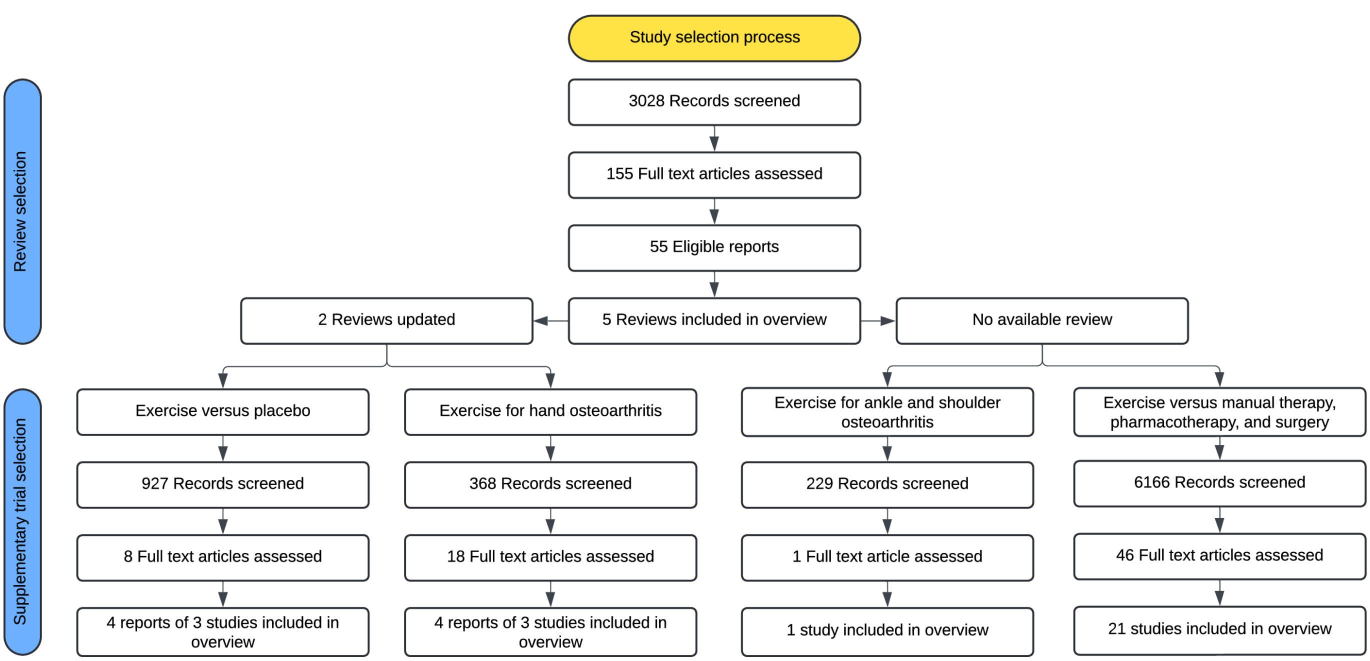

Figure 1 Modified Preferred Reporting Items for Systematic Reviews and Meta-Analysis diagram of the systematic review and supplementary trial selection process. The overview employed a sequential search strategy: (1) identification of eligible systematic reviews (Boxes 1–5), which were prioritised based on comprehensiveness, currency and adequacy of reported results and (2) supplementary searches for randomised controlled trials (Boxes 6–9) when included reviews were incomplete, required updating or did not address certain osteoarthritis subtypes or comparator interventions. In step 1, five reviews were included, of which two were considered incomplete (exercise vs placebo; exercise for hand osteoarthritis). No eligible reviews were identified for exercise in ankle or shoulder osteoarthritis or for comparisons with manual therapy, pharmacotherapy or surgery, which prompted step 2, the supplementary search for primary trials.

was characterised as slight (CCA 0.2%) (online supplemental data 14, figure S4).

Supplementary trial selection and characteristics

We performed supplementary RCT searches to update two of the included reviews[76 77] and to fill areas lacking an eligible review (exercise vs manual therapy, pharmacotherapy and surgery; exercise for ankle and shoulder osteoarthritis) (online supplemental data 14, table S33). Overall, the search for RCTs yielded 9425 records. After screening, we included 30 reports from 28 studies[79–108] (figure 1; Reasons for exclusion in online supplemental data 15, table S34), involving twelve different treatment comparisons (exercise vs no treatment κ=3, placebo κ=3, manual therapy κ=3, NSAIDs κ=2, corticosteroids κ=2, viscosupplements κ=5, PRP κ=2, prolotherapy κ=1, arthroscopy κ=2, knee osteotomy κ=1, total knee replacement (TKR) κ=2 and total hip replacement (THR) κ=1) with a total of 4360 patients (median sample size was 105) across four different types of osteoarthritis (knee κ=20, hip κ=2, knee and hip κ=1, hand κ=3, ankle κ=2). The mean age across all the trials was 63, the mean proportion of female participants was 65%, the mean body mass index was 29, the median baseline pain intensity and physical function (on a standardised 0–100 scale; lower values are better) were 44 and 38, respectively. Most patients in non-surgical trials had mild to moderate radiographic severity of osteoarthritis (Kellgren-Lawrence grades 2–3), whereas patients included in surgical trials were selected for medial-compartment involvement or severe osteoarthritis (table 2). One study is awaiting classification due to pending author response.[109] Two additional studies[110 111] were identified through experts but could not be included due to being published in non-English languages only (online supplemental data 15, tables S35 and 36). Five ongoing studies were identified through the search of trial registries, four of which compare exercise to surgery (knee osteotomy κ=1, TKR κ=1, shoulder arthroplasty κ=1, wrist denervation κ=1) (online supplemental data 15, table S37).

Risk of bias in systematic reviews

Information on the AMSTAR-2 ratings for the included reviews can be found in online supplemental data 16. Overall, no review was rated as high or moderate quality; two were rated as being of low quality[74 76] and three reviews[75 77 78] as being of critically low quality. Common critical issues were the use of inappropriate methods for statistical analysis (item 11) and the inadequate assessment of heterogeneity (item 14).

Risk of bias in RCTs

Information on the risk of bias assessments can be found in online supplemental data 17. Overall, no study was rated as low risk of bias, 20 (71%) studies were at some concerns, and eight (29%) studies were rated as high risk of bias. Common sources of bias were the measurement of the outcome (in none of the studies were patients blinded) and the selection of the reported results (eight trials were retrospectively registered).

Results of syntheses

Data handling and synthesis

For the re-analyses, the shared statistical code of two reviews[74 76] and the raw data presented in the forest plot of another review[78] were used. Two reviews[75 77] only presented precalculated effect sizes (final effect size with 95% CI) of which the raw data could be obtained from the authors of one review.[75] For the review by Dean *et al*,[77] the data were extracted from included primary studies. During the re-evaluation of the data presented by the original reviews, we identified some inconsistencies that namely involved ignoring outliers (eg, SE instead of SD),[77 78] use of percentage scores in meta-analysis,[77] use of baseline rather than post-treatment scores,[75] inappropriate exclusion of studies with available data (eg, data available in figures, data in precalculated effect size format)[75 77 78] and minor data extraction or conversion errors.[74 78] After contacting the authors, one review published a correction.[75 112] Detailed information on the data management for the re-analyses including the supplementary RCTs is in online supplemental data 18 and 19. For conciseness, we report effect sizes for exercise compared with no intervention, non-pharmacological comparators and NSAIDs for the short term, intra-articular interventions for the short or medium term (depending on where the evidence could be meaningfully synthesised) and surgical comparators for the longest follow-up term, unless stated otherwise. Forest plots for each comparison can be found in online supplemental data 20. An overview of the main and the sensitivity results is presented in figures 2 and 3 and the corresponding GRADE ratings with the results for each time point can be found in online supplemental data 21.

Exercise therapy versus placebo, sham or attention controls

Mixed osteoarthritis (Knee κ=6, hip κ=2 and hand joint κ=3)

When compared with placebo, exercise therapy may result in a small reduction in pain (mean difference −10.8, 95% CI −19.1 to −2.6, 95% PI −37.5 to 15.8, n=1037, κ=11; very-low certainty) but may not increase physical function (MD −4.3, 95% CI −9.7 to 1.2, 95% PI −17.1 to 8.4, n=945, κ=9; very-low certainty) in the short term but the evidence is very uncertain. Exercise probably results in little to no difference in pain (high certainty) and function (low certainty) in the long term.

Exercise therapy versus no intervention

Knee osteoarthritis

When compared with no intervention, exercise therapy may result in a small effect in pain (MD −12.4, 95% CI −15.6 to −9.2, 95% PI −34.5 to 9.7, n=4184, κ=56; very-low certainty) and physical function (MD −10.0, 95% CI −12.8 to −7.2, 95% PI −29.0 to 9.0, n=4245, κ=54; very-low certainty) in the short term but the evidence is very uncertain. Exercise may result in little to no difference in

Table 1 Characteristics of the included systematic reviews (SRs)

Study	Review type	Protocol/ registered	N databases	Last search	OA type	Intervention	Comparison	N trials included in SR	Baseline characteristics	Outcomes assessed	Time points assessed
Dean *et al*[77]	Non-Cochrane	Yes	4	March 2021	Knee, hip and hand (clinical/ radiological)	Land-based exercise therapy alone or in combination with physical therapy	Placebo (interventions with no plausible therapeutic effect)	13	Age: 67 years Female (%): 79 BMI: 29 kg/m^2 Pain: 44±16 Function: 38±15	Pain Physical function	Short Long
Lawford *et al*[74]	Cochrane	Yes	5	January 2024	Knee (clinical/self-report)	Land-based exercise therapy	No treatment, usual care or limited education	139 (60 eligible)	Age: 41–81 years Female (%): 36–100 BMI: 22–38 kg/m^2 Pain: 51±19 Function: 42±17	Pain Physical function	Short Long
Goff *et al*[78]	Non-Cochrane	Yes	5	April 2020	Knee (clinical/ radiological)	Any form of patient education	Exercise therapy alone	28 (eight eligible)	Age: 67 years Female (%): 74 BMI: 29 kg/m^2 Pain: 35±16 Function: 38±17	Pain Physical function	Short Medium Long
Østerås *et al*[76]	Cochrane	Yes	6	September 2015	Hand (clinical/ radiological)	Exercise therapy alone or in combination with other treatment modalities	No exercise (eg, usual care, waiting list)	7 (six eligible)	Age: 66 years Female (%): 86 BMI: 28 kg/m^2 Pain: 45±19 Function: 36±16	Pain Physical function	Short Medium Long
Teirlinck *et al*[75]	Non-Cochrane	No	6	September 2021	Hip (clinical/ radiological)	Standalone exercise therapy under supervision of (physical) therapist	No treatment, usual care or waiting list	18	Age: 64 years Female (%): 63 BMI: 28 kg/m^2 Pain: 38±17 Function: 40±16	Pain Physical function	Short Long

BMI, body mass index; OA, osteoarthritis.

Table 2 Characteristics of included supplementary trials

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Exercise therapy versus placebo, sham or attention controls for mixed osteoarthritis										
Bandak *et al*; Henriksen *et al*[79 80]	Denmark	Knee OA; Clinical (ACR) and X-ray (KL≥2)	I: group-based exercise and education (GLA:D) for 8 weeks (2× per week; 1 hour) C: IA saline injections (4× over 7 weeks)	206	46	69±8	27±4	II=28% III=32% IV=41%	KOOS-P KOOS-F	S, L
Messier *et al*[81]	USA	Knee OA; X-ray (KL≥2)	I1: high-intensity strength training I2: low intensity strength training C: attention control	377	40	65±8	31±6	II=49% II=41% IV=10%	Womac-P Womac-F	L
Messier *et al*[82]	USA	Knee OA; ACR criteria	I: exercise and intensive dietary programme C: attention control	823	78	65±8	37±7	NA	Womac-P Womac-F	L
Exercise therapy versus usual care (eg, advice) or waiting list for hand osteoarthritis										
Gravås *et al*; Tveter *et al*[83 84]	Norway	Hand OA; Surgical waiting list	I: one session exercise instruction, education and orthoses C: usual care	180	79	63±8	NA	Median: 3	NRS QuickDASH	S
Deveza *et al*[85]	Australia	Hand OA; Clinical and X-ray (KL≥2)	I: two individualised sessions containing exercises, education, splint and diclofenac sodium C: education	204	76	66±9	29±7	II=43% III=39% IV=18%	VAS FIHOA	S
Magni *et al*[86]	New Zealand	Hand OA; Clinical and X-ray (KL≥1)	I1: high intensity training and advice I2: blood flow restriction training and advice C: advice	59	83	72±10	26±5	≥2 = 63%	NRS DASH	S
Exercise therapy versus usual care (eg, advice) or waiting list for ankle osteoarthritis										

Continued

Table 2 Continued

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Smith *et al*[87]	Australia	Ankle OA; Clinical and X-ray (KL≥2)	I: group education plus exercise (12 sessions of 1 hour each over 7 weeks)	30	67	66±12	NA	Median: 3	NRS FAOS ADL	S
			C: general advice							
Exercise therapy versus manual therapy for knee and hip osteoarthritis										
Abbott *et al*[88]	New Zealand	Knee and hip OA; Clinical (ACR)	I: exercise therapy including strengthening, aerobic and range of motion exercises for 9 weeks (and two booster sessions after 16 weeks)	105	58	67±9	29±6	NA	NRS	L
			C: manual therapy							
Dwyer *et al*[89]	South Africa/ USA	Knee OA; Clinical	I: exercise (12 sessions of 20 min), education and manual therapy	52	66	63±11	30±6	I=74% II: 27%	Womac-P Womac-F	S
			C: manual therapy (six sessions over 4 weeks)							
Hoeksma *et al*[90]	Netherlands	Hip OA; Clinical (ACR)	I: supervised exercise therapy for 5 weeks (nine sessions)	109	70	72±7	NA	0=8% I=12% II=39% III=41%	VAS SF-36 physical function	S, M
			C: manual therapy for 5 weeks (nine sessions)							
Exercise therapy versus NSAIDs for knee osteoarthritis										
Doi *et al*[91]	Japan	Knee OA; Clinical (ACR) and X-ray	I: home-based quadriceps exercises	142	76	68±17	25±4	NA	VAS Womac-F	S
			C: NSAIDs (three times per day)							
Yang *et al*[92]	China	Knee OA	I: neuromuscular exercise 3× per week for 12 weeks plus NSAIDs	60	71	61±6	28±3	II=55% III=45%	VAS KOOS-F	S, M
			C: NSAIDs (Celecoxib 200 mg; once per day for 12 weeks)							
Exercise therapy versus PRP injections for knee osteoarthritis										

Continued

Table 2 Continued

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Angoorani *et al*[93]	Iran	Knee OA; Clinical and X-ray (KL 1–3)	I: home-based resistance exercise and 10 sessions electrophysical modalities	54	88	62±10	29±4	NA	KOOS-P KOOS-F	S
			C: IA PRP injections (2×4 weeks apart)							
Argut *et al*[94]	Turkey	Knee OA; Clinical and X-ray (KL 2–3)	I: individual supervised exercise programme (2× per week for 6 weeks, 40 min per session, 12 sessions overall)	56	75	57±7	28±4	II=45% III=56%	NRS Womac-F	S, M
			C: IA PRP injections (3×1 week apart)							
Exercise therapy versus corticosteroid injections for knee osteoarthritis										
Deyle *et al*[95]	USA	Knee OA; Clinical and X-ray (KL≥1)	I: exercise and manual therapy (eight sessions over 4–6 weeks (+1–3 booster sessions after 4 months and 9 months)	156	48	56±9	32±6	I=4% II=44% III=38% IV=15%	Womac-P Womac-F	S, M, L
			C: IA corticosteroid injections (maximum of three injections over trial period)							
Lee *et al*[96]	China	Knee OA; Clinical and X-ray	I: exercise and kinesio taping (4–5× per week for 6 weeks)	287	42	58±4	28±3	I=28% II=48% III=19% IV=7%	Womac-P Womac-F	S, M, L
			C: IA corticosteroid injections (maximum of three injections over trial period)							
Exercise therapy versus viscosupplementation for knee and ankle osteoarthritis										

Continued

Table 2 Continued

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Karatosun *et al*[97]	Turkey	Knee OA; Clinical and X-ray (KL 3)	I: home-based flexibility and resistance exercises directed to the knee for 6 weeks	105	86	57±13	28±6	III=100%	VAS during activity	S, M, L
			C: IA hyaluronic acid injection (3×1 week apart)							
Kawasaki *et al*[98]	Japan	Knee OA; Clinical	I: home-based flexibility and isometric exercises	87	100	71±8	26±4	NA	VAS	M
			C: intra-articular hyaluronic acid injection (1× per week for 5 weeks; then 1 per month for 19 weeks)							
Saccomanno *et al*[99]	Italy	Knee OA; Clinical and X-ray	I: knee exercise programme (5× per week)	104	73	62±12	28	I=67% II=19% III=15%	Womac-P Womac-F	S, M
			C: IA hyaluronic acid injection (three injections 2 weeks apart)							
Kapoor and Mudi [100]	India	Knee OA; X-ray (KL 1–3)	I: quadriceps and hamstring strengthening exercises	248	NA	54±8	23	I=14% II=67% III=20%	VAS Womac-F	S, M
			C: IA hyaluronic acid injection (single injections)							
Karatosun *et al*[101]	Turkey	Ankle OA; Clinical and X-ray (KL 3)	I: home-based flexibility & resistance exercises directed to the ankle for 6 weeks	30	70	55±12	NA	III=100%	VAS during activity AOFAS	S, M, L
			C: IA hyaluronic acid injection (three injections 1 week apart)							
Exercise therapy versus dextrose injections for knee osteoarthritis										

Continued

Table 2 Continued

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Rabago *et al*[102]	USA	Knee OA; Clinical and X-ray	I: home exercise programme (10 exercises, 1–3× per week, 1–3 per day)	61	66	57±8		I-II=42% III - IV=58%	Womac-P Womac-F	S, M, L
			C: intra-articular dextrose injections (3× over 9 weeks with two optional injections at 13 weeks and 17 weeks)							
Exercise therapy versus arthroscopic surgery for knee osteoarthritis										
Kirkley *et al*[103]	Canada	Knee OA; X-ray (KL≥2)	I: supervised physical therapy including aerobic and strengthening exercise and education for 12 weeks (1× per week; 1 hour)	178 (X-over=0%)	62	60±10	31±7	II=44% III=51% IV=5%	Womac-P Womac-F	S, M, L
			C: arthroscopic surgery and postoperative physical therapy							
Katz *et al*[104]	USA	Knee OA; Clinical and X-ray/MRI	I: supervised physical therapy including exercise, manual therapy and electrophysical modalities for 6 weeks	351 (X-over=35%)	57	59±8	30±6	0=21% I=19% II=23% III=26%	Womac-P Womac-F	S, M, L
			C: arthroscopic surgery and postoperative physical therapy							
Exercise therapy versus osteotomy for knee osteoarthritis										

Continued

Table 2 Continued

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Birmingham *et al*[105]	Canada	Knee OA; Clinical and X-ray	I: non-surgical management including exercise, weight management and pain medications over 3 months	71 (X-over=0%)	25	55±7	31±5	I=12% II=34% III=55%	KOOS-P KOOS-F	L
			C: medial opening wedge high tibial osteotomy plus non-surgical management							
Exercise therapy vs joint replacement for knee osteoarthritis										
Skou *et al*[106]	Denmark	Knee OA; X-ray (KL≥2)	I: group-based neuromuscular exercise, education and dietary advice (GLA:D programme) for 12 weeks	100 (X-over=26%)	62	67±9	32±6	II=12% III=42% IV=46%	KOOS-P KOOS-F	S, M, L
			C: total knee replacement plus GLA:D programme							
Rognsvåg *et al*[107]	Norway	Knee OA; X-ray (KL 3–4)	I: education (60 min session), exercise therapy (2×60 min sessions per week for 12 weeks led by trained therapist) and internet delivered cognitive behavioural therapy (10 modules at home)	15	53	64±7	30±4	NA	KOOS-P KOOS-F	S, M, L
			C: total knee replacement plus conservative programme or total knee replacement followed by standard care							
Exercise therapy vs joint replacement for hip osteoarthritis										

Continued

Table 2 Continued

Study ID	Country	Population	Intervention and comparator	N participants	% female	Mean age in years (SD)	Mean BMI (SD)	Radiographic severity (KL)	Pain/function outcome	FU-data availability
Frydendal *et al*[108]	Denmark	Hip OA; Clinical and X-ray	I: resistance training individually delivered by physiotherapists (2× per week; 1 hour) for 12 weeks.	110	50	68±7	28±4	JSN<2 mm	HOOS-P HOOS-F	M, L
			C: total hip replacement							

ACR, American College of Rheumatology; AOFAS, American Orthopaedic Foot and Ankle Society; BMI, body mass index; DASH, Disability of the Arm, Shoulder, and Hand outcome measure; FAOS, Foot and Ankle Outcome Score Activites; FIHOA, Functional Index for Hand Osteoarthritis; FU, follow-up; GLA:D, Good Life with Osteoarthritis in Denmark; HOOS, Hip disability and Osteoarthritis Outcome Score; IA, intra-articular; JSN, joint space narrowing; KL, Kellgren-Lawrence; KOOS-P/F, Knee Injury and Osteoarthritis Outcome Score Pain/Function; L, long term; M, medium term; NRS, Numeric Rating Scale; NSAID, non-steroidal anti-inflammatory drugs; OA, osteoarthritis; PRP, Platelet-rich plasma; S, short term; SF-36, Short form 36 health survey; VAS, Visual Analogue Scale; Womac-P/F, Western Ontario and McMaster Universities Arthritis Index Pain/Function; X-over, cross-over to surgical group at primary time point.

pain (low certainty) and function (moderate certainty) in the long term.

Hip osteoarthritis

When compared with no intervention, exercise therapy probably results in little to no difference in pain (MD −6.7, 95% CI −9.3 to −4.0, 95% PI −13.3 to 0.0, n=1261, κ=16; moderate certainty) and physical function (MD −5.8, 95% CI −8.4 to −3.3, 95% PI −13.0 to 1.4, n=1233, κ=16; moderate certainty) in the short term. Exercise probably results in little to no difference in pain and function in the long term (moderate certainty).

Hand osteoarthritis

When compared with no intervention, exercise therapy probably results in a small reduction in pain (MD −10.0, 95% CI −15.5 to −4.5, 95% PI −24.0 to 4.1, n=804, κ=8; moderate certainty) but probably does not increase physical function (MD −6.2, 95% CI −10.0 to −2.4, 95% PI −15.6 to 2.5, n=792, κ=7; moderate certainty) in the short term. Exercise may result in little to no difference in pain and function in the medium and long term (low certainty).

Exercise therapy versus education

Knee osteoarthritis

When compared with education, exercise therapy probably results in little to no difference in pain (MD −4.6, 95% CI −8.5 to −0.7, 95% PI −14.6 to 4.3, n=399, κ=5; moderate certainty) and physical function (MD −4.0, 95% CI −12.3 to 4.4, τ=1.5, n=230, κ=3; moderate certainty) in the short term. Exercise probably results in little to no difference in pain and function in the medium and long term (moderate certainty). Data from a small pilot trial (n=30)[87] comparing exercise to education for ankle osteoarthritis suggested similar results (low certainty).

Exercise therapy versus manual therapy

Knee and hip osteoarthritis

When exercise was compared with manual therapy, results from single trials[89 90] favoured manual therapy, indicating small reductions in pain in the short and medium term (low certainty). No important effects were found in the long term[88] or for physical function (low certainty).

Exercise therapy versus oral NSAIDs

Knee osteoarthritis

When compared with NSAIDs, exercise therapy may result in little to no difference in pain (MD −5.3, 95% CI −11.4 to 0.7, τ=0, n=177, κ=2; low certainty) and physical function (MD −7.1, 95% CI −12.1 to −2.0, τ=0, n=181, κ=2; low certainty) in the short term. Exercise may result in a small effect in pain (low certainty) and a moderate effect in function (moderate certainty) in the medium term.

Schleimer T, *et al*. *RMD Open* 2026;**12**:e006275. doi:10.1136/rmdopen-2025-006275

Osteoarthritis type and comparator	Trials (N)	MD	95%CI	GRADE
Mixed osteoarthritis				
Placebo (all trials)	11 (1037)	−10.8	(−19.1 to −2.6)	Very−low
Placebo (N >=100 participants)	5 (716)	−3.3	(−6.3 to −0.3)	
Knee osteoarthritis				
No treatment (all trials)	56 (4184)	−12.4	(−15.6 to −9.2)	Very−low
No treatment (N >=100 participants)	13 (2096)	−6.8	(−9.4 to −4.1)	
Therapeutic education	5 (399)	−4.6	(−8.5 to −0.7)	Moderate
Manual therapy	1 (100)	5.6	(−5.9 to 17.1)	Low
NSAIDs	2 (177)	−5.3	(−11.4 to 0.7)	Low
Corticosteroid injections	2 (428)	−4.8	(−8.5 to −1.1)	Moderate
Viscosupplementation	4 (527)	−1.3	(−3.5 to 0.8)	Moderate
PRP injections	1 (50)	6.5	(−3.4 to 16.4)	Very−low
Prolotherapy	1 (55)	7.4	(−2.5 to 17.3)	Low
Arthroscopic surgery	2 (444)	−0.8	(−7.1 to 5.5)	High
Osteotomy	1 (57)	12.4	(4.7 to 20.2)	Low
Total joint replacement	1 (95)	17.1	(10.4 to 23.8)	Moderate
Hip osteoarthritis				
No treatment (all trials)	16 (1261)	−6.7	(−9.3 to −4.0)	Moderate
No treatment (N >=100 participants)	6 (764)	−4.0	(−8.2 to 0.1)	
Manual therapy	1 (94)	12.1	(2.5 to 22.9)	Low
Total joint replacement	1 (103)	24.2	(18.2 to 30.2)	Moderate
Hand osteoarthritis				
No treatment (all trials)	8 (804)	−10.0	(−15.5 to −4.5)	Moderate
No treatment (N >=100 participants)	4 (598)	−8.0	(−16.4 to 0.4)	
Ankle osteoarthritis				
No treatment	1 (26)	−8.0	(−24.8 to 8.8)	Low
Viscosupplementation	1 (15)	−11.0	(−29.7 to 7.7)	Very−low

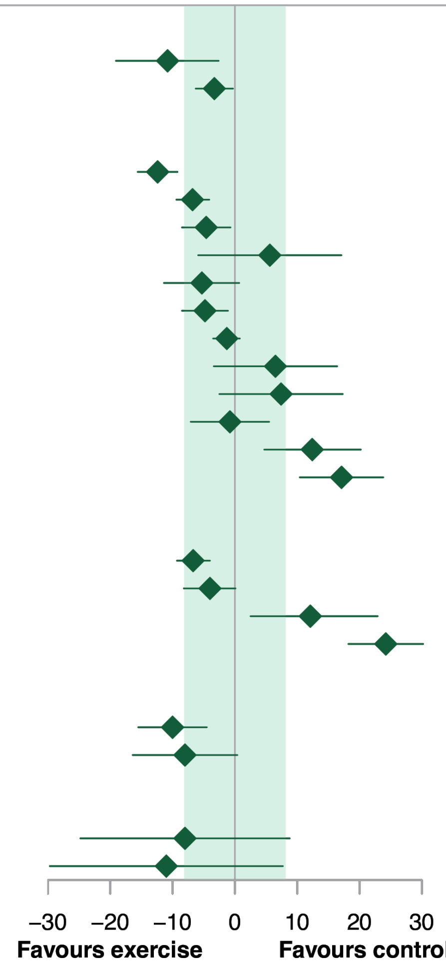

Figure 2 Main and sensitivity analysis (restricting analyses to large trials) for the outcome pain. For comparisons with meta-analyses, time points with the largest number of studies are shown. For comparisons with only narrative synthesis or single trials, the study or time point with the largest number of participants is reported. The shaded area represents the zone of trivial effects (−8< MD < 8). GRADE, Grading of Recommendations Assessment, Development and Evaluation; MD, mean between-group difference; NSAID, non-steroidal anti-inflammatory drug; PRP, Platelet-rich plasma.

Exercise therapy versus intra-articular injections

Knee osteoarthritis

When compared with corticosteroids, exercise therapy probably results in little to no difference in pain (MD −4.8, 95% CI −8.5 to −1.1, τ=0, n=428, κ=2, moderate certainty) and physical function (MD −2.9, 95% CI −7.2 to −1.4, τ=0.8, n=428, κ=2, high certainty) in the short term. When compared with viscosupplementation, exercise therapy probably results in little to no difference in pain (MD −1.3, 95% CI −3.5 to 0.8, τ=0, n=527, κ=4, moderate certainty) in the medium term, while the evidence for physical function is very uncertain. Evidence comparing exercise with prolotherapy, PRP and viscosupplementation (for ankle osteoarthritis) was very uncertain, with heterogeneous findings from single trials indicating moderate, small or no differences between interventions.

Exercise therapy versus arthroscopic surgery

Knee osteoarthritis

When compared with arthroscopy, exercise therapy results in little to no difference in pain (MD −0.8, 95% CI −7.1 to 5.5, τ=0, n=444, κ=2; high certainty) and physical function (MD −0.5, 95% CI −4.3 to 3.4, τ=0, n=441, κ=2; high certainty) in the long term.

Exercise therapy versus osteotomy

Knee osteoarthritis

When exercise was compared with medial opening wedge high tibial osteotomy, small effects in pain (MD 12.4, 95% CI 4.7 to 20.2, n=57, κ=1; low certainty) were found in the long term in favour of osteotomy, but no important effect was observed for physical function (MD 5.8, 95% CI 1.6 to 9.9, n=57, κ=1; low certainty).

Exercise therapy versus total joint replacement

Knee osteoarthritis

When exercise was compared with TKR, moderate effects in pain (MD 17.1, 95% CI 10.4 to 23.8, n=95, κ=1; moderate certainty) and physical function (MD 12.9, 95% CI 6.8 to 19.1, n=95, κ=1; low certainty) were found in the long term in favour of TKR. Data from a small pilot trial (n=15)[107] assessing the feasibility of a large ongoing

Osteoarthritis type and comparator	Trials (N)	MD	95%CI	GRADE
Mixed osteoarthritis				
Placebo (all trials)	9 (945)	−4.3	(−9.7 to 1.2)	Very−low
Placebo (N >=100 participants)	5 (716)	−1.7	(−5.2 to 1.8)	
Knee osteoarthritis				
All trials	54 (4245)	−10.0	(−12.8 to −7.2)	Very−low
>=100 participants	14 (2229)	−7.3	(−10.3 to −4.4)	
Therapeutic education	3 (230)	−4.0	(−12.3 to 4.4)	Moderate
Manual therapy	1 (51(	2.9	(−5.5 to 11.3)	Low
NSAIDs	2 (181)	−7.1	(−12.1 to −2.0)	Low
Corticosteroid injections	2 (428)	−2.9	(−7.2 to 1.4)	High
Viscosupplementation	1 (248)	5.1	(3.8 to 6.4)	Very−low
PRP injections	1 (50)	10.2	(−0.6 to 21.0)	Very−low
Prolotherapy	1 (55)	7.9	(−1.4 to 17.2)	Low
Arthroscopic surgery	2 (441)	−0.5	(−4.3 to 3.4)	High
Osteotomy	1 (57)	5.8	(1.6 to 9.9)	Low
Total joint replacement	1 (95)	12.9	(6.8 to 19.1)	Low
Hip osteoarthritis				
All trials	16 (1233)	−5.8	(−8.4 to −3.3)	Moderate
>=100 participants	6 (764)	−4.4	(−7.9 to −0.9)	
Manual therapy	1 (94)	−2.1	(−11.7 to 7.7)	Low
Total joint replacement	1 (103)	20.8	(15.2 to 26.3)	Moderate
Hand osteoarthritis				
All trials	7 (292)	−6.2	(−10.0 to −2.4)	Moderate
>=100 participants	4 (602)	−5.4	(−11.9 to 1.2)	
Ankle osteoarthritis				
No treatment	1 (26)	4.3	(−3.4 to 12.0)	Low
Viscosupplementation	1 (15)	−3.8	(−13.5 to 5.9)	Very−low

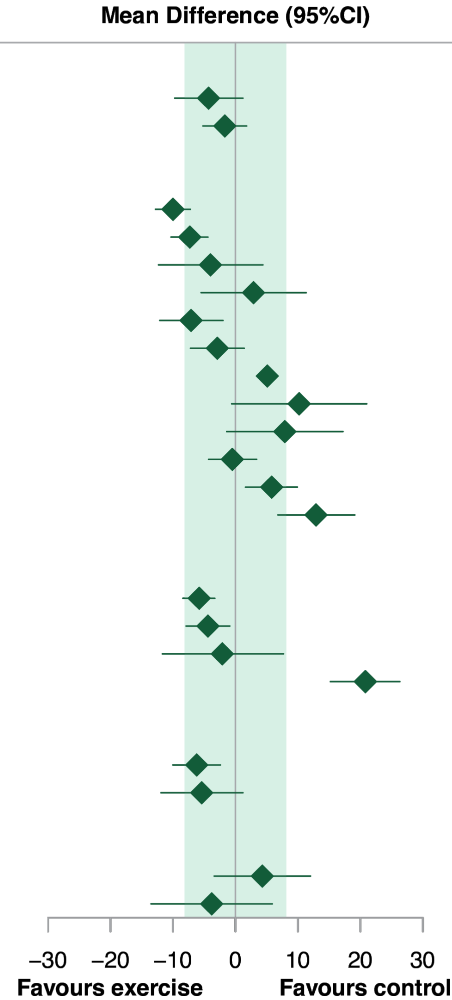

Figure 3 Main and sensitivity analysis (restricting analyses to large trials) for the outcome physical function. For comparisons with meta-analyses, time points with the largest number of studies are shown. For comparisons with only narrative synthesis or single trials, the study or time point with the largest number of participants is reported. The shaded area represents the zone of trivial effects (−8 <MD < 8). GRADE, Grading of Recommendations Assessment, Development and Evaluation; MD, mean between-group difference; NSAID, non-steroidal anti-inflammatory drug; PRP, Platelet-rich plasma.

RCT comparing conservative treatment (including exercise) to TKR suggested similar directions of effect.

Hip osteoarthritis

When exercise was compared with THR, large effects in pain (MD 24.2, 95% CI 18.2 to 30.2, n=103, κ=1; moderate certainty) and physical function (MD 20.8, 95% CI 15.2 to 26.3, n=103, κ=1; moderate certainty) were found in the medium term in favour of THR.

Sensitivity and subgroup analyses

The nine reviews that were not prioritised showed effect sizes mostly similar to the prioritised reviews (online supplemental data 14, table S32). In the subgroup analyses, only the treatment delivery mode in knee osteoarthritis showed significant differences, in favour of individually delivered exercises for short-term physical function improvements (online supplemental data 22). Sensitivity analyses on removing small studies (<100 participants) yielded markedly lower values in short-term pain for the comparisons of exercise vs no intervention and placebo. Use of the standard DL approach to calculate 95% CI tended to increase the precision in analyses, which included few trials (online supplemental data 23).

Reporting bias

Funnel plots and results of Egger's regression tests for analyses with ≥10 trials are presented in online supplemental data 24. Visual and statistical evidence of small-study bias was observed in analyses of short-term pain and physical function in knee osteoarthritis, resulting in a downgrade in the certainty of evidence. As some contrasts included only single studies, the certainty was downgraded due to early positive findings.

An extended funnel plot for short-term pain in knee osteoarthritis was not constructed because of considerable heterogeneity, limiting the influence of any new (large) trial in a RE model. The extended funnel plots for the comparison of exercise therapy versus no intervention for knee (long term) and hip osteoarthritis (short term) suggested that it is implausible that a new study of any size, when added to the respective meta-analyses, could change current conclusions (online supplemental data 25).

DISCUSSION

Main findings

This overview of systematic reviews and supplementary RCTs evaluated evidence on the comparative effectiveness of exercise therapy on pain and physical function across different types of osteoarthritis. We found very low certainty evidence that exercise results in only small effects when compared with placebo and no treatment in predominantly knee osteoarthritis. Analyses of larger trials and longer follow-up time points suggest smaller estimates. Moderate evidence showed negligible effects of exercise in hip and small effects in hand osteoarthritis when compared with no treatment. Evidence of varying certainty and mainly concerned with knee osteoarthritis suggests that exercise may yield outcomes comparable to other conservative non-pharmacological interventions, pharmacotherapy and arthroscopic surgery. We found low to moderate certainty that exercise is less effective than knee osteotomy and total joint replacement of the knee and hip in the longer term in selected osteoarthritis patients.

Strengths and limitations of this overview

This overview was prospectively registered, followed rigorous and up-to-date methodology, applied a prioritisation process to overcome overlapping information among a large volume of reviews, re-analysed and in some instances, updated each meta-analysis presented in the included reviews using recommended meta-analytic methods. Through its broad scope, this overview is comprehensive, offering valuable information on the comparative effectiveness of exercise for decision-makers within a single, accessible document.

Limitations of this overview are that the prioritisation of certain reviews has led to the omission of the results of other relevant reviews. We did a sensitivity analysis by reporting the effect sizes from non-prioritised reviews, showing very similar effect sizes. Given the availability of multiple treatments and limited direct head-to-head comparisons, an NMA might have been considered. However, we deliberately chose not to pursue an NMA due to substantial heterogeneity in the included populations (eg, mild osteoarthritis in exercise trials vs moderate-to-severe osteoarthritis in surgical trials), which likely violates the assumption of joint randomisability across all interventions.[19] Instead, we conducted traditional pairwise meta-analyses, which should be interpreted independently, without drawing indirect comparisons between treatments. Further, the decision to re-analyse certain reviews was based on the certainty of evidence. However, clear criteria for defining conclusive evidence are still lacking[29] and the GRADE approach has been criticised for its subjectivity[113] and low validity in predicting the stability of an effect.[114] Defining conclusiveness of evidence could prevent research waste and shift the focus towards more important research areas.[115] For the GRADE ratings, we applied an MCID of 8 points, representing the smallest worthwhile effect of exercise compared with no treatment. We acknowledge that this threshold may not be directly applicable to comparisons with other active treatments. To address this limitation, we have added a forest plot of our main results in the online supplemental data, illustrating how the interpretation may vary when applying small (8 points), moderate (12.5 points) and large (20 points) MCIDs. Given that many treatment contrasts in our analyses included few trials (<5), the estimates were often imprecise and heterogeneity was not reliably estimable. Bayesian methods incorporating prior information may be better suited for these situations.[116] We also acknowledge that some included trials allowed co-interventions alongside exercise, which may limit the ability to isolate the effect to exercise alone. Where possible, we conducted subgroup analyses, which did not indicate any substantial impact.

Interpretation of the findings

We found very-low certainty evidence that exercise is slightly better than placebo interventions in the short term, with findings not robust to sensitivity analysis (including only larger, higher-quality trials employing more credible controls). Conversely, the review found large effect sizes in favour of exercise,[77] which were not confirmed in our re-analysis because of errors in meta-analysis. Given that many of the included trials do not align with recent reporting standards for mechanistic and efficacy trials[117] and the trials lack a third no treatment group that is required to estimate contextual effects,[118] the findings for this comparison remain inconclusive.

We found largely inconclusive evidence, indicating negligible to small effects of exercise across knee, hip and hand osteoarthritis compared with no treatment, especially in larger and longer-term trials. Our effect sizes were mostly similar to previous reviews.[12 74 76 119 120] A recent individual participant data (IPD) meta-analysis[10] showed negligible effect sizes, which, however, align with our sensitivity analyses restricted to larger trials. Nevertheless, heterogeneity remained high in most analyses (wide 95% PIs), which has also been observed in another review that reported PIs.[121] This suggests imbalance in the distribution of effect modifiers across the trials, which could be related to the 'lumping' of different exercise interventions (eg, different types, modes of delivery and doses), clinical populations (eg, different symptomatic severities) or trials with differing methodological characteristics (eg, small vs very large trials). The limited durability of exercise effects has also been reported in other studies.[10 74] This may be explained by the fact that most interventions were designed for short-term benefit rather than long-term management of a chronic condition, by declining adherence to exercise over time or by the natural fluctuating course of osteoarthritis, in which symptoms may return to baseline following temporary improvement.[122] Lastly, no treatment comparators inadequately control for contextual and specific effects, which may overestimate and obscure the true effects of exercise.[7 123]

We found that exercise yields outcomes comparable to conservative non-pharmacological treatments, pharmacotherapy and arthroscopic surgery. However, this evidence is mostly based on low to very-low certainty. Due to the scarcity of primary evidence, few reviews have yet summarised evidence on these comparators. Two NMAs[124 125] concluded that exercise is similar or even superior to NSAIDs for knee osteoarthritis. They included studies (κ=2,[124] κ=4[125]), which we decided to exclude as they did not adequately prescribe NSAIDs or were in the non-English language. Considering the low quality of these trials and the observed inconsistency in the networks,[124 125] the current evidence is uncertain regarding this treatment contrast. Given the highly controlled nature of pharmacological drug trials[126] and that many patients with osteoarthritis use pain medication concurrently with exercise[127 128] and they are typically prescribed on an as-needed basis,[129] the feasibility and validity of these (in-)direct comparisons can be questioned. Further, our work aligns with the conclusions of existing reviews[130 131] that exercise is comparable to arthroscopic surgery—which itself performs similarly to placebo surgery—for degenerative knee conditions, with our analysis being limited to patients with osteoarthritis only.

We found low to moderate evidence that exercise is less effective than knee osteotomy for medial compartment and total joint replacement of the knee and hip in severe osteoarthritis patients. THR tended to result in larger effects and higher cross-over rates from non-surgical to surgical group (77% after 2 years[108]) compared with replacement of the knee (32%[132]), indicating a strong rationale to opt for a joint replacement instead of prolonged exercise therapy in the presence of severe hip osteoarthritis. For knee osteoarthritis, however, cost-effectiveness data support a prolonged non-surgical treatment strategy with the option of later TKR.[133]

Implications for research

Given the stability of the mean estimates, but large heterogeneity in the observed effects of exercise, future research could focus on identifying subgroups of patients that better respond to certain exercise types or doses.[134 135] This requires consistent reporting of potential effect modifiers, credible subgroup analyses and sharing of IPD for meta-analysis. However, we also note that the publication and inclusion of small studies has led to distorted and heterogeneous effects, which are considerably reduced when these are excluded. This highlights the need for trialists and reviewers to direct research resources on more meaningful questions, informed by methodologically robust larger trials.[136 137] Additionally, answering the causal question "is exercise therapy effective?" requires a carefully designed control condition that to date has been challenging to establish due to the unclear mechanisms of action of exercise for osteoarthritis symptoms.[7] Future pragmatic clinical trials should focus on comparative effectiveness of exercise and physical activity interventions (both against other treatments and across different types/doses of exercise), facilitating NMAs to identify the most effective treatments for osteoarthritis. Lastly, future research should evaluate the effects of exercise for less researched types of osteoarthritis, its role in the early stages of the disease, its effects on other disease-related outcomes (eg, cardiovascular and structural markers) and how the benefits of exercise can be sustained in the long term.

Implications for practice

Our findings question the universal promotion of exercise therapy as the sole focus in first-line treatment to improve pain and physical function in all patients with osteoarthritis. Given that exercise seemingly yields comparable outcomes to many other treatments, doctors and physiotherapists could decide based on the preferences of patients, the possible harms of the treatment and their own expertise. In this context, exercise may be preferable due to its safety,[138] secondary health benefits[139] and lower costs.[140] Further, we acknowledge that some patients may consider exercise worthwhile even if expected effects are small or absent, for example, to prevent worsening of symptoms.[141] However, exercise may not be pursued for everyone, as patients with osteoarthritis often face personal (eg, comorbidity) or social barriers (eg, lack of access) that prevent them from adhering to regular exercise. For instance, in patients with severe hip osteoarthritis, prolonged exercise therapy before joint replacement is unlikely to alleviate symptoms enough to postpone the need for surgery.[108]

CONCLUSION

We found largely inconclusive evidence on exercise for osteoarthritis, suggesting negligible or, at best, short-lasting small effects on pain and function across different types of osteoarthritis compared with placebo or no treatment. These effects appear less pronounced in larger and longer-term trials, questioning the universal promotion of exercise and highlighting the need to revisit research priorities. Varying certainty evidence indicates that exercise yields outcomes comparable to common non-pharmacological and pharmacological treatments and arthroscopic surgery. Clinicians and patients should engage in shared decision-making, weighing the worthwhileness of exercise effects on pain and function alongside secondary health benefits, safety, low-cost profile, care stage and alternative treatment options. Low-to-moderate certainty evidence also suggests that exercise is less effective than knee osteotomy and total knee and hip replacement in selected patients, with more pronounced effects for hip osteoarthritis, aiding informed decision-making between patients and surgeons.

Author affiliations
[1]Bochum University of Applied Sciences, Division of Physiotherapy, 44801 Bochum, Germany
[2]The Parker Institute, Copenhagen University Hospital Bispebjerg Frederiksberg, 2000 Copenhagen, Denmark
[3]Graduate School for Applied Research in North Rhine-Westphalia (Graduate School NRW), Bochum, Germany
[4]GIMBE Foundation, Bologna, Italy
[5]Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, The Netherlands
[6]Department of Orthopaedics, Sahlgrenska University Hospital, Gothenburg, Sweden
[7]Department of Orthopaedics, Institute of Clinical Sciences, University of Gothenburg, Gothenburg, Sweden
[8]Orthopedics and Trauma Surgery, Marien Hospital Soest, Soest, NRW, Germany
[9]German Rheumatism League Federal Association, Bonn, NRW, Germany
[10]Eastern Health Emergency Medicine Program, Melbourne, Victoria, Australia
[11]Eastern Health Clinical School, Monash University, Melbourne, Victoria, Australia
[12]Faculty of Health Sciences Brandenburg, Brandenburg Medical School Theodor Fontane, Institute for Health Services and Health System Research, Ruedersdorf, Germany
[13]Center for Health Services Research, Brandenburg Medical School Theodor Fontane, Ruedersdorf, Germany

Acknowledgements The authors would like to thank Dr. Carolien Teirlinck (PhD student; Department General Practice, Erasmus MC University Medical Center Rotterdam, the Netherlands) for providing feedback to the protocol and sharing their dataset.

Contributors All authors contributed to the conception and design of the study. TS, FT and RD contributed to study selection, data extraction and critical appraisal. TS conducted the statistical analyses. TS drafted the manuscript, and all authors critically revised it and approved the final version. TS is the guarantor of the work. Please see CREDIT statement for further details in online supplemental data 1.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement statement A patient research partner from the German League Against Rheumatism, with lived experience of osteoarthritis, contributed to the interpretation of the results and critically revised the manuscript. The results will be disseminated to patient, public and clinician organisations through the press, social media and scientific conferences.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Part of a topic collection; not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. https://osf.io/wre46/overview?view_only=59f43b5bfd174ac09f9c13c3f095e9da

Author note The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

ORCID iDs
Tim Schleimer https://orcid.org/0009-0004-9085-2588
Florian Teichert https://orcid.org/0000-0003-2211-7974

REFERENCES

1 *Osteoarthritis in over 16s: diagnosis and management*. London: National Institute for Health and Care Excellence (NICE), 2022.
2 Moseng T, Vliet Vlieland TPM, Battista S, *et al*. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis: 2023 update. *Ann Rheum Dis* 2024;83:730–40.
3 Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA:D): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. *BMC Musculoskelet Disord* 2017;18:72.
4 Skou ST, Roos EM. Physical therapy for patients with knee and hip osteoarthritis: supervised, active treatment is current best practice. *Clin Exp Rheumatol* 2019;37 Suppl 120:112–7.
5 Englund M. Bout of the corner men and not the boxers? Contextual effects flex their muscles. *Ann Rheum Dis* 2018;77:159–61.
6 Englund M, Turkiewicz A. The emperor's new clothes? *Osteoarthr Cartil* 2023;31:549–51.
7 Henriksen M, Runhaar J, Turkiewicz A, *et al*. Exercise for knee osteoarthritis pain: Association or causation? *Osteoarthr Cartil* 2024;32:643–8.
8 Bricca A, Lund H, Roos EM, *et al*. When enough is enough - How to determine when the evidence for the effectiveness of a treatment is sufficient? *Osteoarthr Cartil* 2019;27:1253–6.
9 Bartels EM, Juhl CB, Christensen R, *et al*. Aquatic exercise for the treatment of knee and hip osteoarthritis. *Cochrane Database Syst Rev* 2016;3:CD005523.
10 Holden MA, Hattle M, Runhaar J, *et al*. Moderators of the effect of therapeutic exercise for knee and hip osteoarthritis: a systematic review and individual participant data meta-analysis. *Lancet Rheumatol* 2023;5:e386–400.
11 Uthman OA, van der Windt DA, Jordan JL, *et al*. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. *BMJ* 2013;347:f5555.
12 Goh S-L, Persson MSM, Stocks J, *et al*. Efficacy and potential determinants of exercise therapy in knee and hip osteoarthritis: A systematic review and meta-analysis. *Ann Phys Rehabil Med* 2019;62:356–65.
13 Hall M, van der Esch M, Hinman RS, *et al*. How does hip osteoarthritis differ from knee osteoarthritis? *Osteoarthr Cartil* 2022;30:32–41.
14 Kloppenburg M, Namane M, Cicuttini F. Osteoarthritis. *Lancet* 2025;405:71–85.
15 Veronese N, Smith L, Bolzetta F, *et al*. Efficacy of conservative treatments for hand osteoarthritis : An umbrella review of interventional studies. *Wien Klin Wochenschr* 2021;133:234–40.
16 Varongot-Reille C, Barrero-Santiago L, Cuenca-Martínez F, *et al*. Effectiveness of exercise on pain intensity and physical function in patients with knee and hip osteoarthritis: an umbrella and mapping review with meta-meta-analysis. *Disabil Rehabil* 2024;46:3475–89.
17 Chalmers I. What do I want from health research and researchers when I am a patient? *BMJ* 1995;310:1315–8.
18 Gates M, Gates A, Pieper D, *et al*. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. *BMJ* 2022;378:e070849.
19 Dias S, Ades AE, Welton NJ, *et al*. *Network meta-analysis for comparative effectiveness research*. Nashville, TN: John Wiley & Sons, 2018.
20 Belbasis L, Bellou V, Ioannidis JPA. Conducting umbrella reviews. *BMJ Med* 2022;1:e000071.
21 Higgins JPT, Thomas J, Chandler J, *et al*. *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons, 2019.
22 Higgins JPT, Lasserson T, Chandler J, *et al*. *Methodological expectations of cochrane intervention reviews*. 2016.

23 Smith T, Conaghan PG, OMERACT-OARSI Hip and Knee Osteoarthritis Update Working Group. 152 The OMERACT-OARSI core domain set for measurement in clinical trials of hip and/or knee osteoarthritis. *J Rheumatol* 2019;58:981–9.
24 Proposed outcomes. Available: https://musculoskeletal.cochrane.org/proposed-outcomes [Accessed 17 Jun 2023].
25 Kloppenburg M, Bøyesen P, Visser AW, *et al*. Report from the OMERACT Hand Osteoarthritis Working Group: Set of Core Domains and Preliminary Set of Instruments for Use in Clinical Trials and Observational Studies. *J Rheumatol* 2015;42:2190–7.
26 Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010;2010:CD003974.
27 Page MJ, Moher D, Fidler FM, *et al*. The REPRISE project: protocol for an evaluation of REProducibility and Replicability In Syntheses of Evidence. *Syst Rev* 2021;10:112.
28 Krnic Martinic M, Pieper D, Glatt A, *et al*. Definition of a systematic review used in overviews of systematic reviews, meta-epidemiological studies and textbooks. *BMC Med Res Methodol* 2019;19:203.
29 Babić A, Poklepović Peričić T, Pieper D, *et al*. When is the evidence conclusive? Analysis of systematic reviews for which Cochrane declared that conclusions will not change with further studies. *Res Synth Methods* 2022;13:478–88.
30 Lunny C, Brennan SE, McDonald S, *et al*. Toward a comprehensive evidence map of overview of systematic review methods: paper 1-purpose, eligibility, search and data extraction. *Syst Rev* 2017;6:231.
31 Pieper D, Antoine S-L, Neugebauer EAM, *et al*. Up-to-dateness of reviews is often neglected in overviews: a systematic review. *J Clin Epidemiol* 2014;67:1302–8.
32 Goossen K, Hess S, Lunny C, *et al*. Database combinations to retrieve systematic reviews in overviews of reviews: a methodological study. *BMC Med Res Methodol* 2020;20:138.
33 Haddaway NR, Grainger MJ, Gray CT. Citationchaser: A tool for transparent and efficient forward and backward citation chasing in systematic searching. *Res Synth Methods* 2022;13:533–45.
34 Pollock M, Fernandes RM, Newton AS, *et al*. A decision tool to help researchers make decisions about including systematic reviews in overviews of reviews of healthcare interventions. *Syst Rev* 2019;8:29.
35 Lunny C, Pieper D, Thabet P, *et al*. Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews. *BMC Med Res Methodol* 2021;21:140.
36 Mathes T, Prien P, Klerings I, *et al*. Guidance on how to efficiently find, choose, and use available systematic reviews was developed. *J Clin Epidemiol* 2024;173:111466.
37 R Core Team. R: a language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2022.
38 Bougioukas KI, Diakonidis T, Mavromanoli AC, *et al*. ccaR: A package for assessing primary study overlap across systematic reviews in overviews. *Res Synth Methods* 2023;14:443–54.
39 Pieper D, Antoine S-L, Mathes T, *et al*. Systematic review finds overlapping reviews were not mentioned in every other overview. *J Clin Epidemiol* 2014;67:368–75.
40 Mavridis D, White IR. Dealing with missing outcome data in meta-analysis. *Res Synth Methods* 2020;11:2–13.
41 Rohatgi A. Webplotdigitizer: version. 2022.
42 Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
43 Higgins JPT, Altman DG, Gøtzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
44 Sterne JAC, Savović J, Page MJ, *et al*. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
45 Borenstein M, Hedges LV, Higgins JPT, *et al*. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010;1:97–111.
46 Veroniki AA, Jackson D, Viechtbauer W, *et al*. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods* 2016;7:55–79.
47 Veroniki AA, Jackson D, Bender R, *et al*. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. *Res Synth Methods* 2019;10:23–43.
48 Jackson D, Law M, Rücker G, *et al*. The Hartung-Knapp modification for random-effects meta-analysis: A useful refinement but are there any residual concerns? *Stat Med* 2017;36:3923–34.
49 Schulz A, Schürmann C, Skipka G, *et al*. Performing Meta-analyses with Very Few Studies. *Methods Mol Biol* 2022;2345:91–102.
50 Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342.
51 IntHout J, Ioannidis JPA, Rovers MM, *et al*. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016;6:e010247.
52 Nagashima K, Noma H, Furukawa TA. Prediction intervals for random-effects meta-analysis: A confidence distribution approach. *Stat Methods Med Res* 2019;28:1689–702.
53 Donegan S, Williamson P, D'Alessandro U, *et al*. Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods* 2013;4:291–323.
54 Kadlec D, Sainani KL, Nimphius S. With Great Power Comes Great Responsibility: Common Errors in Meta-Analyses and Meta-Regressions in Strength & Conditioning Research. *Sports Med* 2023;53:313–25.
55 Wewege MA, Jones MD, Williams SA, *et al*. Rescaling pain intensity measures for meta-analyses of analgesic medicines for low back pain appears justified: an empirical examination from randomised trials. *BMC Med Res Methodol* 2022;22:285.
56 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.
57 Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *J Stat Softw* 2010;36:1–48.
58 Harrer M, Cuijpers P, Furukawa TA, *et al*. *Doing meta-analysis with R: a hands-on guide*. CRC Press, 2021.
59 Viechtbauer W, Cheung M-L. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods* 2010;1:112–25.
60 Flemyng E, Moore TH, Boutron I, *et al*. Using Risk of Bias 2 to assess results from randomised controlled trials: guidance from Cochrane. *BMJ Evid Based Med* 2023;28:260–6.
61 Nüesch E, Häuser W, Bernardy K, *et al*. Comparative efficacy of pharmacological and non-pharmacological interventions in fibromyalgia syndrome: network meta-analysis. *Ann Rheum Dis* 2013;72:955–62.
62 Langan D, Higgins JPT, Gregory W, *et al*. Graphical augmentations to the funnel plot assess the impact of additional evidence on a meta-analysis. *J Clin Epidemiol* 2012;65:511–9.
63 Ferreira ML, Herbert RD, Crowther MJ, *et al*. When is a further clinical trial justified? *BMJ* 2012;345:e5913.
64 Teichert F. Extfunnel2: create extended funnel plots. R package version 0.1.1. 2023.
65 Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
66 Zeng L, Brignardello-Petersen R, Hultcrantz M, *et al*. GRADE Guidance 34: update on rating imprecision using a minimally contextualized approach. *J Clin Epidemiol* 2022;150:216–24.
67 Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
68 Dekker J, de Boer M, Ostelo R. Minimal important change and difference in health outcome: An overview of approaches, concepts, and methods. *Osteoarthr Cartil* 2024;32:8–17.
69 Christiansen DH, de Vos Andersen N-B, Poulsen PH, *et al*. The smallest worthwhile effect of primary care physiotherapy did not differ across musculoskeletal pain sites. *J Clin Epidemiol* 2018;101:44–52.
70 Nuesch E, Trelle S, Reichenbach S, *et al*. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 2009;339:b3244.
71 Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.
72 Nikolakopoulou A, Higgins JPT, Papakonstantinou T, *et al*. CINeMA: An approach for assessing confidence in the results of a network meta-analysis. *PLoS Med* 2020;17:e1003082.
73 Santesso N, Glenton C, Dahm P, *et al*. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol* 2020;119:126–35.
74 Lawford BJ, Hall M, Hinman RS, *et al*. Exercise for osteoarthritis of the knee. *Cochrane Libr* 2024;2024.
75 Teirlinck CH, Verhagen AP, van Ravesteyn LM, *et al*. Effect of exercise therapy in patients with hip osteoarthritis: A systematic review and cumulative meta-analysis. *Osteoarthr Cartil Open* 2023;5:100338.
76 Østerås N, Kjeken I, Smedslund G, *et al*. Exercise for hand osteoarthritis. *Cochrane Database Syst Rev* 2017;1:CD010388.
77 Dean BJF, Collins J, Thurley N, *et al*. Exercise therapy with or without other physical therapy interventions versus placebo

interventions for osteoarthritis -Systematic review. *Osteoarthr Cartil Open* 2021;3:100195.
78 Goff AJ, De Oliveira Silva D, Merolli M, *et al*. Patient education improves pain and function in people with knee osteoarthritis with better effects when combined with exercise therapy: a systematic review. *J Physiother* 2021;67:177–89.
79 Bandak E, Christensen R, Overgaard A, *et al*. Exercise and education versus saline injections for knee osteoarthritis: a randomised controlled equivalence trial. *Ann Rheum Dis* 2022;81:537–43.
80 Henriksen M, Christensen R, Kristensen LE, *et al*. Exercise and education vs intra-articular saline for knee osteoarthritis: a 1-year follow-up of a randomized trial. *Osteoarthr Cartil* 2023;31:627–35.
81 Messier SP, Mihalko SL, Beavers DP, *et al*. Effect of High-Intensity Strength Training on Knee Pain and Knee Joint Compressive Forces Among Adults With Knee Osteoarthritis: The START Randomized Clinical Trial. *JAMA* 2021;325:646–57.
82 Messier SP, Beavers DP, Queen K, *et al*. Effect of Diet and Exercise on Knee Pain in Patients With Osteoarthritis and Overweight or Obesity: A Randomized Clinical Trial. *JAMA* 2022;328:2242–51.
83 Gravås EMH, Østerås N, Nossum R, *et al*. Does occupational therapy delay or reduce the proportion of patients that receives thumb carpometacarpal joint surgery? A multicentre randomised controlled trial. *RMD Open* 2019;5:e001046.
84 Tveter AT, Østerås N, Nossum R, *et al*. Short-Term Effects of Occupational Therapy on Hand Function and Pain in Patients With Carpometacarpal Osteoarthritis: Secondary Analyses From a Randomized Controlled Trial. *Arthritis Care Res (Hoboken)* 2022;74:955–64.
85 Deveza LA, Robbins SR, Duong V, *et al*. Efficacy of a Combination of Conservative Therapies vs an Education Comparator on Clinical Outcomes in Thumb Base Osteoarthritis. *JAMA Intern Med* 2021;181:429.
86 Magni N, McNair P, Rice D. Six weeks of resistance training (plus advice) vs advice only in hand osteoarthritis: A single-blind, randomised, controlled feasibility trial. *Musculoskelet Sci Pract* 2022;57:102491.
87 Smith MD, Vuvan V, Collins NJ, *et al*. A combined program of education plus exercise versus general advice for ankle osteoarthritis: A feasibility randomised controlled trial. *Musculoskelet Sci Pract* 2024;74:103169.
88 Abbott JH, Robertson MC, Chapple C, *et al*. Manual therapy, exercise therapy, or both, in addition to usual care, for osteoarthritis of the hip or knee: a randomized controlled trial. 1: clinical effectiveness. *Osteoarthr Cartil* 2013;21:525–34.
89 Dwyer L, Parkin-Smith GF, Brantingham JW, *et al*. Manual and manipulative therapy in addition to rehabilitation for osteoarthritis of the knee: assessor-blind randomized pilot trial. *J Manipulative Physiol Ther* 2015;38:1–21.
90 Hoeksma HL, Dekker J, Ronday HK, *et al*. Comparison of manual therapy and exercise therapy in osteoarthritis of the hip: a randomized clinical trial. *Arthritis Rheum* 2004;51:722–9.
91 Doi T, Akai M, Fujino K, *et al*. Effect of Home Exercise of Quadriceps on Knee Osteoarthritis Compared with Nonsteroidal Antiinflammatory Drugs. *Am J Phys Med Rehabil* 2008;87:258–69.
92 Yang X, Song H, Li X, *et al*. Neuromuscular exercise in addition to celecoxib versus celecoxib alone for symptomatic and radiographic knee osteoarthritis: a randomized controlled trial. *BMC Sports Sci Med Rehabil* 2025;17:1–10.
93 Angoorani H, Mazaherinezhad A, Marjomaki O, *et al*. Treatment of knee osteoarthritis with platelet-rich plasma in comparison with transcutaneous electrical nerve stimulation plus exercise: a randomized clinical trial. *Med J Islam Repub Iran* 2015;29:223.
94 Karaborklu Argut S, Celik D, Ergin ON, *et al*. Does the Combination of Platelet-rich Plasma and Supervised Exercise Yield Better Pain Relief and Enhanced Function in Knee Osteoarthritis? A Randomized Controlled Trial. *Clin Orthop Relat Res* 2024;482:1051–61.
95 Deyle GD, Allen CS, Allison SC, *et al*. Physical Therapy versus Glucocorticoid Injection for Osteoarthritis of the Knee. *N Engl J Med* 2020;382:1420–9.
96 Lee M, Jing C, Lee K. Physical therapy vs. glucocorticoid injection in patients with meniscal tears and knee osteoarthritis: a multi-center, randomized, controlled trial. *BMC Med* 2025;23:277.
97 Karatosun V, Unver B, Gocen Z, *et al*. Intra-articular hyaluranic acid compared with progressive knee exercises in osteoarthritis of the knee: a prospective randomized trial with long-term follow-up. *Rheumatol Int* 2006;26:277–84.
98 Kawasaki T, Kurosawa H, Ikeda H, *et al*. Therapeutic home exercise versus intraarticular hyaluronate injection for osteoarthritis of the knee: 6-month prospective randomized open-labeled trial. *J Orthop Sci* 2009;14:182–91.
99 Saccomanno MF, Donati F, Careri S, *et al*. Efficacy of intra-articular hyaluronic acid injections and exercise-based rehabilitation programme, administered as isolated or integrated therapeutic regimens for the treatment of knee osteoarthritis. *Knee Surg Sports Traumatol Arthrosc* 2016;24:1686–94.
100 Kapoor A, Mudi M. A Prospective Randomized Controlled Study for treatment of Osteoarthritis of knee joint with Intra-Articular injection of Hylan GF20 (Single dose 6ml) Versus Conservative Management. *J Clin Orthop Trauma* 2024;57:102586.
101 Karatosun V, Unver B, Ozden A, *et al*. Intra-articular hyaluronic acid compared to exercise therapy in osteoarthritis of the ankle. A prospective randomized trial with long-term follow-up. *Clin Exp Rheumatol* 2008;26:288–94.
102 Rabago D, Patterson JJ, Mundt M, *et al*. Dextrose prolotherapy for knee osteoarthritis: a randomized controlled trial. *Ann Fam Med* 2013;11:229–37.
103 Kirkley A, Birmingham TB, Litchfield RB, *et al*. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2008;359:1097–107.
104 Katz JN, Brophy RH, Chaisson CE, *et al*. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *N Engl J Med* 2013;368:1675–84.
105 Birmingham TB, Primeau CA, Moyer RF, *et al*. High Tibial Osteotomy for Medial Compartment Knee Osteoarthritis : A Randomized Trial With Parallel Preference Arm. *Ann Intern Med* 2025;178:1238–48.
106 Skou ST, Roos EM, Laursen MB, *et al*. A Randomized, Controlled Trial of Total Knee Replacement. *N Engl J Med* 2015;373:1597–606.
107 Rognsvåg T, Bergvad IB, Furnes O, *et al*. Exercise therapy, education, and cognitive behavioral therapy alone, or in combination with total knee arthroplasty, in patients with knee osteoarthritis: a randomized feasibility study. *Pilot Feasibility Stud* 2024;10:43.
108 Frydendal T, Christensen R, Mechlenburg I, *et al*. Total Hip Replacement or Resistance Training for Severe Hip Osteoarthritis. *N Engl J Med* 2024;391:1610–20.
109 Chao J, Jing Z, Xuehua B, *et al*. Effect of Systematic Exercise Rehabilitation on Patients With Knee Osteoarthritis: A Randomized Controlled Trial. *Cartilage* 2021;13:1734S–1740S.
110 Yavuz M, Ataoğlu S, Baki̇ AE, *et al*. Primer Diz Osteoartritinde İzokinetik Egzersiz, Lazer ve Diklofenak İyontoforezi Uygulamalarının Etkilerinin ve Etkinliklerinin Karşılaştırılması. *Duzce Med J* 2013;15:15–21.
111 Jie Z, Ye S, Gu H. Clinical observation of moxibustion combined with knee joint rehabilitation for the treatment of knee osteoarthritis. *J Clin Acupunct Moxibust* 2015;7–9.
112 Teirlinck CH, Verhagen AP, van Ravesteyn LM, *et al*. Corrigendum to “Effect of exercise therapy in patients with hip osteoarthritis: A systematic review and cumulative meta-analysis” [Osteoarthr Cartil Open 5 (2023) 100338]. *Osteoarthr Cartil Open* 2024;6:100447.
113 Norris SL, Bero L. GRADE Methods for Guideline Development: Time to Evolve? *Ann Intern Med* 2016;165:810–1.
114 Gartlehner G, Dobrescu A, Evans TS, *et al*. The predictive validity of quality of evidence grades for the stability of effect estimates was low: a meta-epidemiological study. *J Clin Epidemiol* 2016;70:52–60.
115 Puljak L, Lund H. Definition, harms, and prevention of redundant systematic reviews. *Syst Rev* 2023;12:63.
116 Saueressig T, Pedder H, Bowe SJ, *et al*. Six Meta-analyses on Treatments for Femoroacetabular Impingement Syndrome in a Year and Readers Are None the Wiser: Methods Advice for Researchers Planning Meta-analysis of Data From Fewer Than 5 Trials. *J Orthop Sports Phys Ther* 2021;51:201–3.
117 Hohenschurz-Schmidt D, Vase L, Scott W, *et al*. Recommendations for the development, implementation, and reporting of control interventions in efficacy and mechanistic trials of physical, psychological, and self-management therapies: the CoPPS Statement. *BMJ* 2023;381:e072108.
118 Saueressig T, Pedder H, Owen PJ, *et al*. Contextual effects: how to, and how not to, quantify them. *BMC Med Res Methodol* 2024;24:35.
119 Fransen M, McConnell S, Harmer AR, *et al*. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2015;1:CD004376.
120 Fransen M, McConnell S, Hernandez-Molina G, *et al*. Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev* 2014;2014.
121 Juhl C, Christensen R, Roos EM, *et al*. Impact of exercise type and dose on pain and disability in knee osteoarthritis: a systematic review and meta-regression analysis of randomized controlled trials. *Arthritis Rheumatol* 2014;66:622–36.

122 Englund M, Turkiewicz A. Pain in clinical trials for knee osteoarthritis: estimation of regression to the mean. *Lancet Rheumatol* 2023;5:e309–11.
123 Cashin AG, McAuley JH, Lamb SE, *et al*. Disentangling contextual effects from musculoskeletal treatments. *Osteoarthr Cartil* 2021;29:297–9.
124 Thorlund JB, Simic M, Pihl K, *et al*. Similar Effects of Exercise Therapy, Nonsteroidal Anti-inflammatory Drugs, and Opioids for Knee Osteoarthritis Pain: A Systematic Review with Network Meta-analysis. *J Orthop Sports Phys Ther* 2022;52:207–16.
125 Weng Q, Goh S-L, Wu J, *et al*. Comparative efficacy of exercise therapy and oral non-steroidal anti-inflammatory drugs and paracetamol for knee or hip osteoarthritis: a network meta-analysis of randomised controlled trials. *Br J Sports Med* 2023;57:990–6.
126 Santaguida PL, Helfand M, Raina P. Challenges in Systematic Reviews That Evaluate Drug Efficacy or Effectiveness. *Ann Intern Med* 2005;142:1066–72.
127 Henriksen M, Nielsen SM, Christensen R, *et al*. Who are likely to benefit from the Good Life with osteoArthritis in Denmark (GLAD) exercise and education program? An effect modifier analysis of a randomised controlled trial. *Osteoarthr Cartil* 2023;31:106–14.
128 Koes BW, Chiarotto A, Thorlund JB, *et al*. Does Pain Medication Use Influence the Outcome of 8 Weeks of Education and Exercise Therapy in Patients with Knee or Hip Osteoarthritis? An Observational Study. *Pain Med* 2022;23:1457–63.
129 da Costa BR, Pereira TV, Saadat P, *et al*. Effectiveness and safety of non-steroidal anti-inflammatory drugs and opioid treatment for knee and hip osteoarthritis: network meta-analysis. *BMJ* 2021;375:n2321.
130 O'Connor D, Johnston RV, Brignardello-Petersen R, *et al*. Arthroscopic surgery for degenerative knee disease (osteoarthritis including degenerative meniscal tears). *Cochrane Database Syst Rev* 2022;3:CD014328.
131 Thorlund JB, Juhl CB, Roos EM, *et al*. Arthroscopic surgery for degenerative knee: systematic review and meta-analysis of benefits and harms. *BMJ* 2015;350:h2747.
132 Skou ST, Roos EM, Laursen MB, *et al*. Total knee replacement and non-surgical treatment of knee osteoarthritis: 2-year outcome from two parallel randomized controlled trials. *Osteoarthr Cartil* 2018;26:1170–80.
133 Skou ST, Roos E, Laursen M, *et al*. Cost-effectiveness of total knee replacement in addition to non-surgical treatment: a 2-year outcome from a randomised trial in secondary care in Denmark. *BMJ Open* 2020;10:e033495.
134 Haber T, Lawford BJ, Bennell K, *et al*. Recent highlights and uncertainties in exercise management of knee osteoarthritis. *J Physiother* 2025;71:158–66.
135 Bennell KL, Nelligan RK, Pardo JJ, *et al*. Research priorities for physical activity and exercise management of people with knee and hip osteoarthritis: A multi-stage international priority-setting study from the OARSI Rehabilitation Discussion Group. *Osteoarthr Cartil* 2025;33:1258–67.
136 Chamberlain F, Elliott L, Yarahmadi A, *et al*. Research waste in randomised trials of exercise treatments for chronic low back pain: trial sequential and cumulative meta-analyses by publication date and size. *BMJ Open* 2025;15:e087787.
137 Nuesch E, Trelle S, Reichenbach S, *et al*. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;341:c3515.
138 Niemeijer A, Lund H, Stafne SN, *et al*. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. *Br J Sports Med* 2020;54:1073–80.
139 Schulz JM, Birmingham TB, Atkinson HF, *et al*. Are we missing the target? Are we aiming too low? What are the aerobic exercise prescriptions and their effects on markers of cardiovascular health and systemic inflammation in patients with knee osteoarthritis? A systematic review and meta-analysis. *Br J Sports Med* 2020;54:771–5.
140 Pinto D, Robertson MC, Hansen P, *et al*. Cost-effectiveness of nonpharmacologic, nonsurgical interventions for hip and/or knee osteoarthritis: systematic review. *Value Health* 2012;15:1–12.
141 Longtin C, Salisbury A, Maher CG, *et al*. People Strongly Value Physical Therapies for Low Back Pain Over Doing Nothing, Even When Effects Are Very Small: A Discrete Choice Experiment. *J Orthop Sports Phys Ther* 2025;55:602–10.