Comparative efficacy and safety of exercise modalities in knee osteoarthritis: systematic review and network meta-analysis

Lei Yan,[1,2] Dijun Li,[1,3] Dan Xing,[4] Zijuan Fan,[5] Guangyuan Du,[1] Jingwei Jiu,[6] Xiaoke Li,[6] Janne Estill,[7] Qi Wang,[8] Ahmed Atef Belal,[8] Chen Tian,[9,10] Jiao Jiao Li,[11] Songyan Li,[1] Haifeng Liu,[6] Xuanbo Liu,[6] Yijia Ren,[1] Yiqi Yang,[1] Jinxiu Chen,[12] Yihe Hu,[1] Long Ge,[9,10] Bin Wang[1]

For numbered affiliations see end of the article

Correspondence to: B Wang wangbin_pku@zju.edu.cn; (ORCID 0000-0002-5474-1002)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2025;391:e085242
http://dx.doi.org/10.1136/bmj-2025-085242

Accepted: 16 August 2025

ABSTRACT

OBJECTIVE
To assess the efficacy and safety of various exercise modalities as therapeutic interventions for managing knee osteoarthritis.

DESIGN
Systematic review with network meta-analysis.

DATA SOURCES
PubMed, Embase, Cochrane Library, Web of Science, CINAHL, PsycINFO, AMED, PEDro, Scopus, ClinicalTrials.gov, ICTRP, and ClinicalTrialsRegister.eu from database inception to August 2024.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES
Randomised controlled trials comparing different exercise modalities, including aerobic exercise, flexibility exercise, mind-body exercise, neuromotor exercise, strengthening exercise, mixed exercise, and control group for patients with knee osteoarthritis.

MAIN OUTCOME MEASURES
Primary outcomes included pain, function, gait performance, and quality of life, assessed at short term (four weeks), mid-term (12 weeks), and long term (24 weeks) follow-up. When exact time points were unavailable, data from adjacent time windows were used.

RESULTS
217 randomised controlled trials involving 15 684 participants were included. Moderate certainty evidence showed that, compared with the control group, aerobic exercise probably results in large improvements in pain at short term (standardised mean difference −1.10, 95% confidence interval −1.68 to −0.52) and mid-term follow-up (−1.19, −1.59 to −0.79), function at mid-term (1.78, 1.05 to 2.51), gait performance at mid-term (0.85, 0.55 to 1.14), and quality of life at short term (1.53, 0.47 to 2.59). Mind-body exercise probably results in a large increase in function at short term follow-up (0.88, 0.03 to 1.73; moderate certainty), while neuromotor exercise probably results in a large increase in gait performance at short term follow-up (1.04, 0.51 to 1.57; moderate certainty). Strengthening (0.86, 0.53 to 1.18) and mixed exercise (1.07, 0.68 to 1.46) probably result in a large increase in function at mid-term follow-up, all with moderate certainty evidence. Regarding long term follow-up, flexibility exercise may result in a large reduction in pain (−0.99, −1.63 to −0.36; low certainty); aerobic exercise may result in a large increase in function (0.87, 0.02 to 1.72, low certainty); and mixed exercise may increase function (0.56, 0.26 to 0.86; low certainty) and probably increases gait performance (0.57, 0.21 to 0.92, moderate certainty). Overall, aerobic exercise consistently showed the highest probability of being the best treatment, as reflected by surface under the cumulative ranking curve values (mean 0.72) across outcomes. The safety outcome was reported in a small proportion of studies (40 studies, 18%), and no clear differences were observed between exercise interventions and control.

CONCLUSIONS
In patients with knee osteoarthritis, aerobic exercise is likely the most beneficial exercise modality for improving pain, function, gait performance, and quality of life, with moderate certainty.

SYSTEMATIC REVIEW REGISTRATION
PROSPERO CRD42023469762.

WHAT IS ALREADY KNOWN ON THIS TOPIC
Existing clinical guidelines consistently endorse multimodal treatment approaches, involving exercise, drug, and operative interventions, for managing knee osteoarthritis

Current guidelines provide limited information regarding the specific types of exercise interventions that should be recommended for patients with knee osteoarthritis

Current decision making is usually based on expert opinion and clinical experience

WHAT THIS STUDY ADDS
This network meta-analysis of 217 randomised controlled trials with 15 684 participants evaluates the relative benefit of five exercise modalities for knee osteoarthritis

Aerobic exercise may be the most beneficial exercise modality for improving pain, function, gait performance, and quality of life in knee osteoarthritis

The findings can help to improve the clarity of guideline recommendations regarding the most effective exercise interventions for patients with knee osteoarthritis

Introduction
Osteoarthritis is a common chronic joint disease that results in considerable pain and disability in the global population, particularly older individuals.[1 2] Global trends show that population ageing and rise in obesity rates will continue to drive substantial growth in the incidence, prevalence, and attributable disability rate of osteoarthritis for the foreseeable future.[3 4] The knee is the most common site of osteoarthritis, with nearly 30% of people older than 45 years showing radiological evidence of knee osteoarthritis, half of whom also have severe knee symptoms.[5 6] The pain caused by knee osteoarthritis can substantially affect

an individual's physical function and quality of life, creating a major public health problem worldwide.[7]

Exercise is one of the mainstream treatments for knee osteoarthritis, with potential effects on relieving pain and improving the quality of life.[8] Therapeutic exercise aims to increase lower limb muscle mass, neuromotor control, and joint range of motion.[4] These changes in turn lead to increased lower limb strength, minimising pain, and improving physical function.[9,10] However, the large volume of data available in the literature makes the optimal type of exercise therapy for knee osteoarthritis difficult to determine.[11,12]

Although several recent studies have attempted to compare the effects of different exercise modalities on knee osteoarthritis, major limitations remain in the scope, methodology, and integration of evidence. Many earlier studies failed to capture the full range of exercise types, leading to incomplete comparisons and reduced generalisability.[12-14] Moreover, earlier studies often had methodological limitations: most did not apply the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) framework to assess certainty of evidence, and used the outdated Risk of Bias 1 (RoB 1) tool rather than the more robust Risk of Bias 2 (RoB 2) for risk of bias evaluation.[15,16] Critically, none of the existing network meta-analyses accounted for variation in follow-up duration, because they typically assessed only a single time point. Given the chronic and progressive nature of knee osteoarthritis, it is essential to examine short, medium, and long term outcomes to understand the time dependent effectiveness of exercise interventions. The most recent network meta-analysis in this area was published in 2019, with a literature search ending in 2017.[15] Since then, a growing body of randomised controlled trials on exercise intervention for knee osteoarthritis has emerged. To review these gaps, our study updates and expands on previous work by refining exercise classification, applying updated methodological standards (RoB 2 and GRADE), and evaluating multiple follow-up durations. This approach provides a more comprehensive and clinically meaningful evaluation of the comparative efficacy and safety of exercise interventions for knee osteoarthritis.

Methods

Study design

This systematic review with meta-analysis was designed and performed following the PRISMA-NMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses-Network Meta-Analyses) guidelines[17]; appendix 1) and was registered at PROSPERO (CRD42023469762). Changes made to the protocol during the project are presented in appendix 2. Appendix 3 presents the logic model for exercise intervention in knee osteoarthritis. Along with appendices 4 and 5, this model outlines the key components of study investigation, including population, inputs, outcomes, and assessment, which systematically illustrate the mechanism of exercise intervention.

Literature review

PubMed, Embase, Cochrane Library, Web of Science, CINAHL, PsycINFO, AMED, PEDro, Scopus, ClinicalTrials.gov, ICTRP, and ClinicalTrialsRegister.eu were searched from database inception to August 2024 to identify studies on exercise therapy for knee osteoarthritis.

The detailed search strategy is shown in appendix 6. We assessed all relevant reviews retrieved from the searches to identify potentially eligible studies. Articles were imported into Endnote X9 and duplicates were removed, after which two investigators (DL and XLi) independently screened the titles and abstracts of the studies to identify those eligible for inclusion. We then evaluated the full texts of potentially eligible studies according to the inclusion and exclusion criteria. Another investigator (LY) was consulted to resolve any discrepancies.

Eligibility criteria

Randomised controlled trials that enrolled participants with symptomatic knee osteoarthritis according to findings on clinical, radiological, or magnetic resonance imaging examination were eligible for inclusion. Eligible trials compared exercise therapy with either a control group or alternative forms of exercise therapy (appendix 4). No restrictions were imposed on the dose, duration, or intensity of the exercise interventions. The reported outcomes had to include at least one of the following areas: function, pain, quality of life, gait performance, and adverse events. Trials were excluded if they included patients who had a history of knee surgery or who had knee osteoarthritis and hip osteoarthritis at the same time.

Data extraction

Two independent assessors (DL and XLi) extracted baseline characteristics and outcome data into a standardised form. We extracted details of the exercise interventions from each study, including their frequency, intensity, and duration. To quantify the energy expenditure associated with each intervention, we used the Compendium of Physical Activities,[18] and expressed exercise intensity as metabolic equivalents of task (METs) in minutes per week. The outcome measures of interest were pain, function, gait performance, and quality of life scores reported at baseline and at four, 12, and 24 weeks after initiation of the exercise intervention. If measurements from four weeks were not available, we used the nearest available time point between two and six weeks of follow-up. Similarly, we used data between eight and 16 weeks if data from 12 weeks were unavailable; if data from 24 weeks were unavailable, we used data from a time point after 18 weeks that was closest to 24 weeks.

Quality assessment

We used RoB 2 to assess the methodological quality of the included studies.[19] Using the licensed Excel tool to execute the updated RoB 2 methodology for randomised trials, two reviewers (DL and ZF) independently assessed the risk of bias for the pertinent outcomes, in the studies included in the network meta-analysis. For each included study, we evaluated the risk of bias according to the following five domains: randomisation process, deviations from the intended interventions (assessed specifically for the effect of assignment to intervention), missing outcome data, measurement of the result, and options of the reported outcomes (appendix 7). Any disagreement was settled with the help of a third investigator (LY).

Statistical analysis

Data analysis and synthesis

For outcomes related to pain, function, gait performance, and quality of life, when studies assessed the same outcome using different instruments, we used the standardised mean difference to transform the outcomes to a uniform scale. Risk ratios were used for adverse events. Magnitudes of effect were categorised as follows: trivial (standardised mean difference <0.20), small (≥0.20 to <0.50), moderate (≥0.50 to <0.80), and large (≥0.80).[20] We followed the wording recommendations from GRADE Guidelines 26 to standardise the interpretation and reporting of effect estimates.[21] To ensure a conservative and clinically meaningful interpretation of effectiveness, we used a standardised mean difference of 0.50 as the threshold for the minimal important difference.[22] For pain, higher scores indicate greater pain levels. Conversely, for outcomes such as function, gait performance, and quality of life, higher scores signify better outcomes. In the main text, we present standardised mean differences as the intervention group minus the control group. Therefore, a negative standardised mean difference means that the results favour the intervention for pain, and a positive standardised mean difference indicates that the result favours the intervention for function, gait performance, and quality of life. We calculated standard deviations (SDs) from standard errors, P values, confidence intervals, or plots when SDs were lacking. If none of these methods could be used, we obtained SDs from additional papers included in our network meta-analysis using established inference techniques (appendix 8.1).[23]

R version 4.1.2 and the package netmeta (version 2.1-0) (R Project for Statistical Computing) were used for network meta-analysis, using the frequentist model with a graph theoretical technique. We used the Moore-Penrose pseudoinverse method, which was based on weighted least square regression, in the estimator, and the DerSimonian-Laird random effects modelling to perform a pairwise meta-analysis to assess the variation in study heterogeneity and gather concrete evidence. Because results suggested aerobic exercise might have a larger effect than other exercises, we conducted a post hoc analysis using each type of aerobic exercise as a separate node in the analysis. We used league tables of the respective treatment effect sizes as a visual aid to compare network estimations, and we determined significance for all comparisons using a two tailed threshold of P<0.05. In addition, we used the surface under the cumulative ranking curve to rank the intervention hierarchy in the network meta-analysis.[24]

We assessed transitivity of the indirect comparison hypothesis by comparing the distribution of trial characteristics (year of publication, follow-up period, average percentage of women, mean age, body mass index, exercise dose, exercise duration, exercise intensity, and sample size). For all comparison loops, we used node splitting to evaluate the local inconsistency of direct and indirect results. To assess the stability of conclusions from the main model, we did three sensitivity analyses excluding studies with small sample size (<50 per study), studies with high risk of bias, or studies not reporting standard deviation of the outcome measures. We quantified publication bias by examining the funnel plot asymmetry, and used Egger's regression test, Begg's test, and Thompson–Sharp test to assess asymmetry. To investigate the association between the relevant covariates and each outcome, we conducted a network meta-regression under the assumption of a shared coefficient.

Certainty of evidence

We used the GRADE framework to evaluate the certainty of evidence for each network estimate, which classifies evidence as having high, moderate, low, or very low certainty (appendix 8.2).[25] Certainty was assessed according to the following categories: risk of bias, imprecision, inconsistency, and indirectness. For both direct and indirect comparisons, the starting point for certainty in estimates was high. We assessed the certainty of indirect estimates by examining the dominant lowest order loop, defined as the minimal order evidence circuit (typically a three treatment loop, or a higher order loop if no triangle exists) that contributes the largest share of information to the indirect comparison. We identified this evidence circuit by using the contribution matrix. The matrix quantifies how much each direct comparison contributes to each indirect estimate, which we used to judge the reliability of the indirect evidence. The certainty rating chosen was the lowest of the direct estimates contributing to the indirect comparison. We considered further down rating each indirect comparison for intransitivity if the interventions or populations were dissimilar between the direct comparisons informing the loop that contributed most to the indirect estimate.

We rated the certainty of network estimates according to the highest level of direct estimates and indirect estimates contributing to the network estimates. We considered downgrading the certainty in the network estimate if there was incoherence between indirect and direct estimates, or if there was imprecision (credible interval crossing the invalid line) around the treatment effect.[26]

Patient and public involvement

This study is a meta-analysis of previously published data and did not involve patients or members of the public in its design, conduct, or reporting. However, we acknowledge that one of the peer reviewers was a patient, whose feedback helped inform the final interpretation and framing of results.

Results

Study selection and characteristics

We identified 45 804 potentially relevant studies from database searches. After removing duplicates, we screened the titles, abstracts, and full texts of 14 109 articles, resulting in full texts of 596 potentially eligible studies (fig 1). Appendix 9 shows the reasons for exclusion of 379 articles. Finally, 217 included studies with a total of 15 684 participants satisfied the eligibility criteria. These studies, published between 1990 and 2024, are detailed in appendix 10. Among the 88 intervention arms categorised as mixed exercise, 83% (n=73) combined exactly two components, most commonly strengthening and neuromotor (24%; n=21), strengthening and flexibility (24%; n=21), and strengthening and aerobic (19%; n=17). Notably, 88% (n=77) of mixed exercise interventions included strengthening exercise, indicating that it was commonly used as a core component (appendix 10.3). Considering all 12 outcomes in the risk of bias assessment, the overall judgment indicated that three quarters of comparisons (557, 76%) were at some concerns, 159 (21%) at high risk, and only 16 (2%) at low risk (appendix 12.2).

Overview of studies

Network plots of outcome measures are shown in figure 2. A total of 49 studies (23%) reported a follow-up period of 24 weeks or longer. The US (38 studies, 18%) and China (36 studies, 17%) were the two most

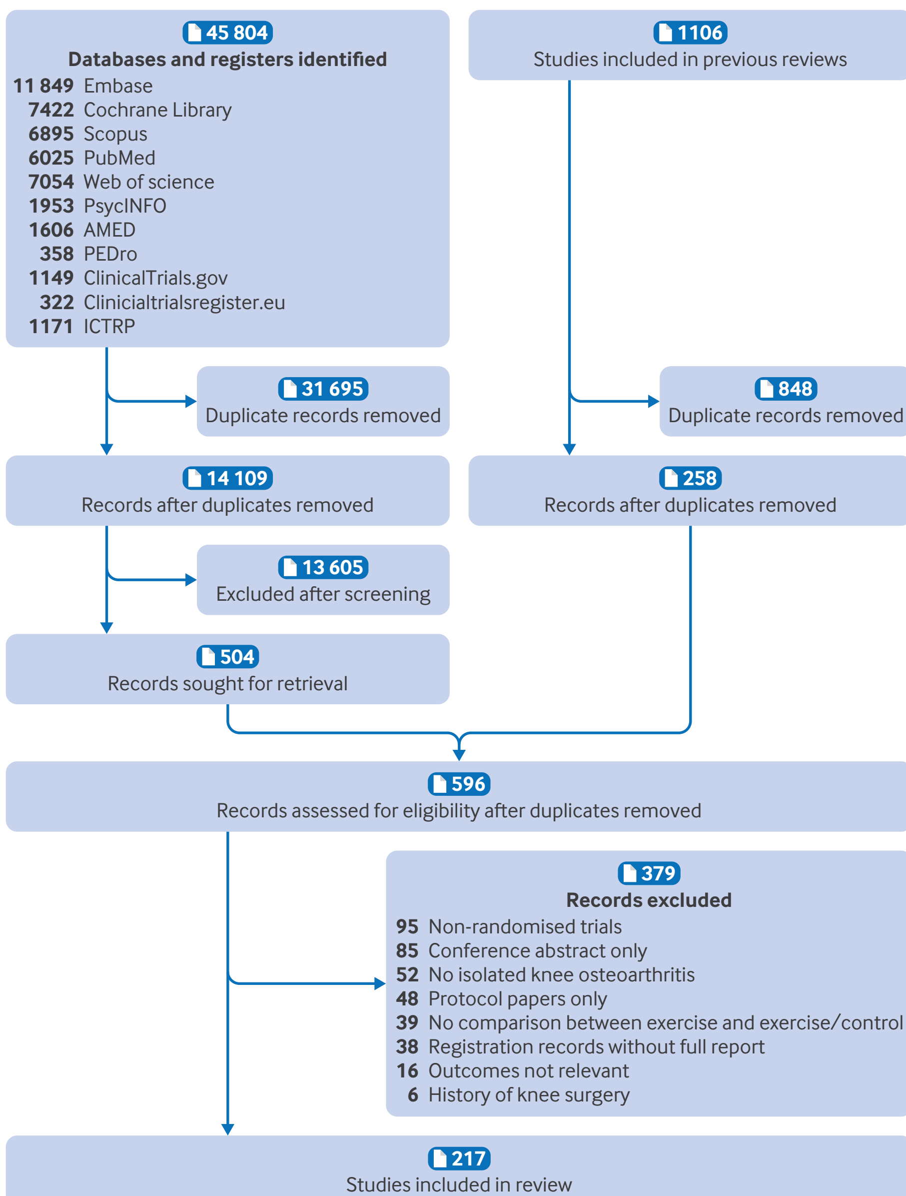

Fig 1 | Flow diagram of the study selection process common countries of origin among the included studies. A total of 137 of 217 studies (63%) focused on comparisons between an exercise intervention and a control group (appendix 10.2).

Outcomes

Figure 3, figure 4, and appendix 11 show the results of network meta-analysis for different outcome measures. According to the GRADE assessment results, imprecision was the most common reason for downgrading certainty of evidence (appendix 12). Funnel plots appeared largely symmetrical on visual inspection, indicating minimal evidence of small study effects in the network meta-analysis overall (appendix 12.6). Formal testing supported this impression for most comparisons: Egger's test was non-significant

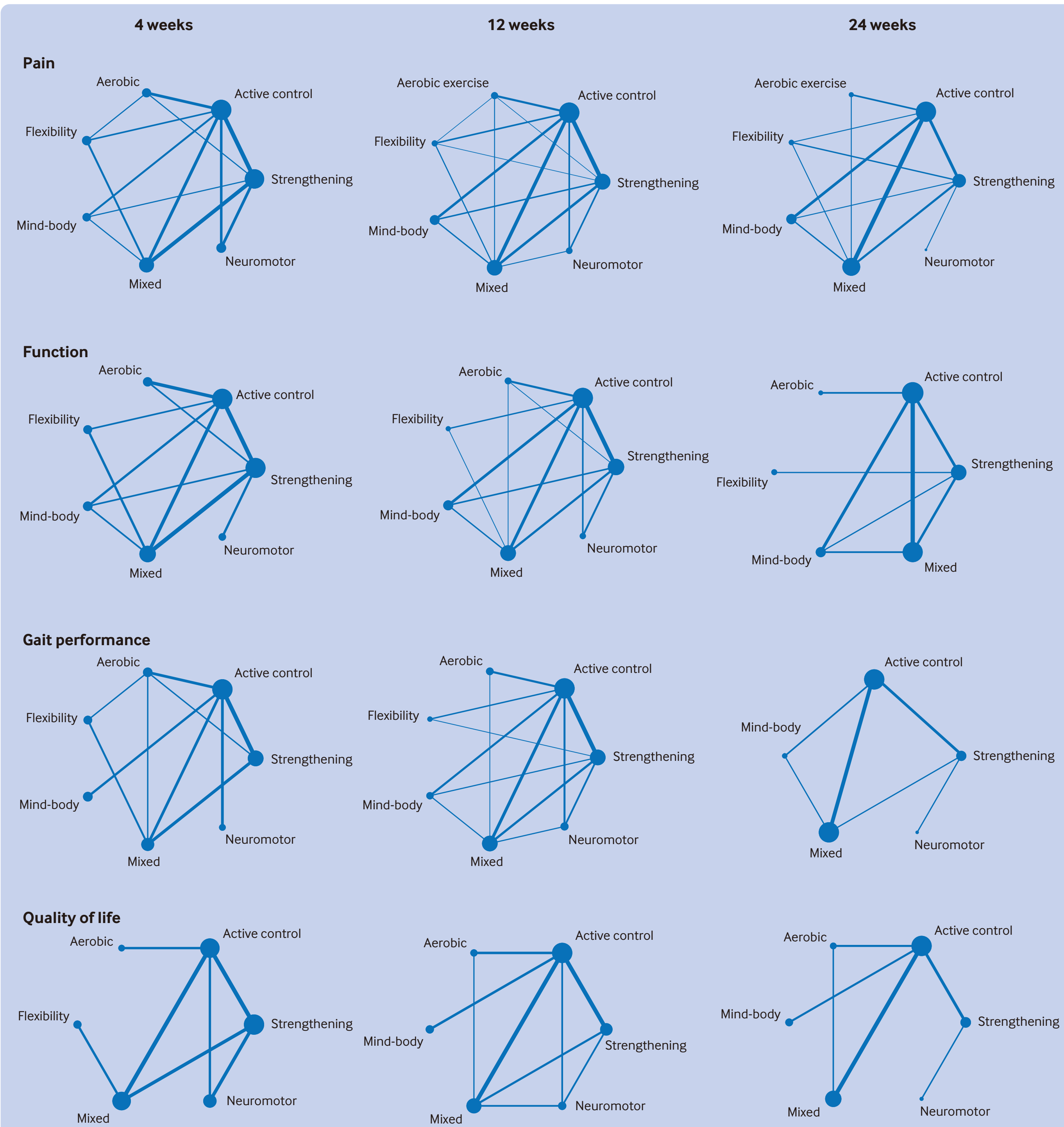

Fig 2 | Network graphs of direct comparisons for the 12 outcome measures. Line width is proportional to the number of studies comparing each pair of treatments, and node size is proportional to the number of participants (sample size)

Outcome and time	No of participants	Standardised mean difference (95% CI)	Certainty of evidence
Pain - 4 weeks			
Aerobic exercise	139	-1.10 (-1.68 to -0.52)	Moderate
Mind-body exercise	119	-0.82 (-1.54 to -0.10)	Very low
Mixed exercise	376	-0.75 (-1.16 to -0.34)	Very low
Flexibility exercise	147	-0.63 (-1.28 to 0.02)	Low
Strengthening exercise	611	-0.21 (-0.57 to 0.14)	Low
Neuromotor exercise	160	0.00 (-0.54 to 0.55)	Low
Pain - 12 weeks			
Aerobic exercise	280	-1.19 (-1.59 to -0.79)	Moderate
Mixed exercise	1554	-0.75 (-0.95 to -0.54)	Moderate
Mind-body exercise	595	-0.74 (-1.04 to -0.44)	Low
Strengthening exercise	1363	-0.64 (-0.83 to -0.46)	Very low
Flexibility exercise	229	-0.62 (-1.05 to -0.20)	Low
Neuromotor exercise	269	-0.50 (-0.88 to -0.11)	Moderate
Pain - 24 weeks			
Flexibility exercise	141	-0.99 (-1.63 to -0.36)	Low
Mixed exercise	1545	-0.41 (-0.62 to -0.2)	Moderate
Mind-body exercise	504	-0.34 (-0.65 to -0.03)	Low
Strengthening exercise	866	-0.22 (-0.49 to 0.05)	Moderate
Aerobic exercise	89	-0.02 (-0.55 to 0.50)	Moderate
Neuromotor exercise	32	-0.01 (-1.03 to 1.01)	Moderate
Function - 4 weeks			
Aerobic exercise	46	1.32 (0.44 to 2.20)	Low
Mind-body exercise	74	0.88 (0.03 to 1.73)	Moderate
Mixed exercise	222	0.42 (-0.13 to 0.96)	Moderate
Strengthening exercise	350	0.33 (-0.16 to 0.83)	Moderate
Neuromotor exercise	70	0.00 (-0.99 to 0.99)	Low
Flexibility exercise	64	-0.16 (-1.10 to 0.79)	Moderate
Function - 12 weeks			
Aerobic exercise	170	1.78 (1.05 to 2.51)	Moderate
Mixed exercise	1053	1.07 (0.68 to 1.46)	Moderate
Strengthening exercise	1147	0.86 (0.53 to 1.18)	Moderate
Mind-body exercise	431	0.75 (0.27 to 1.23)	Moderate
Neuromotor exercise	169	0.69 (-0.01 to 1.40)	Moderate
Flexibility exercise	111	0.45 (-0.49 to 1.39)	Low
Function - 24 weeks			
Aerobic exercise	68	0.87 (0.02 to 1.72)	Low
Mixed exercise	1325	0.56 (0.26 to 0.86)	Low
Mind-body exercise	327	0.19 (-0.24 to 0.61)	Low
Strengthening exercise	784	0.09 (-0.32 to 0.49)	Low
Flexibility exercise	108	0.05 (-1.14 to 1.24)	Low

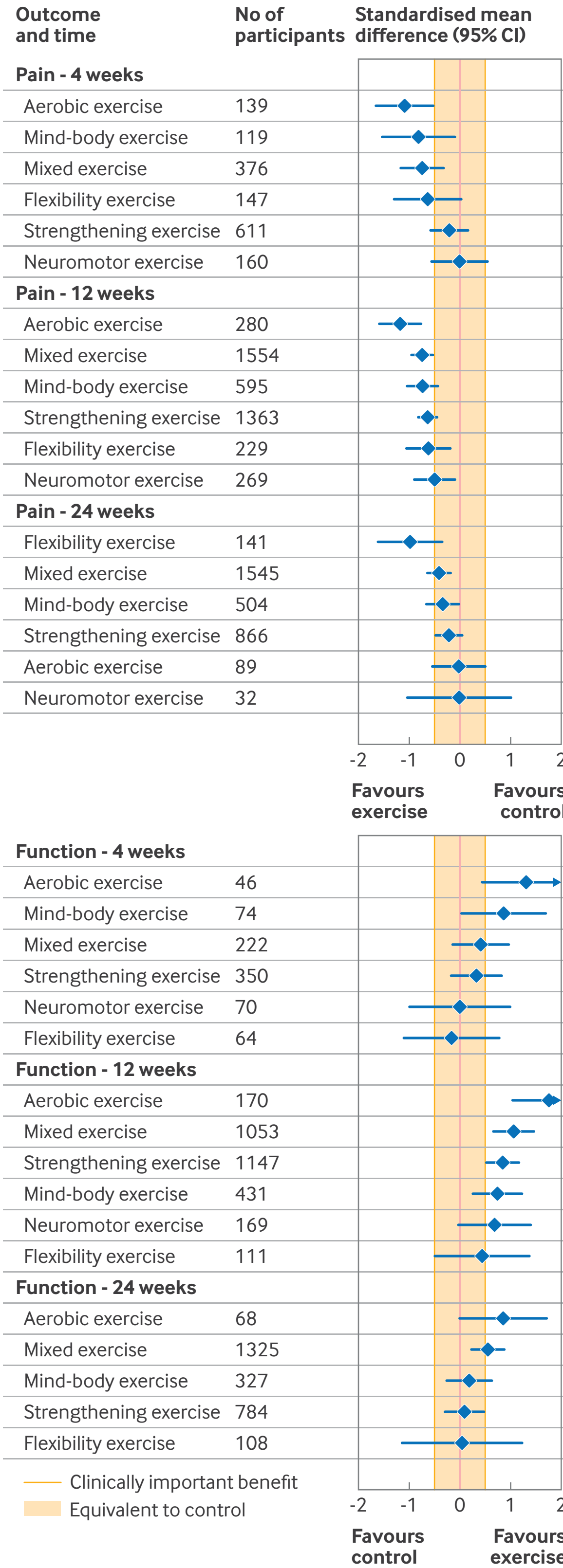

Fig 3 | Result of network meta-analysis of different exercise modalities on pain and function compared with control group

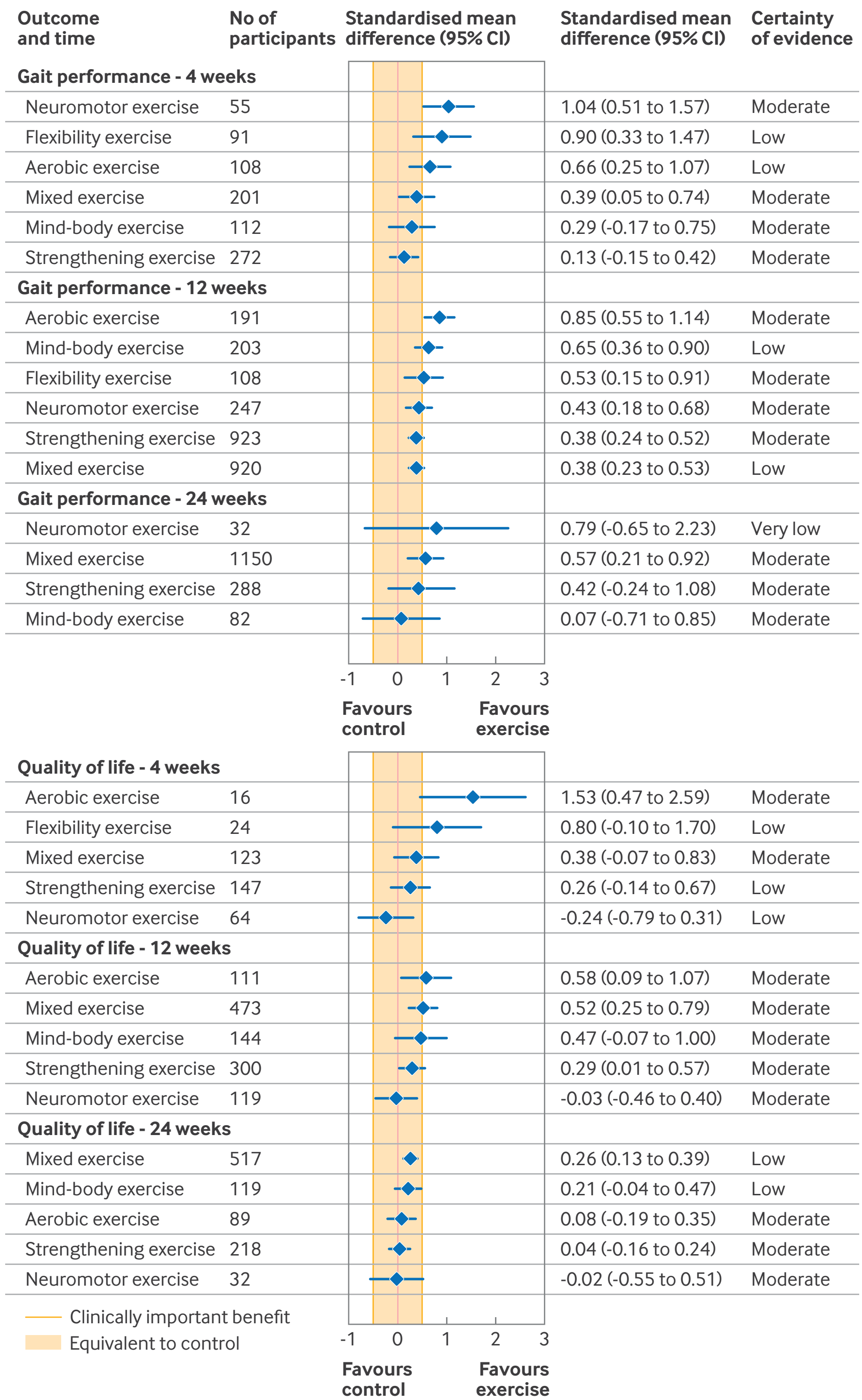

Fig 4 | Result of network meta-analysis of different exercise modalities on gait performance and quality of life compared with control group

for eight of the 12 plots (all P=0.19-0.96). Significant asymmetry was detected only for pain at four weeks (P=0.002) and 12 weeks (P<0.001), and for quality of life at four weeks (P=0.03) and 12 weeks (P=0.005), suggesting potential small study effects in these specific outcomes. We also conducted a corresponding meta-regression analysis (appendix 13), showing that higher dose and intensity of exercise was not associated with the effect of exercise. All sensitivity analyses (appendix 14) confirmed the robustness of the results.

Only comparisons with effects that exceed the minimal important difference, are supported by low or higher certainty of evidence (according to

GRADE assessment), or satisfy both requirements are described here. A more detailed narrative summary of all comparisons presented in figure 3 and figure 4 is available in appendix 15.

Pain

Among the included studies, 190 trials including 13 339 patients with knee osteoarthritis reported on pain. Aerobic exercise probably results in a large reduction in pain at short term follow-up (standardised mean difference −1.10, 95% confidence interval −1.68 to −0.52; moderate certainty). Flexibility exercise may reduce pain at short term follow-up (−0.63, −1.28 to 0.02; low certainty). Aerobic exercise probably results in a large reduction in pain at mid-term follow up (−1.19, −1.59 to −0.79, moderate certainty). Mixed exercise and neuromotor exercise probably reduce pain at mid-term follow-up (mixed exercise −0.75, −0.95 to −0.54; moderate certainty; neuromotor exercise −0.50, −0.88 to −0.11; moderate certainty). Mind-body exercise and flexibility exercise may reduce pain at mid-term follow-up (mind-body exercise −0.74, −1.04 to −0.44; low certainty; flexibility exercise −0.62, −1.05 to −0.20; low certainty). Flexibility exercise may result in a large reduction in pain at long term follow-up (−0.99, −1.63 to −0.36; low certainty).

Function

Function was reported in 127 trials including 9508 patients with knee osteoarthritis. Aerobic exercise may result in a large increase in function at short term follow-up (1.32, 0.44 to 2.20; low certainty). Mind-body exercise probably results in a large increase in function at short term follow-up (0.88, 0.03 to 1.73; moderate certainty). Aerobic exercise, mixed exercise, and strengthening probably result in a large increase in function at mid-term follow-up (aerobic exercise 1.78, 1.05 to 2.51; moderate certainty; mixed exercise 1.07, 0.68 to 1.46; moderate certainty; strengthening exercise 0.86, 0.53 to 1.18; moderate certainty). Mind-body exercise and neuromotor exercise probably increase function at mid-term follow-up (mind-body exercise 0.75, 0.27 to 1.23; moderate certainty; neuromotor exercise 0.69, −0.01 to 1.40; moderate certainty). Aerobic exercise may result in a large increase in function at long term follow-up (0.87, 0.02 to 1.72; low certainty). Mixed exercise may increase function at long term follow-up (0.56, 0.26 to 0.86; low certainty).

Gait performance

Gait performance was reported in 117 trials including 8483 patients with knee osteoarthritis. Neuromotor exercise probably results in a large increase in gait performance at short term follow-up (1.04, 0.51 to 1.57; moderate certainty). Flexibility exercise may result in a large increase in gait performance at short term follow-up (0.90, 0.33 to 1.47; low certainty). Aerobic exercise may increase gait performance at short term follow-up (0.66, 0.25 to 1.07; low certainty). Aerobic exercise probably results in a large increase in gait performance at mid-term follow-up (0.85, 0.55 to 1.14; moderate certainty). Mind-body exercise may increase in gait performance at mid-term follow-up (0.63, 0.36 to 0.90; low certainty). Flexibility exercise probably increases gait performance at mid-term follow-up (0.53, 0.15 to 0.91; moderate certainty). Mixed exercise probably increases gait performance at long term follow-up (0.57, 0.21 to 0.92; moderate certainty).

Quality of life

Quality of life was reported in 56 trials including 3940 patients with knee osteoarthritis. Aerobic exercise probably results in a large increase in quality of life at short term follow-up (1.53, 0.47 to 2.59; moderate certainty). Flexibility exercise may result in a large increase in quality of life at short term follow-up (0.80, −0.10 to 1.70; low certainty). Aerobic exercise and mixed exercise probably increase quality of life at mid-term follow-up (aerobic exercise 0.58, 0.09 to 1.07; moderate certainty; mixed exercise 0.52, 0.25 to 0.79; moderate certainty).

Safety

A total of 40 studies provided sufficient data on adverse effects occurring throughout the intervention period. Compared with control, the risk ratio was 2.33 (95% confidence interval 0.73 to 7.14) for aerobic exercise, 1.82 (0.82 to 4.0) for neuromotor exercise, 1.37 (0.85 to 2.27) for strengthening exercise, 1.35 (0.79 to 2.33) for mixed exercise, 0.94 (0.26 to 3.45) for flexibility exercise, and 0.61 (0.13 to 2.86) for mind-body exercise. All confidence intervals were wide and crossed the null value (risk ratio 1), indicating substantial imprecision and no clear evidence of difference in safety between exercise types and control (appendix 16).

Additional analysis

Results of surface under the cumulative ranking curve values are presented for all outcomes (appendix 17). Specifically, aerobic exercise consistently demonstrated the highest surface under the cumulative ranking curve values across outcomes (mean 0.72), indicating its superior overall efficacy compared with other interventions. Based on the above findings that aerobic exercise had a better overall treatment effect, we further explored whether improved benefits were associated with specific types of aerobic exercise. Post hoc results indicated no clear differences among types of aerobic exercise (appendix 18).

Discussion

Principal findings

This study analysed 217 randomised controlled trials, including 15 684 patients to compare five commonly prescribed exercise treatment regimens for primary osteoarthritis. The findings demonstrated that exercise was effective at improving pain, function, gait performance, and quality of life in individuals with knee osteoarthritis. Compared to the control group, aerobic exercises were found to be beneficial in relieving pain (short term and mid-term follow-up), improving function (short term, mid-term and long term follow-up), gait performance (short term and mid-term follow-up), and quality of life (short term and mid-term follow-up). Mind-body exercise likely results in a large increase in function, whereas neuromotor exercise likely results in a large increase in gait performance at short term follow-up. Strengthening and mixed exercise probably result in a large increase in function at mid-term follow-up, while flexibility exercise may result in a large reduction in pain at long term follow-up. None of the exercise interventions resulted in more adverse safety events than the control group, suggesting that exercise therapy is a safe treatment approach.

The outcomes of therapeutic exercise in patients with knee osteoarthritis can be influenced by a variety of factors, such as adherence to the programme and inherent characteristics of the outcome measures. Evidence suggests that greater adherence is associated with larger improvements in pain and function.[27] However, adherence was not consistently reported across the included trials in our study, which limits further analysis. Additionally, commonly used outcome measures, such as Western Ontario and McMaster Universities Osteoarthritis Index or Knee injury and Osteoarthritis Outcome Score, may lack the required sensitivity to detect early neuromuscular or psychosocial improvements induced by exercise.[28 29]

While our primary analysis used a conservative minimal important difference threshold of 0.5 to reflect clearly meaningful clinical effects, we also considered how the interpretation might change if a more permissive threshold of 0.2 was used, which is often cited as the lower bound for minimal important difference. Given the nature of knee osteoarthritis as a chronic condition, long term outcomes are of particular interest. In our primary analysis, which used a minimal important difference of 0.5, none of the exercise interventions showed clinically meaningful differences compared with the control group at the 24 week follow-up. However, when the minimal important difference was lowered to 0.2, both flexibility and mixed exercise showed clinically relevant improvements in certain long term outcomes. Specifically, flexibility exercise showed clinically relevant improvements in pain, while mixed exercise also exceeded the minimal important difference threshold for improvements in pain, function, and gait performance. The appearance of these long term effects underscores the potential value of sustained engagement in certain types of exercise, particularly those with flexibility based or multimodal components. Our use of the 0.5 minimal important difference thus reflects a conservative approach in our data interpretation, while acknowledging that different thresholds for minimal important difference may reveal a broader scope of benefit.

Comparisons with previous studies

Previously, Uthman et al[16] used network meta-analysis to determine the efficacy of various exercise therapies on osteoarthritis. Following this, Goh et al[15] standardised and broadened the exercises examined in their study and added new outcome measures (performance and quality of life). We compared our findings with the most recent review of exercise interventions for osteoarthritis, conducted by Goh et al in 2019.[15] The review by Goh and colleagues comprised 103 randomised controlled trials that assessed five exercise intervention regimens for osteoarthritis. Their findings suggested that aerobic exercise was the most beneficial for pain and performance, which is consistent with our conclusion. The review also found that strengthening and flexibility exercises had a moderate effect on various outcomes, while mixed exercise was least effective for all outcomes.

Our approach differs substantially to that of Goh and colleagues, allowing a more in-depth analysis of the results. Firstly, Goh et al combined outcome indicators taken at various time points and used these together to reflect on the therapeutic impact of exercise. A potential problem with this approach is that the follow-up time varied greatly among the included studies, but the review used the latest available data point for analysis regardless of the duration of follow-up. In contrast, we stratified the results into three time periods, representing the short term, mid-term, and long term outcomes at approximately four, 12, and 24 weeks of treatment, respectively.

Secondly, we used the most recent version of the risk of bias tool (RoB 2), where the evaluation indicators were more precise and unambiguous than the earlier version (RoB 1).[30-32] The use of RoB 2 yielded more objective evaluations, owing to the standardised questions and response options of RoB 2. Additionally, we used the GRADE system for a rigorous assessment of certainty of evidence, including direct and indirect evidence, and final network evidence.[33 34] Thus, we may be more confident in the dependability of our data and reliability of our final assessment.

Substantial evidence has accrued since Goh et al's review in 2019. Our analysis included studies published between 2019 and 2024, almost doubling the number of randomised controlled trials included in this research. We did not find any high quality network meta-analyses comparing different exercise programmes for knee osteoarthritis published since 2019; therefore, to the best of our understanding, the present analysis is the most comprehensive and up to date.[15]

In addition to the network meta-analysis by Goh et al, several recent systematic reviews and meta-analyses have investigated the effects of various exercise modalities in knee osteoarthritis from different perspectives. Gao and colleagues focused on mind-body exercises such as tai chi, yoga, and pilates, and reported significant improvements in pain, stiffness, physical function, and psychological wellbeing compared to usual care, with tai chi and pilates appearing the most effective.[35] Similarly, Qiao et al found that mind-body exercise resulted in significant improvement in pain, stiffness, function, depression, and walking ability compared with the no exercise control group.[36] Wang et al found evidence for the effect of exercise on balance function, where they identified multicomponent training and high weekly exercise volume (≥180 min/week) as key moderators of efficacy.[37] Meanwhile, an individual participant data meta-analysis by van Middelkoop et al revealed that individuals with shorter symptom duration experienced greater benefit from exercise therapy, suggesting that symptom duration may be an important treatment effect modifier.[38] Xiang et al further demonstrated that telehealth supported exercise programmes led to modest improvements in pain, function, and self-efficacy (ie an individual's confidence in their ability to carry out specific activities), indicating that digital delivery may be a viable option for long term management.[13] While differing in focus and methodology, these studies collectively underscore the importance of exercise for long term knee osteoarthritis management. Our systematic review and network meta-analysis extends the current evidence by providing evidence stratified by time and emphasis on methodological transparency.

Limitations

Some limitations of this systematic review should be noted. Firstly, we only included randomised controlled trials where the participants had knee osteoarthritis. The conclusions drawn from this study may not therefore apply to other types of osteoarthritis. Secondly, this study lacked evidence resulting from direct comparisons between some of the included types of exercise, which may have influenced our findings. Most studies made direct comparisons between exercise intervention and the control group, but in this study, we lacked head-to-head comparisons between different exercise types. Many network results could only be determined through indirect comparisons, and the results should be interpreted with caution.

Thirdly, while most of Egger's tests were statistically non-significant, indicating a low risk of publication bias overall, significant asymmetry was observed in four comparisons, specifically for outcomes relating to pain and quality of life at early time points. This asymmetry suggests potential small study effects that warrant cautious interpretation of results. Finally, poor and inconsistent reporting of data on socioeconomic characteristics hindered the investigation of these factors as potential effect modifiers.

Implications for clinical practice

Both the Osteoarthritis Research Society International and the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases advocate for education, structured exercise, and weight loss as primary treatments for osteoarthritis.[39 40] However, current guidelines on the relative usefulness and therapeutic efficacy of different exercise therapies that should be prescribed to patients with osteoarthritis are ambiguous. Limited data are available to assist health professionals and patients in determining which exercise interventions should be prioritised.[39 41] For example, current guidelines prioritise aerobic, strengthening, and mind-body exercises and temporarily omit aquatic sports from the core set of exercises. Our review contributes additional evidence supporting the effectiveness of aerobic exercises as treatments for knee osteoarthritis. We found that aerobic exercise offered superior combined therapeutic benefits. Given these findings, we provide the following recommendations for clinical practice.

We recommend aerobic exercise as a first line intervention for knee osteoarthritis management, particularly when the aim is to improve functional capacity and reduce pain. While other exercises may offer complementary benefits to patients, they should not replace aerobic exercise as the primary strategy. Exercise prescriptions should be tailored based on patient preferences, comorbidities, and mobility levels to maximise adherence to the programme and long term benefits to patients.[42 43]

Patients should engage regularly in structured aerobic activities, such as walking, cycling, or swimming, to optimise symptom management. If aerobic exercise is not possible owing to individual limitations, alternative forms of structured physical activity may still be beneficial. Patients should consult healthcare professionals to develop an individualised exercise plan that aligns with their physical capabilities and treatment goals.

Implications for research

Our study highlights several critical limitations in current trials on exercise interventions for knee osteoarthritis, emphasising the need for higher quality clinical research in the future. One major issue is the high variability in study quality, with small sample sizes in some trials. Additionally, incomplete reporting of exercise parameters, such as intensity, frequency, and duration of exercise, makes it difficult to establish a clear link between exercise interventions and treatment outcomes. Given that knee osteoarthritis is a chronic and progressive condition, studying long term outcomes is particularly important for evaluating the sustained benefits of exercise. However, only a few studies (n=12, 5.5%) have reported follow-up data beyond one year. Therefore, future research should prioritise well designed, adequately powered trials with longer follow-up durations to better assess the durability of exercise effects over time. Finally, Kellgre-Lawrence grading was recorded to classify the severity of knee osteoarthritis in the included studies. However, nearly all studies included participants with different Kellgre-Lawrence grades, making it difficult to distinguish the therapeutic effects of exercise intervention on patients with different degrees of knee osteoarthritis. Future research focusing on disease severity could help clinicians make more informed choices regarding exercise therapy for patients with knee osteoarthritis.

Conclusions

Exercise treatment offers obvious benefits for individuals with knee osteoarthritis. With moderate certainty quality evidence, aerobic exercise might be the best exercise modality for improvements in pain, function, gait performance, and quality of life. This study presents a comprehensive and up-to-date analysis of the effects of exercise intervention as a treatment strategy for knee osteoarthritis, which could assist clinicians in prescribing therapy to improve treatment outcomes for patients.

AUTHOR AFFILIATIONS

[1]Department of Orthopaedic Surgery, First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

[2]Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[3]Department of Orthopaedics, Affiliated Renhe Hospital of China Three Gorges University, Yichang, China

[4]Arthritis Clinic & Research Centre, Peking University People's Hospital, Peking University, Beijing, China

[5]Department of Medical Statistics, School of Public Health, Sun Yat-sen University, Guangzhou, China

[6]Department of Orthopaedics, Second Hospital of Shanxi Medical University, Taiyuan, China

[7]Institute of Global Health, University of Geneva, Geneva, Switzerland

[8]Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada

[9]Evidence Based Social Science Research Centre, School of Public Health, Lanzhou University, Lanzhou, China

[10]Department of Social Medicine and Health Management, School of Public Health, Lanzhou University, Lanzhou, China

[11]School of Biomedical Engineering, Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW, Australia

[12]Department of Rehabilitation Medicine, First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

Contributors: BW, LG, and YH are joint corresponding authors and contributed equally to this work. LY, DL, and DX are joint first authors and contributed equally to this work. BW, LY, DL, DX, LG, and YH contributed to conception and trial design. LY, DL, ZF, GD, JJ, XL, SL, HL, XLiu, YR, and DX conducted screening, extraction, and the risk of bias assessment. LY, DL, DX, BW, YY, and JC provided clinical expertise. DL, GD, DX, and CT conducted statistical analyses. DL, LY, GD, and JJ prepared the graphs and tables. QW, AB, JE, and JJL critically revised the manuscript. BW, YH, and LG provided supervision and mentorship. LY and DL drafted the manuscript. All authors gave crucial feedback on the revised report and approved the final version of the manuscript. BW is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study was supported by the National Natural Science Foundation of China (grant 81802204), Zhejiang Medicine and Health Science and Technology Project (grant 2023RC010), Zhejiang Basic Public Welfare Fund (grant LTGY23H060007), and Zhejiang University First Affiliated Hospital of Zhejiang University School of Medicine Medical Education Research Key Project (grant zyjg202404). The funders had no role in considering the study design or in the collection, analyses, interpretation of data, writing of the report, or decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: support from the National Natural Science Foundation of China, Zhejiang Medicine and Health Science and Technology Project, Zhejiang Basic Public Welfare Fund, and Zhejiang University The First Affiliated Hospital of Zhejiang University School of Medicine Education Research Key Project 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.

Ethical approval: This systematic review and network meta-analysis did not require ethical approval.

Data sharing: The guarantor (BW) is willing to examine all requests for the full dataset after a period of two years from the date of this publication. The statistical codes used for analyses are available in appendix 8.3.

Transparency: The lead author (BW) 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 planned and registered have been explained.

Dissemination to participants and related patient and public communities: Findings will be disseminated via academic conferences, educational sessions, social media, and newsletters. We will collaborate with patient organisations to codevelop plain language summaries and infographics, incorporating insights from individuals with lived experience of knee osteoarthritis to ensure clarity and accessibility.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. *Lancet* 2019;393:1745-59. doi:10.1016/S0140-6736(19)30417-9

2 Glyn-Jones S, Palmer AJ, Agricola R, et al. Osteoarthritis. *Lancet* 2015;386:376-87. doi:10.1016/S0140-6736(14)60802-3

3 Long H, Liu Q, Yin H, et al. Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. *Arthritis Rheumatol* 2022;74:1172-83. doi:10.1002/art.42089

4 Sharma L. Osteoarthritis of the Knee. *N Engl J Med* 2021;384:51-9. doi:10.1056/NEJMcp1903768

5 Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: A review. *JAMA* 2021;325:568-78. doi:10.1001/jama.2020.22171

6 Martel-Pelletier J, Barr AJ, Cicuttini FM, et al. Osteoarthritis. *Nat Rev Dis Primers* 2016;2:16072. doi:10.1038/nrdp.2016.72

7 Shi J, Fan K, Yan L, et al. Cost Effectiveness of Pharmacological Management for Osteoarthritis: A Systematic Review. *Appl Health Econ Health Policy* 2022;20:351-70. doi:10.1007/s40258-022-00717-0

8 Bennell KL, Dobson F, Hinman RS. Exercise in osteoarthritis: moving from prescription to adherence. *Best Pract Res Clin Rheumatol* 2014;28:93-117. doi:10.1016/j.berh.2014.01.009

9 Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. *Br J Sports Med* 2015;49:1554-7. doi:10.1136/bjsports-2015-095424

10 Jansen MJ, Viechtbauer W, Lenssen AF, Hendriks EJ, de Bie RA. Strength training alone, exercise therapy alone, and exercise therapy with passive manual mobilisation each reduce pain and disability in people with knee osteoarthritis: a systematic review. *J Physiother* 2011;57:11-20. doi:10.1016/S1836-9553(11)70002-9

11 Larkin H. National Academies: Expand US Wastewater Disease Surveillance. *JAMA* 2023;329:531. doi:10.1001/jama.2023.0739

12 Juhl C, Christensen R, Roos EM, Zhang W, Lund H. 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. doi:10.1002/art.38290

13 Xiang XN, Wang ZZ, Hu J, et al. Telehealth-Supported Exercise or Physical Activity Programs for Knee Osteoarthritis: Systematic Review and Meta-Analysis. *J Med Internet Res* 2024;26:e54876. doi:10.2196/54876

14 Bell EC, Wallis JA, Goff AJ, Crossley KM, O'Halloran P, Barton CJ. Does land-based exercise-therapy improve physical activity in people with knee osteoarthritis? A systematic review with meta-analyses. *Osteoarthritis Cartilage* 2022;30:1420-33. doi:10.1016/j.joca.2022.07.008

15 Goh SL, Persson MSM, Stocks J, et al. Relative Efficacy of Different Exercises for Pain, Function, Performance and Quality of Life in Knee and Hip Osteoarthritis: Systematic Review and Network Meta-Analysis. *Sports Med* 2019;49:743-61. doi:10.1007/s40279-019-01082-0

16 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. doi:10.1136/bmj.f5555

17 Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777-84. doi:10.7326/M14-2385

18 Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc* 2011;43:1575-81. doi:10.1249/MSS.0b013e31821ece12

19 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. doi:10.1136/bmj.l4898

20 Higgins JT, eds. Chapter 15: Interpreting results and drawing conclusions. Cochrane Handbook Systematic Reviews of Interventions version 6: Cochrane 2021.

21 Santesso N, Glenton C, Dahm P, et al, GRADE Working Group. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol* 2020;119:126-35. doi:10.1016/j.jclinepi.2019.10.014

22 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582-92. doi:10.1097/01.MLR.0000062554.74615.4C

23 Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006;59:7-10. doi:10.1016/j.jclinepi.2005.06.006

24 Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J Clin Epidemiol* 2016;76:193-9. doi:10.1016/j.jclinepi.2016.02.016

25 Izcovich A, Chu DK, Mustafa RA, Guyatt G, Brignardello-Petersen R. A guide and pragmatic considerations for applying GRADE to network meta-analysis. *BMJ* 2023;381:e074495. doi:10.1136/bmj-2022-074495

26 Guyatt GH, Oxman AD, Vist GE, et al, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD

27 Pisters MF, Veenhof C, Schellevis FG, Twisk JW, Dekker J, De Bakker DH. Exercise adherence improving long-term patient outcome in patients with osteoarthritis of the hip and/or knee. *Arthritis Care Res (Hoboken)* 2010;62:1087-94. doi:10.1002/acr.20182

28 Parkes MJ, Callaghan MJ, O'Neill TW, Forsythe LM, Lunt M, Felson DT. Sensitivity to Change of Patient-Preference Measures for Pain in Patients With Knee Osteoarthritis: Data From Two Trials. *Arthritis Care Res (Hoboken)* 2016;68:1224-31. doi:10.1002/acr.22823

29 Ageberg E, Roos EM. Neuromuscular exercise as treatment of degenerative knee disease. *Exerc Sport Sci Rev* 2015;43:14-22. doi:10.1249/JES.0000000000000030

30 Loef M, Walach H, Schmidt S. Interrater reliability of ROB2 - an alternative measure and way of categorization. *J Clin Epidemiol* 2022;142:326-7. doi:10.1016/j.jclinepi.2021.09.003

31 Minozzi S, Gonzalez-Lorenzo M, Cinquini M, et al, University of Milan Post Graduate Course on Systematic Review Working Group. Adherence of systematic reviews to Cochrane RoB2 guidance was frequently poor: a meta epidemiological study. *J Clin Epidemiol* 2022;152:47-55. doi:10.1016/j.jclinepi.2022.09.003

32 Minozzi S, Dwan K, Borrelli F, Filippini G. Reliability of the revised Cochrane risk-of-bias tool for randomised trials (RoB2) improved with the use of implementation instruction. *J Clin Epidemiol* 2022;141:99-105. doi:10.1016/j.jclinepi.2021.09.021

33 Harvey LA. GRADE the evidence. *J Physiother* 2013;59:5. doi:10.1016/S1836-9553(13)70139-5

34 Xie CX, Machado GC. Clinimetrics: Grading of Recommendations, Assessment, Development and Evaluation (GRADE). *J Physiother* 2021;67:66. doi:10.1016/j.jphys.2020.07.003

35 Gao K, Tao J, Liang G, Gong C, Wang L, Wang Y. Comparative efficacy of mind-body exercise for pain, function, quality of life in knee osteoarthritis: a systematic review and network meta-analysis. *J Orthop Surg Res* 2025;20:384. doi:10.1186/s13018-025-05682-7

36 Qiao H, Hao X, Wang G. Effects of mind-body exercise on knee osteoarthritis: a systematic review and meta-analysis of randomized controlled trials. *BMC Musculoskelet Disord* 2024;25:229. doi:10.1186/s12891-024-07278-4

37 Wang X, Chen Z, Liang Y, et al. Effects of Exercise on Balance Function in People with Knee Osteoarthritis: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Healthcare (Basel)* 2025;13:1312. doi:10.3390/healthcare13111312

38 van Middelkoop M, Schiphof D, Hattle M, et al. People with short symptom duration of knee osteoarthritis benefit more from exercise therapy than people with longer symptom duration: An individual participant data meta-analysis from the OA trial bank. *Osteoarthritis Cartilage* 2024;32:1620-7. doi:10.1016/j.joca.2024.07.007

39 Arden NK, Perry TA, Bannuru RR, et al. Non-surgical management of knee osteoarthritis: comparison of ESCEO and OARSI 2019 guidelines. *Nat Rev Rheumatol* 2021;17:59-66. doi:10.1038/s41584-020-00523-9

40 Rausch Osthoff AK, Niedermann K, Braun J, et al. 2018 EULAR recommendations for physical activity in people with inflammatory arthritis and osteoarthritis. *Ann Rheum Dis* 2018;77:1251-60. doi:10.1136/annrheumdis-2018-213585

41 Geenen R, Overman CL, Christensen R, et al. EULAR recommendations for the health professional's approach to pain management in inflammatory arthritis and osteoarthritis. *Ann Rheum Dis* 2018;77:797-807. doi:10.1136/annrheumdis-2017-212662

42 Lee FI, Lee TD, So WK. Effects of a tailor-made exercise program on exercise adherence and health outcomes in patients with knee osteoarthritis: a mixed-methods pilot study. *Clin Interv Aging* 2016;11:1391-402. doi:10.2147/CIA.S111002

43 Zmerly H, Milanese C, El Ghoch M, et al. Personalized Physical Activity Programs for the Management of Knee Osteoarthritis in Individuals with Obesity: A Patient-Centered Approach. *Diseases* 2023;11:182. doi:10.3390/diseases11040182

Web appendix: Supplementary files