OPEN ACCESS

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

Qianlin Weng,[1] Siew-Li Goh [2,3] Jing Wu,[4] Monica S M Persson,[5,6] Jie Wei,[4,7] Aliya Sarmanova,[5,6] Xiaoxiao Li,[4] Michelle Hall,[6,8,9] Michael Doherty,[5,6,9] Ting Jiang,[1,5,6,10] Chao Zeng [1,4,11] Guanghua Lei [1,4,11,12] Weiya Zhang[5,6,9]

ABSTRACT

Objective Clinical guidelines recommend exercise as a core treatment for knee or hip osteoarthritis (OA). However, how its analgesic effect compares to analgesics, for example, oral non-steroidal anti-inflammatory drugs (NSAIDs) and paracetamol—the most commonly used analgesics for OA, remains unknown.

Design Network meta-analysis.

Data sources PubMed, Embase, Scopus, Cochrane Library and Web of Science from database inception to January 2022.

Eligibility criteria for selecting studies Randomised controlled trials (RCTs) comparing exercise therapy with oral NSAIDs and paracetamol directly or indirectly in knee or hip OA.

Results A total of n=152 RCTs (17 431 participants) were included. For pain relief, there was no difference between exercise and oral NSAIDs and paracetamol at or nearest to 4 (standardised mean difference (SMD)=−0.12, 95% credibility interval (CrI) −1.74 to 1.50; n=47 RCTs), 8 (SMD=0.22, 95% CrI −0.05 to 0.49; n=2 RCTs) and 24 weeks (SMD=0.17, 95% CrI −0.77 to 1.12; n=9 RCTs). Similarly, there was no difference between exercise and oral NSAIDs and paracetamol in functional improvement at or nearest to 4 (SMD=0.09, 95% CrI −1.69 to 1.85; n=40 RCTs), 8 (SMD=0.06, 95% CrI −0.20 to 0.33; n=2 RCTs) and 24 weeks (SMD=0.05, 95% CrI −1.15 to 1.24; n=9 RCTs).

Conclusions Exercise has similar effects on pain and function to that of oral NSAIDs and paracetamol. Given its excellent safety profile, exercise should be given more prominence in clinical care, especially in older people with comorbidity or at higher risk of adverse events related to NSAIDs and paracetamol.

TRIAL REGISTRATION NUMBER
CRD42019135166

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Exercise is an effective treatment for osteoarthritis. It has been recommended as a core therapy by National Institution for Health and Care Excellence and other treatment guidelines for osteoarthritis because of its favourable safety profile.

⇒ However, whether the analgesic effect of exercise is comparable to that of oral non-steroidal anti-inflammatory drugs (NSAIDs) and paracetamol, the most common analgesic given for osteoarthritis, remains unknown.

WHAT ARE THE NEW FINDINGS

⇒ Exercise has been compared with oral NSAIDs and paracetamol through a network meta-analysis of 152 randomised controlled trials (17 431 participants) for knee or hip osteoarthritis.

⇒ Exercise is indeed a medicine and its analgesic effect is similar to that of oral NSAIDs and paracetamol.

INTRODUCTION

Osteoarthritis (OA) is the most common form of joint disease and the leading cause of pain in older people.[1] Pain symptoms associated with knee or hip OA result in increased physical and walking disability and an increased risk of all-cause mortality.[2] The main management goal in OA is to relieve pain without increasing treatment-related adverse effects.

Oral non-steroidal anti-inflammatory drugs (NSAIDs) and paracetamol (or acetaminophen) are the most frequently prescribed analgesics to control pain and improve physical function of OA,[3 4] with 10% to 35% of the OA population reporting the use of oral NSAIDs or paracetamol.[5 6] However, oral NSAIDs and paracetamol are associated with gastrointestinal or cardiovascular complications and even increased risk of death, especially in older people with comorbidities.[6–8] Current National Institution for Health and Care Excellence and international guidelines strongly recommend exercise as a core therapy for management of knee or hip OA.[9–12] However, implementation of exercise in clinical practice remains limited and suboptimal, in part due to the time commitment required by health practitioners, absence of agreed standard protocols, lack of confidence in capability to exercise and concerns about joint overloading among people with OA.[13–15] Also, it is still unclear whether exercise has an analgesic effect equivalent to that from analgesics such as oral NSAIDs and paracetamol. The current evidence on the direct comparison between these treatments is sparse.[16–19] The

majority of exercise trials used usual care as a control whereas the majority of drug trials used placebo as a control; therefore, the effect size from exercise trials is not comparable to that from oral NSAIDs and paracetamol trials. This may further preclude the uptake of exercise as a core therapy in clinical practice.[20–22] The comparative efficacy of exercise and oral NSAIDs and paracetamol will help to confirm the analgesic benefit of exercise. Such information may enhance public awareness of exercise as an effective treatment for OA rather than just a physical activity for general health. It will inform patient–practitioner discussion and shared decision-making and encourage patients to be more proactive about including exercise in their individualised OA management plan.[23]

However, few head-to-head randomised controlled trials (RCTs) are available,[16–19] and the results were discordant. In these four previous RCTs, one study (n=48) reported a more beneficial effect of oral NSAIDs,[16] two studies (n=141 and 142, respectively) showed no difference between the effect of exercise and oral NSAIDs and paracetamol[17 18] and one study found more benefit from exercise.[19] We, therefore, undertook this network meta-analysis (NMA) to gather all RCTs, which either directly compared exercise with oral NSAIDs and paracetamol, or indirectly compared these two treatments via a common comparator in an NMA (e.g., usual care). We estimated the comparative effect size for the two common outcomes in OA trials, that is, pain and function in people with knee or hip OA.

METHODS

Protocol and registration

The reporting of this NMA followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses[24] and the NMA protocol was registered with PROSPERO.

Eligibility criteria

The inclusion criteria were as follows: (1) RCTs; (2) studies on participants with knee or hip OA; (3) studies comparing exercise therapy with oral NSAIDs and paracetamol, or studies comparing exercise therapy with any common comparator that may be shared with oral NSAIDs and paracetamol (usual care/ no treatment/waiting list control, glucosamine sulphate/chondroitin sulphate, intra-articular hyaluronic acid, topical NSAIDs, acupuncture); (4) studies reporting pain or function; (5) studies published in any language.

The following studies were excluded: (1) secondary analyses, such as combined data analyses of published RCTs; (2) studies with less than 1 week of follow-up; (3) studies using a cross-over design; (4) exercise therapy or oral NSAIDs and paracetamol or any common comparator combined with other active interventions; (5) studies for postoperative pain; (6) abstract only (insufficient data).

Throughout this text, the term 'exercise' was used to refer to 'exercise therapy'. Exercise therapy is defined as a planned, structured, repetitive and purposeful physical activity for the improvement or maintenance of a specific health condition (or disease).[25] Exercise therapy encompasses aerobic, muscle strengthening, flexibility/neuromotor skills training (flexibility/ skill) or mind–body exercise (e.g., tai chi, yoga).[26] Studies were classified as mixed exercise when they included more than one exercise type mentioned above, or when the authors did not specify it as a single component exercise. Any form of exercise therapy was eligible, regardless of content, duration, frequency or intensity.

'Usual care' control was classified based on the report. In 'usual care', participants were expected to continue their routine general care. Control groups that were not given any specific intervention such as 'waiting list', usual physical activity or no treatment, or where the authors did not specify the nature of the control, were also classified as 'usual care'. 'Waiting-list' controls were given an active intervention after the trial period, with no new intervention being delivered during the trial period.[27]

Literature search

Systematic literature searches were conducted using PubMed, Embase, Scopus, Cochrane Library and Web of Science up to January 2022 (online supplemental appendix 1). Additionally, the references of relevant reviews and selected articles were examined for potentially relevant trials.

Study selection

After removing duplicates, all titles and abstracts were screened independently for potentially eligible studies by two reviewers (QW and JW), and relevant RCTs were identified. Reports of studies considered potentially relevant by at least one reviewer were retrieved in full text. The eligibility of the retrieved full-text articles for final inclusion was assessed independently by two reviewers. Disagreement was resolved through discussion, and if no consensus was reached, a third reviewer (CZ) was involved to make the final decision.

Data extraction

Baseline characteristics and outcome data were extracted into a standardised form by two independent assessors (QW and JW). The outcome measures of interest were pain and function scores reported at baseline, 4, 8 and 24 weeks. When 4-week data were not available, we used the data reported at the closest time point from 2 to 6 weeks of follow-up. All types of exercise and oral NSAIDs and paracetamol were included during data extraction. Change-from-baseline pain scores were extracted or calculated. If a study reported multiple pain scales, the scale with the highest sensitivity to change was used.[28] The function subscale of the Western Ontario and McMaster Universities Arthritis Index (WOMAC) was used for the assessment of functional improvement. If a study did not measure or report WOMAC function, the Lequesne Index or one of the other functional measurement scales was used instead. Corresponding authors of studies with missing data were contacted through ResearchGate or email with a request for the data.

Quality assessment

The Cochrane risk of bias assessment tool[29] was used to assess the methodological quality of RCTs. The following seven domains were evaluated in each included study: random sequence generation; allocation concealment; blinding of participant and personnel; blinding of outcome assessment; incomplete outcome data; selective reporting and other biases. Each domain was assigned a judgement of low risk, high risk or unclear risk of bias. Because it is not possible to truly blind health practitioners and participants in any study related to exercise treatment, this was not included in the overall risk of bias assessment of each study. In addition, the quality level of this NMA was evaluated by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach.[30 31] The results were evaluated for evidence quality, which was classified into high, moderate, low or very low levels (online supplemental appendix 2). The summary of findings of outcome was presented using the

template of the GRADE NMA summary of findings (NMA-SoF) table for exercise compared with oral NSAIDs and paracetamol (online supplemental appendix 2).[32]

Statistical analysis

For pain and functional outcomes, the standardised mean difference (SMD) was used to standardise the results to a uniform scale when the studies assessed the same outcome with different instruments.[29] Two SMDs were calculated, one for exercise versus usual care to confirm whether exercise was effective (better than usual care) and the other for exercise versus oral NSAIDs and paracetamol to examine whether exercise was as effective as oral NSAIDs and paracetamol. For both SMDs, exercise was the intervention and usual care or oral NSAIDs and paracetamol were the control; hence, a positive value favours exercise, whereas a negative value favours usual care or oral NSAIDs and paracetamol. It was calculated by dividing the difference in mean values between treatment groups by the pooled standard deviation (SD) of change-from-baseline. If SDs were not reported, we calculated them from standard errors (SEs) or confidence intervals (CIs). When change-from-baseline SDs were not presented, we calculated the missing SDs using the formula: $SD_{change} = \sqrt{SD_{baseline}^2 + SD_{follow-up}^2 - (2*0.5*SD_{baseline}*SD_{follow-up})}$.[29] Those studies which did not report mean scores at baseline and endpoint or change-from-baseline scores with SDs, SEs or CIs were not included in the analyses. For studies with multiple intervention groups, we will combine all relevant groups to create a single pairwise comparison.[29] Clinically, an SMD of 0.20 was considered as a small effect, 0.50 a moderate effect and 0.80 a large effect, according to Cohen.[33]

Bayesian NMA methods were used to assess the comparative efficacy of exercise versus oral NSAIDs and paracetamol. Bayesian NMA methods pool direct and indirect evidence on relative treatment effects.[34 35] The Markov Chain Monte Carlo method was used to estimate posterior densities for unknown variables.[35 36] A random effects model was adopted as the most appropriate and conservative method to account for differences between RCTs. Two Markov chains ran simultaneously with different initial values.[36 37] Bayesian random effects model was based on the Dias model and uninformative prior probability distributions were used for all parameters.[38 39] A total of 50 000 simulations were generated for each of the two sets of initial values and the first 10 000 simulations were discarded as the burn-in period. The WinBUGS codes are available at http://www.bristol.ac.uk/social-community-medicine/projects/mpes/.

The pooled SMDs were generated from the median of the posterior distribution.[37 40] The 2.5th and 97.5th percentiles of the posterior distribution were considered the lower and upper limit, respectively, of the 95% credible interval (CrI). A 95% CrI may be interpreted as there being a 95% probability that the parameter takes a true value in the specified range.[41] Heterogeneity was defined as the variability of results across trials, with $\tau^2 < 0.04$ indicating a low level and $\tau^2 > 0.4$ a high level.[42–44] The fit of the model to the data was measured by calculating the posterior mean residual deviance.[38] If the mean of the residual deviance is similar to the number of data points of the model, it indicates that the model fits the data adequately.[38] We estimated the inconsistency between the direct and indirect evidence. The global inconsistency of the entire network was assessed with the design-by-treatment interaction model,[45] and the local inconsistency in the network was estimated with the node-splitting method.[46] In order to facilitate the clinical interpretation, we assessed the probability that the exercise intervention would

be likely to reach the minimum clinically important difference (MCID). We prespecified a MCID of 0.37 SD units, corresponding to 0.9 cm on a 10 cm visual analogue scale.[2] This threshold of 0.37 SD units was based on the median MCID found in recent studies of people with OA.[2 47]

All statistical analyses were conducted using WinBUGS software (V.1.4.3, MRC Biostatistics Unit, Cambridge, UK) and STATA software (V.11.0, Stata, College Station, Texas, USA).

Additional analysis

The transitivity assumption was met as we used the same definition for the common comparator, for example, usual care or glucosamine sulphate/chondroitin sulphate, for both exercise and oral NSAIDs and paracetamol. It was also assessed by comparing the distribution of trial characteristics (publication year, percentage female, mean age, baseline pain and function score) across studies grouped by comparison. To assess the robustness of the results obtained by the primary model, we performed several sensitivity analyses on the primary outcomes of pain and function to explore potential causes for heterogeneity. Four sensitivity analyses were conducted according to sample size (≥ 30/arm), low risk of allocation concealment, intervention without prescribing paracetamol and studies without outliers (effect size >5).[27] We estimated publication bias by visual assessment of funnel plot asymmetry (online supplemental appendix 3).[48] We then assessed the probability that each intervention could be ranked as the most effective treatment for pain relief or functional improvement. We obtained a hierarchy of the competing interventions using the surface under the cumulative ranking curve (SUCRA) and mean ranks. SUCRA values were expressed as percentages, the higher value representing the higher probability of being the best option (online supplemental appendix 4).[49]

RESULTS

Study selection and characteristics

The search strategy retrieved 46 635 related articles after duplicates were removed. After title or abstract screening, the full texts of 2738 potentially eligible articles were reviewed (figure 1). Of these, 152 studies (17 431 participants) met the inclusion criteria for NMA. The network of all treatment comparisons analysed for pain and function is presented in figure 2 and online supplemental appendix 5. The assumption of transitivity was met as we used the same definition of the common comparator. It was confirmed by comparing distributions of baseline characteristics where no variability was observed in the study (online supplemental appendix 6).

The characteristics of the included studies are shown in table 1 and online supplemental appendix 7. Of the 152 trials, 132 (15 005 participants) reported pain-related outcomes and 125 (12 929 participants) reported physical function outcomes. Apart from four trials comparing exercise with oral NSAIDs and paracetamol directly, there were 49 studies that had data available at or nearest to 4 weeks, two studies had data available at 8 weeks and nine studies had data available at 24 weeks. Most of the trials recruited participants with knee OA (n=127, 83.6%), 12 studies investigated hip OA (7.9%) and 13 studies (8.6%) recruited a mix of participants with knee or hip OA. In the current study, a total of 95 articles were finally extracted for the analysis, 59 articles were included for the comparative efficacy of exercise versus oral NSAIDs and paracetamol and 83 were used for the comparative efficacy of exercise versus usual care, each contributed outcomes at different time points.
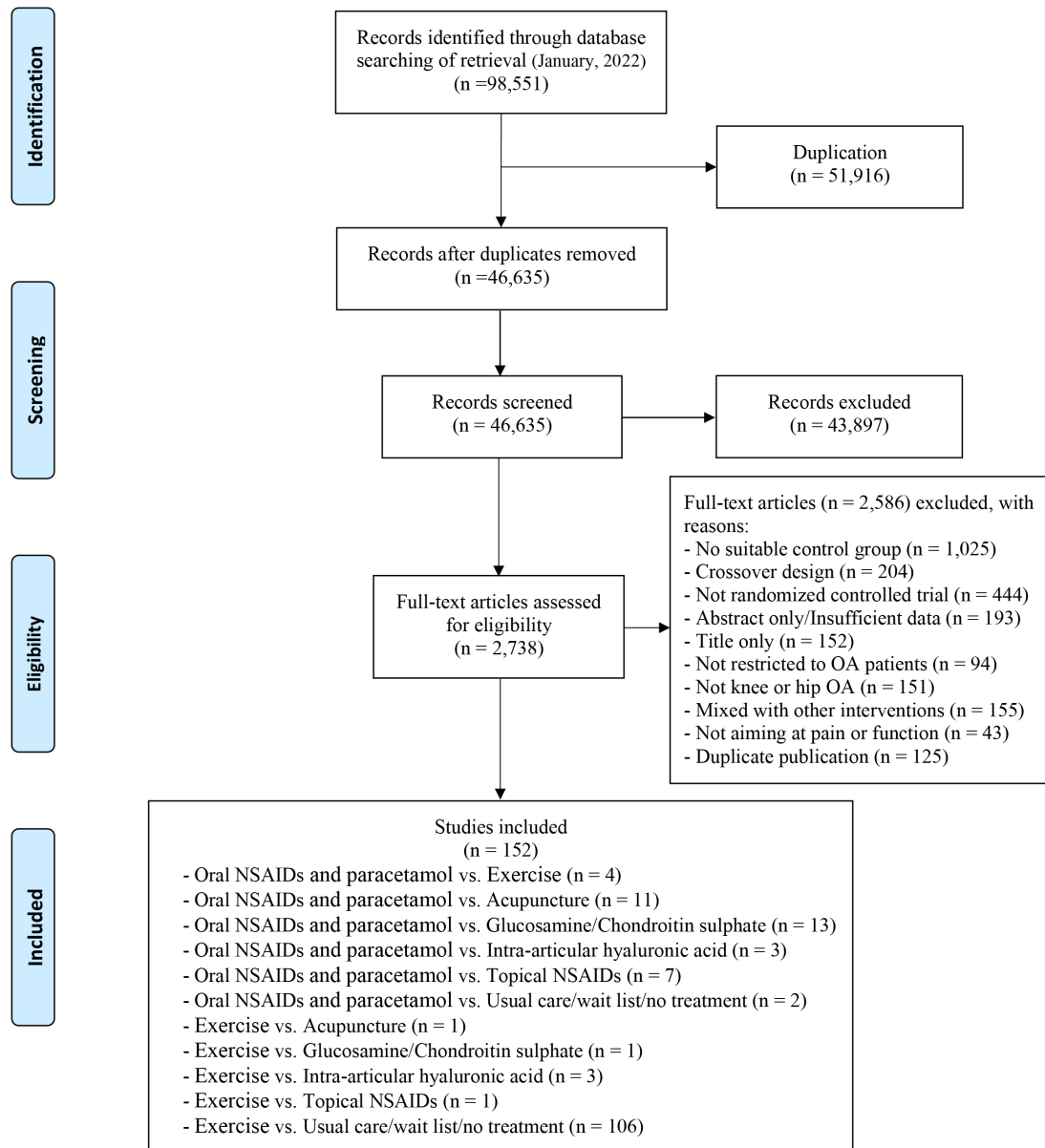
Figure 1 Summary of studies identification and selection according to the PRISMA flow diagram. NSAIDs, non-steroidal anti-inflammatory drugs; OA, osteoarthritis; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

The methodological quality was evaluated for all included trials (online supplemental appendix 8. The generation of the allocation sequence was adequate in most trials (n=105, 69.1%). Allocation concealment was adequate in almost half of the trials (n=73, 48.0%) and 95 trials (62.5%) masked outcome assessors to treatment allocation. The potential risk of bias likely to be introduced by incomplete data was high in 16 trials (10.5%). The risk of selective reporting bias was low in most trials (n=144, 94.7%) and 19 (12.5%) trials were commercially funded.

Efficacy of exercise over usual care

For pain relief, exercise was more effective than usual care at or nearest to four (SMD=1.31, 95% CrI 0.61 to 2.01), eight (SMD=0.78, 95% CrI 0.58 to 0.98) and 24 weeks (SMD=0.19, 95% CrI 0.00 to 0.38). For functional improvement, exercise was also more effective than usual care at or nearest to 4 (SMD=1.08, 95% CrI 0.29 to 1.88), 8 (SMD=0.94, 95% CrI 0.70 to 1.18) and 24 weeks (SMD=0.20, 95% CrI 0.04 to 0.37).

The effect size for function exceeded the prespecified MCID of 0.37 at or nearest to 4 weeks. Similarly, there was enough evidence to support a MCID treatment effect at or nearest to 4 (pain) and 8 weeks (pain and function), with the probability that the effect size for exercise compared with usual care was 0.37 or higher being >95%.

Comparative efficacy between exercise versus oral NSAIDs and paracetamol

As shown in table 2, there was no statistical difference between exercise and oral NSAIDs and paracetamol in pain relief at or nearest to 4 (SMD=−0.12, 95% CrI −1.74 to 1.50), 8 (SMD=0.22, 95% CrI −0.05 to 0.49) and 24 weeks (SMD=0.17, 95% CrI −0.77 to 1.12). For functional improvement, no statistical difference between exercise and oral NSAIDs and paracetamol was evident at or nearest to 4 (SMD=0.09, 95% CrI −1.69 to 1.85), 8 (SMD=0.06, 95% CrI −0.20 to
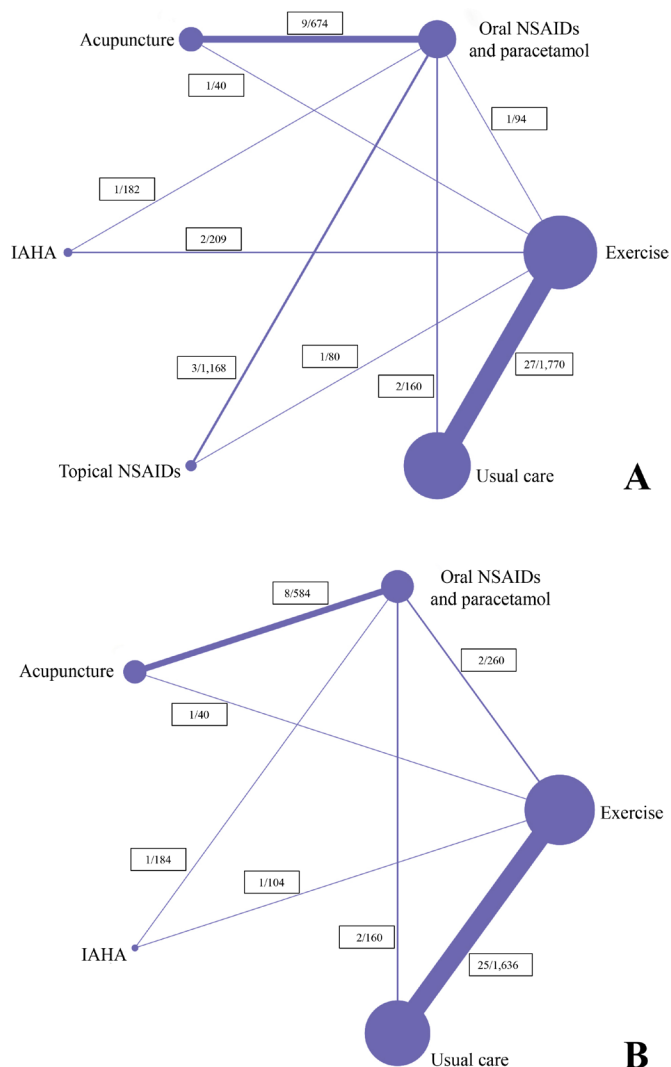
Figure 2 Structure of network formed by interventions (A) pain relief at or nearest to 4 weeks; (B) functional improvement at or nearest to 4 weeks. NSAIDs, non-steroidal anti-inflammatory drugs; IAHA, intra-articular hyaluronic acid.

Table 2 Comparative efficacy for exercise versus oral NSAIDs and paracetamol among knee or hip OA

	Number of RCTs	Number of participants	SMD* (95% CrI)
Efficacy at/nearest to 4 weeks			
Pain	47	4377	−0.12 (−1.74 to 1.50)
Function	40	2968	0.09 (−1.69 to 1.85)
Efficacy at 8 weeks			
Pain	2	210	0.22 (−0.05 to 0.49)
Function	2	214	0.06 (−0.20 to 0.33)
Efficacy at 24 weeks			
Pain	9	2141	0.17 (−0.77 to 1.12)
Function	9	2141	0.05 (−1.15 to 1.24)

*For SMD, negative value favours oral NSAIDs and paracetamol, whereas the positive value favours of exercise.
CrI, credible interval; n, number; NSAIDs, non-steroidal anti-inflammatory drugs; OA, osteoarthritis; RCTs, randomised controlled trials; SMD, standardised mean difference.

0.33) and 24 weeks (SMD=0.05, 95% CrI −1.15 to 1.24). The league table presenting all results of the NMA is available in online supplemental appendix 9.

Model fit, heterogeneity and sensitivity analysis in NMA
Evaluation of the goodness of fit demonstrated a good fit with a posterior mean residual deviance of 94.6 (94 data points) for pain at or nearest to 4 weeks, 3.8 (4 data points) for pain at 8 weeks, 18.1 (18 data points) for pain at 24 weeks, 80.1 (80 data points) for function at or nearest to 4 weeks, 3.9 (4 data points) for function at 8 weeks and 18.2 (18 data points) for function at 24 weeks. At or nearest to 4 weeks, we did not find any inconsistency between evidence derived from direct and indirect comparisons on exercise and oral NSAIDs and paracetamol for pain and function using the design-by-treatment inconsistency model (p>0.05, online supplemental appendix 10); also, the node-splitting analysis did not show any significant difference between direct and indirect evidence (p>0.05, figure 3). Test for inconsistency was not available at 8 weeks and 24 weeks due to the absence of indirect evidence at these

Table 1 Basic characteristics of included randomised controlled trials (n=152 studies)

Comparison	No. of trials	No. of pts	Target joint	Mean age (range)	Female %	Mean FU periods (weeks, range)
Exercise vs. oral NSAIDs and paracetamol	4	495	Knee	61.4 (55.8–68.9)	71.4	6.5 (4–8)
Exercise vs. usual care	106	9497	Knee or hip	65.6 (39.8–86.1)	70.7	13.1 (1–72)
Oral NSAIDs and paracetamol vs. usual care	2	160	Knee	56.6 (52.0–58.9)	75.0	2.0 (NA)
Exercise vs. acupuncture	1	40	Knee	53.0 (50.4–55.6)	NA	4.0 (NA)
Oral NSAIDs and paracetamol vs. acupuncture	11	990	Knee or hip	59.6 (44.5–65.4)	62.7	18.1 (2–108)
Exercise vs. IAHA	3	296	Knee	62.5 (55.3–71.2)	85.1	11.3 (4–24)
Oral NSAIDs and paracetamol vs. IAHA	3	317	Knee	65.4 (56.6–68.5)	66.9	24.3 (5–52)
Exercise vs. GS/CS	1	70	Knee	56.4 (55.9–56.8)	78.6	24.0 (NA)
Oral NSAIDs and paracetamol vs. GS/CS	13	2992	Knee	59.5 (54.0–70.0)	69.0	26.9 (4–96)
Exercise vs. topical NSAIDs	1	80	Knee	59.4 (58.9–59.9)	75.0	6.0 (NA)
Oral NSAIDs and paracetamol vs. topical NSAIDs	7	2494	Knee	61.9 (56.9–64.0)	62.5	7.7 (4–12)

CS, chondroitin sulphate; FU, follow-up; GS, glucosamine sulphate; IAHA, intra-articular hyaluronic acid; NA, not available; NSAIDs, non-steroidal anti-inflammatory drugs; No. of pts, number of participants included; No. of trials, number of trials included.
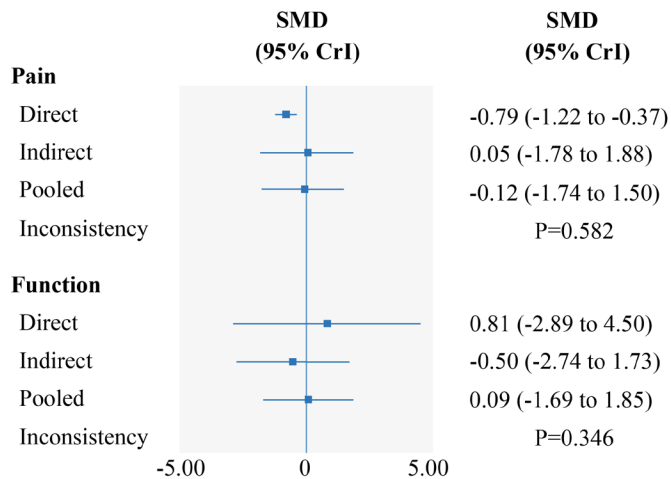
Figure 3 Forest plots depicting estimates from direct and indirect comparison for exercise to oral NSAIDs and paracetamol. CrI, credible interval; NSAIDs, non-steroidal anti-inflammatory drugs; SMD, standardised mean difference.

two time points. τ^2 estimates suggested low statistical heterogeneity for both pain ($\tau^2=0.00$) and physical function ($\tau^2=0.00$) at 8 weeks. However, there was significant heterogeneity across the trials (pain at or nearest to 4 weeks: $\tau^2=1.84$; pain at 24 weeks: 0.45; function at or nearest to 4 weeks: $\tau^2=2.04$; function at 24 weeks: 0.60). As summarised in online supplemental appendix 11, several sensitivity analyses for pain and function suggested that the results obtained were robust.

DISCUSSION
Principal findings
This NMA is based on 152 RCTs, which included 17 431 participants, to compare the efficacy between exercise and oral NSAIDs and paracetamol for knee or hip OA. The results showed that exercise was indeed a clinically effective treatment (better than usual care) in reducing pain and improving physical function in people with knee or hip OA. While these effect sizes of exercise gradually decreased over a period of time, they were not different from those obtained from oral NSAIDs and paracetamol at short (4 weeks), medium (8 weeks) or long (24 weeks) term periods of treatment.

Comparisons with previous studies
To date, there is limited NMA and one conventional meta-analysis comparing the relative effect of exercise versus analgesics (opioids and NSAIDs) on knee pain.[20 21] A recent NMA integrated two direct evidence and 91 indirect evidence to confirm that exercise provides superior analgesia compared with NSAIDs (0.54, 95% CI 0.19 to 0.89).[21] Nevertheless, in that NMA, the NSAIDs group consisted of both oral and topical interventions, the control group for analgesics contained placebo and the control group for exercise included usual care, education, ultrasound therapy and physiotherapy.[21] As a treatment intervention, placebo was treated the same as other above control groups, such as usual care, during the analysis in that NMA. These study designs may potentially affect the structure of the network (due to inconsistency and heterogeneity) and, therefore, affect the intervention outcomes. Another conventional meta-analysis included six Cochrane reviews (four pharmacology, two exercise) with 9806 participants (5627 pharmacology and 4179 exercise). The pooled effect size was 0.41

(95% CI 0.23 to 0.59) for pharmacological treatments and 0.46 (95% CI 0.34 to 0.59) for exercise. The authors concluded that the effects of exercise on knee pain were similar to the effects of analgesics.[20] However, the effect size of analgesics in both meta-analyses is not comparable with the effect size of exercise as the former was the effect over placebo, whereas the latter was the effect over usual care, education, ultrasound therapy and physiotherapy.[20 21] In addition, it is well known that placebo is more effective than usual care or no treatment,[50] hence an intervention may not be superior to placebo, for example, acupuncture versus sham acupuncture, but may still be superior to usual care.[51] The only way to examine the difference between exercise and oral NSAIDs and paracetamol is to run a head-to-head comparison, either through an RCT or an NMA of RCTs, where both interventions are placed in the same context, for example, no blinding. Our NMA fulfils this context, where exercise was compared with oral NSAIDs and paracetamol directly within a trial or indirectly through a common comparator between trials, both being randomised but not blinded.

Only a few direct comparison RCTs have assessed the efficacy of exercise versus oral NSAIDs and paracetamol for OA and the results were inconsistent.[16–19] In a recent RCT, oral NSAIDs (n=48) were reported more beneficial than exercise (n=46) in people with knee OA at or nearest to 4 weeks.[16] Additionally, two randomised trials (n=141 and 142, respectively) in people with knee OA compared 8 weeks of exercise and oral NSAIDs and paracetamol and found reduction in pain and improvement in function from baseline in each group but no differences between groups.[17 18] However, another RCT (n=166) showed that exercise was more effective than oral NSAIDs on pain in people with knee OA over a 12-week follow-up period.[19] These studies might be limited to draw sufficient evidence due to the low sample size of the great majority of the studies and the differences in methodological quality. The use of an NMA, representing the most comprehensive RCT evidence, allows for greater power and greater precision to confirm the comparative effect between exercise and oral NSAIDs and paracetamol.[52]

Limitations
Several limitations of this study should be noted. First, we were unable to fully explore the reasons for heterogeneity because many covariates for exercise effect in OA are normally not recorded in trials.[27 41 53] Second, the inconsistency between direct and indirect evidence at 8 and 24 weeks could not be examined due to the lack of head-to-head comparisons, so caution must be taken when interpreting the results at these two time points. Moreover, the numerical difference between the direct and indirect evidence was large at or nearest to 4 weeks, which may be in part due to fewer trials and a relatively small sample size included in direct evidence (one RCT with 94 participants for pain, two RCTs with 210 participants for function), and to the lack of allocation concealment.[54] However, there was no local or global inconsistency between the indirect estimates and the available direct evidence; thus, the pooled results in the current study seem to be reliable.[35 37] Third, the included RCTs and participants in this study were restricted to knee or hip OA. The conclusions may not be generalisable to OA at other joints. Fourth, there were insufficient data to perform a subgroup analysis according to the type of exercise. Therefore, we could not clarify which types, frequency or intensity of exercise are comparable to oral NSAIDs and paracetamol. Fifth, the effect sizes for exercise at or nearest to 4 weeks compared with usual care are very large. This may be because we used a more stringent

definition for 'usual care' control, and we pooled both direct and indirect evidence which increased the variations between trials and amplified the effect sizes by contextual effects.[55] Sixth, the lack of placebo and blinding in this NMA may overestimate the true effect of both exercise and oral NSAIDs and paracetamol efficacy. However, as discussed before, it did make the two treatments comparable in the same context without placebo and blinding and allowed us to estimate the relative efficacy between these two treatments. Furthermore, we included all trials in the main analysis irrespective of their risk of bias. The inclusion of studies with small sample size may lead to the small study effect, that is, a smaller study will inevitably result in a larger SMD.[56 57] However, we undertook two sensitivity analyses, one based on the allocation concealment and the other based on sample size (≥ 30 per arm, online supplemental appendix 11), and the results did not differ between these and the main analysis. As the sample size in exercise trials was in general smaller than drug trials, we were unable to perform sensitivity analysis for studies with sample size of ≥ 100/arm. The small study effect remains a caveat for this NMA. However, this may not affect the conclusion as the difference would become even smaller if the small study effect was removed. Finally, like other types of meta-analysis, it is prone to publication bias and other risk of bias and is limited to the information reported in the paper. The methodology of combining direct and indirect evidence may increase heterogeneity. Moreover, due to the large number of comparisons in the NMA, multiplicity may have increased the rate of false positives for the statistically significant results (type I error).[58]

Clinical and research implications

This study has confirmed that exercise is a medicine. Its analgesic effect is similar to that obtained from the most commonly used analgesics, oral NSAIDs and paracetamol, without serious side effects as those associated with oral NSAIDs and paracetamol.[41] The findings support the current recommendation of using exercise as a core therapy for OA. It also suggests that exercise may be used as an analgesic replacement therapy for older people with comorbidity or multimorbidity and people at higher risk of adverse events related to NSAIDs and paracetamol.[59] However, it is worth emphasising that although there is no direct evidence that exercise has significant side effects in the treatment of knee or hip OA, inadequate type and intensity of exercise might aggravate the symptoms and progression of OA.[60]

CONCLUSIONS

Exercise is effective for pain and function due to knee and hip OA. Its effect is similar to that of oral NSAIDs and paracetamol at short (4 weeks), median (8 weeks) and long-term (24 weeks) follow-up. However, this conclusion is based mainly on indirect comparisons. Further direct evidence for OA outcomes and the long-term benefits of exercise over oral NSAIDs and paracetamol for other outcomes such as comorbidities are still needed.

Author affiliations
[1]Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan, China
[2]Centre for Epidemiology and Evidence-Based Practice, University of Malaya, Kuala Lumpur, Malaysia
[3]Sports and Exercise Medicine Research and Education Group, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[4]Hunan Key Laboratory of Joint Degeneration and Injury, Changsha, Hunan, China
[5]Academic Rheumatology, School of Medicine, University of Nottingham, Nottingham, UK
[6]Pain Centre Versus Arthritis, University of Nottingham, Nottingham, UK
[7]Health Management Center, Xiangya Hospital, Central South University, Changsha, Hunan, China
[8]Division of Physiotherapy Rehabilitation Sciences Education, University of Nottingham, Nottingham, UK
[9]Versus Arthritis Centre for Sport, Exercise and Osteoarthritis Research, University of Nottingham, Nottingham, UK
[10]Department of Ultrasonography, Xiangya Hospital, Central South University, Changsha, Hunan, China
[11]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan, China
[12]Hunan Engineering Research Center of Osteoarthritis, Changsha, Hunan, China

ORCID iDs
Siew-Li Goh http://orcid.org/0000-0001-5898-1196
Chao Zeng http://orcid.org/0000-0003-4889-5731
Guanghua Lei http://orcid.org/0000-0003-2987-138X

REFERENCES

1 Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: a review. *JAMA* 2021;325:568–78.
2 da Costa BR, Reichenbach S, Keller N, *et al*. Effectiveness of non-steroidal anti-inflammatory drugs for the treatment of pain in knee and hip osteoarthritis: a network meta-analysis. *Lancet* 2017;390:e21–33.
3 Bannuru RR, Schmid CH, Kent DM, *et al*. Comparative effectiveness of pharmacologic interventions for knee osteoarthritis: a systematic review and network meta-analysis. *Ann Intern Med* 2015;162:46–54.
4 Zeng C, Doherty M, Persson MSM, *et al*. Comparative efficacy and safety of acetaminophen, topical and oral non-steroidal anti-inflammatory drugs for knee osteoarthritis: evidence from a network meta-analysis of randomized controlled trials and real-world data. *Osteoarthritis Cartilage* 2021;29:1242–51.
5 Kingsbury SR, Gross HJ, Isherwood G, *et al*. Osteoarthritis in Europe: impact on health status, work productivity and use of pharmacotherapies in five European countries. *Rheumatology* 2014;53:937–47.
6 Zeng C, Zhang W, Doherty M, *et al*. Initial analgesic prescriptions for osteoarthritis in the United Kingdom, 2000-2016. *Rheumatology* 2021;60:147–59.

7 Malfait A-M, Schnitzer TJ. Towards a mechanism-based approach to pain management in osteoarthritis. *Nat Rev Rheumatol* 2013;9:654–64.

8 Schjerning Olsen A-M, Fosbøl EL, Lindhardsen J, *et al*. Duration of treatment with nonsteroidal anti-inflammatory drugs and impact on risk of death and recurrent myocardial infarction in patients with prior myocardial infarction: a nationwide cohort study. *Circulation* 2011;123:2226–35.

9 National Clinical Guideline Centre (UK). Osteoarthritis: care and management. In: *National Institute of health and clinical excellence*, 2014. https://www.nice.org.uk/guidance/CG177

10 Kolasinski SL, Neogi T, Hochberg MC, *et al*. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. *Arthritis Rheumatol* 2020;72:220–33.

11 American Academy of Orthopaedic Surgeons. *Treatment of osteoarthritis of the knee: evidence-based guideline*. 2nd edn, 2013. http://www.aaos.org/research/guidelines/TreatmentofOsteoarthritisoftheKneeGuideline.pdf

12 Bannuru RR, Osani MC, Vaysbrot EE, *et al*. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthritis Cartilage* 2019;27:1578–89.

13 Sharma L. Osteoarthritis of the knee. *N Engl J Med* 2021;384:51–9.

14 Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. *Lancet* 2019;393:1745–59.

15 Nelligan RK, Hinman RS, Kasza J, *et al*. Effects of a Self-directed web-based strengthening exercise and physical activity program supported by automated text messages for people with knee osteoarthritis: a randomized clinical trial. *JAMA Intern Med* 2021;181:776–85.

16 Zhang J, Chen Y, Huang G. Clinical observation of moxibustion combined with knee joint rehabilitation for the treatment of knee osteoarthritis. *zhen jiu lin chuang za zhi* 2015.

17 Doi T, Akai M, Fujino K, *et al*. Effect of home exercise of quadriceps on knee osteoarthritis compared with nonsteroidal antiinflammatory drugs: a randomized controlled trial. *Am J Phys Med Rehabil* 2008;87:258–69.

18 Holsgaard-Larsen A, Christensen R, Clausen B, *et al*. One year effectiveness of neuromuscular exercise compared with instruction in analgesic use on knee function in patients with early knee osteoarthritis: the EXERPHARMA randomized trial. *Osteoarthritis Cartilage* 2018;26:28–33.

19 Chao J, Jing Z, Xuehua B. Effect of systematic exercise rehabilitation on patients with knee osteoarthritis: a randomized controlled trial. *Cartilage* 2020;1947603520903443.

20 Henriksen M, Hansen JB, Klokker L, *et al*. Comparable effects of exercise and analgesics for pain secondary to knee osteoarthritis: a meta-analysis of trials included in Cochrane systematic reviews. *J Comp Eff Res* 2016;5:417–31.

21 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.

22 Dobson F, Bennell KL, French SD, *et al*. Barriers and facilitators to exercise participation in people with hip and/or knee osteoarthritis: synthesis of the literature using behavior change theory. *Am J Phys Med Rehabil* 2016;95:372–89.

23 Miller KJ, Gonçalves-Bradley DC, Areerob P, *et al*. Comparative effectiveness of three exercise types to treat clinical depression in older adults: a systematic review and network meta-analysis of randomised controlled trials. *Ageing Res Rev* 2020;58:100999.

24 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.

25 World Health Organization. *Global recommendations on physical activity for health*. Geneva, 2010. https://www.ncbi.nlm.nih.gov/pubmed/26180873

26 Medicine ACoS. *ACSM's guidelines for exercise testing and prescription*. In: Lippincott williams & wilkins, 2013.

27 Goh S-L, 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.

28 Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. *Best Pract Res Clin Rheumatol* 2006;20:721–40.

29 Higgins J, Green SE. Cochrane Handbook for systematic reviews of interventions, 2011. Available: http://www.cochrane-handbook.org [Accessed 26 May 2017].

30 Salanti G, Del Giovane C, Chaimani A, *et al*. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014;9:e99682.

31 Puhan MA, Schünemann HJ, Murad MH, *et al*. A grade Working group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014;349:g5630.

32 Yepes-Nuñez JJ, Li S-A, Guyatt G, *et al*. Development of the summary of findings table for network meta-analysis. *J Clin Epidemiol* 2019;115:1–13.

33 Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Earlbaum, 1988.

34 Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897–900.

35 Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105–24.

36 Brémaud P. *Markov chains: Gibbs fields, Monte Carlo simulation, and Queues / P.* Brémaud: Springer Science & Business Media, 1999.

37 Dias S, Welton N, Sutton A. Nice dsu technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials, 2011. Available: http://www.nicedsu.org.uk [Accessed 15 May 2017].

38 Dias S, Sutton AJ, Ades AE, *et al*. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33:607–17.

39 Dias S, Ades AE, Welton NJ. *Network meta-analysis for decision-making*. In: John Wiley & Sons, 2018.

40 Warn DE, Thompson SG, Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. *Stat Med* 2002;21:1601–23.

41 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.

42 da Costa BR, Juni P. Systematic reviews and meta-analyses of randomized trials: principles and pitfalls. *Eur Heart J* 2014;35:3336–45.

43 Jackson D, Bowden J. Confidence intervals for the between-study variance in random-effects meta-analysis using generalised heterogeneity statistics: should we use unequal tails? *BMC Med Res Methodol* 2016;16:118.

44 Fei Y, Lam CK, Cheung BMY. Efficacy and safety of newer $P2Y_{12}$ inhibitors for acute coronary syndrome: a network meta-analysis. *Sci Rep* 2020;10:16794.

45 Higgins JPT, Jackson D, Barrett JK, *et al*. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012;3:98–110.

46 Dias S, Welton NJ, Caldwell DM, *et al*. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932–44.

47 Wandel S, Jüni P, Tendal B, *et al*. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;341:c4675.

48 Chaimani A, Higgins JPT, Mavridis D, *et al*. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8:e76654.

49 Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64:163–71.

50 Zhang W, Robertson J, Jones AC, *et al*. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2008;67:1716–23.

51 Manheimer E, Linde K, Lao L, *et al*. Meta-analysis: acupuncture for osteoarthritis of the knee. *Ann Intern Med* 2007;146:868–77.

52 Mills EJ, Lester R, Thorlund K, *et al*. Interventions to promote adherence to antiretroviral therapy in Africa: a network meta-analysis. *Lancet HIV* 2014;1:e104–11.

53 Bartholdy C, Juhl C, Christensen R, *et al*. The role of muscle strengthening in exercise therapy for knee osteoarthritis: a systematic review and meta-regression analysis of randomized trials. *Semin Arthritis Rheum* 2017;47:9–21.

54 Song F, Xiong T, Parekh-Bhurke S, *et al*. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ* 2011;343:d4909.

55 Zou K, Wong J, Abdullah N, *et al*. Examination of overall treatment effect and the proportion attributable to contextual effect in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2016;75:1964–70.

56 Nüesch E, Trelle S, Reichenbach S, *et al*. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;341:c3515.

57 Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. *PLoS One* 2018;13:e0204056.

58 Tricco AC, Thomas SM, Veroniki AA, *et al*. Comparisons of interventions for preventing falls in older adults: a systematic review and meta-analysis. *JAMA* 2017;318:1687–99.

59 Dekker J, Buurman BM, van der Leeden M. Exercise in people with comorbidity or multimorbidity. *Health Psychol* 2019;38:822–30.

60 Bennell K, Hinman RS, Wrigley TV, *et al*. Exercise and osteoarthritis: cause and effects. *Compr Physiol* 2011;1:1943–2008.