Journal of Physiotherapy 67 (2021) 252–262

Journal of
PHYSIOTHERAPY

journal homepage: www.elsevier.com/locate/jphys

Research

Some types of exercise are more effective than others in people with chronic low back pain: a network meta-analysis

Jill A Hayden [a], Jenna Ellis [a], Rachel Ogilvie [a], Samuel A Stewart [a], Matthew K Bagg [b,c], Sanja Stanojevic [a], Tiê P Yamato [d,e], Bruno T Saragiotto [d,e]

[a] Department of Community Health & Epidemiology, Dalhousie University, Halifax, Canada; [b] Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia; [c] WA SportsMed Physiotherapy, Perth, Australia; [d] Musculoskeletal Health Sydney, School of Public Health, University of Sydney, Sydney, Australia; [e] Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de São Paulo, São Paulo, Brazil

KEY WORDS

Chronic low back pain
Functional limitation
Exercise
Network meta-analysis
Physical therapy

ABSTRACT

Question: What are the effects of specific types of exercise treatments on pain intensity and functional limitation outcomes for adults with chronic low back pain? **Design**: Systematic review with network meta-analysis of randomised controlled trials. **Participants**: Adults with non-specific low back pain for ≥ 12 weeks. **Intervention**: Exercise treatments prescribed or planned by a health professional that involved conducting specific activities, postures and/or movements with a goal to improve low back pain outcomes. **Outcome measures**: Pain intensity (eg, visual analogue scale or numerical rating scale) and back-related functional limitations (eg, Roland Morris Disability Questionnaire or Oswestry Disability Index), each standardised to range from 0 to 100. **Results**: This review included 217 randomised controlled trials with 20,969 participants and 507 treatment groups. Most exercise types were more effective than minimal treatment for pain and functional limitation outcomes. Network meta-analysis results were compatible with moderate to clinically important treatment effects for Pilates, McKenzie therapy, and functional restoration (pain only) and flexibility exercises (function only) compared with minimal treatment, other effective treatments and other exercise types. The estimated mean differences for these exercise types compared with minimal treatment ranged from −15 to −19 for pain and from −10 to −12 for functional limitation. **Conclusion**: This review found evidence that Pilates, McKenzie therapy and functional restoration were more effective than other types of exercise treatment for reducing pain intensity and functional limitations. Nevertheless, people with chronic low back pain should be encouraged to perform the exercise that they enjoy to promote adherence. **Registration**: DOI:10.1002/14651858.CD009790. **[Hayden JA, Ellis J, Ogilvie R, Stewart SA, Bagg MK, Stanojevic S, Yamato TP, Saragiotto BT (2021) Some types of exercise are more effective than others in people with chronic low back pain: a network meta-analysis. *Journal of Physiotherapy* 67:252–262]**

© 2021 Australian Physiotherapy Association. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Background

Exercise is a common treatment approach for chronic low back pain recommended by clinical practice guidelines as a first line of care.[1,2] However, there is limited evidence to support the use of one type of exercise or program characteristic over another. Our recent Cochrane review including 249 randomised clinical trials found consistent, moderately strong evidence that exercise treatment was more effective than no treatment or usual care for the management of chronic low pain.[3] Exercise treatments investigated in the included trials were heterogeneous and varied in specific exercise types, program design, dose, delivery format and whether they were combined with other conservative treatments.[3]

Traditional methods for meta-analysis cannot answer important questions about which treatment works best,[4–6] and as they compare only two treatments at one time, do not allow full analysis of trials investigating multiple treatment groups within studies. Traditional meta-analysis methods synthesise important overarching questions, but generally do not include all study information and often ignore or are unable to account for important treatment heterogeneity in design and delivery characteristics.[7] Network meta-analysis methods, comparing multiple treatments simultaneously and considering other potential sources of heterogeneity, have the potential to better identify the best approach for low back pain management.

Therefore, the specific research question for this systematic review was:

> What are the effects of specific types of exercise treatments on pain intensity and functional limitation outcomes for adults with chronic low back pain?

Methods

This study is reported according to the Network Meta-Analysis extension of the Preferred Reporting Items for Systematic Reviews

https://doi.org/10.1016/j.jphys.2021.09.004
1836-9553/© 2021 Australian Physiotherapy Association. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

and Meta-Analyses (PRISMA-NMA) reporting guidelines.[8] The PRISMA-NMA checklist is presented in Appendix 1 on the eAddenda.

Identification and selection of studies

This study builds upon a recent Cochrane review that examined exercise therapy for chronic low back pain.[3] The literature search, which is described fully in the published Cochrane review,[3] included a comprehensive electronic search of nine databases (Appendix 2 on the eAddenda). Selection criteria and operationalised definitions for population, intervention, comparison, outcome characteristics and study design are described (Box 1). We followed the standard protocol for study selection and data extraction as recommended by Cochrane Back and Neck Group Methods Guidelines.[9] Further description of the study selection and data extraction processes can be found in the published Cochrane review.[3]

Assessment of characteristics of studies

Participants

Data were extracted on participant setting, age, sex, duration of low back pain episode, presence of radicular symptoms, and pain and functional limitations at baseline.

Exercise treatments

Exercise treatments were described by the type of exercise, program design, delivery approach, dose (intensity and duration) and inclusion of additional interventions (Appendix 3 on the eAddenda). We included trials that allocated participants to any of 11 categories of exercise treatments listed in Box 1. In cases where the intervention involved several types of exercise, it was judged whether: a single type dominated, in which case that exercise type was assigned as the main category for the primary analyses; or no type of exercise clearly dominated, in which case this intervention was assigned to the mixed category. These 11 categories were used as nodes in the primary analyses.

Comparison interventions

We also included trials that allocated participants to any of the exercise treatments and to at least one of nine comparison interventions shown in Box 1. In the related Cochrane review, we further categorised these comparison interventions into three groups: no treatment/usual care (including placebo and education), ineffective interventions (electrotherapy) and other conservative treatment (psychological therapy, anti-inflammatory/analgesics, relaxation, manual therapy, physiotherapy, back school), according to guideline recommendations.[10,11] For this network meta-analysis, we combined no treatment/usual care and ineffective interventions into one group referred to as 'minimal treatment' node, which was used with the 'other conservative treatment' node in the primary analyses.

Outcomes

The primary analyses in this study assessed the effect of treatments on pain intensity (eg, measured with a visual analogue scale or numeric rating scale) and back-related functional limitations (eg, measured with the Roland Morris Disability Questionnaire[12] or Oswestry Disability Index[13]). Pain and functional limitation outcomes were self-reported as continuous measures and analysed on the continuous scale. To facilitate synthesis across trials and interpretability of the results, each trial's pain and functional limitation outcome data were converted to standardised 0 to 100 (maximum) scales. Existing outcome data for all available follow-up time points were collected. The primary meta-analysis and meta-regression analyses used the most complete data available for each outcome separately. Outcomes at the available follow-up period closest to short-term (post-treatment time-period $\geq$ 4 weeks, closest to 3 months) were used for these analyses.

When trial authors did not respond to a request for missing data, missing variance scores were imputed using the mean variance from trials with similar populations of people with low back pain. Where data were reported as a median and interquartile range (IQR), the median was used to estimate the mean. If sample size information was unavailable for follow-up, the sample size was carried forward from baseline.

Box 1. Systematic review selection criteria, and additional criteria[a] for inclusion and exclusion in this network meta-analysis.

Study design
- Randomised controlled trials
- Completed and published

Population
- Adults
- Chronic non-specific low back pain ($\geq$ 12 weeks)

Interventions
- Exercise treatments prescribed or planned by a health professional
- Involved specific activities, postures and/or movements with a goal to improve low back pain outcomes
- Categorised as:
  - Core strengthening/motor control
  - Mixed exercise types
  - General strengthening
  - Aerobic exercises
  - Pilates
  - Stretching
  - Yoga
  - Functional restoration
  - McKenzie therapy
  - Flexibility
  - Other specific exercises

Comparators
- Placebo, no treatment or usual care
- Education
- Manual therapy
- Back school
- Electrotherapy
- Mixed physiotherapy (not involving exercise)
- Psychological therapy
- Anti-inflammatory/analgesics
- Relaxation

Outcomes
- Pain intensity (eg, visual analogue scale, numerical rating scale)
- Back-related functional limitations (eg, Roland Morris Disability Questionnaire, Oswestry Disability Index)

Integrity
- Trials were excluded from the review if they were judged to be either plagiarised or published in a presumed predatory journal in addition to at least one other research integrity concern (eg, high risk of bias, inadequate reporting)

Data availability[a]
- Trials were excluded from the analysis if they did not have any outcome follow-up of $\geq$ 4 weeks
- Trials were excluded from the analysis if they did not have data for analysis in either pain or functional limitations (allowing for standard deviation imputation and carry-through of sample size from baseline)

Risk of bias

The risk of bias assessment for trials was conducted using the criteria in the Cochrane Collaboration Risk of Bias Tool (version 1).[14,15] We assessed potential bias related to: randomisation; treatment allocation concealment; blinding of participants, care providers and outcome assessors; drop-out rate; intention to treat; seletive outcome reporting; similarity at baseline; avoidance of co-interventions; compliance; and similar timing of outcome assessment. Individual criteria were scored as 'high risk', 'low risk' or 'unclear risk'. A trial with a low risk of bias was defined as fulfilling six or more of the 12 criteria items, and with no other fatal flaws. For additional detail regarding risk of bias assessment see the published Cochrane review.[3]

Data analysis

The primary analyses included descriptive analyses, pairwise meta-analyses and network meta-analyses for the 11 specified exercise types, and two comparison categories: minimal treatment and other conservative treatment. We used two software packages[a,b] for data preparation and analyses. Statistical code is provided in Appendix 4 on the eAddenda.

Descriptive analyses

The individual and summary characteristics of all included trials were described using the appropriate descriptive statistics, and reasons for exclusion of trials that were included in the related Cochrane review were documented.

Effectiveness of exercise types

Meta-analyses were conducted for each pairwise comparison of treatments for which there were data available. DerSimonian and Laird random-effects models were used to estimate the pooled effects of intervention (MD and 95% CI) and measures of heterogeneity (τ^2, I^2) for each comparison. The sample size of the repeated group was evenly split across instances where a comparison group contributed to multiple observations within the same meta-analysis (ie, in comparisons involving trials with more than two groups). Egger's test was used for asymmetry when $\geq$ 10 intervention groups contributed to a comparison.

Network meta-analysis

We conducted network meta-analyses to estimate the effects of the interventions on pain and functional limitation outcomes separately. A frequentist inconsistency model[16] was fitted using contrast-based linear mixed-effects modelling. The models included: fixed-effect parameters for the effects of intervention; baseline outcome value and their interaction; and random-effects terms to account for correlation between observed effects in trials with more than two groups. We specified random-effects terms using a compound symmetric covariance structure with rho = 0.5.[17]

Assessment of network transitivity

We assumed that all participants in the included trials were equally likely to be randomised to any of the interventions in the observed trials (ie, we assumed that the transitivity assumption was plausible). Nonetheless, we considered participant setting, duration of the low back pain episode, radicular symptoms and baseline outcome values as modifiers of treatment effects. We also considered exercise treatment intervention dose/intensity, delivery format and presence of additional interventions to be effect modifiers. Accordingly, the distribution of the population variables across network comparisons was assessed (Appendix 5 on the eAddenda) and each treatment effect modifier covariate was modelled (Appendix 6 on the eAddenda). This assessment of network transitivity and considering theoretical mechanisms of modification led us to adjust for baseline outcome values in the primary analyses, and explore the impact of exercise dose and additional co-interventions in sensitivity analyses.

Assessment of incoherence

We evaluated incoherence (ie, agreement between direct and indirect evidence) of the pain and function outcome networks globally and evaluated local incoherence for each treatment comparison using side-splitting[18] and by evaluating statistical incoherence of the network separately in every closed loop.[19] In the loop-specific approach, loops formed only by multi-arm trials were excluded and correlation induced by multi-arm trials was mitigated by dropping the direct comparison with the largest number of trials from multi-arm trials when it appeared in a particular loop. Local incoherence was considered to be statistically significant if loop-specific 95% confidence intervals did not include zero.

Sensitivity analyses

Four types of sensitivity analyses were conducted to explore the impact of methodological decisions and to further explore residual heterogeneity and incoherence in the primary analyses. First, exploratory network meta-analyses were conducted with the nine specific comparison treatments uncategorised. Next, included trials were restricted to more homogeneous measures of pain intensity (including trials using the visual analogue scale or numerical rating scale only), and functional limitations (including trials using the Roland Morris Disability Questionnaire[12] or Oswestry Disability Index[13] only). Third, we excluded trials judged as having high risk of bias. Fourth, we excluded trials judged to have an improbable or outlying mean outcome if the absolute difference between any exercise group and any comparison group over all available follow-ups was greater than a predetermined threshold of 30/100 for pain and 20/100 for functional limitations, selected based on clinical judgement.

For all sensitivity and exploratory analyses, we assessed changes in the interpretation of pairwise treatment effects for the analysis compared to the primary network meta-analysis and changes in overall model heterogeneity and incoherence observed.

Certainty of the evidence

The CINeMA web application (which adapts GRADE domains to network meta-analysis) was used to evaluate confidence in findings from the primary network meta-analyses due to: risk of bias within comparisons, publication bias, indirectness, imprecision, heterogeneity and incoherence.[20] A detailed description of the assessment process is provided in Appendix 7 on the eAddenda.

A clinically important difference in outcome between low back pain treatments was interpreted as a difference in pain of 15 points out of 100, and difference in function of 10 points. These were calculated as the smallest worthwhile effects based on a 30% reduction in outcome,[21] from the average baseline pain (50.9, 95% CI 49.1 to 52.8), and average baseline functional limitations (38.9, 95% CI 35.8 to 42.0) for included trials. Differences were considered statistically significant at the 5% level. We defined five categories of results (categories 1 to 4 favouring one treatment): 1. Clinically important difference; 2. Moderate difference compatible with a clinically important difference; 3. Moderate difference; 4. Small difference; and 5. No difference. We assessed changes in interpretation of results for sensitivity analyses with an algorithm to identify changes in the interpretation of the effect direction, size and compatibility with a clinically important difference (Appendix 8 on the eAddenda). We defined important changes in interpretation of the results for sensitivity analyses from primary analysis results based on the number of changes in the results category: no or one change = no concerns; two or three changes = some concerns; and four or more changes = major concerns.

Results

This review provides an up-to-date assessment of the effectiveness of exercise treatment of chronic non-specific low back pain. It included 249 randomised trials (24,486 participants), with 217 of these trials (87%) providing sufficient data for meta-analysis (20,969 participants at baseline) (Figure 1). Citations for the included studies are presented in Appendix 9 on the eAddenda. The reasons that trials were excluded from this study are presented in Appendix 10 on the eAddenda. There was a total of 507 treatment groups in the included trials. In total, 126 trials compared exercise to non-exercise comparisons, and 91 compared only groups receiving different types of exercise interventions (Figure 2).

Characteristics of included trials

Table 1 describes summary characteristics of included trial populations and Appendix 11 (on the eAddenda) provides detailed information about each trial included in this analysis. The included trials were mostly conducted in healthcare settings (56%, 122 trials); 65 trials (30%) were from general populations or mixed settings and

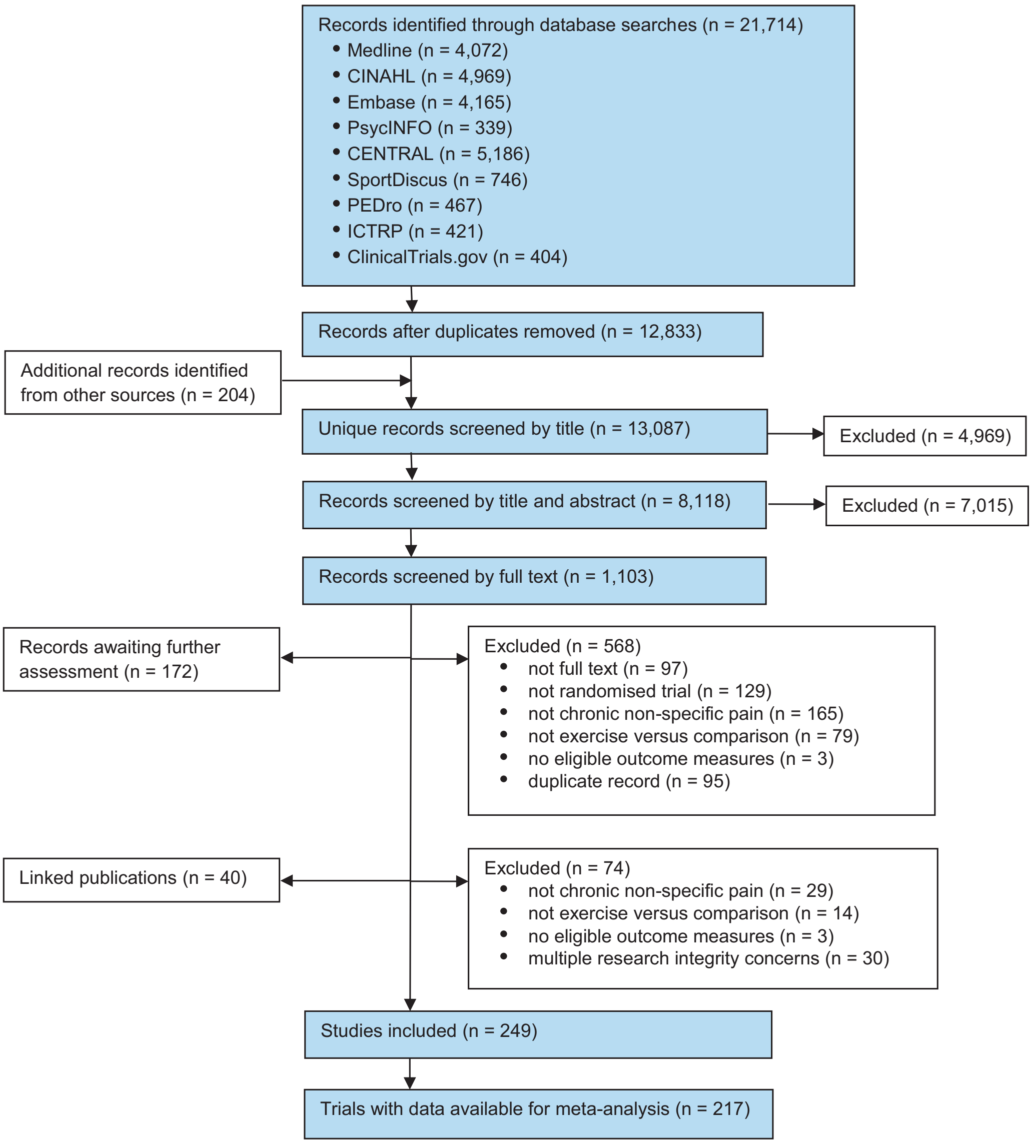

Figure 1. Flow of trials through the review.

21 trials (10%) were from occupational settings. The number of participants in included trials ranged from 10 to 722, with a median sample size of 70 participants (IQR 43 to 120). The trial participants reported a mean pain severity at baseline of 44/100 (95% CI 43 to 45), with 67 included trials (31%) describing participants with chronic low back pain of moderate symptom duration (mean duration 12 weeks to 3 years) and the same number including participants with longer duration of chronic low back pain of > 3 years (83 trials did not specify the duration of chronic low back pain). The median time point of the short-term outcome follow-up reported in this project was 12 weeks (IQR 7 to 12).

Of the 217 included trials, 108 (50%) were judged to be at risk of selection bias, 172 (79%) were judged to be at risk of performance bias, 158 (73%) were judged to be at risk of detection bias, 38 (18%) were judged to be at risk of attrition bias, 12 (6%) were judged to be at risk of reporting bias and 150 (69%) were assessed as susceptible to other potential sources of bias. The full risk of bias assessments for each included trial are provided in Appendix 12 on the eAddenda.

Characteristics of treatments

Included trials had between one and four exercise treatment groups, totalling 369 exercise groups across the 217 trials. Table 2 describes the summary characteristics of the exercise treatment groups. There were between five and 110 exercise groups available for each of the predefined 11 types of exercise categories. The most common type of exercise investigated was core strengthening (30%), followed by treatment groups comprising mixed types of exercises (ie, three or more types) (26%) and general strengthening exercises (12%). We classified 45% of exercise treatments as 'back specific' and 29% as 'whole body' exercises. The exercise program design was classified as individualised for 19%, partially individualised for 36% and standardised for 45%. The exercise programs that were investigated were mostly delivered in a supervised group setting (40%) or with individual healthcare provider supervision (39%). The programs had a median treatment time of 12 hours (IQR 8 to 20) delivered over a median period of 8 weeks (IQR 6 to 12). Most of the exercise treatment groups included other additional interventions (57%), the most common being advice/education (31%).

The 138 comparison groups included in the trials were categorised as minimal treatment (62%) and effective treatment/unclear effectiveness (38%). The minimal treatment category included comparison groups that provided placebo, no treatment or usual care (62%), education only (33%) or electrotherapy (6%). The effective or unclear effectiveness category of comparisons included treatments such as manual therapy (27%), mixed non-exercise physical therapy (48%) and back school (12%). A detailed description of comparison groups is provided in Appendix 11.

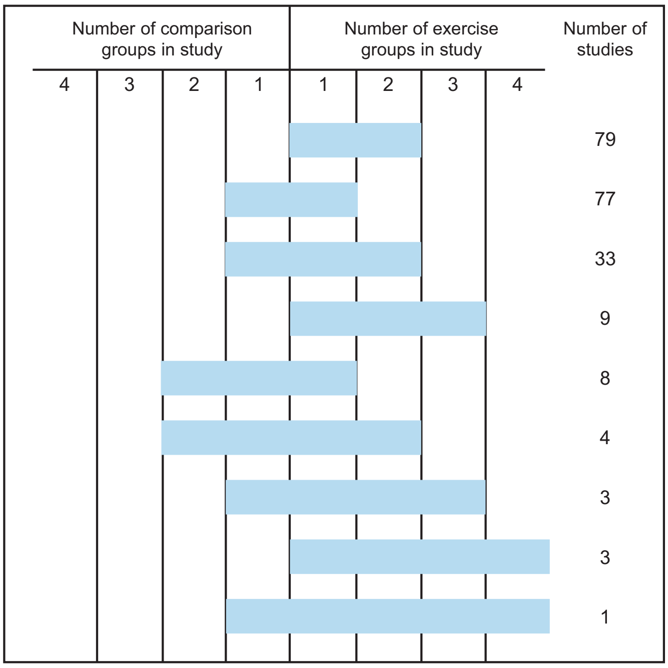

Figure 2. Breakdown of exercise treatment and comparison treatment group numbers for trials available for the meta-analyses presented in this study (217 included trials).

Table 1
Characteristics of included trial populations.

Characteristic	All trials (n = 217)
Population source, n (%)	
healthcare	122 (56)
occupational	21 (10)
general or mixed	65 (30)
other	1 (< 1)
not specified	8 (4)
Total participants at baseline *(n)*	
pooled	20,969
per study, median (IQR; range)	70 (43 to 120; 10 to 722)
Age of participants *(y)*, mean (95% CI)	44 (43 to 45)
Male participants *(%)*, mean (95% CI)	44 (41 to 47)
Category of low back pain, n (%)	
chronic	189 (87)
mixed chronic	27 (12)
not specified	1 (< 1)
Recurrent pain, n (%)	
yes	17 (8)
no	200 (92)
Pain duration, n (%)	
moderate	67 (31)
long	67 (31)
not specified	83 (38)
Baseline pain severity *(0 to 100)*, mean (95% CI)	51 (49 to 53)
Leg pain or neurological symptoms, n (%)	
all	0 (0)
some	90 (42)
none	73 (34)
not specified	54 (25)
Treatment groups *(n)*	
pooled	507
per study, median (IQR; range)[a]	2 (2 to 3; 2 to 5)
exercise groups, n (%)[a]	369 (73)
comparison groups, n (%)[a]	138 (27)
Outcomes assessed, n (%)[b]	
pain intensity	202 (93)
functional limitations	199 (92)
work status	30 (14)
health-related quality of life	67 (31)
adverse outcomes	66 (30)
global perceived recovery	40 (18)
Follow-up periods, n (%)	
immediate (< 6 weeks)	34 (16)
short term (6 to 12 weeks)	168 (77)
moderate (13 to 47 weeks)	116 (54)
long term (≥ 48 weeks)	63 (29)

[a] For primary network meta-analyses, exercise groups of the same exercise type and comparison groups of the same type were combined, giving 422 treatment groups (290 exercise groups, 132 comparison groups).
[b] Trial reported outcome measurement.

Comparative effectiveness of exercise treatment types

The overall certainty of the evidence for pain intensity outcomes was judged to be low for 64% of pairwise comparisons (49 of 77) and moderate for 36% of comparisons (28 of 77). For functional limitation outcomes, the overall certainty of the evidence was low for 16% of pairwise comparisons (12 of 77), moderate for 82% of comparisons (63 of 77) and high for 3% of comparisons (2 of 77).

Pain intensity outcomes

In total, 198 trials measured pain intensity outcomes (including 466 treatment groups, 17,534 participants), with 166 providing data for meta-analyses (399 groups, 15,553 participants). This meant that direct evidence was available for 52 of the possible 77 treatment comparisons in the network. The number of trials available for each comparison ranged from one (24 participants) to 18 (1,739 participants). Two exercise types – core strengthening and mixed exercise type – and the minimal treatment comparison type had direct comparison pairings available with all other treatment types. Egger's test suggested possible publication bias in 29% of meta-analyses with ≥ 10 trials available (two of seven meta-analyses).

Direct pairwise meta-analyses: Pairwise meta-analyses for all comparisons are provided with detailed individual trial-level information, summarised in Appendix 5 and presented in forest plots in Appendix 13 on the eAddenda. The data available for pain outcomes directly comparing each of the 11 exercise types with minimal treatment ranged from one trial (110 participants) for flexibility exercises to 17 trials (1,614 participants) for mixed exercises. The mean difference in pain intensity at short-term follow-up from direct (head-to-head) evidence favoured each exercise type compared with minimal treatment for all types other than flexibility exercises (MD 5.0, 95% CI −8.9 to 18.9; one trial favouring minimal treatment). The largest mean differences in pain intensity from direct evidence (for comparisons with more than one trial providing data) were observed for Pilates (MD −21.8, 95% CI −29.6 to −14.1, 11 trials, 800 participants, I^2 = 91.1), McKenzie therapy (MD −14.1, 95% CI −27.7 to −0.4, two trials, 170 participants, I^2 = 71.2), stretching (MD −14.0, 95% CI −21.1 to −6.8, six trials, 354 participants, I^2 = 54.6), general strengthening (MD −13.4, 95% CI −20.6 to −6.2, nine trials, 433 participants, I^2 = 79.8) and core strengthening exercises (MD −12.8, 95% CI −17.8 to −7.9, 17 trials, 1,545 participants, I^2 = 82.3).

Direct pairwise evidence was available for nine of the 11 exercise types compared with other effective treatments, with evidence ranging from one trial (75 participants) for functional restoration to 18 trials (1,429 participants) for core strengthening exercises. Three or fewer trials provided direct evidence for all except core strengthening, mixed exercise type, general strengthening and aerobic exercises compared with other effective treatments. The mean difference in pain intensity at short-term follow-up compared with other effective treatments from direct evidence favoured Pilates (MD −18.3, 95% CI −23.4 to −13.1, two trials, 161 participants), McKenzie therapy (MD −17.6, 95% CI −33.6 to −1.5, two trials, 170 participants, I^2 = 86.6) and functional restoration (MD −29.0, 95% CI −41.0 to −17.0, one trial, 75 participants). Moderate to no differences in treatment effect were observed for other exercise types (Appendix 5 and Appendix 13).

Network meta-analysis (primary analyses): The network was well connected, with 66% of comparisons having direct evidence (Figure 3). The network meta-analysis results comparing exercise types and minimal treatment are presented in Figure 4. The mean difference in pain intensity at short-term follow-up favoured each exercise type compared with minimal treatment, with mean treatment effects ranging from −6.8 to −18.7 (decreased pain intensity with exercise treatment). Outcomes for the following exercise types were compatible with a clinically important treatment effect

Table 2
Description of exercise group characteristics in included trials (217 trials, 368 exercise groups).

Characteristic	Exercise groups (n = 368)[a]
Types of exercise, n (%)	
core strengthening	110 (30)
mixed exercise types	96 (26)
general strengthening	44 (12)
aerobic	25 (7)
Pilates	24 (7)
stretching	17 (5)
other specific exercises	15 (4)
yoga	13 (4)
functional restoration	10 (3)
McKenzie therapy	9 (3)
flexibility	5 (1)
Specificity of exercise, n (%)	
whole body	108 (29)
back specific	167 (45)
both	64 (17)
not specified	29 (8)
Exercise program design, n (%)	
individualised	69 (19)
partially individualised	131 (36)
standardised	166 (45)
not specified	2 (< 1)
Primary delivery format, n (%)	
independent exercise	19 (5)
independent exercise with follow-up	25 (7)
group supervision	147 (40)
individual supervision	143 (39)
not specified	34 (9)
Duration of intervention *(hr)*, median (IQR; range) (284 of 368)	12.0 (8.0 to 20.0; 0.3 to 156.0)
Duration of intervention *(wk)*, median (IQR; range) (361 of 368)	8 (6 to 12; 1 to 52)
Dose of intervention, n (%)	
high dose (≥ 20 hours)	120 (33)
low dose (< 20 hours)	246 (67)
not specified	2 (< 1)
Other additional interventions, n (%)	
yes	208 (57)
no	136 (37)
not specified	24 (7)

[a] For the primary network meta-analyses, exercise groups of the same exercise were combined giving 290 exercise groups.

compared with minimal treatment (ie, 95% CI suggesting > 15-point improvement in pain outcomes): Pilates, McKenzie therapy, functional restoration, core strengthening and other specific exercise types. The network meta-analysis mean difference estimate for Pilates (MD −18.7, 95% CI −24.4 to −13.1) was most likely to be compatible with clinically important improvement compared with minimal treatment. The certainty of this evidence was judged to be moderate overall.

The mean difference in pain intensity at short-term follow-up compared with other effective treatments from network meta-analysis was compatible with a clinically important difference for Pilates (MD −11.2, 95% CI −17.2 to −5.3) and moderate treatment effect for McKenzie therapy (MD −7.3, 95% CI −14.1 to −0.5) and functional restoration (MD −7.2, 95% CI −13.9 to −0.4). Small to no treatment effects were observed for other exercise types compared with other effective treatments. The certainty of this evidence was judged to be low to moderate overall.

Comparing the various exercise types with each other, Pilates, core strengthening, McKenzie therapy and functional restoration exercises had larger improvement in pain intensity compared with several other exercise types. Pilates was more effective than all other exercise types, compatible with a clinically important difference in pain intensity, compared with stretching exercises (MD −11.9, 95% CI −18.2 to −5.6), aerobic exercises (MD −11.8, 95% CI −17.6 to −6.0), flexibility exercises (MD −11.3, 95% CI −22.5 to 0.0), yoga (MD −10.7, 95% CI −17.4 to −3.9) and mixed exercises (MD −9.9, 95% CI −15.2 to −4.7), with low to moderate certainty evidence overall. Core strengthening exercises were moderately more effective than stretching exercises (MD −6.6, 95% CI −11.8 to −1.3), aerobic exercises (MD −6.5, 95% CI −10.6 to −2.5) and flexibility exercises (MD −5.9, 95% CI −16.2 to 4.3), with low certainty evidence overall. McKenzie therapy was moderately more effective than stretching (MD −8.0, 95% CI −15.8 to −0.2), aerobic exercises (MD −7.9, 95% CI −15.0 to −0.9) and flexibility exercises (MD −7.4, 95% CI −19.1 to 4.4), with moderate certainty evidence overall. Functional restoration exercises were moderately more effective than stretching (MD −7.8, 95% CI −15.5 to −0.2), aerobic exercises (MD −7.8, 95% CI −14.2 to −1.3) and flexibility exercises (MD −7.4, 95% CI −19.1 to 4.4), with low to moderate certainty evidence. There was a small to moderate difference in pain outcomes comparing Pilates to McKenzie therapy, core strengthening exercises and functional restoration, with Pilates appearing more effective (Figure 5).

Network meta-analysis (exploring all comparisons): The network meta-analysis comparing the effectiveness of all 11 exercise types and all nine specific comparison treatment types included 162 pairwise comparisons with direct and indirect data from 166 trials (399 groups, 15,553 participants) for pain intensity outcomes (Appendix 14 on the eAddenda). There was direct evidence for 81 of the possible treatment comparisons (50%). Eleven of 41 direct comparisons with two or more trials (27%) had I^2 values > 75%. Sparseness of the network led to wide confidence intervals for some treatment comparisons; however, we observed moderate differences compatible with a clinically important difference in pain outcomes for Pilates compared with all other conservative treatments (most likely MD range −9.3 to −17.5).

Functional limitation outcomes

The network for this study included data about functional limitation outcomes from 187 trials (433 treatment groups, 16,926 participants). Direct and indirect pairwise data from 149 trials (355 treatment groups, 14,220 participants) about functional limitation outcomes were available. This represented direct evidence for 49 of the possible 77 treatment comparisons. The number of trials available for pairwise comparisons ranged from one (24 participants) to 16 (1,749 participants). Two exercise types – core strengthening and mixed exercise type – and minimal treatment comparison type had direct comparison pairings available with all other treatment types. Direct meta-analyses that included two or more trials (57%, 28 of 49 direct meta-analyses) had statistical heterogeneity ranging from 0 to 93.1% measured with the I^2 statistic. Eleven pairwise meta-analyses had I^2 values < 50% and nine had values > 75%. Egger's test suggested possible publication bias in 13% of meta-analyses with ≥ 10 trials available (one of eight meta-analyses).

Direct pairwise meta-analyses: Pairwise meta-analyses for each exercise type compared with minimal treatment are presented in Figure 4. For functional limitation outcomes, data available for each of the 11 exercise types compared with minimal treatment ranged from one trial (90 participants) for functional restoration exercises to 16 trials (1,749 participants) for mixed exercise type. The mean difference in functional limitations at short-term follow-up from direct evidence favoured each exercise type compared with minimal treatment. The largest reduction in functional limitations from direct evidence was observed for Pilates (MD −13.1, 95% CI −18.6 to −7.7, 10 trials, 780 participants, I^2 = 88.3).

Direct pairwise evidence was available for nine of the 11 exercise types compared with other effective treatments, with evidence ranging from one trial (28 participants) for 'other' specific exercises to 14 trials (1,716 participants) for mixed exercise type. Four or fewer trials provided direct evidence for all except core strengthening and mixed exercise type compared with other effective treatments. The mean difference in functional limitations at short-term follow-up compared with other effective treatments from direct evidence favoured functional restoration (MD −25.3, 95% CI −38.1 to −12.5, one trial, 75 participants), McKenzie therapy (MD −16.1, 95% CI −19.5 to −12.8, one trial, 45 participants) and Pilates (MD −8.3, 95% CI −20.9 to 4.4, two trials, 161 participants, I^2 = 72.9). Small to no differences in treatment effect were observed for other exercise types. Pairwise meta-analyses for all comparisons are summarised in

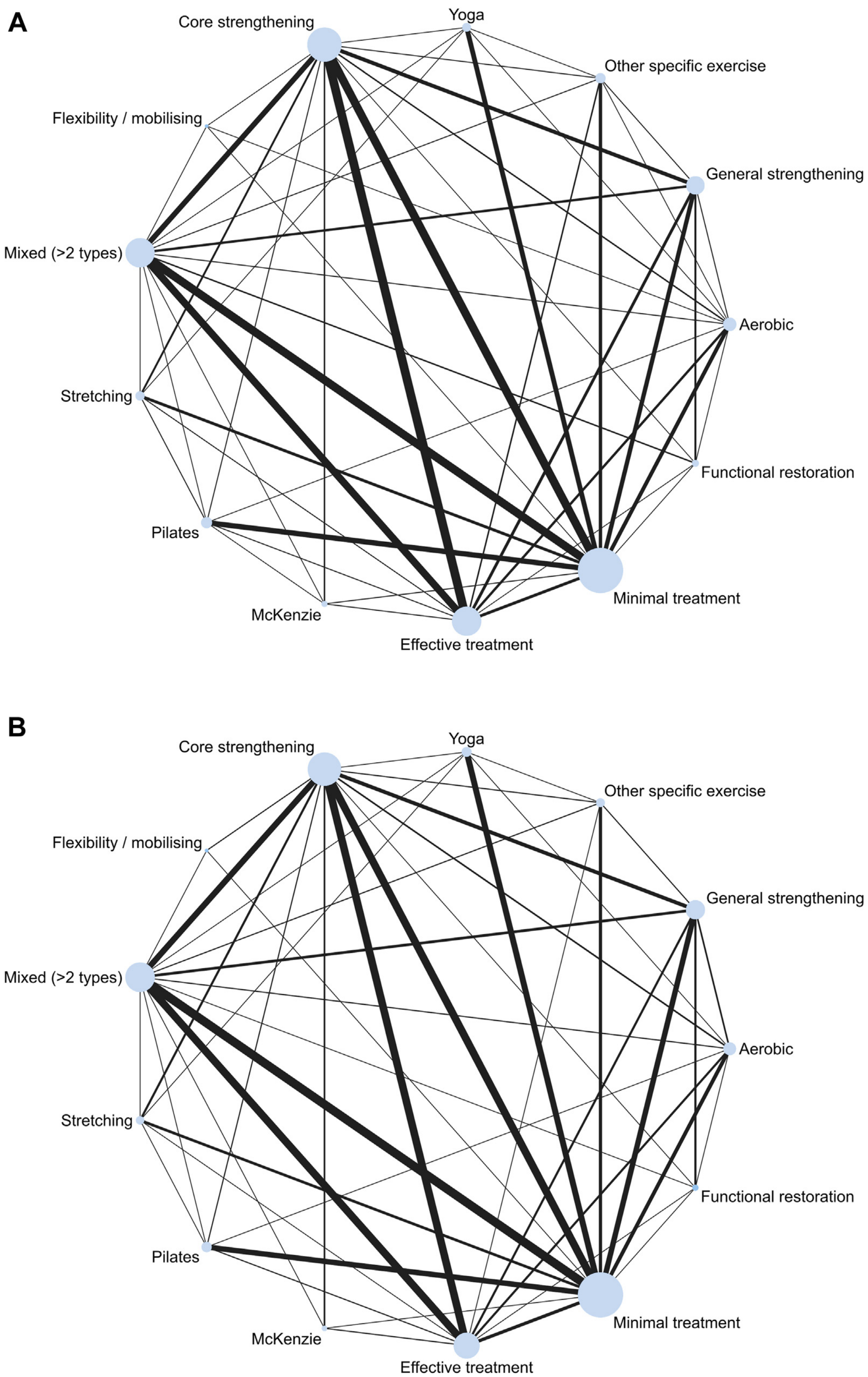

Figure 3. Network plot presenting the trial data contributing evidence comparing exercise treatment types for short-term outcomes: **A.** pain intensity: 166 trials, 359 groups, 15,553 participants. **B.** functional limitations: 149 trials, 322 groups, 14,220 participants. The size of the nodes represents how many times the exercise appears in any comparison about that treatment and the width of the edges represents the total sample size in the comparisons it connects.

Appendix 5 and provided with detailed individual trial-level information in Appendix 13.

Network meta-analysis (primary analysis): The network plot for functional limitation outcomes primary analysis was well connected, with 62% of comparisons having direct evidence and no disconnected nodes (Figure 3). The network meta-analysis results comparing exercise types and minimal treatment found mean difference in functional limitations at short-term follow-up favoured each exercise type compared with minimal treatment, with mean treatment effects ranging from −3.4 to −11.7 (decreased functional limitations with exercise treatment), with treatment effects for four of 11 exercise types most compatible with a clinically important effect (ie, 95% CI

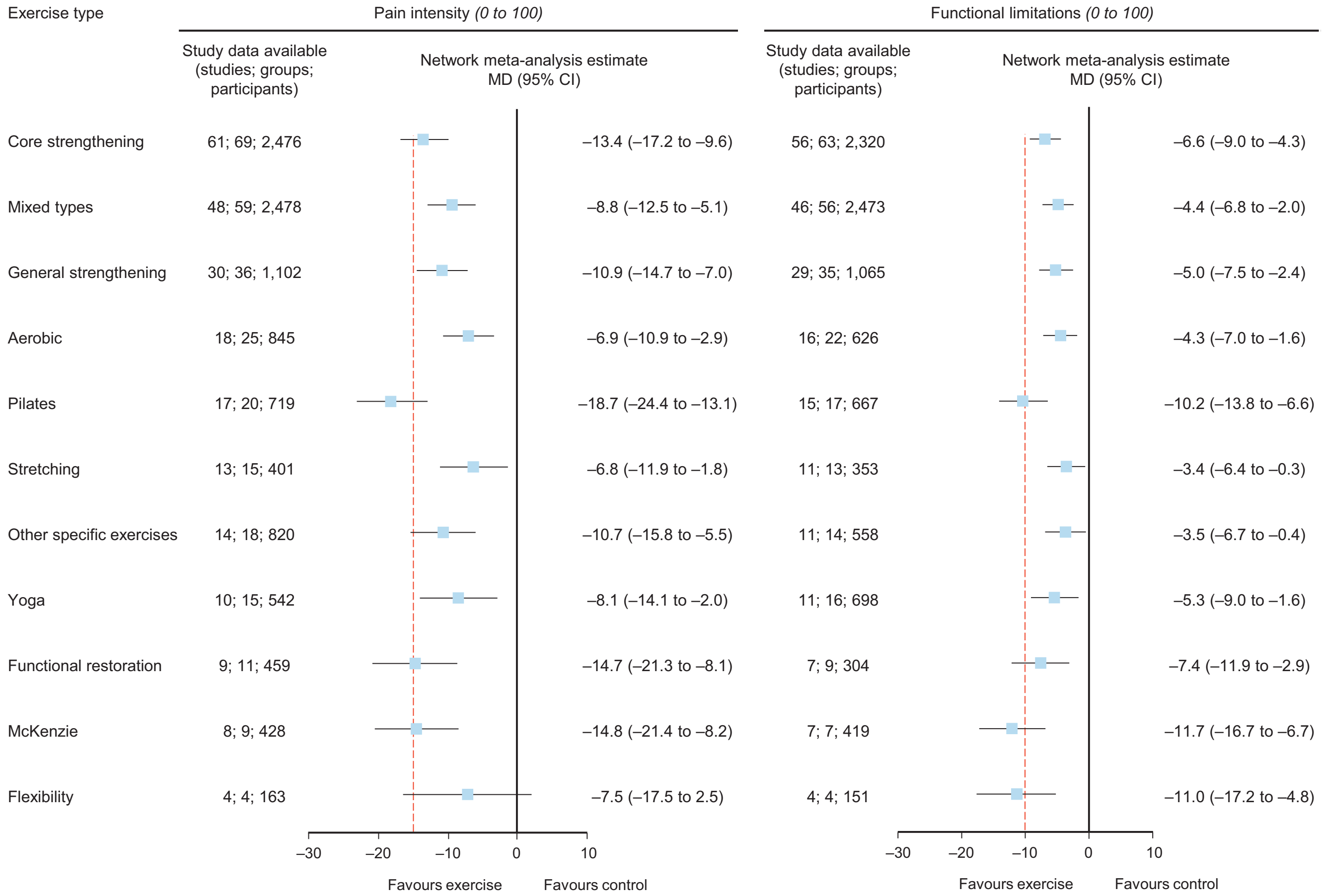

Figure 4. Summary network meta-analysis results for each exercise type compared with minimal treatment comparisons for short-term outcomes: pain intensity, functional limitations. The hashed line indicates clinically important difference.

suggesting > 10-point improvement in functional limitation outcomes): McKenzie therapy (MD −11.7, 95% CI −16.7 to −6.7), flexibility exercises (MD −11.0, 95% CI −17.2 to −4.8), Pilates (MD −10.2, 95% CI −13.8 to −6.6) and functional restoration exercise (MD −7.4, 95% CI −11.9 to −2.9) (Figure 4). The certainty of this evidence was judged to be moderate to high overall.

The mean difference in functional limitations at short-term follow-up compared with other effective treatments from network meta-analysis was compatible with a clinically important difference for McKenzie therapy (MD −7.1, 95% CI −12.1 to −2.1) and functional restoration (MD −6.4, 95% CI −12.7 to −0.1), and moderate treatment effect for Pilates (MD −5.5, 95% CI −9.4 to −1.7). The certainty of this evidence was judged to be moderate. Small to no treatment effects were observed for other exercise types compared with other effective treatments.

Comparing the various exercise types with each other, McKenzie therapy, Pilates and flexibility exercises had larger improvement in functional limitations, compatible with a clinically important difference, compared with other exercise types investigated. There were no observed differences in functional limitation outcomes between these three exercise types at short-term follow-up (Figure 5).

Network meta-analysis (exploring all comparisons): Network meta-analysis comparing the effectiveness of the 11 exercise types and nine specific comparison treatment types included 162 pairwise comparisons of direct and indirect data (Appendix 14). Functional limitation outcome data from 149 trials (355 treatment groups, 14,220 participants) provided evidence for 74 of the possible treatment comparisons (46%). Ten of 37 direct comparisons with two or more trials (27%) had I^2 values > 75%. Sparseness of the network led to imprecise estimates for some treatment comparisons; however, moderate differences compatible with clinically important differences were observed in functional limitation outcomes for Pilates, McKenzie therapy and flexibility exercises compared with most other conservative treatments (most likely mean difference range −5.3 to −12.4).

Sensitivity analyses

We summarised sensitivity analysis results for exercise types compared with minimal treatment (Appendix 15 on the eAddenda; full output for sensitivity analyses is available on request). There were minimal changes in the interpretation of pain and functional limitation outcomes for exercise types compared with minimal treatment for each of the sensitivity analyses (all 66 sensitivity analysis results judged to have no concerns). Overall, for network meta-analyses for pairwise sets of two treatments there were small changes to the size of the network meta-analytic effects (higher or lower); however, interpretation of results was unchanged for 98% of pain sensitivity analyses (226 of 231 pairwise comparison network meta-analysis results) and 99% of functional limitation sensitivity analyses (228 of 231 pairwise comparison network meta-analysis results).

Exploring potential effect modifying characteristics (Appendix 6), there were no clear differences in the interactions of symptom duration and inclusion of participants with leg pain or neurological symptoms (covariates) with intervention effects across comparisons. A high dose of most exercise treatments appeared to reduce pain and functional limitation outcomes more than low dose; more so for Pilates compared with minimal treatment than for other types, although the observed confidence intervals overlapped. Co-interventions appeared to improve effectiveness of most exercise types for pain and function but not for stretching exercises. Network meta-analysis models including adjustment for dose and additional co-interventions as treatment effect modifying covariates were con-

A

Core strengthening											
−4.6 (−7.7 to −1.5)	Mixed (> 2 types)										
−2.5 (−6.2 to 1.1)	2.1 (−1.7 to 5.8)	General strengthening									
−6.5 (−10.6 to −2.5)	−1.9 (−6.0 to 2.3)	−4.0 (−8.3 to 0.4)	Aerobic								
5.3 (0.4 to 10.2)	9.9 (4.7 to 15.2)	7.9 (2.2 to 13.6)	11.8 (6.0 to 17.6)	Pilates							
−6.6 (−11.8 to −1.3)	−2.0 (−7.4 to 3.5)	−4.0 (−9.7 to 1.7)	−0.1 (−5.9 to 5.7)	−11.9 (−18.2 to −5.6)	Stretching						
−2.8 (−7.7 to 2.2)	1.9 (−3.2 to 6.9)	−0.2 (−5.4 to 5.0)	3.8 (−1.6 to 9.1)	−8.1 (−14.5 to −1.6)	3.8 (−2.7 to 10.4)	Other specific exercises					
−5.4 (−10.9 to 0.2)	−0.7 (−6.4 to 4.9)	−2.8 (−8.9 to 3.3)	1.2 (−4.9 to 7.2)	−10.7 (−17.4 to −3.9)	1.2 (−5.5 to 8.0)	−2.6 (−9.0 to 3.8)	Yoga				
1.3 (−4.8 to 7.3)	5.9 (−0.1 to 11.9)	3.8 (−2.1 to 9.8)	7.8 (1.3 to 14.2)	−4.1 (−11.4 to 3.3)	7.8 (0.2 to 15.5)	4.0 (−3.2 to 11.2)	6.6 (−1.0 to 14.3)	Functional restoration			
1.4 (−5.0 to 7.9)	6.0 (−0.6 to 12.7)	4.0 (−3.0 to 10.9)	7.9 (0.9 to 15.0)	−3.9 (−11.2 to 3.3)	8.0 (0.2 to 15.8)	4.2 (−3.5 to 11.8)	6.8 (−1.3 to 14.9)	0.1 (−8.3 to 8.6)	McKenzie		
−5.9 (−16.2 to 4.3)	−1.3 (−11.6 to 8.9)	−3.4 (−13.8 to 7.0)	0.6 (−9.7 to 10.8)	−11.3 (−22.5 to 0.0)	0.6 (−10.4 to 11.7)	−3.2 (−14.2 to 7.8)	−0.6 (−12.0 to 10.9)	−7.2 (−18.9 to 4.5)	−7.4 (−19.1 to 4.4)	Flexibility/ mobilising	
−13.4 (−17.2 to −9.6)	−8.8 (−12.5 to −5.1)	−10.9 (−14.7 to −7.0)	−6.9 (−10.9 to −2.9)	−18.7 (−24.4 to −13.1)	−6.8 (−11.9 to −1.8)	−10.7 (−15.8 to −5.5)	−8.1 (−14.1 to −2.0)	−14.7 (−21.3 to −8.1)	−14.8 (−21.4 to −8.2)	−7.5 (−17.5 to 2.5)	Minimal treatment
−5.9 (−9.8 to −2.0)	−1.3 (−5.2 to 2.6)	−3.3 (−7.4 to 0.7)	0.6 (−3.5 to 4.8)	−11.2 (−17.2 to −5.3)	0.7 (−4.7 to 6.1)	−3.1 (−8.6 to 2.3)	−0.5 (−6.9 to 5.9)	−7.2 (−13.9 to −0.4)	−7.3 (−14.1 to −0.5)	0.1 (−10.1 to 10.2)	Effective treatment

B

Core strengthening											
−2.2 (−4.1 to −0.3)	Mixed (> 2 types)										
−1.7 (−4.0 to 0.6)	0.5 (−1.9 to 2.9)	General strengthening									
−2.3 (−5.0 to 0.3)	−0.1 (−2.9 to 2.6)	−0.7 (−3.6 to 2.3)	Aerobic								
3.5 (0.3 to 6.7)	5.7 (2.3 to 9.1)	5.2 (1.5 to 8.9)	5.9 (2.0 to 9.7)	Pilates							
−3.2 (−6.3 to −0.2)	−1.0 (−4.3 to 2.2)	−1.6 (−5.1 to 1.9)	−0.9 (−4.6 to 2.8)	−6.8 (−10.6 to −3.0)	Stretching						
−3.1 (−6.3 to 0.2)	−0.9 (−4.2 to 2.4)	−1.4 (−4.8 to 2.0)	−0.8 (−4.5 to 3.0)	−6.6 (−10.9 to −2.4)	0.2 (−3.9 to 4.2)	Other specific exercises					
−1.3 (−4.8 to 2.1)	0.9 (−2.6 to 4.3)	0.3 (−3.5 to 4.1)	1.0 (−2.8 to 4.8)	−4.9 (−9.1 to −0.6)	1.9 (−2.1 to 5.9)	1.7 (−2.3 to 5.7)	Yoga				
0.8 (−3.4 to 5.0)	3.0 (−1.3 to 7.2)	2.4 (−1.6 to 6.5)	3.1 (−1.3 to 7.6)	−2.8 (−7.8 to 2.3)	4.0 (−1.0 to 9.0)	3.9 (−1.1 to 8.9)	2.1 (−3.0 to 7.2)	Functional restoration			
5.1 (0.3 to 9.9)	7.3 (2.4 to 12.2)	6.7 (1.6 to 11.9)	7.4 (2.2 to 12.6)	1.5 (−4.1 to 7.2)	8.3 (2.8 to 13.8)	8.2 (2.6 to 13.7)	6.4 (0.7 to 12.2)	4.3 (−1.9 to 10.5)	McKenzie		
4.4 (−1.9 to 10.6)	6.6 (0.3 to 12.8)	6.0 (−0.5 to 12.5)	6.7 (0.1 to 13.3)	0.8 (−6.2 to 7.9)	7.6 (0.9 to 14.4)	7.5 (0.6 to 14.3)	5.7 (−1.3 to 12.7)	3.6 (−3.9 to 11.1)	−0.7 (−8.5 to 7.1)	Flexibility/ mobilising	
−6.6 (−9.0 to −4.3)	−4.4 (−6.8 to −2.0)	−5.0 (−7.5 to −2.4)	−4.3 (−7.0 to −1.6)	−10.2 (−13.8 to −6.6)	−3.4 (−6.4 to −0.3)	−3.5 (−6.7 to −0.4)	−5.3 (−9.0 to −1.6)	−7.4 (−11.9 to −2.9)	−11.7 (−16.7 to −6.7)	−11.0 (−17.2 to −4.8)	Minimal treatment
−2.0 (−4.5 to 0.5)	0.2 (−2.3 to 2.7)	−0.3 (−3.1 to 2.4)	0.3 (−2.5 to 3.1)	−5.5 (−9.4 to −1.7)	1.2 (−2.1 to 4.6)	1.1 (−2.4 to 4.6)	−0.7 (−4.6 to 3.3)	−2.8 (−7.4 to 1.8)	−7.1 (−12.1 to −2.1)	−6.4 (−12.7 to −0.1)	Effective treatment

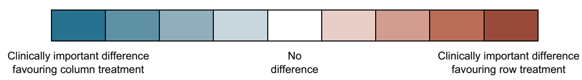

Figure 5. League table of network meta-analysis results for all comparisons between exercise and non-exercise interventions. Effects are expressed as the mean difference (95% CI) between interventions on short-term outcomes **A.** Pain intensity, 0 to 100 (198 trials, 466 groups, 17,534 participants). **B.** Functional limitations, 0 to 100 (187 trials, 433 groups, 16,926 participants). Blue shading indicates that the intervention listed in the column is more effective than that in the row, whereas orange shading indicates the intervention listed in the row is more effective. The depth of shading indicates the size of the treatment effect and likely compatibility with a clinically important effect (Appendix 7 describes the shading algorithm).

ducted. Adjusting for these exercise characteristic covariates did not modify the conclusions about either Pilates or stretching exercises (Appendix 16 on the eAddenda). In these pain and function models, Pilates remained consistently better than other exercise types; core strengthening exercise types were consistently better than other exercise types, McKenzie therapy was not consistently better, and stretching exercises were consistently worse than other types and non-exercise comparisons.

Certainty of the network meta-analysis evidence

The overall certainty of the evidence available for each comparison in the primary network meta-analyses, assessed using the CINeMA framework, is reported in Appendix 17 on the eAddenda. Result characteristics that lowered our certainty in the evidence included major concerns about within-study bias in 31% of comparisons, and some concerns in the remaining 69% of comparisons (53 of 77). Heterogeneity was common, with some concerns or major concerns identified in all pairwise comparisons (40 of 77 had some concerns, 37 had major) for pain outcomes and 91% of functional limitation outcome comparisons (62 of 77 had some concerns, eight had major). Examining the coherence of the network, direct and indirect estimates of treatment effect were found to agree with respect to direction and size; interpretation of direct and indirect meta-analysis results differed for three of 52 comparisons (pain) (one major concern, two some concerns) and eight of 49 comparisons (functional limitations) (all some concerns) with direct data available. Loop-specific coherence was observed for 92 of 101 loops with direct data available for pain intensity outcomes, and 79 of 90 loops with direct data available for functional limitation outcomes (Appendix 18 on the eAddenda). The global test of incoherence did not show significant incoherence for pain outcomes (χ^2 = 49.8, 74 df, p = 0.99) or for functional limitation outcomes (χ^2 = 80.2, 68 df, p = 0.15). There was minimal observed change in heterogeneity or incoherence for any sensitivity analyses (Appendix 19 on the eAddenda).

Discussion

It is believed that this study is the largest network meta-analysis in the field of low back pain. The primary network meta-analysis showed pain and functional limitation outcomes compatible with clinically important effects for Pilates, McKenzie therapy and functional restoration exercise types.

All exercise types, other than stretching exercises when adjusted for dose and additional co-interventions, were consistently more effective than minimal care and most other comparison treatments for reducing pain intensity and improving functional limitations in people with chronic low back pain. Pilates exercises were found to be more effective for pain outcomes than other comparison treatments and other types of exercise treatment. In the primary models, McKenzie therapy and Pilates were found to more effective for functional limitation outcomes than other comparisons and other types of exercise treatment. McKenzie therapy was no longer better than other comparisons when adjusted for dose and additional co-interventions. High dose of most exercise treatments appeared to reduce pain and functional limitation outcomes more than low dose, and the addition of co-interventions appeared to improve effectiveness of most exercise types for pain and functional limitation outcomes.

Comparison with other studies

There are many systematic reviews of specific exercise treatment types for persistent low back pain, including published Cochrane reviews,[22–25] upcoming Cochrane protocols[26–28] and many recent systematic reviews that have been published outside of the Cochrane Library.[29–36] These reviews have included five to 29 trials on treatments for chronic low back pain, a small proportion of the trials included here, with more focused review questions and selection criteria, and some differences in methods. Similar to the current findings, the Cochrane reviews have reported low to moderate quality evidence that the specific exercise type investigated produced small to moderate improvements in outcomes compared with minimal intervention. However, these Cochrane reviews reported limited evidence of important differences compared with other types of exercise and other conservative treatments.

Owen and colleagues (2019) conducted a similar overarching network meta-analysis of exercise treatment for chronic low back pain to compare effectiveness of different exercise types.[37] That review included 89 trials comparing exercise and non-exercise treatments; they used stricter inclusion criteria and a less comprehensive search than our review. The considerable difference in number of included trials makes it challenging to compare findings. However, Owen and colleagues similarly reported low-quality evidence for their network meta-analysis due to risk of bias and heterogeneity, and reported Pilates to be effective for improving pain.

Strengths and limitations

Our results were limited by the quality of the included evidence. The evidence available for most treatment comparisons was judged to be of moderate certainty, due to within-study risk of bias and heterogeneity. Unexplained heterogeneity remained, despite exploration of multiple population, exercise and methodological characteristics; this decreased confidence in the available evidence. Incomplete reporting of trial and population characteristics, differing opinions about treatment type classifications, and potential misclassification of exercise types and population characteristics are additional limitations. It should also be noted that the interventions that appeared to be the most effective were also interventions that are costly to deliver and to 'purchase' for patients. It is possible that our results were conflated with other factors related to higher socioeconomic status in these patient groups (eg, physical labour, other healthcare access and health status).

Our study also had several strengths. It used a comprehensive search, robust selection criteria and detailed data extraction processes. The large number of included trials increased generalisability and provided a dense, well-connected network for analyses. It considered quality/integrity issues when selecting trials, used best methods for analyses, and comprehensively explored methodological issues and potential effect modifiers. The results are reported in a complete and transparent way.

Incoherence, often a concern in network meta-analysis because it relies on the assumption of transitivity, did not seem to be a problem with the evidence available, as evidenced by global and local tests of incoherence. While there was some evidence of exercise characteristic effect modifying covariate differences (dose and additional co-interventions) that may indicate lack of transitivity in the network, adjusting for these covariates did not modify our overall conclusions.

Implications for clinical practice

This network meta-analysis provides further support for exercise treatment being effective and guideline-recommended care. It found that most exercise types are more effective than minimal treatment; to increase dose, patients should be encouraged to perform the exercise that they enjoy and will take part in consistently. We found evidence that Pilates, McKenzie therapy and functional restoration were more effective than other types of exercise treatment for reducing pain intensity and functional limitations; the observed effect of McKenzie therapy may be related to higher dose and/or co-interventions in included trials. If the observed pain and function outcomes align with the patient's goals, it may be appropriate to recommend these types of exercise programs if they are available and financially feasible for the patient.

Implications for research

One considerable advantage of network meta-analysis and meta-regression over traditional aggregate meta-analysis is the ability to consider all evidence from trials. This includes direct and indirect evidence, as well as other population and delivery characteristics. The strengths of network meta-analysis were used to include a large number of trials and borrow strength from indirect evidence. However, potential challenges are similar to traditional meta-analyses where the evidence is limited by the reporting and quality of the trials included. There is a continued goal to improve the planning, conduct and reporting of the trials.

Future trials on exercise treatment should evaluate other relevant patient outcomes aligned with the proposed mechanisms of exercise treatment and cost-effectiveness, as this will guide individuals and clinicians in their choice for the best treatment.

This review explored several sources of heterogeneity, including population and treatment characteristics, outcomes and methodological characteristics. Statistical heterogeneity was moderate to substantial in most analyses, not reduced by sensitivity, subgroup analyses or considering suspected treatment effect modifiers. Examining individual-level patient characteristics and their relationship with effectiveness of exercise with individual participant data was beyond the scope of this review; however, this would be important to consider in future research.

In conclusion, our findings were compatible with moderate to clinically important treatment effect for Pilates and McKenzie therapy compared with minimal treatment, other effective treatments and other exercise types. This analysis should help guide primary care clinicians in their patient management and referral practices.

What was already known on this topic: Clinical practice guidelines recommend exercise as first-line management for chronic low back pain. However, there is limited evidence to support the use of one type of exercise or program characteristic over another.
What this study adds: Most exercise types were more effective than minimal treatment for pain and functional limitation outcomes. Pilates, McKenzie therapy and functional restoration were more effective than other types of exercise treatment for reducing pain intensity and functional limitations.

Footnotes: [a] Stata SE v14.2, Stata Corp, College Station, USA. [b] R software V4.0.3, R Core Team, Vienna, Austria.

eAddenda: Appendices 1 to 19 can be found online at https://doi.org/10.1016/j.jphys.2021.09.004.

Ethics approval: Not applicable.

Competing interests: Nil.

Source(s) of support: Funding was received from Research Nova Scotia to carry out analyses related to the Cochrane review.

Acknowledgements: Dr. Yamato is supported by the São Paulo Research Foundation, FAPESP (2019/10330-4), São Paulo, Brazil. Dr. Saragiotto is supported by the São Paulo Research Foundation, FAPESP (2020/14231-8), São Paulo, Brazil.

Provenance: Not invited. Peer reviewed.

Correspondence: Jill A Hayden, Department of Community Health & Epidemiology, Dalhousie University, Halifax, Canada. Email: jhayden@dal.ca

References

1. Chou R, Qaseem A, Snow V, Casey D, Cross JT, Shekelle P, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*. 2007;147:478–491.
2. Toward Optimized Practice (TOP) Low Back Pain Working Group. Evidence-informed primary care management of low back pain: Clinical practice guideline. 3rd edition; 2017. Edmonton, AB: Toward Optimized Practice. Available from: https://actt.albertadoctors.org/CPGs/Lists/CPGDocumentList/LBP-guideline.pdf. Accessed 1 September, 2021.
3. Hayden J, Ellis J, Ogilvie R, Malmivaara A, van Tulder M. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev*. 2021;CD009790:In Press.
4. Ferreira ML, Smeets RJ, Kamper SJ, Ferreira PH, Machado LA. Can we explain heterogeneity among randomized clinical trials of exercise for chronic back pain? A meta-regression analysis of randomized controlled trials. *Phys Ther*. 2010;90:1383–1403.
5. Crequit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer. *BMC Med*. 2016;14:8.
6. Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JP, Mavergames C, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med*. 2014;11:e1001603.
7. Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med*. 2005;142:776–785.
8. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, 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;162:777–784.
9. Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 Updated method guideline for systematic reviews in the Cochrane Back and Neck Group. *Spine*. 2015;40:1660–1673.
10. de Campos TF. Low back pain and sciatica in over 16s: assessment and management NICE Guideline [NG59]. *J Physiother*. 2017;63:120.
11. Hauk L. Low Back Pain: American College of Physicians Practice Guideline on Noninvasive Treatments. *Am Fam Physician*. 2017;96:407–408.
12. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine*. 2000;25:3115–3124.
13. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980;66:271–273.
14. Higgins J. *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011]. 2021. The Cochrane Collaboration. www cochrane-handbook org
15. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
16. Law M, Jackson D, Turner R, Rhodes K, Viechtbauer W. Two new methods to fit models for network meta-analysis with random inconsistency effects. *BMC Med Res Methodol*. 2016;16:1–14.
17. Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med*. 1996;15:2733–2749.
18. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med*. 2010;29:932–944.
19. Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. *Stata J*. 2015;15:905–950.
20. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, et al. CINeMA: An approach for assessing confidence in the results of a network meta-analysis. *PLoS Med*. 2020;17:e1003082.
21. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Grotle M, et al. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. *J Clin Epidemiol*. 2013;66:1397–1404.
22. Saragiotto BT, Maher CG, Yamato TP, Costa LO, Menezes Costa LC, Ostelo R, et al. Motor control exercise for chronic non-specific low-back pain. *Cochrane Database Syst Rev*. 2016.
23. Schaafsma FG, Whelan K, van der Beek AJ, van der Es-Lambeek LC, Ojajärvi A, Verbeek JH. Physical conditioning as part of a return to work strategy to reduce sickness absence for workers with back pain. *Cochrane Database Syst Rev*. 2013.
24. Wieland LS, Skoetz N, Pilkington K, Vempati R, D'Adamo CR, Berman BM. Yoga treatment for chronic non-specific low back pain. *Cochrane Database Syst Rev*. 2017.
25. Yamato TP, Maher CG, Saragiotto BT, Hancock MJ, Ostelo R, Cabral CMN, et al. Pilates for low back pain. *Cochrane Database Syst Rev*. 2015.
26. Liu H, Yao K, Zhang J, Li L, Wu T, Brox JI, et al. Sling exercise therapy for chronic low-back pain. *Cochrane Database Syst Rev*. 2013.
27. Machado L, Lin CWC, Clare H, van Tulder MW. The McKenzie method for (sub)acute non-specific low-back pain (protocol). *Cochrane Database Syst Rev*. 2012.
28. Machado L, van Tulder MW, Lin CWC, Clare H, Hayden JA. The McKenzie method for chronic non-specific low-back pain (protocol). *Cochrane Database Syst Rev*. 2012.
29. Alhakami AM, Davis S, Qasheesh M, Shaphe A, Chahal A. Effects of McKenzie and stabilization exercises in reducing pain intensity and functional disability in individuals with nonspecific chronic low back pain: a systematic review. *J Phys Ther Sci*. 2019;31:590–597.
30. Namnaqani FI, Mashabi AS, Yaseen KM, Alshehri MA. The effectiveness of McKenzie method compared to manual therapy for treating chronic low back pain: a systematic review. *J Musculoskelet Neuronal Interact*. 2019;19:492–499.
31. Niederer D, Mueller J. Sustainability effects of motor control stabilisation exercises on pain and function in chronic nonspecific low back pain patients: a systematic review with meta-analysis and meta-regression. *PLoS One*. 2020;15:e0227423.
32. Shi Z, Zhou H, Lu L, Pan B, Wei Z, Yao X, et al. Aquatic exercises in the treatment of low back pain: a systematic review of the literature and meta-analysis of eight studies. *Am J Phys Med Rehabil*. 2018;97:116–122.
33. Vanti C, Andreatta S, Borghi S, Guccione AA, Pillastrini P, Bertozzi L. The effectiveness of walking versus exercise on pain and function in chronic low back pain: a systematic review and meta-analysis of randomized trials. *Disabil Rehabil*. 2019;41:622–632.
34. Wewege MA, Booth J, Parmenter BJ. Aerobic vs. resistance exercise for chronic non-specific low back pain: a systematic review and meta-analysis. *J Back Musculoskelet Rehabil*. 2018;31:889–899.
35. Zhang Y, Loprinzi PD, Yang L, Liu J, Liu S, Zou L. The beneficial effects of traditional Chinese exercises for adults with low back pain: a meta-analysis of randomized controlled trials. *Medicina*. 2019;55:118–141.
36. Zou L, Zhang Y, Yang L, Loprinzi PD, Yeung AS, Kong J, et al. Are mindful exercises safe and beneficial for treating chronic lower back pain? a systematic review and meta-analysis of randomized controlled trials. *J Clin Med*. 2019;8:628–654.
37. Owen PJ, Miller CT, Mundell NL, Verswijveren SJ, Tagliaferri SD, Brisby H, et al. Which specific modes of exercise training are most effective for treating low back pain? Network meta-analysis. *Br J Sports Med*. 2020;54:1–12.