BMJ Open Sport & Exercise Medicine

DOSage of Exercise for chronic low back pain disorders (DOSE): protocol for a systematic review with dose-response network meta-analysis

Nitin Kumar Arora ,[1,2] Lars Donath ,[2] Patrick J Owen ,[3,4] Clint T Miller ,[5] Svenja Kaczorowski ,[1,2] Tobias Saueressig ,[6] Hugo Pedder ,[7] Niamh L Mundell,[5] Scott D Tagliaferri ,[8,9] Ashish Diwan ,[10] Xiaolong Chen ,[10] Xiaohui Zhao ,[11] Eva-Maria Huessler ,[12] Katja Ehrenbrusthoff ,[1] Jon J Ford ,[13,14] Andrew J. Hahne ,[13] Ludwig Hammel,[15] Heike Norda,[16] Daniel L Belavy [1]

To cite: Arora NK, Donath L, Owen PJ, *et al*. DOSage of Exercise for chronic low back pain disorders (DOSE): protocol for a systematic review with dose-response network meta-analysis. *BMJ Open Sport & Exercise Medicine* 2024;**10**:e002108. doi:10.1136/bmjsem-2024-002108

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjsem-2024-002108).

Accepted 5 July 2024

For numbered affiliations see end of article.

Correspondence to
Professor Daniel L Belavy; belavy@gmail.com

ABSTRACT

Chronic low back disorders are the leading cause of direct and indirect healthcare burden globally. Exercise training improves pain intensity, mental health and physical function. However, the optimal prescription variables are unknown. We aim to compare the efficacy of various exercise dosages for chronic low back disorders to identify the optimal prescription variables. Six databases (Medline, SPORTDiscus, CINAHL, PsycINFO, EMBASE and CENTRAL), trial registries (ClinicalTrials.gov and WHO International Clinical Trials Registry Platform) and reference lists of prior systematic reviews will be searched, and we will conduct forward and backward citation tracking. We will include peer-reviewed randomised controlled trials (individual, cluster or cross-over trials) published in English or German language comparing exercise training to other exercise training or non-exercise training interventions (conservative, non-surgical, non-pharmacological, non-invasive treatments, placebo, sham, usual/standard care, no-treatment control, waitlist control) in adults with chronic low back disorders. Outcomes will include pain intensity, disability, mental health, adverse events, adherence rate, dropout rate and work capacity. Version 2 of the Cochrane risk-of-bias tool will be employed. The dose will be categorised as cumulative dose (total and weekly minutes of exercise training) and individual dose prescription variables (intervention duration, session duration, frequency and intensity). Dose-response model-based network meta-analysis will be used to assess the comparative efficacy of different exercise doses to determine a dose–response relationship. The certainty of evidence will be assessed using the Grading of Recommendations Assessment, Development and Evaluation. Information about optimal exercise training dosage will help in enhancing treatment outcomes.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Exercise training can improve pain intensity, mental health and disability in adults with chronic low back disorders, yet the optimal prescription variables are unknown.

WHAT THIS STUDY ADDS

⇒ This study will identify the optimal exercise training prescription variables for treatment outcomes in adults with chronic low back disorders.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Findings may inform clinical practice guidelines on optimal exercise training prescription variables for chronic low back disorders. In the long term, we expect findings to improve healthcare resource allocation by identifying the minimal exercise dosage, thereby reducing the overprescription and underprescription of exercise therapy.

INTRODUCTION

Low back pain is the leading cause of disability and work absenteeism worldwide.[1] The Global Burden of Disease estimates show that 619 million people were affected by low back pain in 2020, contributing to 69 million years lived with disability.[1] Around 70%–80% of the global population experience low back pain at least once in their lifetime.[2] Women and older individuals tend to have a higher prevalence rate of chronic low back pain as compared with men and younger adults.[3] With more than 50% of the global population comprising adults, the incidence of chronic low back disorders (CLBDs) is expected to increase with an ageing population.[4]

Low back pain can be acute, subacute or chronic. As acute cases often resolve without intervention, the majority of exercise-related research is targeted at CLBDs.[5 6] CLBDs are defined as low back pain (between the 12th rib and upper part of the inferior gluteal fold,

with or without leg pain) that persists for more than 12 weeks from onset.[7] CLBDs encompass various diagnoses related to pain in the lower back (eg, disc herniation, spondylolisthesis, hemivertebrae and spondylosis) and pain radiating to lower extremities (sciatica, radicular pain, lumbar radiculopathy, spinal stenosis, neurological claudication).[7 8] Though the diagnoses encompassing CLBDs differ pathoanatomically, current evidence-based conservative treatment recommendations remain similar for these diagnoses.[9]

International clinical practice guidelines[10] and systematic reviews with meta-analyses consistently recommend exercise training for CLBDs (chronic non-specific low back pain and radicular syndromes).[11] While exercise training is effective for CLBDs, a key evidence gap is the most efficacious 'dose' of exercise training. This pertains to the efficacy of individual exercise prescription variables, such as frequency, duration and intensity, and of a collective load (weekly and total minutes of exercise training) based on these prescription variables. A recent systematic review and network meta-analysis (NMA) assessed different modes of exercise training for chronic non-specific low back pain and showed that the optimal exercise training dose is still unclear.[12] For example, the randomised controlled trials (RCTs) included in Owen *et al*[12] had an exercise intervention ranging from 4 to 24 weeks in duration and from one to seven sessions per week. When implemented in clinical practice, this wide variation in total exercise volume, or 'dose', would likely lead to different efficacy and represent a very broad difference in related healthcare costs.

In a recent umbrella review, we evaluated systematic reviews of exercise for chronic musculoskeletal pain conditions.[13] We showed that 94.2% (258/274) of systematic reviews on the topic had 'low or critically low' methodological quality per AMSTAR-2.[14] 93.4% (256/274) of the reviews reported data on different exercise prescription variables (duration, frequency, intensity and volume) but only around 11% (30/274) analysed the relationship between different dosages of exercise prescription variables (frequency, duration or intensity) and outcomes in musculoskeletal pain. We also noted that only eleven reviews assessed, quantitatively in a meta-analysis, the relationship between exercise dose and intervention efficacy in CLBDs. All 11 reviews focused on pain and physical function, with only one of these reviews addressing mental health. None of these reviews considered adverse events or adherence rates, despite the potential influence of exercise dose on all three aforementioned outcome variables. Also, most of the included reviews performed only subgroup analysis or meta-regression with one type of exercise training. Specifically, we identified the following evidence gaps: (a) examining all forms of exercise while performing relevant subgroup analyses, (b) encompassing CLBDs more widely to inform clinical practice while performing relevant subgroup analyses for population (non-specific low back pain and radicular syndromes), (c) performing a high-quality review per AMSTAR 2 criteria, (d) considering mental health outcomes, adverse events and adherence rates and (e) assessing the reporting quality of exercise training.

Given the existing gaps in evidence and the disease burden, our objective is to conduct a high-quality systematic review utilising dose-response NMA of RCTs that examines exercise training in various doses (including intervention duration in weeks, length of individual exercise sessions (session duration), the number of sessions per week (exercise frequency), exercise intensity, weekly and total dose (minutes)) for alteration of pain, disability, mental health, adverse events, adherence rate, drop-out rate or work capacity (the number of sick leaves or return to work) in individuals with CLBDs.

METHODS

This review will be conducted and reported in line with the current update of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses[15] guidelines and PRISMA extension for NMAs.[16] The PRISMA-Protocols checklist was used to draft this protocol report.[17] Furthermore, we incorporate the methodological recommendations outlined in the technical support documents of the Decision Support Unit at the National Institute for Health and Care Excellence.[18 19] We registered the protocol for the current review on the Open Science Framework.[20]

Eligibility criteria

A list of exclusion criteria is available in online supplemental data 1.

Population: Adults (aged 18 years and over) with CLBDs (>12 weeks of pain and discomfort, localised below the costal margin and above the inferior gluteal folds, with or without leg pain).[8] Back pain with or without leg pain and without any specific spinal pathologies (ie, vertebral fracture, malignancy, spinal infection, axial Spondyloarthritis, cauda equina syndrome)[11] will be included. Hence, non-specific low back pain (ie, spondylolisthesis, spondylosis, disc herniation, disc degeneration, scoliosis, failed back surgery syndrome and hemivertebrae) and radicular syndromes (ie, radicular pain, sciatica, radiculopathy, spinal stenosis) will be included.[11] No restrictions will be placed on sex or race. Recurrent low back pain, defined as low back pain less than 12 weeks with a minimum 6 months pain-free duration,[8] will be excluded. Detailed information about the population is available in online supplemental data 2.

Intervention: Exercise training is 'a series of specific movements to train or develop the body by a routine practice or as physical training to promote good physical health'.[21] All modes/types of exercises[22] will be included under the following subclasses: (a) resistance, (b) stabilisation/motor control, (c) Pilates, yoga, traditional eastern approaches, (d) aerobic, (e) stretching and (f) other and water-based exercises. Detailed information about the types of exercises to be included in this review

can be found in online supplemental data 3. Studies where a clinician or study investigator instructed or prescribed exercises to patients to improve the low back disorder will be included.

We will focus on exercise-only studies. To approach this pragmatically, we require exercise training as the primary treatment component. Exercise training should constitute more than 50% of the assigned intervention (per the extractor judgements; see online supplemental data 4) to be considered as the primary treatment component. Further, exercise interventions with a secondary treatment component (ie, a cointervention, such as a large education component) will be excluded. To qualify as a secondary treatment component, it needed to account for at least 20% of the total intervention (as determined by the extractor). The primary and secondary intervention criteria are available in online supplemental data 4. Studies that only include multidisciplinary interventions and no exercise group, where the focus is on multidisciplinary treatment, which may have included exercise will not be included. RCTs that involve an exercise prescription by a therapist and advice on practising the exercise routine at home (using books, videos, notes and educational material) will be included. Education or advice to do exercise will not be considered exercise training. RCTs examining the effects of a single bout of exercise will be excluded.

Comparator: Eligible comparators include exercise interventions, non-exercise conservative treatments (manual therapies and manipulation, McKenzie method or McKenzie exercise, psychological therapies, electrophysical agents, education, physical therapy (otherwise not falling into specific treatment combination), massage, multidisciplinary (multidisciplinary pain management)), true control (no treatment or waitlist control), usual or standard care, placebo or sham treatment. Studies that compare only the same type of exercise at the same dose (intensity, frequency, intervention duration, session duration) will not be included. In studies where all groups perform the same type of exercise if certain exercise dose components are not reported but other dose variables are the same across all groups, we will assume the missing dose variable is the same in both groups, and such studies will be excluded. A detailed list of comparators is available in online supplemental data 5.

Outcomes

Primary outcomes: Disability (eg, Roland Morris Disability Questionnaire (RMDQ), pain intensity (eg, Visual Analogue Scale, Numerical Rating Scale) or mental health (eg,SF-36 (short form-36) mental health summary scale).

Secondary outcomes: The number of subjects with adverse events (related or unrelated to intervention),[23] number of drop-outs (overall and discontinuation due to adverse events), adherence rate (number of sessions attended out of the total number of intended exercise sessions), work-capacity (return to work (yes/no), sick leave (yes/no)).

We will include studies that contain at least one of the primary outcomes of interest, and the secondary outcomes will be sought within those studies, if present. A detailed list of included outcomes scales and priority order for continuous outcomes is available in online supplemental data 6.

Study design: We will include peer-reviewed RCTs (individual, cluster randomised or cross-over design) reported in English or German. Previous research demonstrated that excluding non-English articles does not greatly affect point estimates but can lead to a narrower CI.[24] We pragmatically chose to include articles in languages where the author team is fluent.

Information sources and search strategy

The database search strategy of a prior review[7] will be updated to the current date and specified for exercise interventions (using the search terms described in online supplemental data 7. We will search the electronic databases of PubMed, EMBASE, CINAHL, CENTRAL, SPORTDiscus and PsycINFO for peer-reviewed and published RCTs. In addition to this search update, we will further search ClinicalTrials.gov and WHO ICTRP to identify ongoing and completed trials relevant to our review. As part of the prior project,[7] we screened references from 285 published English language systematic reviews on CLBDs and 17 Cochrane reviews published from January 1990 to July 2019. Further, we will incorporate the reference lists of systematic reviews published after 2019 on exercise and chronic low back pain identified in our recent umbrella review.[13] Furthermore, we will conduct backward and forward citation tracking to ensure the inclusion of relevant records in our study.

Selection process

As part of quality assurance, the screeners will conduct pilot title abstract and full-text screening independently with 100 and 10 articles, respectively. Based on the refined criteria following pilot screening, duplicate and independent title abstracts and full texts screening will be performed (online supplemental data 8). Exercise studies identified in a prior review[7] will be further subjected to the additional exercise-focused criteria of this review and screened independently in duplicate. Conflicts at the title abstract and full-text screening stage will be discussed, and an adjudicator will be contacted for a final decision in case the disagreement persists.

Data collection process

Extractors (online supplemental data 8) will pilot data extraction for 10 studies included at the full-text screening stage, and the extracted data will be compared with identify themes of disagreement. Following refinement of the extraction template based on the pilot extraction, duplicate data extraction will be performed in a custom-made Excel sheet for all the included studies. Conflicts

in extraction will be resolved through discussion, and an adjudicator will be contacted if the disagreement persists. An additional person will independently conduct random cross-checking of 5% of full-text screened records and 5% of extracted studies.

Data items

Data will be extracted for publication details (author information, year of publication), population (chronic low back pain condition, mean duration of symptoms, mean age of participants, number of participants in each group, number of females), interventions and comparator information (type of exercise interventions, comparators), outcome details (mean/median, SD/IQR/ranges and number of participants for the outcome variables in intervention and control groups at multiple follow-ups (short term (>1 day (d) but ≤3 months), intermediate term (>3 months but <12 months) and long-term (≥12 months)), trial funding and type of RCT as detailed in Belavy *et al*.[7] If the SD is not available in the study record, we will calculate it based on alternate measures that contain information about the variability, for example, median, IQR, SE and CI.[25] When data are presented only in figures, we will use ImageJ to manually extract the data by measuring the axis and data point length (in pixels).[26]

The following exercise training characteristics will be extracted as per the Consensus on Exercise Reporting Template (CERT) items: exercise training frequency (times per week), time (in minutes per week), intensity (high/moderate/low),[27] duration of intervention (in weeks), session duration (in minutes per session), the number of daily and total sessions (online supplemental data 9) delivery of exercise training, the number of sets and repetitions, rest period between repetitions and sets, exercise setting, trainer experience, supervision, adherence measures. Dichotomous (yes/no) responses will be extracted for the following CERT items: description of exercises (photograph, video or illustrations), equipment used, exercise training progression, home exercise programme, individually tailored prescription, starting exercise level and motivation level and fidelity of planned and executed intervention[28] (online supplemental data 10).

Data coding and management

In line with the recommendations by Pedder *et al* 2021,[29] we will organise the datasets for dose-response analysis across various follow-up time points. This will involve extracting and classifying the interventions as follows:

1. Active versus inactive: classifying interventions as exercise or one of the included comparators (non-exercise conservative treatments, true/waitlist control, usual or standard care, or placebo/sham treatment interventions (see online supplemental data 5).
2. Dose level: classifying interventions by an overall 'cumulative dose' of the exercise (weekly and total minutes of exercise training)
3. Component level: classifying interventions based on the following:
   a. Intensity: low/moderate/high.
   b. Frequency: Number of sessions per week.
   c. Duration: intervention and session duration.

Missing data

If we are unable to extract outcome data (mean, median, SD, IQR, range, N responders, N randomised) directly from the record or calculate it from the data provided, we will contact authors (three repeated contact attempts over 4 weeks). If solely SD data are missing, we will impute these using the pooled SD (on the log-scale) from studies for which SD is available for a given outcome.[30] If data for mean or N are missing, the study will be excluded from meta-analysis.

Risk of bias assessment

Two reviewers will independently assess the risk of bias (RoB) in the included studies using version 2 of the Cochrane risk-of-bias tool for randomised trials (RoB 2).[31] Any conflicts in the RoB ratings will be discussed, and an adjudicator will be contacted if they persist. Cochrane RoB 2 is assessed on a per-outcome/study basis. The outcomes included in the current study are typically, with the exception of dropouts, patient reported and we expect the RoB rating to not deviate importantly for these outcomes. For feasibility, we will assess Cochrane RoB 2 for one of the outcomes in the following priority order: pain, disability and mental health. Where trials report multiple time points, Cochrane guidance is not available as to which to choose.[31] We will assess the time point closest to the end of the exercise-intervention phase. The RoB ratings will be visualised in R[32] using the robvis package.[33]

Data synthesis and analysis

NMAs will be performed in a Bayesian framework using JAGS[34] and Stan.[35] R statistical computing environment and R packages MBNMAdose,[36] netmeta,[37] robvis,[33] multinma,[38] metafor[39] will be used for statistical analyses and data visualisation.

For continuous outcomes, we will use the standardised mean difference (SMD) as the effect measure (using internal reference SDs for standardisation rather than within-trial SDs[40]) as the standardisation allows pooling and comparison of the results from different outcome scales.[41] The outcome is binary for outcomes such as the number of people with adverse events or sick leave, but we expect that studies with longer follow-ups will have more events. Using a cloglog link function and assuming an exponential (constant) hazard rate, we will estimate HRs for the treatment effects, allowing studies with differing follow-up times to inform this.[19] In the case of continuous subjective outcomes that are reverse scaled (higher values indicating better results instead of lower values), we will multiply the mean values in each group by −1 per the recommendations of the Cochrane

Handbook.[25] The American Guidelines on Treatment of low back pain[42] recommend a minimum threshold for a 'small' effect on pain to be 5 points (0–100 VAS scale; 0.5 points on a 0–10 scale) and for function, 5 points on the Oswestry Disability Index (0–100 scale) and 1 point on the RMDQ (0–24 scale). We will take these minimum thresholds to determine whether different doses of exercise show at least a 'small' difference from other exercise doses. The effect sizes for the different comparisons will be back-transformed to the original scales by multiplying the meta-analysis SMD results with SDs for each scale following the guidance by Murad *et al* 2019.[43]

Stages of data analysis

We will perform meta-analyses of increasing complexity levels to investigate the impact on heterogeneity. We will report results for models at the most complex stage that can be reliably estimated based on the available data for each outcome.

1. Stage 1: Perform a random effects pairwise meta-analysis for exercise vs control that 'pools' together different doses of exercise
2. Stage 2: Perform, if possible, a dose-response model-based NMA[44 45] (MBNMA) that allows the estimation of a dose–response relationship for different doses of exercise. This method allows for fitting a dose–response relationship between different doses to add statistical power and help link doses that might otherwise be disconnected in standard NMA.[46] Several dose-response functions will be explored to determine the most parsimonious relationship (see section on Model selection). Standard NMA will also be performed on the 'split' network, which assumes that each dose of exercise is an independent node in the network, to investigate the impact of assuming a dose–response relationship.[42]
3. Stage 3: Perform a dose-response components NMA that allows for estimation of the effects of different dose components (intensity, duration, frequency, weekly dose and total dose).[47] A dose–response relationship will be explored for each continuous dose component, and different dose-response functions will be fitted to determine the most parsimonious relationship.
4. Stage 4: If data allow, we will explore interactions between different dose components and their dose–response relationships identified in stage 3. This will involve fitting an additional interaction parameter between each dose-response parameter for the different dose components. Therefore, such a model will be highly parameterised and require a large amount of data at different doses. As such, it may not be identifiable in our planned analysis.

Handling of exercise dose data

We anticipate that it will be possible to perform a dose-response model based on NMA using cumulative exercise dose (weekly and total duration of intervention) and individual exercise prescription variables (intervention duration, frequency, intensity and session duration). Continuous aspects of different exercise prescription variables will be explored via different dose-response functions using MBNMA models.

Assessment of homogeneity/similarity

A qualitative assessment by a table of study characteristics that organises trials by treatment comparison and box plots per treatment comparisons will be first performed to evaluate the balance in the distribution of potential effect modifiers (pain duration, baseline pain intensity, baseline disability and type of low back pain) across treatment comparisons in the network.[48 49] If there are any initial concerns regarding balance in effect modifiers, we will consider subgrouping by levels of the effect modifier.

The network will be assumed to have a common heterogeneity variance. The between-study SD (tau) and its corresponding 95% credible interval will be estimated for each meta-analysis and used to assess the degree of heterogeneity. In the presence of high heterogeneity, sensitivity, subgroup and meta-regression analyses will be conducted to identify the cause and impact of potential effect modifiers (see the section on Meta-regression, subgroups and sensitivity analyses).

Assumption of transitivity/consistency

Transitivity is a property of evidence loops in a network that denote the agreement between direct and indirect evidence.[50] Balance in effect modifiers between comparisons is required for the assumption of transitivity to be valid, and we will perform a qualitative assessment of these as described in the section above (see the section on Assessment of homogeneity/similarity).

We will test the assumption of transitivity in NMA models by conducting a global assessment of inconsistency for the collected data by comparing the model fit statistics between NMA models (both standard and dose-response models[45]) and corresponding unrelated mean effects (UME) models.[51] Given that dose-response fit and transitivity are linked, we will follow the steps described in Pedder *et al*[45] to investigate transitivity in dose-response MBNMA models. The differences in these models will be assessed using dev-dev plots, between-study SD and model fit characteristics.[52]

If concerns are raised within the UME model following the assessment of global inconsistency, node splitting will be used to compare the difference between direct and indirect evidence where possible. It should be noted that methods have not yet been developed for more complex and novel dose-response component models to investigate transitivity in them. Therefore, we will only explore transitivity in simpler models (standard NMA and dose-response MBNMA). However, if we find the transitivity assumption valid in the simpler models then it is likely to be valid in the dose-component models that make less strong assumptions regarding the similarity of interventions.

If we discover any data points or studies contributing to node inconsistencies, we will review these studies for data errors or differences in study characteristics. After correcting any errors, we will repeat the previously described process.

Model estimation

NMA models will be fitted in a Bayesian framework. The convergence of models will be assessed using Gelman-Rubin statistics (R-hat values) and by inspecting the trace plots. For random effects models with insufficient information to reliably estimate between-study SD, informative prior distributions developed in a previous project will be used.[53] For SMD, we will use a log-Student-t prior distribution for tau with a location parameter (mean) of −1.51 and scale parameter (SD) of 1.135. For the OR, we will use a log-normal prior distribution for tau^2 with a location parameter (mean) of −1.445 and a scale parameter (SD) of 0.955.

Model selection

Model selection will be performed by comparing model fit statistics (residual deviance, between-study SD, effective number of parameters, deviance information criterion (DIC)) between candidate models, with a difference of 3–5 points on DIC being considered to be a meaningful difference.[54] The simpler model will be preferred where DIC is similar between models.

A wide range of dose-response functions (log-linear, Emax, polynomial, fractional polynomial, spline) will be explored to identify the most parsimonious model. We will approach dose-response model selection and exploration using the following steps:

- Fit fixed treatment effects models and estimate the best-fitting dose-response function.
- Compare random versus fixed treatment effects.
- Investigate potential inconsistency in the selected model.

Handling different trial designs

We will follow the Cochrane guidelines for including data from cluster and cross-over RCTs.[55] To assess the robustness of our findings, we will perform a sensitivity analysis with different intracluster correlation coefficients for cluster-RCTs.[55] In the case of cross-over RCTs, we will use the estimated relative treatment effect if the study tests and finds no evidence of a carryover effect. If the study authors do not provide information about the carryover effect or for more complex dose-response models that require data to be specified in an arm-based format, we will use the data from the first cross-over period.

Secondary analyses: meta-regression, subgroup and sensitivity analyses

We will assess the following baseline characteristics as potential effect modifiers that may influence the outcome in CLBDs: pain duration, baseline pain intensity, baseline disability and type of low back pain (eg, radicular syndromes vs non-specific low back pain).[56 57] We will assess the forest plots, statistical heterogeneity (between study SD: τ) and 95% prediction interval to evaluate the validity of the homogeneity assumption. If these effect modifiers are not balanced between treatment comparisons, we will conduct a pairwise random-effects meta-regression to examine the potential impact of effect modification.

In cases where we identify effect modification in pairwise comparisons, we will also explore the potential for effect modification in selected MBNMA models using NMA.[58] Where possible, we will conduct a sensitivity analysis to assess the impact of RoB by separately analysing the studies with a low RoB. Subgroup analyses will be considered for different types of low back pain. Additionally, we will perform a sensitivity analysis by excluding studies where the SD was imputed.

Reporting bias

We will assess potential publication bias and small study effects using statistical and non-statistical methods.[59] We will use funnel plots for the pairwise random effects meta-analysis to evaluate small study effects in our meta-analysis.[60] We will include a bias adjustment in the MBNMA if statistically significant small study effects are identified. This adjustment will assume that biases will favour exercise interventions over control interventions, with no bias between different types or doses of exercise.[61]

Certainty assessment and confidence in recommendations

The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach[62] will be used by two independent assessors to assess the certainty of evidence from the NMA at discrete time points results. The GRADE approach is described in online supplemental data 11.

Equity, diversity and inclusion of patient representatives

We included patient representatives from the Deutsche Vereinigung Morbus Bechterew e.V. (DVMB) and Unabhängige Vereinigung aktiver Schmerzpatienten in Deutschland (UVSD SchmerzLOS e.V.) during the steering committee meetings for review planning and execution. Our committee composition encompasses early, mid-career and senior researchers spanning diverse fields such as physiotherapy, orthopaedic surgery, biostatistics and exercise physiology. Additionally, our team represented various nationalities, including Germany, India, Australia, China and the UK.

DISCUSSION

This systematic review will be the first to employ a dose-response MBNMA to evaluate the relative efficacy of various exercise intervention doses for CLBDs. Given that exercise is consistently recommended for managing CLBDs,[9 63] our dose prescription approach aims to optimise care for patients with these conditions.[21] By doing so, we will address clinically relevant gaps in existing

literature, and the findings can also be used to develop high-quality clinical practice guidelines for implementation in pragmatic healthcare settings. Additionally, our work will identify key gaps in the primary research literature on exercise dosage and CLBDs, helping guide future high-quality RCTs where needed. The findings from this work will be directly applicable to clinical healthcare. They will promote physical health, mental health, pain reduction, adverse events reduction and improvement of exercise adherence in individuals with low back pain disorders.

The results of this review will offer appropriate dosage models, which can be customised to individualise exercise treatment strategies, thereby guiding future research in specific domains of CLBDs. Notably, inadequate reporting of exercise interventions often hinders the translation of research findings into clinical practice.[28] This review aims to enhance the clinical relevance of systematic review evidence through high-quality reporting and targeted analysis to determine how much exercise should be used to reduce back pain. Considering the growing socioeconomic and healthcare burden associated with CLBDs,[1] understanding exercise dosage will also enable optimising time and resource allocation in healthcare. This can lead to more efficient healthcare systems by avoiding underprescription and overprescription exercise interventions for specific CLBDs.

This will be one of the first comprehensive reviews of the entire literature spectrum instead of solely focusing on chronic, non-specific low back pain. Furthermore, this review will incorporate the state-of-the-art dose-response MBNMA analysis,[29] and extend the methodology by incorporating a components approach, alongside other commonly used analytical techniques such as meta-regression (including univariate and multivariate linear regression modelling) and subgroup analysis for individual pain conditions and exercise categories.[7] This multifaceted approach addresses previously unanswered questions concerning exercise dose in the context of low back pain. The dose-response MBNMA serves as a crucial foundation for avoiding the violation of the consistency assumption by preventing the need for pooling different exercise doses and thus reducing inconsistency in the analysis. Beyond addressing our primary research question, the operationalisation of an intervention as a function of multiple-dose components may provide a valuable novel approach for exploring other non-pharmacological treatments (eg, psychotherapy) for various disease areas.

Our review possesses additional strengths, including using diverse information sources such as trial registries and prior systematic reviews. It is important to recognise that our review process has some potential limitations. First, comparing exercise dosages across various types of exercise training may pose challenges, especially considering the lack of literature on specific exercise types and dosages that might hinder our planned comparisons. However, our chosen approach for meta-analysis, namely dose-response MBNMA, has the potential to provide precise estimates based on the different dose comparisons available in the data. This can help make assumptions for the missing dose variable. Second, categorising all low back pain conditions under the umbrella term 'CLBDs' might introduce variability in the treatment effects. Nevertheless, we plan to employ several analytical approaches, including subgroup analysis, sensitivity analysis and meta-regression techniques, to assess and account for the potential confounders that might be contributing to the heterogeneity in the included studies.

In summary, this systematic review and dose-response MBNMA will generate knowledge regarding the optimal exercise prescription variables for treating adults with CLBD. Findings will inform future primary studies and contribute to the development of clinical practice guidelines.

Author affiliations
[1]Hochschule für Gesundheit, Department of Applied Health Sciences, Division of Physiotherapy, University of Applied Sciences, Bochum, Germany
[2]Department of Intervention Research in Exercise Training, German Sport University Cologne, Koln, Germany
[3]Eastern Health Emergency Medicine Program, Melbourne, Victoria, Australia
[4]Eastern Health Clinical School, Monash University, Melbourne, Victoria, Australia
[5]Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences, Deakin University, Geelong, Victoria, Australia
[6]Physio Meets Science GmbH, Leimen, Germany
[7]Population Health Sciences, University of Bristol, Bristol, UK
[8]Centre for Youth Mental Health, The University of Melbourne, Melbourne, Victoria, Australia
[9]Orygen Ltd, Parkville, Victoria, Australia
[10]St. George Hospital, Department of Orthopaedic Surgery, Spine Service, University of New South Wales, Sydney, New South Wales, Australia
[11]Xi'an University of Architecture & Technology, Beilin, China
[12]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Essen, Germany
[13]Discipline of Physiotherapy, School of Allied Health, Human Services & Sport, La Trobe University, Bundoora, Victoria, Australia
[14]Advance Healthcare, Victoria, Victoria, Australia
[15]Deutsche Vereinigung Morbus Bechterew e.V. (DVMB), 97421 Schweinfurt, Germany
[16]UVSD SchmerzLOS e.V. (Unabhängige Vereinigung aktiver Schmerzpatienten in Deutschland), 24534 Neumünster, Germany

X Nitin Kumar Arora @nitintarika, Lars Donath @larsdonath, Patrick J Owen @PatrickOwenPhD, Clint T Miller @_clintmiller, Svenja Kaczorowski @svenjakacz, Tobias Saueressig @tobi_saueressig, Hugo Pedder @HugoPedder, Niamh L Mundell @NiamhMundell, Scott D Tagliaferri @ScottTags, Ashish Diwan @DrAshishDiwan and Daniel L Belavy @belavyprof

Contributors Please refer to supplemental data 8: CRediT author statement, Guarantor: DLB and NKA.

Funding This systematic review is funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung; grant number 01KG2313).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

ORCID iDs
Nitin Kumar Arora http://orcid.org/0000-0002-8841-2549
Lars Donath http://orcid.org/0000-0001-6039-0141
Patrick J Owen http://orcid.org/0000-0003-3924-9375
Clint T Miller http://orcid.org/0000-0001-7743-6986
Svenja Kaczorowski http://orcid.org/0000-0001-8794-3252
Tobias Saueressig http://orcid.org/0000-0002-3332-3856
Hugo Pedder http://orcid.org/0000-0002-7813-3749
Scott D Tagliaferri http://orcid.org/0000-0003-3669-4131
Ashish Diwan http://orcid.org/0000-0003-1037-8421
Xiaolong Chen http://orcid.org/0000-0003-2083-0549
Xiaohui Zhao http://orcid.org/0000-0002-0406-9342
Eva-Maria Huessler http://orcid.org/0000-0002-8144-5428
Katja Ehrenbrusthoff http://orcid.org/0000-0003-1667-8943
Jon J Ford http://orcid.org/0000-0001-5524-190X
Andrew J. Hahne http://orcid.org/0000-0002-6227-5339
Daniel L Belavy http://orcid.org/0000-0002-9307-832X

REFERENCES

1 GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the global burden of disease study 2021. *Lancet Rheumatol* 2023;5:e316–29.
2 Hoy D, Brooks P, Blyth F, *et al*. The epidemiology of low back pain. *Best Pract Res Clin Rheumatol* 2010;24:769–81.
3 Jones GT, Macfarlane GJ. Epidemiology of low back pain in children and adolescents. *Arch Dis Child* 2005;90:312–6.
4 Hoy D, Bain C, Williams G, *et al*. A systematic review of the global prevalence of low back pain. *Arthritis Rheum* 2012;64:2028–37.
5 Heuch I, Foss IS. Acute low back usually resolves quickly but persistent low back pain often persists. *J Physiother* 2013;59:127.
6 Pengel LHM, Herbert RD, Maher CG, *et al*. Acute low back pain: systematic review of its prognosis. *BMJ* 2003;327:323.
7 Belavy DL, Diwan AD, Ford J, *et al*. Network meta-analysis for comparative effectiveness of treatments for chronic low back pain disorders: systematic review protocol. *BMJ Open* 2021;11:e057112.
8 van Tulder M, Becker A, Bekkering T, *et al*. Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 2006;15:S169–91.
9 Oliveira CB, Maher CG, Pinto RZ, *et al*. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *Eur Spine J* 2018;27:2791–803.
10 Lin I, Wiles L, Waller R, *et al*. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: systematic review. *Br J Sports Med* 2020;54:79–86.
11 Bardin LD, King P, Maher CG. Diagnostic triage for low back pain: a practical approach for primary care. *Med J Aust* 2017;206:268–73.
12 Owen PJ, Miller CT, Mundell NL, *et al*. Which specific modes of exercise training are most effective for treating low back pain? Network meta-analysis. *Br J Sports Med* 2020;54:1279–87.
13 Arora NK, Donath L, Owen PJ, *et al*. The impact of exercise prescription variables on intervention outcomes in musculoskeletal pain. *Sports Med* 2024;54:711–25.
14 Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
15 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
16 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.
17 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
18 Dias S, Welton NJ, Sutton AJ, *et al*. *NICE DSU technical support document 4: inconsistency in networks of evidence based on randomised controlled trials*. London: National Institute for Health and Care Excellence (NICE), 2014.
19 Dias S, Welton NJ, Sutton AJ, *et al*. *NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials [Internet]*. London: National Institute for Health and Care Excellence (NICE), 2014.
20 Arora NK, Donath L, Owen PJ, *et al*. Dosage of exercise for chronic low back pain disorders (DOSE): a systematic review with dose-response network meta-analysis [Published Online First]. 2024.
21 Hayden JA, Ellis J, Ogilvie R, *et al*. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev* 2021;9:CD009790.
22 Geneen LJ, Moore RA, Clarke C, *et al*. Physical activity and exercise for chronic pain in adults: an overview of cochrane reviews. *Cochrane Database Syst Rev* 2017;1:CD011279.
23 Niemeijer A, Lund H, Stafne SN, *et al*. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. *Br J Sports Med* 2020;54:1073–80.
24 Morrison A, Polisena J, Husereau D, *et al*. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care* 2012;28:138–44.
25 Higgins JPT, Li T, Deeks JJ. Chapter 6: choosing effect measures and computing estimates of effect. In: *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022)*. Cochrane, 2022.
26 Vucic K, Jelicic Kadic A, Puljak L. Survey of cochrane protocols found methods for data extraction from figures not mentioned or unclear. *J Clin Epidemiol* 2015;68:1161–4.
27 Bidonde J, Busch AJ, Webber SC, *et al*. Aquatic exercise training for fibromyalgia. *Cochrane Database Syst Rev* 2014.
28 Slade SC, Dionne CE, Underwood M, *et al*. Consensus on exercise reporting template (CERT): explanation and elaboration statement. *Br J Sports Med* 2016;50:1428–37.
29 Pedder H. Mbnmadose: an R package for incorporating dose-response information into network meta-analysis. Evidence Synthesis and Meta-Analysis in R Conference 2021; 2021
30 Marinho VC, Higgins JP, Sheiham A, *et al*. Fluoride toothpastes for preventing dental caries in children and adolescents. *Cochrane Database Syst Rev* 2003.
31 Sterne JAC, Savović J, Page MJ, *et al*. Rob 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
32 R Core Team. R: A language and environment for statistical computing. 2023.
33 McGuinness LA, Higgins JPT. Risk-of-bias visualization (Robvis): an R package and shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods* 2021;12:55–61.
34 Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. 2003.
35 Stan Development Team. Stan modeling language users guide and reference manual, version 2.33.0. 2023. Available: stan-dev.github.io. //mc-stan.org/users/documentation/ [Accessed 1 Jan 2024].
36 Pedder H. Mbnmadose: dose-response model-based network meta-analysis. the comprehensive R archive network [Published Online First]. 2022.
37 Schwarzer G, Carpenter JR, Rücker G. Meta-analysis with R. In: *Meta-analysis with*. Cham: Springer International Publishing, 2015. Available: https://link.springer.com/10.1007/978-3-319-21416-0
38 PhillippoD. Multinma: Bayesian network meta-analysis of individual and aggregate data. 2023.
39 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36.
40 Daly C, Welton NJ, Dias S, *et al*. *Meta-analysis of continuous outcomes: guideline methodology document 2*. NICE Guidelines Technical Support Unit, 2021.
41 Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;53:1119–29.

42 Chou R, Qaseem A, Snow V, *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.
43 Murad MH, Wang Z, Chu H, *et al*. When continuous outcomes are measured using different scales: guide for meta-analysis and interpretation. *BMJ* 2019;364:k4817.
44 Mawdsley D, Bennetts M, Dias S, *et al*. Model-based network meta-analysis: a framework for evidence synthesis of clinical trial data. *CPT Pharmacometrics Syst Pharmacol* 2016;5:393–401.
45 Pedder H, Dias S, Boucher M, *et al*. Methods to assess evidence consistency in dose-response model based network meta-analysis. *Stat Med* 2022;41:625–44.
46 Pedder H, Dias S, Bennetts M, *et al*. Joining the dots: linking disconnected networks of evidence using dose-response model-based network meta-analysis. *Med Decis Making* 2021;41:194–208.
47 Welton NJ, Caldwell DM, Adamopoulos E, *et al*. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *Am J Epidemiol* 2009;169:1158–65.
48 Chaimani A, Caldwell DM, Li T, *et al*. Additional considerations are required when preparing a protocol for a systematic review with multiple interventions. *J Clin Epidemiol* 2017;83:65–74.
49 Spineli LM, Kalyvas C, Yepes-Nuñez JJ, *et al*. Low awareness of the transitivity assumption in complex networks of interventions: a systematic survey from 721 network meta-analyses. *BMC Med* 2024;22.
50 Salanti G, Higgins JPT, Ades AE, *et al*. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279–301.
51 Dias S, Welton NJ, Sutton AJ, *et al*. Evidence synthesis for decision making 4. *Med Decis Making* 2013;33:597–606.
52 Daly C, Downing BC, Welton NJ. A practical guide to inconsistency checks in Bayesian network meta-analysis. 2022.
53 Rhodes KM, Turner RM, Higgins JPT. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *J Clin Epidemiol* 2015;68:52–60.
54 Spiegelhalter DJ, Best NG, Carlin BP, *et al*. The deviance information criterion: 12 years on. *J R Stat Soc Series B Stat Methodol* 2014;76:485–93.
55 Higgins JP, Eldridge S, Li T, *et al*. Chapter 23: including variants on randomized trials. In: Higgins J, Thomas J, Chandler J, et al, eds. *Cochrane handbook for systematic reviews of interventions*. Cochrane, 2021.
56 Boonstra AM, Reneman MF, Waaksma BR, *et al*. Predictors of multidisciplinary treatment outcome in patients with chronic musculoskeletal pain. *Disabil Rehabil* 2015;37:1242–50.
57 Cecchi F, Pasquini G, Paperini A, *et al*. Predictors of response to exercise therapy for chronic low back pain: result of a prospective study with one year follow-up. *Eur J Phys Rehabil Med* 2014;50:143–51.
58 Donegan S, Williamson P, D'Alessandro U, *et al*. Assessing key assumptions of network meta-analysis: a review of methods: assessing assumptions of network meta-analysis. *Res Synth Methods* 2013;4:291–323.
59 Nikolakopoulou A, Higgins JPT, Papakonstantinou T, *et al*. Cinema: an approach for assessing confidence in the results of a network meta-analysis. *PLOS Med* 2020;17:e1003082.
60 Chaimani A, Salanti G. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. *Res Synth Methods* 2012;3:161–76.
61 Dias S, Welton NJ, Marinho VCC, *et al*. Estimation and adjustment of bias in randomized evidence by using mixed treatment comparison meta-analysis. *J R Stat Soc Ser A Stat Soc* 2010;173:613–29.
62 Salanti G, Del Giovane C, Chaimani A, *et al*. Evaluating the quality of evidence from a network meta-analysis. *PLoS ONE* 2014;9:e99682.
63 Khorami AK, Oliveira CB, Maher CG, *et al*. Recommendations for diagnosis and treatment of lumbosacral radicular pain: a systematic review of clinical practice guidelines. *J Clin Med* 2021;10:2482.