PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Relative effectiveness of different forms of exercises for treatment of chronic low back pain: protocol for a systematic review incorporating Bayesian network meta-analysis
AUTHORS	GAO, CHENGFEI; Chen, Guanghui; Yang, Hui; Hua, Zhen; Xu, Peng; Wong, Mansang; He, Chengqi

VERSION 1 - REVIEW

REVIEWER	Nefyn Williams University of Liverpool UK
REVIEW RETURNED	19-Sep-2018

GENERAL COMMENTS	General Comments This is a well-written good quality protocol for a network meta-analysis (NMA) for an important clinical problem. The justification for network meta-analysis is good. The authors have followed the guidance of the PRISMA protocol. However, I have some additional queries from the PRISMA NMA checklist. I have made some minor grammatical suggestions. Title I would prefer “Relative effectiveness of different forms of exercise for…” Methods Network geometry – please describe how you will provide a brief overview of the treatment network, including gaps and potential biases. Please describe how you will handle: a) Handling of multi-arm trials b) Selection of prior distributions in Bayesian analysis c) Assessment of model fit Assessment of inconsistency - Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network studied. Discussion Please comment on: a) The assumption of transitivity for NMA b) How Bayesian methods can be criticised for their perceived complexity and subjectivity due to the choice of a prior distribution c) The assumption of consistency between direct and indirect evidence d) Network geometry and considerations for bias e) The tendency for rankings to exaggerate small differences in relative effects, especially if they are based on limited information. Typos/grammar P3 line 16 pain relief and improvement in function or disability of all interventions P4 line 1 This will be the first systematic review to use network meta-analysis to investigate the relative effectiveness…

	P5 last line and p6 line 1 Network meta-analysis (NMA) allows for simultaneous consideration of the relative … P6 Objective line 10 …to compare the effectiveness… P6 line 11 …and investigate which form of exercise is best… P7 line 6 …passive movement exercises, as well as any combination of interventions, will be excluded. P7 line 8 …included in the current study. P7 Comparators line 10 and line 11 (defined as… P7 line 12 participants who did not participate in any form of organized exercise or physical activity, except their activities of daily living) will be included. P9 line 1 passing through this initial screening stage… P9 line 3 relevant articles after reviewing the full texts. The level of agreement… P9 Data extraction line 12 drop-out rate; P9 line 15 outcome data of interest P9 line 20 pain relief and improvement in function or disability for all interventions. Exercise interventions are usually considered to be most effective… P9 line 25 When change values… P9 line 26 the mean values and SDs post-intervention… P10 Statistical analysis line 11 For the purpose of this study, P10 line 21 relative effectiveness… P11 line 6 Nodes in network geometry will represent different forms of exercise, the lines between nodes will indicate… P11 line 8 The size of the nodes and the thickness of the lines will be detrmined by the sample size of the interventions and the number of included studies, P11 line 10 The probability of each… P11 line 16 (but will not be limited… P11 line 20 …will be explored, by excluding… P12 Discussion line 13 will be based on whatever…the included studies report. P12 line 18 NMA is a method… P12 line 23 meta-analysis, heterogeneity will exist among the… P12 line 29 identify sources of heterogeneity.

REVIEWER	Christopher Carroll University of Sheffield
REVIEW RETURNED	05-Oct-2018

GENERAL COMMENTS	In many ways, this is an excellent, high quality systematic review protocol. Some minor revisions are needed. The manuscript needs to be very carefully proof-read by someone with a very high standard of written English. There are multiple small errors of grammar and syntax. Just a few are highlighted below: p.2, l.28: a superfluous ‘the’: ‘the pain relief’. Delete ‘the’. p.4, l.8: reword: ‘creating a large economic and societal burden on society’, should be ‘creating a large economic and societal burden’. p.4, 11.12-13: reword: ‘will never full-recovery and develop into chronic low back pain’ should be ‘will never fully-recover and can develop into chronic low back pain’’

	p.4, l.14: reword: 'The mechanism of the pain' should be 'The mechanism of pain' etc. pp.10-11: A sentence or two should be added at the end of the 'Assessment of risk of bias' section (p.10, ll.15-16) to clarify exactly how the findings of this process are to be used in the analysis. The statistical analysis only makes mention of 'excluding' high risk of bias studies (p.11, l.36). p.11: The statistical analysis mentions 'excluding' high risk of bias studies (p.11, l.36). It is more conventional to subgroup studies by risk of bias rather than excluding them altogether. Are studies at 'Unclear' risk of bias therefore to be treated the same as those judged to be at 'Low' risk of bias? This needs to be justified. It is also unclear here whether this represents an overall assessment of 'risk of bias' (if so, how is this to be determined?), or whether these analyses will be conducted separately using the risk of bias findings for each domain. Risk of bias might also be included as a variable in the proposed meta-regression. p.21, ll.4-13: There are several published, validated RCT filters (e.g. for MEDLINE), one of which should be used here. If what has been applied is a published filter, its reference needs to be cited. If it is not, then a validated filter should be used, or justification provided for not doing so. It would have been interesting to know, from the scoping searches, how large the evidence base is likely to be, as this then informs the potential for the NMA.

REVIEWER	Pietro Ferrari IARC, France
REVIEW RETURNED	28-Nov-2018

GENERAL COMMENTS	Comment to the Editor This is an interesting description of the study protocol where several analytical steps are described. The question is whether BMJ Open is interested in publishing this sort of study descriptions. If this is the case, the current manuscript provides a very accurate description of the study design of an interesting research question, i.e. whether physical exercise is beneficial to control chronic low back pain. On the other hand no details are provided on the number of studies that will contribute to the final study, thus limiting to fully appreciate the impact of the evaluation. However, the protocol is very thorough. General comments: This is well written and comprehensive study design description of an evaluation of a meta-analysis or intervention studies implemented to evaluate the role of physical exercise on the chronic low back pain (LBP). The description of the study is exhaustive and meets high standards in terms of several analytical steps, including quality control checks, study selection, study evaluations, procedures to control for bias. The described protocol will certainly allow a complex evaluation to be completed in an accurate way. Details are provided on a classical meta-analysis that will be complemented by a Bayesian network meta-analysis,

	which will be followed by meta-regression to highlight factors that influence the relationship between physical exercise and cLBP. Detailed comments Title: It is not necessarily clear what the Author mean with the expression “relative efficacy” of different forms of physical activity. As the text is prone to subjective interpretations, possibly consider revision. In the description of the steps of the meta-analysis, it is stated that random effects will be used if results will be compatible the hypothesis of homogeneous results across studies. Rather, the Dersimonian and Laird approach for random effects is advocated in case of heterogeneity of the exposure/outcome relationship. Limited detailed are provided on the Bayseian network meta-analysis. Primarily, about the rationale of using a Network meta-analysis rather than ordinary analysis (this is only quickly mentioned). Second, no details on whether the Bayesian analysis will make use of informative or non-informative priors, and which distributional assumption these will involve.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Nefyn Williams

Institution and Country: University of Liverpool UK

Please state any competing interests or state ‘None declared’: None

General Comments

This is a well-written good quality protocol for a network meta-analysis (NMA) for an important clinical problem. The justification for network meta-analysis is good. The authors have followed the guidance of the PRISMA protocol. However, I have some additional queries from the PRISMA NMA checklist. I have made some minor grammatical suggestions.

Title

1. I would prefer “Relative effectiveness of different forms of exercise for…”

Response:

We agree with the reviewer, and the title has been revised to “Relative effectiveness of different forms of exercise for treatment of chronic low back pain: protocol for a systematic review incorporating Bayesian network meta-analysis”

Methods

2. Network geometry – please describe how you will provide a brief overview of the treatment network, including gaps and potential biases.

Response:

We add the following sentence to answer reviewer's comment:

"Qualitative description of network geometry will be provided by evaluating the diversity (number of treatments and how often they are tested) and co-occurrence (whether particular treatments and comparisons are preferred or avoided) of the treatment network.42" (page11, line19-25)

Reference:

Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. Ann Intern Med. 2008;148(7):544–53.

3. Please describe how you will handle:

a) Handling of multi-arm trials:

Response:

To answer reviewer's comment, the following sentences were added to page9, line9-16.

"For multi-arm studies, data from any arm that meets the inclusion criteria will be extracted. Multi-arm studies will be treated as multiple independent two-arm studies in pairwise meta-analyses. In the NMA, the correlations between effect sizes induced by multi-arm studies will be accounted for using a multivariate approach.19"

Reference:

Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool[J]. Research synthesis methods,2012,3(2):80-97. DOI:10.1002/jrsm.1037

b) Selection of prior distributions in Bayesian analysis

Response:

To answer reviewer's comment, the following sentences were added to page10, line25-29.

"As previous knowledge about exercise efficacy on cLBP is inconclusive, the non-informative priors for the treatment effect and between-study variance will be used in Bayesian analysis. Posterior distributions of the model parameters will be utilized to present the results of the NMA."

c) Assessment of model fit

Response:

To answer reviewer's comment, the following sentences were added to page10, line35-37.

"Model fitness will be assessed by comparing the posterior mean of the residual deviance with the number of unconstrained data points.13"

Reference:

Dias S, Welton NJ, Caldwell DM. Checking consistency in mixed treatment comparison meta-analysis[J].Statistics in medicine,2010,29(7-8):932-944.DOI:10.1002/sim.3767

4. Assessment of inconsistency - Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network studied.

Response:

To answer reviewer's comment, the following sentences were added to page10, line 37-43

"The inconsistency between the direct and indirect evidence will be assessed locally by using the node-splitting method,13 and globally by running the design-by-treatment interaction model.41"

Reference:

Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010;29(7-8):932–44. doi:10.1002/sim.3767

Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods. 2012;3(2):98–110. doi:10.1002/jrsm.1044.

Based on the above comments, we rewritten the section of "Bayesian analysis" as follows:

"A Bayesian framework using the Markov Chains Monte Carlo method will be conducted to compare the relative effectiveness of different exercise interventions simultaneously.38 The selection between fixed-effects and random-effects models will be based on the deviance information criterion (DIC) of each model. The model with the lower DIC will be selected (with difference >5 indicating a significant difference in fit).39 As previous knowledge on exercise efficacy is inconclusive, a non-informative prior will be used in Bayesian analysis. Posterior distributions of the model parameters will be utilized to present the results of the NMA. Four Markov chains will be run simultaneously with different arbitrarily chosen initial values. The first 5000 simulations will be discarded, and the posterior summaries will be based on 50000 simulations. Convergence of the simulation will be checked with the Gelman-Rubin-Brooks method.40 Model fitness will be assessed by comparing the posterior mean of the residual deviance with the number of unconstrained data points13. The inconsistency between the direct and indirect evidence will be assessed locally by using the node-splitting method,13 and globally by running the design-by-treatment interaction model.41" (page10, line16-43)

Discussion

5. Please comment on:

a) The assumption of transitivity for NMA

Response:

To answer reviewer's comment, the following sentences were added to page12, line24-34.

"The assumption of transitivity is the principle for network meta-analysis. In the case that transitivity is violated, the validity of indirect and mixed treatment effect estimates is compromised.19 The distribution of clinical and methodological variables that could act as effect modifiers across treatment comparisons will be evaluated to assure the similarity between the included studies are sufficiently comparable to allow for reliable data synthesis."

Reference:

Salanti G. Indirect and mixed-treatment comparison, network, or multiple treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods. 2012;3(2): 80–97. doi:10.1002/jrsm.1037.

b) How Bayesian methods can be criticized for their perceived complexity and subjectivity due to the choice of a prior distribution？

Response:

To answer reviewer's comment, the following sentences were added to page13, line28-41.

"Thus far, quite a few published NMA have applied a Bayesian method, which offered more flexibility for statistical modeling than traditional methods. However, Bayesian method has been criticized for its subjectivity introduced by the choice of a prior distribution. Different prior distributions can be used which can generate different results, and therefore, a sensitivity analysis is always required. There are also practical difficulties in implementation of Bayesian methodology due to its mathematical complexity. Further work is needed to resolve this difficulty, particularly when computing large hierarchical models with extremely large number of parameters."

c) The assumption of consistency between direct and indirect evidence

Response:

To answer reviewer's comment, the following sentences were added to page12, line33-39.

"Consistency is the extension of transitivity across a closed 'loop of evidence', where both direct and indirect evidence are available for a given treatment comparison. To check the assumption of consistency in the entire work, locally and globally will be applied."

d) Network geometry and considerations for bias

Response:

To answer reviewer's comment, the following sentences were added to page12, line53-page12, line6.

"The geometry of a network can provide the wider clinical context of the evidence, which can help to identify gaps of evidence in the treatment network. While the network structure may be shaped by various preference biases other than rational choices for treatment comparators. These biases may have important implications for the strength of interpretation of the evidence.42 Evaluation of network geometry will be performed in this study to seek out such biases."

Reference:

Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. Ann Intern Med. 2008;148(7):544–53.

e) The tendency for rankings to exaggerate small differences in relative effects, especially if they are based on limited information.

Response:

To answer reviewer's comment, the following sentences were added to page13, line10-23.

"One strength of NMA is that it can provide ranking information about all evaluated interventions for studied outcomes. However, it should avoid ranking treatments solely on the basis of the probability for each treatment of being the best. Because the probability of being the best does not account for the uncertainty in the relative treatment effects and can spuriously give higher ranks to treatment for which sparse data is available.44 To minimize potential biases, ranking information will be reported accompanied with effect sizes of pairwise comparison (such as means and 95% credible intervals) in this study."

Reference:

Salanti G, Ades AE, Ioannidis JP (2011) Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 64: 163–171.

We have re-written the discussion part based on the reviewer's above comments. (page12, line21-page13, line53)

6. Typos/grammar

P3 line 16 pain relief and improvement in function or disability of all interventions

P4 line 1 This will be the first systematic review to use network meta-analysis to investigate the relative effectiveness…

P5 last line and p6 line 1 Network meta-analysis (NMA) allows for simultaneous consideration of the relative …

P6 Objective line 10 …to compare the effectiveness…

P6 line 11 …and investigate which form of exercise is best…

P7 line 6 …passive movement exercises, as well as any combination of interventions, will be excluded.

P7 line 8 …included in the current study.

P7 Comparators line 10 and line 11 (defined as…

P7 line 12 participants who did not participate in any form of organized exercise or physical activity, except their activities of daily living) will be included.

P9 line 1 passing through this initial screening stage…

P9 line 3 relevant articles after reviewing the full texts. The level of agreement…

P9 Data extraction line 12 drop-out rate;

P9 line 15 outcome data of interest

P9 line 20 pain relief and improvement in function or disability for all interventions. Exercise interventions are usually considered to be most effective…

P9 line 25 When change values…

P9 line 26 the mean values and SDs post-intervention…

P10 Statistical analysis line 11 For the purpose of this study,

P10 line 21 relative effectiveness…

P11 line 6 Nodes in network geometry will represent different forms of exercise, the lines between nodes will indicate…

P11 line 8 The size of the nodes and the thickness of the lines will be determined by the

sample size of the interventions and the number of included studies,

P11 line 10 The probability of each…

P11 line 16 (but will not be limited…

P11 line 20 …will be explored, by excluding…

P12 Discussion line 13 will be based on whatever…the included studies report.

P12 line 18 NMA is a method…

P12 line 23 meta-analysis, heterogeneity will exist among the…

P12 line 29 identify sources of heterogeneity.

Response:

The above typos and grammar errors raised by the reviewer were corrected accordingly, and the English language of whole manuscript has been checked and improved by a native English speaker.

Reviewer: 2

Reviewer Name: Christopher Carroll

Institution and Country: University of Sheffield

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

In many ways, this is an excellent, high quality systematic review protocol.

1. Some minor revisions are needed.

The manuscript needs to be very carefully proof-read by someone with a very high standard of written English. There are multiple small errors of grammar and syntax. Just a few are highlighted below:

p.2, l.28: a superfluous 'the': 'the pain relief'. Delete 'the'.

p.4, l.8: reword: 'creating a large economic and societal burden on society', should be 'creating a large economic and societal burden'.

p.4, 11.12-13: reword: 'will never full-recovery and develop into chronic low back pain' should be 'will never fully-recover and can develop into chronic low back pain''

p.4, l.14: reword: 'The mechanism of the pain' should be 'The mechanism of pain' etc.

Response:

The grammar/syntax errors raised by the reviewer were corrected accordingly, and the English language of whole manuscript has been checked and improved by a native English speaker.

2. pp.10-11: A sentence or two should be added at the end of the 'Assessment of risk of bias' section (p.10, ll.15-16) to clarify exactly how the findings of this process are to be used in the analysis. The statistical analysis only makes mention of 'excluding' high risk of bias studies (p.11, l.36).

Response:

As suggested by the reviewer, the following sentence was added to page9, line39-41.

"Results from these appraisals will be considered as criteria for the subsequent meta-regression and subgroup analysis."

3. p.11: The statistical analysis mentions 'excluding' high risk of bias studies (p.11, l.36). It is more conventional to subgroup studies by risk of bias rather than excluding them altogether. Are studies at 'Unclear' risk of bias therefore to be treated the same as those judged to be at 'Low' risk of bias? This needs to be justified. It is also unclear here whether this represents an overall assessment of 'risk of bias' (if so, how is this to be determined?), or whether these analyses will be conducted separately using the risk of bias findings for each domain. Risk of bias might also be included as a variable in the proposed meta-regression.

Response:

We will first assess the risk of bias for each domain of individual studies. After that, studies will be classified according to the following categories: studies with "low risk of bias" in all of the above domains will be considered as "studies with low risk of bias", studies with "uncertain risk of bias" or "high risk of bias" in one or more of the above domains will be considered as "studies with high risk of bias". (page9, line32-41)

We have made correction according to the reviewer's comments.

Subgroup analysis and meta-regression based on the risk of bias will be conducted to evaluate the effect of this moderator on the results of NMA.

4. p.21, ll.4-13: There are several published, validated RCT filters (e.g. for MEDLINE), one of which should be used here. If what has been applied is a published filter, its reference needs to be cited. If it is not, then a validated filter should be used, or justification provided for not doing so.

Response:

The strategies for limiting searches to randomized controlled trials of this study were made based on Cochrane Handbook for Systematic Reviews of Interventions. The reference has been added. (see Supplementary file 2) (page43, line14-21)

5. It would have been interesting to know, from the scoping searches, how large the evidence base is likely to be, as this then informs the potential for the NMA.

Response:

As suggested by the reviewer, the search results of each item were list in Supplementary file 2. (page41-43)

Reviewer: 3

Reviewer Name: Pietro Ferrari

Institution and Country: IARC, France

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

General comments:

This is well written and comprehensive study design description of an evaluation of a meta-analysis or intervention studies implemented to evaluate the role of physical exercise on the chronic low back pain (LBP). The description of the study is exhaustive and meets high standards in terms of several analytical steps, including quality control checks, study selection, study evaluations, procedures to control for bias. The described protocol will certainly allow a complex evaluation to be completed in an accurate way. Details are provided on a classical meta-analysis that will be complemented by a Bayesian network meta-analysis, which will be followed by meta-regression to highlight factors that influence the relationship between physical exercise and cLBP.

Detailed comments

1. Title: It is not necessarily clear what the Author mean with the expression “relative efficacy” of different forms of physical activity. As the text is prone to subjective interpretations, possibly consider revision.

Response:

We agree with the reviewer, and the title has been revised to “Relative effectiveness of different forms of exercise for treatment of chronic low back pain: protocol for a systematic review incorporating Bayesian network meta-analysis”

2. In the description of the steps of the meta-analysis, it is stated that random effects will be used if results will be compatible the hypothesis of homogeneous results across studies. Rather, the Dersimonian and Laird approach for random effects is advocated in case of heterogeneity of the exposure/outcome relationship.

Response:

This section was rewritten as follows:

“While the I2 >50% will be considered to represent significant heterogeneity, the random effects model (DerSimonian and Laird method) will be used for pooling the results.37 The source of heterogeneity will be explored using meta-regression and sensitivity analyses” (page10, line9-14)

Reference:

Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. Bmj 2003;327(7414):557-60.

3. Limited detailed are provided on the Bayesian network meta-analysis. Primarily, about the rationale of using a Network meta-analysis rather than ordinary analysis (this is only quickly mentioned). Second, no details on whether the Bayesian analysis will make use of informative or non-informative priors, and which distributional assumption these will involve.

Response:

Additional rational of using a NMA rather than traditional pairwise meta-analyses has been added as follows:

“Moreover, both indirect and direct evidence can be used together in NMA, which can acquire a higher degree of precision in the estimation of effectiveness of different exercises compared with pairwise meta-analyses.” (page5, line7-13)

The selection of prior distribution has been added as follows:

“As previous knowledge on exercise efficacy is inconclusive, a non-informative prior will be used in Bayesian analysis. Posterior distributions of the model parameters will be utilized to present the results of the NMA.” (page10, line25-29)

VERSION 2 – REVIEW

REVIEWER	Nefyn Williams University of Liverpool United Kingdom
REVIEW RETURNED	08-Jan-2019

GENERAL COMMENTS	All of my comments have been addressed