BMJ Open

Management of acute musculoskeletal pain (excluding low back pain): protocol for a systematic review and network meta-analysis of randomised trials

Jason W Busse,[1,2,3] Samantha Craigie,[3] Behnam Sadeghirad,[2,3] Rachel Couban,[2] Patrick Hong,[4] Yvgeniy Oparin,[5] Curtis May,[6] Annie Lok,[2] Gordon H Guyatt[3]

ABSTRACT

Introduction Acute, non-low back-related musculoskeletal pain is common and associated with significant socioeconomic costs. No review has evaluated all interventional studies for acute musculoskeletal pain, which limits attempts to make inferences regarding the relative effectiveness of treatments.

Methods and analysis We will conduct a systematic review of all randomised controlled trials evaluating therapies for acute musculoskeletal pain (excluding low back pain). We will identify eligible, English-language, trials by a systematic search of the Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, Medline, Physiotherapy Evidence Database (PEDro) and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to February 2018. Eligible trials will: (1) enrol patients presenting with acute, non-low back-related musculoskeletal pain (duration of pain ≤4 weeks), and (2) randomise patients to alternative interventions or an intervention and a placebo/sham arm. Fractures will be considered ineligible, unless they are non-surgical and therapy is directed at pain relief. Pairs of reviewers will, independently and in duplicate, screen titles and abstracts of identified citations, review the full texts of potentially eligible trials and extract information from eligible trials. We will use a modified Cochrane instrument to evaluate risk of bias. Disagreements will be resolved through discussion to achieve consensus. We will use the Grading of Recommendations Assessment, Development and Evaluation approach to evaluate the quality of evidence supporting treatment effects. When possible, we will conduct: (1) in direct comparisons, a random-effect meta-analysis to establish the effectiveness of therapeutic interventions on patient-important outcomes; and (2) multiple treatment comparison meta-analysis to assess the relative effects of treatments. We will use a priori hypotheses to explain heterogeneity between studies. We will use STATA V.14.2 for all analyses.

Ethics and dissemination No research ethics approval is required for this systematic review, as no confidential patient data will be used. The results of this systematic review will be disseminated through publication in a peer-reviewed journal, conference presentations and will inform a clinical practice guideline.

PROSPERO registration number CRD42018094412.

Strengths and limitations of this study

► Our broad study eligibility criteria will increase generalisability of our results.
► We will use the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to evaluate our confidence in treatment effects.
► We will optimise interpretability by presenting risk differences and measures of relative effect for all outcomes reported, and by presenting our findings with GRADE evidence profiles.
► Findings from our review will inform a clinical practice guideline and identify key areas for future research.
► Our results will be limited by possible shortcomings of the primary studies.

INTRODUCTION

Acute non-low back-related musculoskeletal (MSK) pain, which include strains and sprains, dislocations and whiplash present for less than 1 month,[1] is associated with considerable morbidity in North America. *The Burden of Musculoskeletal Diseases in the United States*, third edition, found that sprains and strains are the most frequent injury type for which medical care is sought and the majority of MSK injuries occur among working-age adults.[2] In 2013, there were 2 807 880 emergency department visits for sports-related injuries in the USA,[3] and over 70% of visits to the emergency department are because of pain-related problems.[4] Management often yields suboptimal outcomes: a survey of 842 patients with acute pain at 20 USA and Canadian hospitals found that 40% reported their pain did not change or increased after

visiting the emergency department, and 74% of patients were discharged in moderate-to-severe pain.[5]

Despite the prevalence of acute MSK problems in primary care, a number of studies have highlighted inadequate pain education among physicians.[6–8] The availability of numerous pharmacological and non-pharmacological therapies further complicates management decisions. Currently available treatments include opioids, non-steroidal anti-inflammatory drugs, muscle relaxants, acetaminophen, exercise, supervised rehabilitation, joint manipulation and mobilisation, massage, acupuncture and acupressure, ultrasound, low-level laser therapy and transcutaneous electrical nerve stimulation.[9–12]

A recently published series of 24 systematic reviews of treatment for common acute MSK injuries[13] suffers from many important limitations, including: (1) selection bias due, in part, to exclusion of studies deemed to be at high risk of bias without empirically exploring if effect estimates differed from studies assigned low risk of bias, (2) interpreting results from individual studies as effective if the mean difference in a given outcome met the minimally important difference (MID) and ineffective if not—an interpretation that relies on the unlikely assumption that all patients will experience the same degree of improvement, and fails to consider the distribution around the mean and the proportion of patients that achieve the MID, (3) no appraisal of the overall quality (confidence, certainty) of the evidence, and (4) no statistical pooling of treatment effects. We propose to conduct a systematic review of randomised trials to assess the comparative effectiveness of available non-surgical treatments for acute MSK pain (excluding low back pain) and assess quality of evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology.

METHODS

Protocol registration
This review was commissioned by the American Academy of Family Physicians and American College of Physicians and sponsored by the National Safety Council.

Standardised reporting
We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols checklist when writing our report.[14]

Information sources
We will identify eligible, English language, randomised clinical trials (RCTs) through a systematic search of Medline, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Physiotherapy Evidence Database (PEDro) and Cochrane Central Register of Controlled Trials (CENTRAL), from inception of each database through February 2018. Our search has been refined for individual databases by an experienced medical librarian (see online supplementary appendix).

We will scan reference lists of included studies and relevant systematic reviews for additional eligible articles.

Eligibility criteria and study selection
We will include therapeutic trials that: (1) enrol adult patients (≥18 years) presenting with acute, non-low back-related MSK pain (pain with duration <4 weeks or defined by authors as 'acute') in an outpatient setting, and (2) randomise them to currently available, non-surgical, alternative interventions directed at pain relief (pharmacological or non-pharmacological) or a currently available, non-surgical, intervention directed at pain relief and a placebo/sham arm. Eligible MSK conditions include sprains and strains. Surgical fractures, non-surgical fractures unless therapy is directed at pain relief, low back pain, neuropathic pain, acute cancer pain, acute postoperative pain, acute dental pain, pain associated with labour and delivery, visceral pain, pain due to infection and headaches will be excluded. We will exclude interventions targeted at the treatment of acute low back pain on request by the study funder, as they have previously commissioned evidence syntheses on this topic.[9 10]

Ten teams of trained reviewers will work independently in pairs to screen titles and abstracts of identified citations, using standardised, pilot-tested forms in DistillerSR, an online systematic review software (Evidence Partners, Ottawa, Canada; http://systematic-review.net/). The same teams of reviewers will screen full texts of any articles judged as potentially eligible. Reviewers will discuss disagreements to come to consensus, referring to an adjudicator if necessary. We will measure agreement between reviewers by calculating kappa (κ) values to assess the reliability of full-text review, and interpret them using the following thresholds: <0.20 as slight agreement, 0.21–0.40 as fair agreement, 0.41–0.60 as moderate agreement, 0.61–0.80 as substantial agreement and >0.80 as almost perfect agreement.[15]

Data abstraction
We designed standardised data abstraction forms and a detailed instruction video (accessible at: https://www.youtube.com/watch?v=1nwFJ61K3sQ). We will conduct calibration exercises prior to beginning data abstraction to ensure consistency and accuracy of extractions. Seven teams of reviewers will extract data independently and in duplicate. We will extract the following data from eligible trials into a standardised spreadsheet: study characteristics (eg, the first author, publication year, country of origin and funding source), participant and trial characteristics (eg, sample size, mean age of participants, clinical condition, type and severity of injury, proportion with known chronic pain/condition prior to acute injury, and the proportion receiving opioids at the time of enrolment), characteristics of interventions and comparators, patient-important outcomes (pain, function, health-related quality of life, patient satisfaction, return to work, proportion of patients with relief, reinjury and all reported adverse events). We will extract pain at any

time point, whereas for other outcomes, we will use the longest follow-up reported.[16]

Risk of bias assessment

Among eligible studies, we will independently assess the following risk of bias issues: (1) random sequence generation, (2) allocation concealment, (3) blinding of study participants, personnel and outcome assessors, (4) incomplete outcome data (≥20% missing data will be considered at high risk of bias) and (5) other sources of bias.[17] To assess the risk of bias, we will use a modified version of the Cochrane risk of bias instrument.[18] Our instrument will use the following responses: 'definitely yes' or 'probably yes' (considered as low risk of bias), or 'definitely no' or 'probably no' (considered as high risk of bias). These response options have published evidence of validity for assessing blinding.[18] Any discrepancy in the assessment of risk of bias will be resolved by discussion, or third party adjudication if needed. We will contact authors for missing information regarding risk of bias issues and unpublished data (eg, effect estimates without accompanying estimates of precision).

Data synthesis

MSK problems are increasingly being considered together as risk factors,[19] prognosis,[20] and treatments are often similar in guideline recommendations.[21 22] For the purposes of statistical pooling, we will explore treatment effects of interventions across all MSK problems eligible for this review; however, we will also explore if treatment effects differ by clinical condition or injury severity. Clustering strategies for clinical condition and injury severity will be informed by the trials eligible for our review, which will be reviewed by a technical expert panel, blinded to study results. Treatment effects will be pooled using the longest follow-up time reported, except for pain, which will be pooled at the most commonly reported short-term, medium-term and long-term follow-up times reported by trials eligible for our review. For our review, these categories will be 30–120 min post treatment (short), 1–7 days post treatment (medium) and 3–12 weeks post treatment (long). As such, a single trial could contribute to up to three time points for our pooled results for pain relief. Alternately, trials that reported pain relief at time points outside of these timeframes would not contribute data for our analyses of pain relief.

Methods for direct comparisons

We will pool dichotomous outcomes that are reported by >1 RCT and calculate the relative risk and risk difference (RD), using baseline risk estimates from the placebo arm of eligible studies, and the associated 95% CIs. We will pool all continuous outcomes that are reported by >1 study and calculate the weighted mean differences (WMDs) and associated 95% CIs. We will employ methods described in Cochrane Handbook to estimate the mean and SD when median, range and sample size are reported, and to impute the SD if the SE or SD for the differences are

not reported.[23] For continuous outcomes, when studies report effect estimates using different measurement instruments that tap into a common construct (eg, pain), we will first transform all outcomes to a common instrument on a domain-by-domain basis.[24] We will use change scores from baseline to end of follow-up rather than end-of-study scores, in order to account for interpatient variability. If authors do not report change scores, we will calculate them using the baseline and end-of-study score and a correlation coefficient derived from the largest trial at lowest risk of bias in the meta-analysis that does report a change score. We will use DerSimonian-Laird random-effect models[25] for all pairwise comparisons.

Interpreting effect estimates for continuous outcomes is challenging,[26] and we will present the MID for all pooled effect estimates. The MID is the smallest amount of improvement in a treatment outcome that a patient would recognise as important.[27] If we find multiple MID estimates are available, we will use the smallest difference that has been validated.

However, simply presenting the MID risks interpreting all mean effects that fall below the MID as unimportant, whereas we have found that, typically, mean differences of 1/2 the MID result in RDs of about 10%—a potential benefit that many patients may consider important.[28] Thus, concluding that an effect is unimportant requires confidence that the mean difference is less than 1/2 the MID (and perhaps less). To optimise interpretability, we will calculate the RD of achieving the MID for all statistically significant WMDs. Specifically, for each individual study, we will assume that the SDs of outcome measurements are the same in both the treatment and control groups, and that change scores in both groups are normally distributed. We will use the median or mean, and SD of the control group, with the established MID for the outcome in question to estimate the probability of achieving ≥MID in the control group. We will use the pooled mean difference to estimate the mean in the treatment group and calculate the probability of achieving ≥MID in the treatment group. Finally, we will use risks in both groups to acquire the RD for achieving ≥MID.

Methods for network meta-analysis

We will perform network meta-analysis to synthesise the available evidence from the entire network of trials by integrating direct and indirect estimates for each comparison into a single summary treatment effect. We will use a Frequentist random-effect model using the methodology of multivariable meta-analysis to assess the comparative effectiveness of eligible interventions.[29 30]

Although the assumptions for network meta-analysis are similar to conventional meta-analysis, key extra assumptions are transitivity (there are no effect modifiers influencing the indirect comparisons) and coherence (direct and indirect effect estimates being similar).[31] We will identify issues of incoherence by comparing direct evidence (ie, estimates from pairwise comparisons) with indirect evidence (ie, estimates from network meta-analysis) using

the node-splitting method.[32 33] In this approach, incoherence is assessed locally by evaluating the consistency assumption in each closed loop of the network separately as the difference between direct and indirect estimates for a specific comparison in the loop. We will assume a common heterogeneity estimate within each loop. We will also confirm the coherence assumption in the entire network using a 'design-by-treatment' model.[34] In case we find significant incoherence in the network (highly significant p value from the design-by-treatment model), we will perform network meta-analysis using an inconsistency model. If using an inconsistency model results in non-sensical results, we will explore the network for the source(s) of incoherence and further expand (disintegrating interventions based on differences in population or intervention characteristics) or exclude the node(s) introducing incoherence into the network (eg, excluding node(s) with less than 20 events for binary outcomes or comparisons with only one trial with very few participants for continuous outcomes).

We will report our findings with probability statements of intervention effects. Probability rankings allow us to report a chance percentage of which interventions rank higher[35]; however, simplifying the results of a network down to probabilities can lead to misinterpretations, specifically, when particular comparisons (ie, nodes) are not well connected or when the quality of evidence varies between comparisons.[36 37] Following display of the rank probabilities using rankogram, we will use the surface under the cumulative ranking (SUCRA) line to aid in the interpretation of relative effect of the interventions. An intervention with a SUCRA value of 100 is certain to be the most effective, whereas an intervention with 0 is certain to be the least effective.[35] We will use STATA (StataCorp, Release V.14.2, College Station, Texas, USA) for all statistical analyses. All comparisons will be two tailed using a threshold p≤0.05.

Subgroup analyses, meta-regression and sensitivity analysis

We will use Cochran's Q statistic and I^2 to determine statistical heterogeneity for direct meta-analysis.[32] We have developed five hypotheses to explain heterogeneity between trials: (1) different clinical conditions will show different treatment effects; (2) more severe injuries will show smaller treatment effects than less severe injuries (eg, higher grades of strains and sprains vs lower grades); (3) older patients will show smaller treatment effects than younger patients; (4) longer follow-up will show smaller treatment effects than shorter follow-up; and (5) higher dose/intensity of treatment will show larger treatment effects. We will perform subgroup analyses regardless of heterogeneity estimates, if there are at least two trials in each subgroup. Moreover, we will explore the effect of risk of bias on treatment effects.

Assessing quality of the evidence

We will use the GRADE approach to assess the quality of evidence on an outcome-by-outcome basis. The starting point for quality of evidence for RCTs is high, but may be rated down based on the limitations in risk of bias, imprecision, inconsistency and indirectness and publication bias.[38] When there are at least 10 studies for meta-analysis,[39 40] we will assess publication bias by visual assessment of asymmetry of the funnel plot and calculation of Begg's test.

We will also use the GRADE approach to assess quality of evidence for indirect and network (mixed) effect estimates.[41 42] Indirect effect estimates are calculated from available 'loops' of evidence, which includes first-order loops (based on a single common comparator treatment, the difference between the treatments A and B is based on comparisons of A and C as well as B and C) or higher order loops (more than one intervening treatment connecting the two interventions). We will visually examine the network map and where first-order loops are available for indirect comparisons, the quality of evidence will be the lower of the ratings for the two direct estimates contributing to the first-order loop. In the absence of a first-order loop, a higher order loop will be used to rate the quality of evidence. We may rate down quality of evidence further for intransitivity.[42] The transitivity assumption implies similarity of trials in terms of population, intervention, settings and trial methodology.[43]

It is very rare for a network meta-analysis to establish a single treatment option as clearly superior to all others. Thus, we will categorise interventions according to three categories: (1) those that are clearly superior, (2) those with intermediate effectiveness and (3) those that are inferior. Treatments no better than placebo will be in the lowest tier, those better than placebo in tier 1 (likely intermediate); those superior to at least 1 tier 1 treatment will be judged superior. Treatments will be further categorised according to quality of evidence supporting those estimates (high and moderate vs low or very low). Interventions with moderate-quality or high-quality evidence will be ranked as either 'among the most effective', 'inferior to the most effective/superior to the least effective' or 'among the least effective'. Interventions supported by low or very low quality evidence will be ranked into the same three categories but prefaced with 'may be' to acknowledge the reduced confidence in supporting evidence (eg, 'may be among the most effective') and will be presented separately from those supported by moderate-quality or high-quality evidence.

Patient and public involvement

No patients were involved in setting the research question, in developing plans for design, interpretation, reporting or implementation of the study. We plan to disseminate the results of this study to organisations supporting patients with acute MSK pain.

DISCUSSION

With the high prevalence of acute non-low back MSK pain, the associated socioeconomic burden and the paucity of evidence on the comparative effectiveness of treatment options, there is an urgent need for a high-quality systematic

review to inform evidence-based management of these problems.

Our proposed review has several strengths in relation to existing reviews. First, we will explore all currently available non-pharmacological and pharmacological treatment options for all acute MSK problems (excluding low back pain) reported among eligible trials. It is plausible that individual acute MSK problems respond similarly to similar interventions, and thus by pooling across individual injuries, it may be possible to provide a more precise estimate of treatment effect. Second, we will update the search to present date. Third, we will use the GRADE approach to evaluate the quality of evidence supporting treatment effects. Fourth, we will ensure interpretability by presenting RDs and measures of relative effect for all outcomes reported, and by presenting our findings with GRADE evidence profiles.

A potential limitation will be the nature of available treatment comparisons to build robust networks for our analyses. The findings of our review will help inform patients with acute non-low back-related MSK pain about their therapeutic options, identify key areas for research and facilitate a clinical practice guideline for the management of acute, non-low back-related MSK pain.

Author affiliations
[1]Department of Anesthesia, McMaster University, Hamilton, Ontario, Canada
[2]Michael G. DeGroote Institute for Pain Research and Care, McMaster University, Hamilton, Ontario, Canada
[3]Department of Health Research Methods, Evidence, and Impact (HEI), McMaster University, Hamilton, Ontario, Canada
[4]Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[5]Department of Medicine, McMaster University, Hamilton, Ontario, Canada
[6]Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

REFERENCES

1. Olsen MF, Bjerre E, Hansen MD, *et al*. Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. *BMC Med* 2017;15:35.
2. United States Bone and Joint Initiative. The Burden of Musculoskeletal Diseases in the United States (BMUS) Rosemont, IL. Third Edition, 2014. Available: http://www.boneandjointburden.org (Accessed 3 Mar 2017).
3. Agency for Healthcare Research and Quality (AHRQ), Center for Delivery Organization and Markets Healthcare Cost and Utilization Project (HCUP). National Inpatient Sample (NIS) and Nationwide Emergency Department Sample(NEDS). 2013.
4. Todd KH, Miner JR. Pain in the emergency room. In: Fishman SM, Ballantyne JC, Rathmell JP, eds. *Bonica's Management of Pain*. Philadelphia, PA, USA: Lippincott, Williams and Wilkins, 2010:1576–87.
5. Todd KH, Ducharme J, Choiniere M, *et al*. Pain in the emergency department: results of the pain and emergency medicine initiative (PEMI) multicenter study. *J Pain* 2007;8:460–6.
6. Vadivelu N, Mitra S, Hines R, *et al*. Acute pain in undergraduate medical education: an unfinished chapter!. *Pain Pract* 2012;12:663–71.
7. Fishman SM, Young HM, Lucas Arwood E, *et al*. Core competencies for pain management: results of an interprofessional consensus summit. *Pain Med* 2013;14:971–81.
8. Watt-Watson J, Murinson BB. Current challenges in pain education. *Pain Manag* 2013;3:351–7.
9. Chou R, Deyo R, Friedly J, *et al*. Nonpharmacologic Therapies for Low Back Pain: A Systematic Review for an American College of Physicians Clinical Practice Guideline. *Ann Intern Med* 2017;166:493–505.
10. Chou R, Deyo R, Friedly J, *et al*. Systemic Pharmacologic Therapies for Low Back Pain: A Systematic Review for an American College of Physicians Clinical Practice Guideline. *Ann Intern Med* 2017;166:480–92.
11. Qaseem A, Wilt TJ, McLean RM, *et al*. Clinical Guidelines Committee of the American College of Physicians. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline from the American College of Physicians. *Ann Intern Med* 2017;166:514–30.
12. Motov S, Strayer R, Hayes BD, *et al*. The treatment of acute pain in the emergency department: a white paper position statement prepared for the american academy of emergency medicine. *J Emerg Med* 2018;54:731–6.
13. Côté P, Shearer H, Ameis A, *et al*. *Enabling recovery from common traffic injuries: a focus on the injured person*: UOIT-CMCC Centre for the Study of Disability Prevention and Rehabilitation, 2015.
14. 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.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
16. Tendal B, Nüesch E, Higgins JP, *et al*. Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. *BMJ* 2011;343:d4829.
17. Higgins JP, Altman DG, Gøtzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
18. Higgins JPT, Sterne JAC, Savović J, *et al*. A revised tool for assessing risk of bias in randomized trials In: Chandler J, McKenzie J, Boutron I, Welch V (editors). Cochrane Methods. *Cochrane Database Syst Rev* 2016;10.
19. Taanila H, Suni JH, Kannus P, *et al*. Risk factors of acute and overuse musculoskeletal injuries among young conscripts: a population-based cohort study. *BMC Musculoskelet Disord* 2015;16:104.
20. de Vos Andersen NB, Kent P, Hjort J, *et al*. Clinical course and prognosis of musculoskeletal pain in patients referred for physiotherapy: does pain site matter? *BMC Musculoskelet Disord* 2017;18:130.
21. Australian Acute Musculoskeletal Pain Guidelines Group. Evidence-based management of acute musculoskeletal pain. 2003:ISBN1875378499.
22. Babatunde OO, Jordan JL, Van der Windt DA, *et al*. Effective treatment options for musculoskeletal pain in primary care: a systematic overview of current evidence. *PLoS One* 2017;12:e0178621.
23. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions, version [5.1.0] (updated March 2011)*: The Cochrane Collaboration, 2011.

24. Thorlund K, Walter SD, Johnston BC, *et al*. Pooling health-related quality of life outcomes in meta-analysis-a tutorial and review of methods for enhancing interpretability. *Res Synth Methods* 2011;2:188–203.
25. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
26. Busse JW, Bartlett SJ, Dougados M, *et al*. Optimal strategies for reporting pain in clinical trials and systematic reviews: recommendations from an OMERACT 12 Workshop. *J Rheumatol* 2015;42:1962–70.
27. Schünemann HJ, Guyatt GH. Commentary–goodbye M(C)ID! Hello MID, where do you come from? *Health Serv Res* 2005;40:593–7.
28. Busse JW, Wang L, Kamaleldin M, *et al*. Opioids for chronic noncancer pain: a systematic review and meta-analysis. *JAMA* 2018;320:2448–60.
29. White IR. Network Meta-analysis. *Stata J* 2015;15:951–85.
30. White IR, Barrett JK, Jackson D, *et al*. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods* 2012;3:111–25.
31. Donegan S, Williamson P, D'Alessandro U, *et al*. Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods* 2013;4:291–323.
32. Higgins JP, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
33. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006;101:447–59.
34. Higgins JP, Jackson D, Barrett JK, *et al*. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012;3:98–110.
35. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64:163–71.
36. Mbuagbaw L, Rochwerg B, Jaeschke R, *et al*. Approaches to interpreting and choosing the best treatments in network meta-analyses. *Syst Rev* 2017;6:79.
37. Trinquart L, Attiche N, Bafeta A, *et al*. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. *Ann Intern Med* 2016;164:666–73.
38. Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
39. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011. Available: http://handbook-5-1.cochrane.org/ (Accessed 25 Oct 2018).
40. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.
41. Puhan MA, Schunemann HJ, Murad MH, *et al*. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014;349:g5630.
42. Brignardello-Petersen R, Bonner A, Alexander PE, *et al*. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *J Clin Epidemiol* 2018;93:36–44.
43. Salanti G, Higgins JP, Ades AE, *et al*. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279–301.