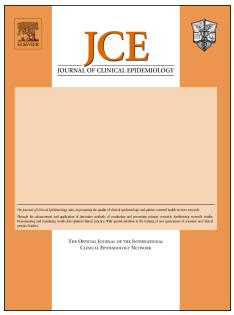
Accepted Manuscript



Additional considerations are required when preparing a protocol for a systematic review with multiple interventions

Anna Chaimani, Deborah M. Caldwell, Tianjing Li, Julian PT. Higgins, Georgia Salanti

PII: S0895-4356(16)30775-2

DOI: 10.1016/j.jclinepi.2016.11.015

Reference: JCE 9284

To appear in: *Journal of Clinical Epidemiology*

Received Date: 1 April 2016

Revised Date: 21 November 2016

Accepted Date: 30 November 2016

Please cite this article as: Chaimani A, Caldwell DM, Li T, Higgins JP, Salanti G, Additional considerations are required when preparing a protocol for a systematic review with multiple interventions, *Journal of Clinical Epidemiology* (2017), doi: 10.1016/j.jclinepi.2016.11.015.

Additional considerations are required when preparing a protocol for a systematic review with multiple interventions

Anna Chaimani[a]*, Deborah M Caldwell[b], Tianjing Li[c], Julian PT Higgins[b], Georgia Salanti[d,e,a]

[a] Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

[b] School of Social and Community Medicine, University of Bristol, Bristol, UK

[c] Center for Clinical Trials and Evidence Synthesis, Department for Epidemiology, John Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[d] Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

[e] Bern Institute of Primary Care (BIHAM), University of Bern, Bern, Switzerland

*Corresponding author

Email: achaiman@cc.uoi.gr

Department of Hygiene and Epidemiology

University of Ioannina School of Medicine

University Campus

Ioannina 45110, Greece

Tel: +30 26510 07844

ABSTRACT

Objectives: The number of systematic reviews that aim to compare multiple interventions using network meta-analysis is increasing. In this paper, we highlight aspects of a standard systematic review protocol that may need modification when multiple interventions are to be compared.

Study Design and setting: We take the protocol format suggested by Cochrane for a standard systematic review as our reference, and compare the considerations for a pairwise review with those required for a valid comparison of multiple interventions. We suggest new sections for protocols of systematic reviews including network meta-analyses with a focus on how to evaluate their assumptions. We provide example text from published protocols to exemplify the considerations.

Results and Conclusion: Standard systematic review protocols for pairwise meta-analyses need extensions to accommodate the increased complexity of network meta-analysis. Our suggested modifications are widely applicable to both Cochrane and non-Cochrane systematic reviews involving network meta-analyses.

Keywords: comparative effectiveness review; eligibility criteria; transitivity; network meta-analysis; indirect comparison; mixed treatment comparison

Running Title: Preparing a protocol for a network meta-analysis

Word Count: 4988

What is new?

- The standard format of systematic review protocols can be extended for a comparative effectiveness systematic review that aims to compare three or more interventions
- Consideration of the underlying assumptions of network meta-analysis is important when planning a comparative effectiveness systematic review to ensure a consistent evidence base
- Systematic reviewers and journal editors should consider the suggestions herein and publish protocols that include methodological considerations pertinent to comparative effectiveness systematic reviews and network meta-analyses

1 Introduction

The number of systematic reviews that simultaneously compare multiple interventions surges in the medical literature (1,2). Such reviews, which we refer to as comparative effectiveness reviews (CER), may include a data synthesis component, network meta-analysis (NMA). The popularity of the technique prompted the rapid methodological development of NMA, the availability of technical guidance, statistical tutorials (3–7), and recommendations for the proper conduct and reporting of CER (8–12). Good practice requirements of a standard systematic review apply equally to a CER, although several practices need to be adapted to accommodate the complexity imposed by comparing more than two interventions at a time.

A protocol for a systematic review provides a pre-specified outline of the research question under consideration as well as the detailed methods of the review (13). Review authors are encouraged to register the protocol using an international repository such as *PROSPERO* (14) to ensure transparency and reproducibility. As a key player in producing high-quality systematic reviews, Cochrane requires publication of the review protocol following an established methodology and outline (13).

In this paper, we aim to highlight parts of a protocol for a systematic review that may need special consideration when three or more competing interventions are to be compared. The protocol items that generally do not need modifications include the background, search strategy, the definition of the outcomes of interest and the risk of bias assessment for the included studies. Most items, though, require modifications. We organize our guidance following the protocol format suggested by Cochrane for a standard systematic review and elaborate on sections that need adjustment to address methodological issues pertinent to CER and NMA. We have previously prepared a protocol template for Cochrane authors (available on the website of the Cochrane Comparing Multiple Interventions Group http://methods.cochrane.org/cmi/) (15). Following feedback from users of this template, we elaborate on our previous work accordingly. The present paper replaces the template and aims to disseminate our recommendations to the wider community of systematic reviewers. We assume that readers are familiar with the basic requirements and statistical methods underpinning NMA, and direct interested readers to published tutorials for further guidance (16,17). Throughout we provide example text from published protocols to exemplify the considerations.

2 Illustrative CER protocols

We use three published CER protocols including NMA as our examples: (1) comparative effectiveness and safety of different treatments for panic disorder in adults (18); (2) safety of anti-epileptic drugs (AEDs) on infants and children exposed in utero or during breastfeeding (19); and (3) efficacy and acceptability of first- and second-generation antidepressants in the acute treatment of major depression (20).

3 Sections of a CER protocol and considerations for NMA

3.1 Setting the rationale for the review: background and objectives

Systematic reviews are intended to summarize evidence to inform decision-making. In any protocol for a systematic review on the effects of interventions, authors should describe the rationale for the review. In the background section of CER protocols, authors should specify their rationale for undertaking a CER and justify their intention to use NMA (9). The presence of several competing interventions and multiple independent pairwise comparisons may persuade researchers to synthesize simultaneously all the available data via NMA, so that the comparative effectiveness and safety of all possible comparisons between pairs of interventions can be estimated (21,22). In the absence of head-to-head trials comparing the active interventions of interest directly, when direct comparisons are limited, or when the aim is to assess the relative ranking of interventions across a range of clinical outcomes, a CER using NMA is the best available approach (17). However, as NMA rests on assumptions that can be difficult to evaluate, researchers should clarify which review questions are best achieved using NMA.

Review authors should also define in their objectives the health condition, population(s), setting(s), interventions, and outcome(s) of interest (13). The main difference for a NMA is that the list of interventions of interest is likely to be longer, and the distinction between interventions and comparators is not obvious: an intervention might be used as the experimental treatment in one study

and the control comparator in another. Hence, "IC" in the PICO acronym should be interpreted as a list of all the interventions to be compared.

For example, in the anti-epileptics protocol the authors justified the need for their review by stating that (19): *"Some AEDs have been associated with increased risk of harm to the fetus and infants. For example, exposure to valproate has led to increased risk of major congenital malformations* (23)*, cognitive delay, and minor congenital abnormalities* (24–27)*. Phenobarbital has been associated with minor congenital abnormalities and developmental delay* (23,28)*. Carbamazepine and lamotrigine have been associated with minor congenital abnormalities* (29–32)*. However, other than studies of the use of valproate, many studies have produced inconsistent findings regarding harm to the fetus and infant with use of other agents* (33)*. As such, our objective is to evaluate the comparative safety of AEDs for infants and children who were exposed in utero or during breastfeeding through a systematic review and network meta-analysis"*

3.2 Specifying the eligibility criteria and the network of possible comparisons

A review protocol will specify the eligibility criteria for studies and participants in light of the research question and the risk of important heterogeneity. This means that eligible studies should be similar enough in terms of design and participant characteristics to ensure a sufficiently homogeneous evidence base that can be validly synthesized. For a standard systematic review, the eligible interventions are defined by describing the pairwise comparison(s) of interest, such as the comparison of a new treatment versus an existing intervention.

In a CER with NMA, the eligibility criteria should clearly consider the transitivity (or exchangeability) assumption (21,22). This requires sufficient clinical and methodological comparability across all direct comparisons in the whole network. It implies that one can validly compare two interventions via a connected indirect route involving one or more intermediate comparators (21,22). When defining the population and interventions of interest the review author should assess the plausibility of this assumption, which requires all included interventions to be legitimate alternatives and therefore considered jointly randomizable. Joint randomizability means that a large multi-arm trial including all eligible interventions is conceptually meaningful even if practically not feasible (22). Authors should state in the protocol whether they believe that this might

be a reasonable setting for the expected network of interventions, for instance, by reporting that *"We assume that patients who fulfil the inclusion criteria are equally eligible to be randomised to any of the interventions we plan to compare"*. If such an assumption is not likely to hold a priori, it might be necessary to set narrower eligibility criteria for included interventions and populations. More extensive discussions on the plausibility of transitivity and examples of situations where this assumption might be violated can be found elsewhere (17,34,35).

Another difference between a standard systematic review protocol and a CER with NMA protocol is that review authors may distinguish between interventions of direct interest for forming recommendations for practice (also called the *decision comparator set* (12)) and interventions included to supplement the analysis. Additional interventions (such as placebo or no treatment) may provide useful information for the treatments of interest via indirect evidence. These additional interventions are expected to increase the precision of the results (36), connect a sparse network or help estimate heterogeneity, among other reasons. The concept of joint randomizability should apply to all treatments included in the network. In other words, an individual should be eligible to be allocated to any of the available interventions irrespective of whether it is of direct interest or supplementary. As there may be a diminishing return of including supplemental interventions in an analysis (35), a CER protocol should provide rationale for including them. The synthesis set cannot always be specified a priori, because the authors of the review may not be aware of all eligible interventions. Thus the protocol should also include information on whether unspecified interventions will be considered for post-hoc inclusion in the network within the context of jointly randomizable interventions, and how such decisions will be made.

Different eligibility criteria for interventions will result in different collections of evidence in the synthesis, and due to the interrelationships across direct and indirect evidence this can lead to different effect estimates and relative rankings (37). Restrictions to eligibility should be given a clear rationale. A graphical illustration of the anticipated network of competing interventions is highly desirable to present, in a comprehensive way, the definition of the nodes in the network and present any intended grouping or splitting of interventions as part of secondary analyses (e.g. by distinguishing different dosages of the same drug) (38). In the depression example, the authors list 21 drugs and placebo as eligible interventions and then they explain why they selected this set of interventions (20).

Finally, in this section the authors could describe how they plan to deal with different intervention doses, modalities, administration frequencies and routes as well as with co-interventions (39–41). Interventions should be fully described to the extent possible. Review authors sometimes split the nodes of the network to be able to explore differences within intervention nodes. It is possible, though, that they will eventually follow a lumping approach due to limited availability of data to explore such differences. To date no unanimously optimal approach has been suggested to define the nodes of the network as this depends on the clinical context. A decision on lumping or splitting the nodes of a network should be formed on the basis of the research question of the review and the outcomes of interest, as well as considering the underlying assumptions (i.e. merging insufficiently similar interventions might violate transitivity).

In the depression network, the authors addressed the issue of the different drug doses in their inclusion criteria and considering a sensitivity analysis (20): *"We will include only study arms randomising patients to drugs within the licensed dose. Both fixed-dose flexible-dose designs will be allowed (42). There is a possibility that some trials compare one agent at the upper limit of its therapeutic range with another agent at the lower limit of its therapeutic range within the same study. We plan to capture this study characteristic by adding a dichotomous variable indicating whether dosages are comparable, and use this information for a sensitivity analysis."*.

3.3 Specifying the outcomes of interest

A systematic review protocol should define a priori all primary and secondary outcomes of interest (43). It also should explain the preferred ways of measuring, aggregating, and analyzing outcomes to reduce the risk of selective inclusion of outcomes and results on the part of the systematic reviewer. However, it is common in CERs to perform NMA only for a subset of the pre-specified outcomes, due to lack of data. Review authors should be aware that in a CER different interventions may work with different mechanisms across multiple time points, and therefore combining outcomes measured at different follow-up times might be inappropriate. In addition, different interventions might lead to different adverse events and therefore review authors should focus on adverse outcomes that are potentially relevant to all or at least most of the interventions.

The reporting of the outcomes under investigation should not be considered as a criterion for including studies. The number and size of studies that do not report the outcome of interest will be considered when interpreting the summary estimates. It is possible that unreported outcomes are not favourable for the studied interventions, and therefore the summary estimate would have been different had these missing outcomes being observed. CER protocols should mention that network diagrams will be produced to show the amount of evidence and the topology for all selected studies and for each subset of studies contributing to each outcome.

It is also recommended that in this section of the protocol authors should specify for which outcomes, if at all, the authors plan to estimate the relative ranking of the competing interventions. For example, the protocol of the depression network reports (20): *"[…] we will conduct a random effects NMA to synthesise all evidence for each outcome, and obtain a comprehensive ranking of all treatments."*.

3.4 Searching and selecting eligible studies

A detailed description of the intended search strategy and selection process should be included in any systematic review protocol to ensure that authors will make the best effort to identify all potentially eligible studies for the review. In NMA, all studies comparing at least two of the competing interventions in the synthesis set are eligible for inclusion. Hawkins et al. (44) have suggested methodology for building iterative search algorithms with the aim of identifying studies that may provide useful indirect evidence. Constructing the evidence base for a CER with NMA is usually a complex procedure. Involvement of a knowledgeable librarian is recommended.

3.5 Describing the data extraction process and risk of bias assessment

In general, a protocol will specify the information that review authors plan to extract from the included studies. For example, this may include study design and setting, studied population, dosage of interventions, outcome data, and risk of bias assessments (45). For standard systematic reviews, the extraction of study or patient characteristics that potentially act as effect modifiers is usually associated with subgroup analyses or meta-regression for the investigation of heterogeneity. The strategy to compute missing statistics (such as standard deviations) is often described. Risk of bias assessments for NMAs will typically be similar to those for standard systematic reviews. However, it

is possible that different judgements might be made about risks of bias for different pair-wise comparisons within a multi-arm trial.

Similar to standard systematic review protocols, a CER protocol including NMA requires to specify whether arm-level or contrast-level (i.e. effect sizes and standard errors comparing a pair of interventions) data will be extracted when both are available, since this has implications for the choice of statistical model. Also, in a NMA the impact of potential effect modifiers needs to be evaluated in the context of the transitivity assumption. The plausibility of transitivity can be judged by comparing the distribution of the potential effect modifiers across the available direct comparisons in the network (10,22,46). Therefore, data extraction on these characteristics is necessary rather than optional. Authors should also describe how these could be related to the violation of the transitivity assumption.

For example, in the panic disorder protocol the authors report (18): *"From each included study we will extract data on the following study, interventions and population characteristics that may act as effect modifiers:*

1. *Methods: study design, randomization (individual or cluster), total duration of study, number of study centres and location, study setting, withdrawals, and date of study.*
2. *Participants: number, setting, diagnostic criteria, presence or absence of medical and psychiatric comorbidities, presence or absence of elderly participants, percentage of patients with agoraphobia, percentage of patients with baseline depression, inclusion criteria, and exclusion criteria.*
3. *Interventions: medication dose, medication dose range, use of rescue medication.*
4. *Outcomes: primary and secondary outcomes specified and collected, and time points reported. Where possible we will extract data at the arm level, not summary effects.*
5. *Notes: sponsorship/funding for trial, and notable conflicts of interest of trial authors"*

3.6 Selecting effect measures

Every protocol should report the effect measure(s) the authors intend to use for each outcome of interest. It should specify both the effect measures to be used in the statistical analyses and the effect measures to be used in the presentation of the findings, as these will not necessarily be the same. For

example, ratio scales are often preferred to analyse the data as they are associated with better mathematic properties, but presentation of results using absolute measures (such as the risk difference) is desirable as these measures are easier to interpret. Also, the use of multiple effect measures should be justified and the way any discrepancies between them will be handled should be described.

The additional consideration for a CER with NMA protocol is whether authors will also report the relative ranking of the competing interventions for one or more of the outcomes. This relates to the aims of the CER; they often include statements like *"we will obtain a comprehensive ranking of all treatments"* (from the depression protocol (20)). In such cases, unlike standard systematic review protocols, protocols including NMA should specify a priori the measure that will be used to rank the competing interventions (e.g. cumulative ranking curves, SUCRAs, mean ranks, or median ranks) as well as how the uncertainty of ranking will be reflected in the conclusions (47,48). For example, graphical tools such as the rankograms reflect visually the uncertainty in the ranking probabilities (38,49,50). Reviewers should refrain from using the probabilities of being the best as a measure of relative ranking because these are known to yield misleading intervention hierarchies (51).

In the panic disorder CER protocol the authors considered the following measures for relative ranking (18): *"We will also estimate the ranking probabilities for all treatments of being at each possible rank. We will also obtain a treatment hierarchy using the surface under the cumulative ranking curve (SUCRA) and mean ranks* (49)*. SUCRA can also be re-expressed as a percentage interpreted as the percentage of effectiveness/acceptability of an intervention that would be ranked first without uncertainty."*

3.7 Unit of analysis issues and missing outcome data

Protocols should explicitly describe how authors plan to deal with cluster-randomized and crossover trials. Ignoring the different design of such studies in the analysis can yield biased findings or findings with inappropriate precision. NMAs are also subject to the same biases that may arise when special study designs or missing data are treated incorrectly.

An additional threat to the validity of results is the presence of missing data. Authors should describe their strategy for obtaining missing information as well as any methods they will use, and the

underlying assumptions made, in studies with missing data in the analysis. When missing outcome data are believed to be an important threat to the validity of study results, the amount of and reasons for missingness need to be considered when drawing conclusions about the risk of attrition bias and the robustness of the inference. While the same considerations are applicable to CER, making assumptions about how informative is missingness become increasingly complex as one needs to consider the reasons for missingness in each intervention arm. For more information about the methodology on this issue the reader may consult elsewhere (52,53).

3.8 Qualitative assessment of the appropriateness of a synthesis

The underlying assumption in a standard meta-analysis is that studies evaluating the same comparison have enough clinical and methodological similarities to make their synthesis meaningful. This assumption refers to the comparability/similarity in terms of population characteristics as well as study design, intervention, outcomes and settings (54). In a standard systematic review protocol, authors should state how they plan to examine this assumption, for example by generating descriptive statistics to assess whether characteristics are comparable across all studies that inform a pairwise comparison.

The same ideas translate to the whole network of studies in a NMA. The validity of NMA relies on the transitivity assumption. As discussed above, the assumption of transitivity can be assessed by comparing the distribution of the potential effect modifiers across the direct comparisons in the network (10,46). However, the utility of this approach is sometimes limited by the fact that effect modifiers are under-reported, or that very few studies are available for each direct comparison. A strategy for assessing both clinical and methodological homogeneity and transitivity should be considered.

In the depression example the authors plan to evaluate transitivity using the following approach (20): "*The clinical features, which have been demonstrated to date to moderate efficacy of antidepressants include bipolarity* (55), *psychotic features* (56), *and subthreshold depression* (57). *We have assured transitivity in our network with regard to these variables by limiting our samples to participants with non-psychotic unipolar major depression. Other clinical or methodological variables that may influence our primary outcomes of antidepressant efficacy or acceptability include: age, depressive*

severity at baseline (58,59), *and the dosing schedule* (60). *We will investigate if these variables are similarly distributed across studies grouped by comparison. The inclusion of placebo and concerns about its potential to violate the transitivity assumption have been highlighted in general* (17,61) *and particularly in depression studies* (62,63). *Consequently, the comparability of placebo-controlled studies with those that provide head-to-head evidence will be examined carefully*."

3.9 Describing the planned statistical analyses

CERs can involve two types of analyses: a series of independent pairwise meta-analyses and NMA. Review authors should state in the protocol whether they intend to use both analyses. Additionally, as in standard systematic review protocols, the description of the planned statistical analysis should include technical details of the statistical model and a justification for model choice.

Authors should report the specific statistical model they will use to fit NMA (e.g. as multivariate meta-analysis (64) or a hierarchical model (3,65)). It is highly recommended to specify the framework (Bayesian vs frequentist) and the software to be used for analysis. If a Bayesian framework is to be used, additional technical details such as prior distributions and convergence of Markov chain Monte Carlo approaches should be described. Finally, assumptions about the heterogeneity variance and the method for estimating it should be reported; specifically whether a single heterogeneity parameter will be assumed for all comparisons or different, comparison-specific parameters will be employed. Reviewers should also report whether the method they plan to use to synthesize data correctly accounts for the correlated nature of data from multi-arm studies (66). The complexity of the methodology makes the involvement of a knowledgeable statistician with experience in NMA critical.

For example, in the depression protocol both direct meta-analyses and NMA are described (20): *"For each pair-wise comparison, we will synthesise data to obtain summary standardised mean differences (SMD, Cohen' s d) for continuous outcomes or ORs for dichotomous outcomes, both with 95% Credible Intervals (CrI). […] If the collected studies appear to be sufficiently similar with respect to the distribution of effect modifiers (refer Assessment of transitivity assumption section), we will conduct a random effects NMA to synthesise all evidence for each outcome, and obtain a comprehensive ranking of all treatments.[…]"*

In a standard systematic review protocol, authors should describe the methods they plan to use for the evaluation of statistical heterogeneity (e.g. via a chi-squared test or the I^2 measure) if they expect to synthesize a set of comparable studies.

There are statistical means to examine heterogeneity in a NMA, such as an I-square, and the estimate of the common heterogeneity variance (6,67,68). Protocols including NMA should describe the intended strategy to infer about statistical inconsistency which, when present, reflects the risk of important intransitivity. The different available methods are often classified into local and global approaches (68). Local approaches (e.g. contrasting direct evidence with indirect from a specific loop (termed loop-specific approach (69)) or with indirect from the entire network (termed node-splitting (70)) aim to identify pairwise comparisons or loops of evidence that might introduce important inconsistency in the network. Global approaches (e.g. inconsistency models (71,72)) assess the potential for inconsistency in the entire network. Each approach may lead to different conclusions with respect to the presence or the magnitude of inconsistency and thus the use of both local and global approaches is highly recommended (73).

In the anti-epileptic drugs example the authors consider the following approaches for evaluating inconsistency (19): *"We will ensure the following factors are present prior to conducting network meta-analysis: […] ii) consistency between direct and indirect data, which will be examined locally (i.e., in certain paths of the network) using the loop-specific method* (74,75) *and the node-splitting method* (70)*, and globally (i.e., evaluating the network as a whole), using the design-by-treatment interaction model* (72)*; and iii) we will quantify the amount of variability attributed to heterogeneity and inconsistency rather than sampling error, by calculating the* I^2 (67)*. […] We will compare the magnitude of heterogeneity between consistency and inconsistency models to determine how much heterogeneity will be explained by inconsistency. We will first use the design-by-treatment model for the evaluation of inconsistency in a network as a whole and then, if inconsistency is detected, we will employ the loop-specific and node-splitting methods to identify which piece of evidence is responsible for inconsistency"*.

3.10 Assessing reporting biases

Reporting biases, such as publication bias and selective outcome reporting bias, can occur in systematic review even after a well-designed and comprehensive search strategy. To assess the impact of possible reporting biases in standard systematic review, protocols usually consider the use of methods for small-study effects (funnel plots, Egger's test, etc.) (76) and rarely more sophisticated methods such as selection models (77).

These approaches have been extended to the context of NMA (38,78–80) and should be considered when drafting the respective section of a CER protocol. In particular, the comparison-adjusted funnel plot is an extension of the conventional graph that accounts for the fact that different studies estimate different effects (38,80). However, their application often requires additional assumptions for the direction of potential bias particularly when head-to-head comparisons of active interventions are included in the network. Review authors should, when possible, describe in their protocol the assumptions they will make about reporting biases and the rationale behind these choices.

For example, the depression protocol (20) reports: *"We will use the comparison-adjusted (38) and contour-enhanced* (81) *funnel plots to investigate whether results in imprecise trials differ from those in more precise trials. We will also run network meta-regression models to detect associations between study size and effect size* (80). *If an important association is found and publication bias is suspected, we will attempt to explore the possibility that funnel plot asymmetry is due to publication bias by employing a selection model* (82). *"*. The authors here should have ideally also pre-specified the assumptions about the direction of small-study effects within the network.

3.11 Investigating heterogeneity and inconsistency

Heterogeneity and inconsistency are caused by differences in populations or other study characteristics that modify the intervention effects within and across comparisons respectively (83). These characteristics may be included in the list of known potential effect modifiers or might be other unknown factors.

Standard systematic review protocols describe any additional analyses the authors plan to perform to explain anticipated heterogeneity in the data. Similarly in protocols including NMA researchers

should report which characteristics they consider as possible sources of heterogeneity and inconsistency. The impact of such characteristics on the relative effects may be explored via pre-specified subgroup or meta-regression analyses. Interpreting subgroup analyses in NMA might be challenging since subgroups of studies can lead to exclusion of interventions; in such case a comparison of the results between the different subgroups or with the primary analysis can be inappropriate (37).

When network meta-regression analysis is planned, further detail might be necessary regarding the assumptions related to the estimation of the regression coefficients (e.g. assuming consistency or not across coefficients) (84) and the directionality of the effect of covariates. For example, when controlling for risk of bias items in a full network of interventions, the direction of bias is straightforward in placebo-controlled trials (active treatments are expected to be favored) but not in head-to-head trials where stronger assumptions might be necessary (e.g. assuming that newer treatments are favored) (80). If the meta-regression model will be fitted in a Bayesian framework, information on the prior distributions for the regression coefficients should also be given. Finally, review authors also may pre-specify sensitivity analyses to investigate the robustness of findings to particular choices of statistical or non-statistical approach to the NMA.

In the depression network the authors suggest six possible sources of non-random variation (20): *"We will explore whether treatment effects for the two primary outcomes are robust in subgroup analyses and network meta-regression using the following characteristics: (1) study year; (2) sponsorship; (3) depressive severity at baseline; (4) dosing schedule; (5) response to placebo; (6) proportion of participants allocated to placebo; number of recruiting centres (single-centre vs multicentric studies)* (85,86). ".

3.12 Credibility of the evidence and summary of findings table

The most commonly used system to evaluate the credibility of the evidence in a standard systematic review is the Grade of Recommendation, Assessment, Development and Evaluation (GRADE) system, which considers that our confidence to the available evidence might be downgraded with respect to five domains (study limitations, indirectness, inconsistency, imprecision, publication bias)

(87). Typically, the ratings according to GRADE along with the intervention effects for the most important outcomes of the systematic review are presented in a summary of findings table.

Two approaches extending the GRADE system into NMA are currently available and can be considered when preparing a protocol for a CER including NMA (68,88). For large networks, it might be challenging to present the credibility of evidence for every comparison. Therefore, a selected set of the most important comparisons and outcomes should be defined in the protocol for inclusion in the summary of findings table. The full table and the GRADE judgements could be presented in an appendix.

4 Conclusions

The most important considerations for CER protocols are summarised in a Box. As in the case of conventional systematic review, deviations from the protocol may occur in the final review. This phenomenon might be common in the conduct of CERs as decisions about NMA are particularly dependent on the data. Specifically, deviations from the protocol are often required due to data availability and the subsequent trade-offs that are needed between precision and connectedness on the one hand and adequate discrimination of interventions on the other. In this situation the authors should provide a clear description of the reasons that imposed making changes to the initial design of their review (13). In conclusion, we urge review authors of NMA to prepare and share their protocols to minimize post-hoc decisions and to promote the culture of open science. We believe the guidance offered in this tutorial is widely applicable and is likely to improve the quality of NMA.

Acknowledgements: The authors would like to thank Drs. Lisa Hartling and Lorne Becker for their helpful comments and suggestions on an earlier version of the article. The work of AC and GS is supported by the OPERAM project from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 634238. This project is supported by the Methods Innovation Fund Program of the Cochrane Collaboration (MIF1) under the project "Methods for comparing multiple interventions in Intervention reviews and Overviews of reviews".

DMC is supported by a Medical Research Council Population Health Scientist fellowship award G0902118. This work was also undertaken with the support of the Centre for the Development and Evaluation of Complex interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence. Joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economist and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

References

1. Lee AW. Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. J Clin Epidemiol. 2014;67(2):138–43.
2. Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH, Salanti G. Characteristics of Networks of Interventions: A Description of a Database of 186 Published Networks. PLoS ONE. 2014;9(1):e86754.
3. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. Stat Methods Med Res. 2008;17:279–301.
4. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Med Decis Making. 2013;33:641–56.
5. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013;33:607–17.
6. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity--subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making. 2013;33:618–40.
7. Caldwell DM. An overview of conducting systematic reviews with network meta-analysis. Syst Rev. 2014;3(1):109.
8. Hutton B, Salanti G, Chaimani A, Caldwell DM, Schmid C, Thorlund K, et al. The quality of reporting methods and results in network meta-analyses: an overview of reviews and suggestions for improvement. PloS One. 2014;9(3):e92508.
9. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-Analyses of Healthcare Interventions: Checklist and Explanations. Ann Intern Med. 2015 in press;

10. Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R, et al. Indirect Treatment Comparison/Network Meta-Analysis Study Questionnaire to Assess Relevance and Credibility to Inform Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report. Value Health. 2014 Mar;17(2):157–73.
11. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011 Jun;14:417–28.
12. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence synthesis for decision making 7: a reviewer's checklist. Med Decis Making. 2013 Jul;33:679–91.
13. Higgins JP, Green S. Guide to the Contents of a Cochrane Protocol and Review. In: Fellow JPHSSV, Director SGF, editors. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons, Ltd; 2008. p. 51–79.
14. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. PROSPERO at one year: an evaluation of its utility. Syst Rev. 2013 Jan 15;2(1):4.
15. Protocol template for a Cochrane intervention review that compares multiple interventions. Prepared by Anna Chaimani and Georgia Salanti and revised by Lorne Becker, Debbi Caldwell, Julian Higgins and Tianjing Li. Available from: http://methods.cochrane.org/cmi/sites/methods.cochrane.org.cmi/files/uploads/Protocol%20for%20Cochrane%20Reviews%20with%20Multiple%20Interventions.pdf
16. Li T, Puhan MA, Vedula SS, Singh S, Dickersin K, $author.lastName $author firstName. Network meta-analysis-highly attractive but more methodological research is needed. BMC Med. 2011 Jun 27;9(1):79.
17. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. Ann Intern Med. 2013 Jul 16;159:130–7.
18. Guaiana G, Barbui C, Caldwell D, Davies S, Furukawa T, Imai H, et al. Antidepressants, benzodiazepines and azapirones for panic disorder in adults: a network meta-analysis. Cochrane Database Syst Rev. 2015;.
19. Tricco AC, Cogo E, Angeliki VA, Soobiah C, Hutton B, Hemmelgarn BR, et al. Comparative safety of anti-epileptic drugs among infants and children exposed in utero or during breastfeeding: protocol for a systematic review and network meta-analysis. Syst Rev. 2014;3:68.
20. Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ Open. 2016 Jul 1;6(7):e010919.
21. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ. 2005 Oct 15;331:897–900.

22. 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 Meth. 2012;3(2):80–97.
23. Harden CL, Meador KJ, Pennell PB, Allen Hauser W, Gronseth GS, French JA, et al. Management issues for women with epilepsy—Focus on pregnancy (an evidence-based review): II. Teratogenesis and perinatal outcomes. Epilepsia. 2009 May 1;50(5):1237–46.
24. Adab N, Jacoby A, Smith D, Chadwick D. Additional educational needs in children born to mothers with epilepsy. J Neurol Neurosurg Psychiatry. 2001 Jan;70(1):15–21.
25. Adab N, Kini U, Vinten J, Ayres J, Baker G, Clayton-Smith J, et al. The longer term outcome of children born to mothers with epilepsy. J Neurol Neurosurg Psychiatry. 2004 Nov;75(11):1575–83.
26. Gaily E, Kantola-Sorsa E, Hiilesmaa V, Isoaho M, Matila R, Kotila M, et al. Normal intelligence in children with prenatal exposure to carbamazepine. Neurology. 2004 Jan 13;62(1):28–32.
27. Meador KJ, Baker GA, Browning N, Clayton-Smith J, Combs-Cantrell DT, Cohen M, et al. Cognitive Function at 3 Years of Age after Fetal Exposure to Antiepileptic Drugs. N Engl J Med. 2009 Apr 16;360(16):1597–605.
28. Reinisch JM, Sanders SA, Mortensen EL, Rubin DB. In utero exposure to phenobarbital and intelligence deficits in adult men. JAMA. 1995 Nov 15;274(19):1518–25.
29. Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, et al. Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. J Neurol Neurosurg Psychiatry. 2006 Feb;77(2):193–8.
30. Holmes LB, Baldwin EJ, Smith CR, Habecker E, Glassman L, Wong SL, et al. Increased frequency of isolated cleft palate in infants exposed to lamotrigine during pregnancy. Neurology. 2008 May 27;70:2152–8.
31. Meador KJ, Baker GA, Finnell RH, Kalayjian LA, Liporace JD, Loring DW, et al. In utero antiepileptic drug exposure: fetal death and malformations. Neurology. 2006 Aug 8;67(3):407–12.
32. Vajda FJE, Hitchcock A, Graham J, Solinas C, O'Brien TJ, Lander CM, et al. Foetal malformations and seizure control: 52 months data of the Australian Pregnancy Registry. Eur J Neurol Off J Eur Fed Neurol Soc. 2006 Jun;13(6):645–54.
33. Meador KJ, Penovich P, Baker GA, Pennell PB, Bromfield E, Pack A, et al. Antiepileptic drug use in women of childbearing age. Epilepsy Behav EB. 2009 Jul;15(3):339–43.
34. 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 Meth. 2012;3(2):80–97.
35. Caldwell DM, Dias S, Welton NJ. Extending Treatment Networks in Health Technology Assessment: How Far Should We Go? Value Health. 2015 Jul;18(5):673–81.

36. Cooper NJ, Peters J, Lai MC, Juni P, Wandel S, Palmer S, et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? Value Health. 2011 Mar;14:371–80.
37. Mills EJ, Kanters S, Thorlund K, Chaimani A, Veroniki AA, Ioannidis JP. The effects of excluding treatments from network meta-analyses: survey. BMJ. 2013;347:f5195.
38. Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. PLoS One. 2013;8(10):e76654.
39. Giovane CD, Vacchi L, Mavridis D, Filippini G, Salanti G. Network meta-analysis models to account for variability in treatment definitions: application to dose effects. Stat Med. 2013 J;32:25–39.
40. Madan J, Chen Y-F, Aveyard P, Wang D, Yahaya I, Munafo M, et al. Synthesis of evidence on heterogeneous interventions with multiple outcomes recorded over multiple follow-up times reported inconsistently: a smoking cessation case-study. J R Stat Soc Ser A. 2014;177(1):295–314.
41. Mosseri J, Trinquart L, Nizard R, Ravaud P. Meta-Analysis of a Complex Network of Non-Pharmacological Interventions: The Example of Femoral Neck Fracture. PLOS ONE. 2016 Jan 6;11(1):e0146336.
42. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009;373(9665):746–58.
43. Saldanha IJ, Dickersin K, Wang X, Li T. Outcomes in Cochrane systematic reviews addressing four common eye conditions: an evaluation of completeness and comparability. PloS One. 2014;9(10):e109400.
44. Hawkins N, Scott DA, Woods B. How far do you go? Efficient searching for indirect evidence. MedDecisMaking. 2009;29:273–81.
45. Li T, Vedula SS, Hadar N, Parkin C, Lau J, Dickersin K. Innovations in data collection, management, and archiving for systematic reviews. Ann Intern Med. 2015;162(4):287–94.
46. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. BMC Med. 2013;11:159.
47. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. BMC Med Res Methodol. 2015;15:58.
48. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in treatment rankings: Reanalysis of network meta-analyses of randomized trials. Ann Intern Med. 2016;164(10):666–73.
49. 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.
50. Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentational approaches were developed for reporting network meta-analysis. J Clin Epidemiol. 2014; 67(6):672-80

51. Kibret T, Richer D, Beyene J. Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. Clin Epidemiol. 2014;6:451–60.
52. Franchini AJ, Dias S, Ades AE, Jansen JP, Welton NJ. Accounting for correlation in network meta-analysis with multi-arm trials. Res Synth Meth. 2012;3(2):142–60.
53. Spineli LM, Higgins JP, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. Clin Trials. 2013 Jan 15;(1740-7753.
54. Mavridis D, White IR, Higgins JPT, Cipriani A, Salanti G. Allowing for uncertainty due to missing continuous outcome data in pairwise and network meta-analysis. Stat Med. 2015 Feb 28;34(5):721–41.
55. Deeks JJ, Higgins JP, Altman DG. Analysing Data and Undertaking Meta-Analyses. In: Fellow JPHSSV, Director SGF, editors. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons, Ltd; 2008:243–96.
56. Taylor DM, Cornelius V, Smith L, Young AH. Comparative efficacy and acceptability of drug treatments for bipolar depression: a multiple-treatments meta-analysis. Acta Psychiatr Scand. 2014 Dec;130(6):452–69.
57. Wijkstra J, Lijmer J, Burger H, Geddes J, Nolen WA. Pharmacological treatment for psychotic depression. Cochrane Database Syst Rev. 2013;11:CD004044.
58. Barbui C, Cipriani A, Patel V, Ayuso-Mateos JL, van Ommeren M. Efficacy of antidepressants and benzodiazepines in minor depression: systematic review and meta-analysis. Br J Psychiatry J Ment Sci. 2011 Jan;198(1):11–6, sup 1.
59. Fournier JC, De Rubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: A patient-level meta-analysis. JAMA. 2010 Jan 6;303(1):47–53.
60. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. Arch Gen Psychiatry. 2012 Jun;69(6):572–9.
61. Khan A, Kolts RL, Thase ME, Krishnan KRR, Brown W. Research design features and patient characteristics associated with the outcome of antidepressant clinical trials. Am J Psychiatry. 2004 Nov;161(11):2045–9.
62. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. J Clin Epidemiol. 2009;62:857–64.
63. Rutherford BR, Sneed JR, Roose SP. Does Study Design Influence Outcome? The Effects of Placebo Control and Treatment Duration in Antidepressant Trials. Psychother Psychosom. 2009;78(3):172–81.
64. Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, et al. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. J Clin Psychiatry. 2010;71(3):270–9.
65. White IR, Barrett JK, Jackson D, Higgins JPT. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. Res Synth Meth. 2012;3(2):111–25.

66. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004;23(20):3105–24.
67. Jackson D, Barrett JK, Rice S, White IR, Higgins JPT. A design-by-treatment interaction model for network meta-analysis with random inconsistency effects. Stat Med. 2014;33(21):3639–54.
68. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the Quality of Evidence from a Network Meta-Analysis. PLoS ONE 2014;9(7).
69. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50:683–91.
70. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010;29:932–44.
71. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. J Amer Statist Assoc. 2006;101(474):447–59.
72. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and insconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Meth. 2012;3(2):98–110.
73. Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Res Synth Methods. 2013;4(4):291–323.
74. Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. BMJ. 2011;343:d4909.
75. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. IntJ Epidemiol. 2013;42:332–45.
76. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.
77. Copas J. What works?: selectivity models and meta-analysis. J R Stat Soc Ser A Stat Soc. 1999;162(1):95–109.
78. Mavridis D, Sutton A, Cipriani A, Salanti G. A fully Bayesian application of the Copas selection model for publication bias extended to network meta-analysis. Stat Med. 2013;32:51–66.
79. Mavridis D, Welton NJ, Sutton A, Salanti G. A selection model for accounting for publication bias in a full network meta-analysis. Stat Med. 2014 Dec 30;33(30):5399–412.
80. Chaimani A, Salanti G. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. Res Synth Meth. 2012;3(2):161–76.
81. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. J Clin Epidemiol. 2008;61(10):991–6.
82. Mavridis D, Welton NJ, Sutton A, Salanti G. A selection model for accounting for publication bias in a full network meta-analysis. Stat Med. 2014;33(30):5399–412.

83. Chaimani A, Vasiliadis HS, Pandis N, Schmid CH, Welton NJ, Salanti G. Effects of study precision and risk of bias in networks of interventions: a network meta-epidemiological study. IntJEpidemiol. 2013;42:1120–31.
84. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Stat Med. 2009 Jun 30;28:1861–81.
85. Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? J Clin Psychiatry. 2003;64(12):1465–75.
86. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. J Eur Coll Neuropsychopharmacol. 2009;19(1):34–40.
87. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383–94.
88. Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ. 2014;349:g5630.

Box. Important considerations for comparative effectiveness review protocols with network meta-analysis in comparison with standard systematic review protocols

- Review authors should describe early on in the protocol the reasons that suggest a) the need and b) the appropriateness of a comparative effectiveness review with network meta-analysis in relation to the research question (e.g. absence or scarcity of direct evidence, obtaining a treatment hierarchy).
- Inclusion criteria for studies, participants, and interventions should be defined in the light of transitivity.
- Eligible interventions can be categorized into interventions of direct interest for recommendations for practice and interventions that provide indirect evidence (e.g. legacy treatments, placebo etc). Although the latter cannot always be pre-specified in the protocol, authors should give a clear rationale for their selection of interventions in both groups.
- All potential effect modifiers need to be defined in comparative effectiveness review protocols as data on such characteristics are necessary for the evaluation of transitivity.
- Measures for relative ranking should be specified in addition to relative effect measures.
- Search strategy and study selection process should be designed such that any study that compares at least two of the eligible interventions and meets all other inclusion criteria be included in the network.
- Authors should report the approach to network meta-analysis (e.g. multivariate meta-analysis, hierarchical model) they will use and any assumptions they make about heterogeneity variances.
- A clear strategy for the assessment of statistical inconsistency should be given in every comparative effectiveness review with network meta-analysis protocol.
- Methods to assess potential reporting biases for comparative effectiveness reviews with network meta-analysis have been developed and should be considered in the protocol.
- Possible sources of important heterogeneity and inconsistency should be specified in the protocol and used, if possible, in additional analyses to explore the impact of these variables on the findings.