Author's response to reviews

Title: Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: Challenges and opportunities

Authors:

Chris G Cameron (cgcamero@gmail.com)
Bruce Fireman (bruce.fireman@kp.org)
Tammy Clifford (tammyc@cadth.ca)
Doug Coyle (dougc@uottawa.ca)
Brian Hutton (bhutton@ohri.ca)
George Wells (gawells@ottawaheart.ca)
Colin Dormuth (colin.dormuth@ti.ubc.ca)
Robert Platt (robert.platt@mcgill.ca)
Sengwee Toh (darren_toh@harvardpilgrim.org)

Version: 3 Date: 10 August 2015

Author's response to reviews: see over
Dear Systematic Review Editors:

Thank you very much for giving us the opportunity to resubmit our manuscript “Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: Challenges and opportunities”. We appreciate your helpful comments and suggestions. Based on the reviewers’ suggestions, we have revised the manuscript as a Commentary paper and reduced the word count considerably. Below are our responses (in bold) to address the concerns raised by the reviewers. We hope you find them satisfactory. We have attached a revised version of our manuscript with the outlined changes.

Editor’s specific comments:
Please include a method section. This should include the design of the study, the setting, the type of participants or materials involved, a clear description of all interventions and comparisons, and the type of analysis used, including a power calculation if appropriate.

This is not primary research. As such we do not include a methods section. Reviewer #1 noted that this work would be more suited as a “Commentary” paper and we tend to agree.
Reviewer #1:
1. The authors address an interesting and timely topic, the inclusion of non-randomized studies in network meta-analysis. A few of these studies have already appeared, and a treatment of the assumptions underlying such analyses, as well as the various options available for conducting them, is very welcome. I appreciate the balanced account of assumptions underlying network meta-analysis as well as the impact of confounding. However, as a "Methodology" paper the current account seems rather thin. I would suggest that the authors add much more detail on the methods they describe, as well as worked examples and/or simulation studies to contrast them if they wish to present this as a "Methodology" article. Otherwise, I think the work would be more suited as a "Commentary" paper. I think it could be an excellent commentary on the current state of the art, if my other comments below are addressed.

Thank you. We agree with the reviewer's comment and have revised the paper as a commentary paper.

2. The manuscript contains many confusing and/or grammatically incorrect sentences. The authors need to carefully proofread their work and correct it to avoid confusion.

Examples:
- "Conduct of network meta-analysis of non-randomized studies", paragraph 1 - "although the methods have been largely applied to RCTs" does not fit the sentence.
- "Conduct of network meta-analysis of non-randomized studies", paragraph 2 - "but with additional risk of bias given these studies have may still also be limited" does not make sense.
- "Conduct of network meta-analysis of non-randomized studies", paragraph 3 - "Some non-randomized study designs and some may be better suited for network meta-analysis than others." does not make sense.
- "Comparing and combining findings from non-randomized studies with RCTs", paragraph 1 - "to evaluate the comparative performance nonrandomized studies and" does not make sense.
- "Future directions", paragraph 1 - "the likelihood that patients are more similar among compared with inclusion of" does not make sense.

Thank you. We have revised the manuscript carefully and corrected the errors.

3. "What is network meta-analysis?", paragraph 1 - "The key assumption underlying any meta-analysis is transitivity of the studies." - it would be more accurate to say the key assumption is "exchangeability of the studies", and the following two sentences do indeed characterize exchangeability, not transitivity. This assumption is also present for pair-wise meta-analysis. The difference is that in pair-wise meta-analysis, the violation of that assumption can only lead to heterogeneity, whereas in network meta-analysis it can lead to non-transitivity of the relative treatment effects (which is loosely the same thing as inconsistency). I think the observations made at the end of the paragraph
regarding pair-wise meta-analysis would also be more clear if framed in this way. The work by Jansen & Naci (2013) that the authors cite provides a similar framework (as do several more technical papers by Lu & Ades).

Thank you. We have revised the paragraph per reviewer’s suggestion.

4. "Conduct of network meta-analysis of non-randomized studies", paragraph 3 - "The network meta-analytic methods chosen will depend on the study designs of non-randomized studies." - This statement is entirely vacuous. The authors do not specify which methods of network meta-analysis are to be considered, nor which characteristics of the design of non-randomized studies should impact the choice, or how. The remainder of the paragraph seems to deal with methods to adjust for confounding in observational studies, which is a different issue.

Thank you. We have removed the sentence and revised the paragraph per reviewer’s suggestion.

5. "Comparing and combining findings from non-randomized studies with RCTs", paragraph 2 - "Alternatively, non-randomized studies could be incorporated as prior information if a Bayesian approach is used." - Please explain how that is different from naive pooling?

We have expanded this paragraph to provide more detail.

6. "Comparing and combining findings from non-randomized studies with RCTs", paragraph 2 - none of the alternative approaches to incorporating non-randomized evidence are discussed in any detail. This section needs to be expanded greatly to enable the reader to understand the issues at hand. Especially the bias adjustment models need further description.

We have expanded this paragraph to provide more detail.

7. "Future directions", paragraph 1 - "albeit at a potential cost of reduced precision." - It is unclear to me whether this would result in an undesirable loss of precision, or the (desirable) elimination of double counting.

We have removed this sentence and revised the paragraph to make it clearer.

8. "Future directions", paragraph 2 - "has used pair-wise meta-analysis to combine data within CNODES" - please explain what this means exactly? Are results from CNODES themselves the result of meta-analysis? Would that not jeopardize the ability to correct for confounding in a consistent manner? What about the possibility of double counting of participants (as mentioned in the previous comment) if pair-wise meta-analysis is used?
In CNODES, each site performs the analysis locally based on a common protocol. Each site adjusts for confounders and other biases. The site-specific, adjusted results are then pooled via pair-wise meta-analysis. The probability of double counting is low and will only happen if individuals move from one province to another and have their information recorded in two databases.

Minor Essential Revisions
------------------------------------
9. "What is network meta-analysis?" paragraph 1 - "all treatments included in the network meta-analysis could have been included in a trial" - would be clearer to say "a single trial" or "the same trial" (especially given that observational studies are also considered). Perhaps it would be useful to explicitly use "jointly randomizable".

We have revised the paragraph based on the reviewer’s suggestion.

10. "Conduct of network meta-analysis of non-randomized studies", paragraph 1 - "The validity ... comparisons." - this characterization seems better suited to the section introducing network meta-analysis assumptions. I don’t have a problem with the reminding the reader of this assumption and that it is especially problematic for non-randomized studies, but this is not the place to introduce the concept.

We have revised the paragraph based on the reviewer’s suggestion.

11. "Box 1" - "The application of network meta-analysis in non-randomized studies is more complex and less understood than traditional approaches for stakeholders." - the phrase "for stakeholders" doesn't connect well with the remainder of the sentence. Please rephrase. It is also unclear if "traditional approaches" refers to pair-wise meta-analysis incorporating non-randomized studies, or to network meta-analysis with only randomized studies.

We have revised this section based on the reviewer’s suggestion.

12. "Box 1" - point 4 under disadvantages seems redundant / overlapping with point 1. Please distinguish them more clearly.

We have removed point 4 per reviewer’s suggestion.

Discretionary Revisions
13. "What is network meta-analysis?" - I would prefer a section heading that is not a question.

We have revised the section heading based on the reviewer’s suggestion.
Reviewer #2
Major Compulsory Revisions

14. Observational studies are particularly prone to publication bias. There are well defined methods for evaluating publication bias in pair-wise meta-analysis, and the authors should note that NMA practitioners should apply these methods to each of their direct comparisons (see Puhan et. al. BMJ, 2014).

We now discuss publication bias in the revised manuscript.

15. It would be pertinent to have some discussion on how authors can rate the certainty of effect estimates based on NMAs with observational data. The GRADE working group has made some recommendations, that should be helpful (Puhan, BMJ, 2014). The authors should note that the level of certainty (quality of evidence) should be addressed for each pair-wise comparison in the network (ideally for the direct, indirect, and network estimates). The approach would involve, for observational studies, starting at low confidence and then rating down (risk of bias, precision, consistency, directness and publication bias) or up (large magnitude of effect or dose-response gradient).

We have revised the manuscript to include GRADE and other guidelines that provide guidance on the evaluation of the quality of evidence from non-randomized studies.

16. Do the authors believe that both cohort and case-control studies would be appropriate to use in such analyses? They should address this issue.

This is an interesting question. The inclusion of other study designs in network meta-analysis is possible but has its own unique challenges that are beyond the scope of our paper, which is now focused on non-randomized comparative cohort studies.

17. Line 127-133: The authors discuss the underlying assumption of transitivity in NMA and acknowledge the concern that observational studies may introduce unmeasured bias, which may invalidate the study. Two ways for evaluating the transitivity assumption are given: 1) close inspection of the studies for similarity, and 2) compare baseline event rates in the common treatments. The first suggestion is intuitive, but is inevitably very subjective. With respect to the second suggestion, meta-analysis of relative effects assumes relative effects between treatments are similar (for example, treatment A is twice as effective as treatment B in the population being studied, regardless of absolute rates). Therefore, the transitivity assumption may remain valid in cases where study populations may differ in characteristics that are not effect modifiers. On the other hand, we agree that if event rates are similar in the common treatment arms, this may provide some reassurance that the populations are similar. These limitations to evaluating transitivity should be included.

We have revised the paragraph to include these suggestions.
18. It would be helpful to comment on what sensitivity analyses can be or should be performed in NMA to ensure validity of the NMAs.

This is certainly an important point but we felt that it is more suitable for a paper that talks about NMA in general. We have decided to not include this suggestion to focus on the main scope of the paper but will be happy to include this should the editor determine that this is essential.

Minor Essential Revisions:
19. Line 58: ethnical should be ethical.

Corrected.

20. Line 160-162: The authors state that “the validity of NMA is based on the underlying assumption that there is no imbalance in the distribution of effect modifiers across different types of direct treatment comparisons.” We agree with this statement, however NMA also offers the benefit of allowing adjustment for effect modifiers (eg., see Johnston, JAMA, 2014;12:923-33, where the use of behavioural support and exercise was adjusted for with regression and effect estimates were reported independently of these effect modifiers). While there are limits to study-level data, regression could be used to account for some known effect modifiers.

This is another good point but we felt that it is more suitable for a paper that talks about NMA in general. We have decided to not include this suggestion to focus on the main scope of the paper but will be happy to include this should the editor determine that this is essential.

21. Line 236-238: Are the authors referring to NMAs or primary observational studies? The sentence is difficult to understand.

We have clarified this in the revised manuscript.

22. Lines 240-242: The authors appear to incriminate systematic reviews for resulting in excessive variability of methods. The problem isn’t the review, but the eligibility criteria. One could, for instance, restrict eligibility to studies that included at least particular variables for adjustment, or particular methods (such as propensity matching, for instance). This point needs to be clarified.

We have revised this section to offer a more well-balanced discussion.

23. Lines 261-262: The authors state that it would be difficult to justify excluding observational data from large observational drug safety database because of their large sample sizes and homogenous methods. This needs qualification: for instance, if the drug safety monitoring databases do not capture important prognostic factors for adjustment, whereas other studies do, it may be preferable to exclude these studies.
We have revised this section to offer a more well-balanced discussion.

24. Line 276: Data bases should be 1 word.

Corrected.

25. Line 294: Networks should be network.

Corrected.

Discretionary revisions:
26. Lines 268-271: While true, this is not relevant to the commentary.

We have removed this from the revised manuscript.
Reviewer #3
Overall:
The authors have developed a commentary based on their expertise that addresses an important topic. This is a timely question that requires careful consideration and this paper can help provide the background to educate readers on the relevant issues. However, a more in-depth exploration of the issues surrounding this topic would improve this commentary.

Major revisions:
27. Throughout the paper, the term ‘non-randomized studies’ is used; this is a very broad term that includes many different types of study designs. Although the authors seem to restrict their discussion to comparative (cohort?) studies, this is not clear. Given the target audience of this commentary, it is important to differentiate the types of non-randomized studies that are explicitly considered for inclusion in NMAs by the authors as well as those that are not. For example, the authors have not addressed single-arm studies, which may provide the link for an otherwise disconnected network. Therefore it is critical that the authors provide clear overview of study types and how each is considered (or not considered) in context of NMA.

Thank you. The second paragraph of the Background section now explicitly defines the scope of our paper: “In this paper, we describe network meta-analysis involving non-randomized comparative cohort studies – defined as cohort studies that compare two or more treatment alternatives using observational data.” We have also revised the title of the paper to reflect this.

28. The third paragraph of introduction suggests few NMAs have incorporated non-randomized studies and identified four references. Although this statement is fairly subjective, it should be clarified that the identification of these references was not systematic and that only examples are provided. A review by Verde et al. 2015 has identified three additional clinical applications which should be referenced at a minimum (Combining randomized and nonrandomized evidence in clinical research: a review of methods and applications. Research Synthesis Methods 6(1):45-62). Highlighting the need for a more systematic approach to answer these questions, while acknowledging the challenges of doing so (given differences in terminology of non-randomized studies etc.), would strengthen this paper.

Thank you. We have added these references.

29. The first paragraph focuses on adverse events and refers to Box 1. However, the logic provided in Box 1 is not limited to adverse (or rare) events. The paper should either be broadened to all outcomes or be updated to include a clear rationale for a focus on adverse events. There is already much more acceptance in looking beyond RCTs for adverse events, given the often huge amount of uncertainty in NMA of AEs (especially in more complicated or sparse networks). Therefore, the broader question regarding use of non-randomized trials for all outcomes may provide a more useful
commentary, in additional to providing a more relevant exploration of the costs and benefits of this additional evidence.

**We agree with the reviewer. The scope and content of the entire paper has been revised to include all non-randomized comparative cohort studies that assess the safety and effectiveness of medical treatments.**

30. In the second paragraph of ‘Comparing and combining findings from non-randomized studies with RCTs’ the authors identify the methods of one example where non-randomized studies were included (i.e. Hutton et al.). However, there is no mention of the methods that were used by the other clinical applications. This seems highly pertinent to understand what is currently being done in the identified examples. Since high impact journals such as BMJ are open to analyses incorporating non-randomized evidence, the quality of the existing applications has clear implications regarding importance of alternative methods for including non-randomized evidence.

**After deliberation, we concluded that the inclusion of this paragraph is not essential to the paper and have removed it from the revised manuscript.**

31. The methods by Schmidz et al. are summarized briefly (as was done by Verde et al.); however, these references are not complete (i.e. missing Soares et al. 2014; McCarron et al. 2010/2012). Due consideration for these papers is required at minimum.

**Thank you. We have added these references.**

32. The first paragraph regarding ‘What is NMA’ provides useful introduction for readers not familiar with this type of analysis. However, there are several papers dedicated to assessing the transitivity assumption and Figure 1 does not seem to do justice to the existing literature in this area. Similarly, several key references are missing from the methods regarding identification of inconsistencies. Since this background regarding NMAs is well established in the literature, authors should reference the work more completely, or assume a more informed reader and provide a more in-depth discussion regarding the differences when considering non-randomized evidence.

**Thank you. We have added a few additional references.**

33. In the paragraph addressing ‘Confounding in randomized trials and non-randomized studies’ it is unclear what is meant by ‘unadjusted confounding’. Prior to this sentence authors identify the risk of differences in unmeasured treatment effect modifiers. If the differences are unmeasured, adjustments will not reduce bias unless adjusting for a specific kind type of bias that is expected – therefore this is confusing and needs to be clarified. References on adjusting for bias in NMAs would strengthen this section (if that is the intent).

**Thank you. We have revised the paragraph per the reviewer’s suggestion to avoid confusion.**
34. The third paragraph in ‘Conduct of network meta-analysis of non-randomized studies’ begins to address the assessment of transitivity in non-randomized trials. However, the background regarding NMAs does not properly emphasize the importance of identifying potential treatment effect modifiers a priori, which has been recommended in several guidelines and has very important implications for how non-randomized evidence is considered. Since NMAs have already started to consider non-randomized studies, ensuring the process to do so is robust may be the most important message, and therefore should be addressed more specifically. Along these lines, emphasizing the process to assess quality of non-randomized studies and the challenges in doing (tools to do so etc.) improve the manuscript.

The revised manuscript now discusses treatment effect modifiers and the quality of non-randomized studies in greater detail.

35. It would be interesting for authors to assess the conditions under which exploring non-randomized trials would be expected to provide additional benefits beyond the expected costs. Given the absence of clear guidelines or agreement from HTA agencies, it seems that there is an opportunity to highlight the lack of consensus on how best to handle non-randomized trials in the context of NMAs and to provide suggestions.

We have added a paragraph to discuss this issue.

Minor revisions:
36. The introduction states that RCTs are preferred for methodological reasons. However, it seems these reasons should be outlined more clearly as the subsequent sections assume a fairly naïve reader. The research question really hinges on whether it is worthwhile to explore other evidence despite the ‘well-established methodological reasons’, and therefore this should be clear in the beginning (rather than specifying this later on).

We have rephrased this sentence and now provide more details about the strengths of RCTs in the revised text.

37. The first paragraph regarding ‘Conduct of network meta-analysis of non-randomized studies’ does not seem to provide any information that is specific to including non-randomized trials and could be eliminated.

We have deleted a large portion of this paragraph per the reviewer’s suggestion.

38. The second paragraph in ‘Conduct of network meta-analysis of non-randomized studies’ addresses the advantages of considering non-randomized evidence. This would be more helpful in the introduction to justify the research question.

We have re-organized the manuscript to address this comment.
39. The paragraph addressing ‘Confounding in randomized trials and non-randomized studies’ is important for readers to understand the risk of including non-randomized studies. Figure 2 is useful in this sense, although does not seem to acknowledge possibility that confounding could lead to smaller (biased) treatment effect in non-randomized trials. Literature regarding bias (and direction of bias) would improve this section.

We have added more discussion around biases in non-randomized comparative studies and referenced literature related to these biases.

40. The practical implications of considering non-randomized studies for a systematic review (or HTA) process in terms of additional time, costs, and overall impact on decision-making have not been discussed, which would seem important given the audience.

We have added a paragraph related to this.