**Reviewer's report**

**Title:** Network meta-analysis incorporating non-randomized studies for assessing medical product safety: pitfalls and opportunities

**Version:** 2  **Date:** 12 June 2015

**Reviewer:** Gordon Guyatt

**Reviewer's report:**

1. Is the question posed by the authors new and well defined?

This is a commentary written by the authors on the use of network meta-analysis (NMA) for medical product safety monitoring including primary observational studies. The authors discuss some risks and benefits associated with the use of NMA with observational studies. The query posed is well defined and addresses a current gap in the literature.

2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work?

This is a commentary article, no primary research was performed.

3. Are the data sound and well controlled?

N/A

4. Do the figures appear to be genuine, i.e. without evidence of manipulation?

Yes. Is the data from figure 4 hypothetical or from a meta-analysis? The data source should be made clear.

5. Does the manuscript adhere to the relevant standards for reporting and data deposition?

Yes.

6. Are the discussion and conclusions well balanced and adequately supported by the data?

Yes, the authors conclusions that observational studies might be cautiously be used in NMA is supported.

7. Do the title and abstract accurately convey what has been found?

Yes.

8. Is the writing acceptable?

Yes.

**Major Compulsory Revisions:**
1. 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).

2. 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).

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

4. 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.

5. It would be helpful to comment on what sensitivity analyses can be or should be performed in NMA to ensure validity of the NMAs.

Minor Essential Revisions:

1. Line 58: ethnical should be ethical.

2. 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.
3. Line 236-238: Are the authors referring to NMAs or primary observational studies? The sentence is difficult to understand.

4. 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.

5. 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.

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

7. Line 294: Networks should be network.

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

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests