Reviewer's report

Title: A random effect variance shift model for detecting and accommodating outliers in meta-analysis

Version: 1 Date: 16 November 2010

Reviewer: Katherine Jane Lee

Reviewer's report:

This paper presents a novel method of identification and adjustment for outliers in a meta-analysis which had the potential to be extremely useful when carrying out a meta-analysis. Although the majority of this paper is well written there are areas where further clarification is required.

Major Compulsory Revisions

1. When introducing the RVSOM model I was initially confused by the j subscript and it was only later that it became apparent that the RVSOM is calculated per study, with the j indicating which study has an inflated variance. This needs clarifying on page 5.

2. Following on from this, it is only apparent in the examples that you can allow inflated variance for more than one observation in the same model (example 3?), I wonder whether it would be worth including a more general version of this model making this clearer up front?

3. The approach suggested is described as a parametric bootstrap procedure although there is no bootstrapping mentioned in the procedure described on page 7?

4. (Results, example 1, page 8) The author states that “observation 8 is clearly detected as an outlier”, although from the figure the LRT for observation 8 does not seem to reach any of the thresholds for being an outlier?

5. (Results, examples 2 and 3, page 9) The author talks about the methodology “performing well” in these examples but I am not sure what it means for this approach to perform well – In what context and compared to what? And how is this judged?

6. In the discussion the author makes some attempt at what to recommend in practice, i.e. using this approach as a sensitivity analysis, but it is not quite clear what outliers they recommend inflating the variance for or what results to present. Should you present the results from the standard RE MA and then the RVSOM allowing for inflated variances in all of the observations indentified as outliers by the LRT? What is the advice if the treatment effect differs somewhat in the two?

7. (Page 10, paragraph 3) I was a bit confused by the reference to RSOM. Firstly
what does this stand for “Random shifted outlier model”? Does this make sense in terms of a fixed effects model? Does it make sense to assume homogeneity when some estimates are clearly heterogeneous to others?

8. How do the results from this method of analysis compare to those from other methods of allowing for outliers e.g. refs 4 and 23? What is the advantage of this approach?

Minor Essential Revisions

1. In the introduction (page 2) the author talks about “studies that are not representative of the population of interest” having “unfortunate implications”. Firstly if a study if not representative of the population of interest should it really be included in the meta-analysis? There is quite a difference between not being the right population and being the right population but with a strange result and it would be good to clarify this. Also, what does the author mean by “unfortunate implications”?

2. (Page 2) What is the author referring to in “This type of exercise is highly recommended…”? Looking for outliers and then what?

3. It would be good to elaborate a little more in the introduction about the difference between outliers and influential observations, influential observations seem to be mentioned out of the blue at the bottom of page 2.

4. (Bottom of page 3) “We propose a parametric bootstrap procedure… of these statistics,…” – what statistics?

5. (Bottom of page 3) “And to account for multiple testing” – what multiple testing and how do you account for it?

6. (Bottom of page 4) “Model (1) has he form…” – I think this should be “the form”!

7. (Page 6, paragraph 1) I found some of the sentences in this paragraph confusing and I wasn’t sure what the author was trying to say: “We do not however need to determine what constitutes a large wj2…” – why not? And “It is also possible for a large study to provide a relatively small wj2…” – this could be worded better. Also, I wonder whether some of this might be better in the discussion of the examples?

8. (Page 7, identifying and downweighting outliers) “…provide a larger LRT than the others…” – what others?

9. (Page 7) “If no such explanation can be found…” – sentence needs revising.

10. (Results, example 2, page 9) The author states that “this analysis shows that publication bias and outlying trial results are related concepts” – this seems like quite a big statement based on one example and does not seem relevant to the paper.
11. (Results, example 3, page 9) The author states that observation 38 is an obvious outlier although I am not sure I agree as its CI overlaps quite a lot with other CIs from other studies.

12. (Results, example 3, page 9) What does the author mean by “other results can be seen to be more easily explainable due to the variation provided by the random effects models”?

13. (Results, example 3, page 9) Does the final model presented as M1 it table 2 include 3 wj2 terms in a single model? This could be made explicit.

14. (Page 10, paragraph 1) “…the apparent outliers only had serious implications for the CPD-choline analysis” – this is very vague and could be more specific.

15. (Page 10, paragraph 2) The author states that “Determining which studies might be designated as outliers is difficult from the visual inspection of plots”, however I disagree, all of the outliers picked up by the RVSOM approach are the ones that appear as outliers on the forest plot.

16. (Figures 1, 3 and 5) These figures would be more informative if the mean value was represented by the weight of the study in the standard MA.

17. (Figures 2, 4 and 6) I found the titles of the figures very confusing.
   - there is no b
   - figure 2 talks about study i
   - It would be good to include possible to include CIs, particularly for figures b and c so that we would see how much mu and tau2 differ under the different models.
   - I am not sure I follow what the lines on figure d represent?

18. (Tables) Define M0 and M1

19. More generally there are a number of places where the author uses “this” or “these” and it is unclear what the author is referring too.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.