Author's response to reviews

Title: The importance of adjusting for potential confounders in Bayesian hierarchical models synthesising evidence from randomised and non-randomised studies: an application comparing treatments for abdominal aortic aneurysms

Authors:

C. Elizabeth McCarron (mccarrce@mcmaster.ca)
Eleanor M Pullenayegum (pullena@mcmaster.ca)
Lehana Thabane (ThabanL@mcmaster.ca)
Ron Goeree (goereer@mcmaster.ca)
Jean-Eric Tarride (tarride@mcmaster.ca)

Version: 2 Date: 28 May 2010

Author's response to reviews: see over
Dear BMC Medical Research Methodology:

Enclosed are the responses to the reviewer comments for the manuscript MS:

Changes in the manuscript are identified using Track Changes.

We look forward to your final approval.

Best regards,

C. Elizabeth McCarron, MA, MSc, PhD Candidate

Eleanor M Pullenayegum, PhD

Lehana Thabane, BSc, MSc, PhD

Ron Goeree, MA

Jean-Eric Tarride, PhD
Responses to Reviewer Comments

Richard F macLehose

Comment 1. The data example is somewhat unclear. The authors set up their model to assume they have summary data on each exposed and unexposed groups in each study. In their meta analysis, if there were (for example) age difference in a given study it would seem natural for the authors of the original study to adjust for age and report an adjusted OR. In this sense, it seems artificial to revert back to the crude data. Is it the case that although age differences existed between treatment groups, no adjustment was done?

1. The data were taken from a published systematic review (Hopkins et al. 2008) in which no adjustment was made for covariate imbalances. For clarity, the following sentence has been added on page 12.

   i. “No adjustment was made for imbalances in the original study [12].”

Comment 2. The model adjusts for the average difference in covariate level between treatment and control groups. The presence of this difference doesn’t actually imply confounding, nor is it sufficient to control for only this difference to correct for confounding. The magnitude of confounding depends not only on this imbalance, but also on the prevalence of the confounder and the effect of the confounder on the outcome. In your example, it could be that the age difference between exposed and unexposed is constant across studies but that age causes different amounts of bias because the distribution of age differs across studies.

2. In order to address this comment, we have included the following statement in the discussion section (p.18).

   i. “Also, in its current form the proposed model does not address the extent to which variation in age, gender, and cardiac disease across studies may explain variation in study estimates. Rather, the objective of this study was to propose a method to adjust for differences in patient characteristics within studies, as a way of controlling for potential confounders.”

Comment 3. I’m not convinced your comparisons with Prevost are accurate. As you present her/his model it seems not to control for confounding but to allow the effect to vary based on the some other covariate (age in this case). This is what epidemiologists tend to think of as effect modification. That your methods arrive at different answers may not be terribly surprising given you seem to be attempting to do different things.
3. Prevost’s model as presented on p.7 is included to demonstrate how the author extended the three-level Bayesian model to adjust for covariates to explain differences in mean effects. We used a similar approach to adjust for imbalances in patient characteristics between study arms. We then compared the results adjusted for differences to the results adjusted for aggregate study values (i.e., ‘naive’ covariate adjustment) in order to illustrate that the latter approach does not capture the imbalances that could bias the results. As the reviewer states the results may not be wholly unexpected as the models may be attempting to do different things, but we feel the comparison is still worthwhile. No revisions were made to the text.

**Background:**

**Comment 4. Paragraph 1:** “…the treatment effects estimated from RCTs are expected to be less subject to the potential for confounding…” In expectation, there will not be confounding. In any given study there might be. I would suggest making this sentence a bit more precise, since it includes the word “expected.”

4. Paragraph 1: Sentence has been edited.

   i. Original sentence: “Consequently the treatment effects estimated from RCTs are expected to be less subject to the potential confounding effects of extraneous variables [3].”

   ii. Edited version: “Consequently the treatment effects estimated from RCTs are less subject to the potential confounding effects of extraneous variables [3].”

**Comment 5. Paragraph 1:** I would include the issue of compliance to randomization as another reason why observational studies might actually be preferable to RCTs in some cases.

5. Paragraph 1: Text has been added.

   i. “In some cases, compliance to randomisation, among the randomised studies, might also be an issue.”
Methods:

Comment 6. Page 7, bottom: it might be better to keep the additional covariate as a generic variable, say $x_{ij}$, rather than age$_{ij}$.

6. Page 7, bottom: We have renamed the additional covariate $x_{ij}$.

Comment 7. Page 8, line 4-7: this is not an accurate definition of confounding. Notice that your definition would include exogenous variables that affect only the outcome and not the exposure. Such a variable would not, typically, be considered a confounder in epidemiology. It also leaves open intermediate variables. Modern Epidemiology, 3rd edition will provide a more accurate description.

7. Page 8, line 4-7: Description of confounding has been changed in accordance with the definition provided in Modern Epidemiology 3rd edition.

   i. Original description: “Only when the groups being compared are the same in every respect that determines outcome (other than treatment) can an experimenter be certain that any observed differences between the groups are due to treatment and not the result of the confounding effects of extraneous variables.”

   ii. Edited version: “One potential source of bias is confounding [15], where an extraneous factor is associated with both the exposure under study (e.g., treatment) and the outcome of interest, but is not affected by the exposure or outcome [16]. Only when the groups being compared are balanced in all factors, both those that can be measured and those that cannot, that are associated with exposure and that affect the outcome (other than treatment) will it be certain that any observed differences between the groups are due to treatment and not the result of the confounding effects of extraneous variables.”

Paul Gustafson
Comment 1. p. 9: there is some confusion with notation and equations here. First, eq. (7) seems to be missing a summation sign (over m). Second, in the text m is referred to as a confounder but used as an index (of the M confounders). For instance, I think the intended meaning is that $x_{mTij}$ and $x_{MCij}$ are "the values of the m-th potential confounder..." (same issue arises on p. 10).

1. Page 9: Summation sign has been added to eq. (7). Reference to m as both a confounder and an index has been clarified by replacing m potential confounder with m-th potential confounder.

Comment 2. p. 9 and elsewhere: more interpretable (particularly for non-statistical readers) to report standard deviations instead of variances?

2. We were concerned that though standard deviations may be more interpretable, in the current context the values were such that they might be distracting. For example, a normal prior with a mean zero and variance 1000 would now read a normal prior with mean zero and standard deviation 31.62. Someone reading the paper may wonder why a value of 31.62 would be chosen. Therefore, we reported the variances. No revisions were made to the text.

Comment 3. p. 9 and elsewhere: is "truncated to be positive" clearer than "truncated below zero"?

3. Pages 10,13-14: Truncated below zero has been replaced with truncated to be positive.

Comment 4. p. 10, 2nd sentence of 2.3: what variance? There are many floating around in your set-up.

4. Pages 10-11: To clarify variances the following information has been added:

   i. This approach increased the relative proportion of the between-study-type variance ($\tau^2$) associated with the non-randomised studies compared to the randomised studies.

   ii. Sutton et al. [8] centred their informative prior for the population mean on the non-randomised pooled estimate but used a between-study variance for the non-randomised studies ($\sigma_2^2$) that was four times larger than that of the randomised studies ($\sigma_1^2$).

Comment 5. p. 10: truncating the prior to force $|\mu-th1|<|\mu-th2|$ makes intuitive sense, but also would seem to mess-up the interpretation of mu?
5. Pages 10-11: The following clarification has been added on pages 10-11:

   i. “In so doing the interpretation of $\mu$ is altered. Since the constraint gives more weight to the randomised studies, $\mu$ no longer represents the total population studied.”

Comment 6. p. 12: Can’t write sigma $\sim N(0,???)$ etc. without being explicit that you mean half-normal, not real-normal.

6. Page 13: The distributions $(\sigma_i, \tau \sim N(0,0.26))$ $(\sigma_i \sim N(0,0.125), \tau \sim \text{half-normal}(0,0.033))$ have been replaced by $(\sigma_i, \tau \sim \text{half-normal}(0,0.51^2))$ $(\sigma_i \sim \text{half-normal}(0,0.36^2), \tau \sim \text{half-normal}(0,0.18^2))$.

Comment 7. p. 13 and forward, re: multiple imputation. Here is my biggest methodological concern. First, the standard interpretation of “multiple imputation” is that one imputes $m>1$ datasets, fits the complete-data model to each one, and then synthesizes the $m$ inferences. However, the current manuscript reads as if multiple imputation produced 1 dataset, so I'm confused about what was actually done. Second, in a paper that is ostensibly about evidence synthesis, the 2-stage nature of the procedure (impute one (or more) datasets in R, dump into WinBUGS) feels inelegant. The more natural solution (though admittedly there may be devil in the details) would be to include the unobserved $x$ values along with the unobserved parameters inside the MCMC. Of course this would necessitate modeling the joint distribution of $x$, but multiple imputation requires this as well. At the very least, more clarity and discussion is needed around this issue.

Page 14: The following information has been added to clarify the multiple imputation model:

i. “Multiple imputation was conducted using R 2.9.2 software [20] assuming that the covariates were missing completely at random.”

ii. “This approach implemented the bootstrap method to first impute values for each missing variable by randomly selecting from the observed outcomes of that variable and then generated multiple imputations (three datasets) using iterative regression imputation, looping through until approximate convergence. The result was a single imputed dataset of 79 studies (four randomised and 75 non-randomised) which was then analysed, in WinBUGS, adjusting for imbalances in age, gender, cardiac disease, pulmonary disease, and renal disease.”
Page 19: The following has been added to the discussion section:

ii. “Assuming that the covariates were missing completely at random, the current analysis attempted to impute the missing values, though admittedly the two-stage nature of the current approach may appear inelegant (i.e., using R to impute the data and then analysing the new data in WinBUGS). A more natural solution would be to include the unobserved covariate values along with the unobserved parameters inside the MCMC, although this may add an additional layer of complexity. Due to the focus of the paper being Bayesian hierarchical models for combining randomised and non-randomised studies rather than methods to impute missing data, and for convenience, we decided to generate the missing values using R.”

Comment 8. p. 16: a small point, but phrases like "estimated median value of their posterior distribution" disturb me. The "estimated" makes one think of statistical estimation, when of course you really mean "numerically approximated" (via MCMC). I find a lack of clarity between estimation error and numerical/simulation error is a problem in a lot of reported Bayesian analysis.

7. Page 16: Estimated was replaced by numerically approximated.

Comment 9. p. 18: I agree with the sentiment about better reporting characteristics of study populations. As a vaguely related point, and in the spirit of open-access, reproducible research, etc., I hope you will post your data on the web, so that others can reproduce your findings and try alternate analyses.

8. Page 11: Release of data

We added to the paper that the data are available upon request (the codes are provided in the Appendix). I am currently using this data as part of my PhD thesis and have not yet finished using this dataset. The data will be posted on the web as soon as my PhD work is completed (within 9-12 months).