Reviewer's report

Title: Synthesis of clinical prediction models under different sets of covariates: a multivariate meta-analysis approach

Version: 2
Date: 25 June 2015
Reviewer: Richard Riley

Reviewer's report:

This is an important topic, and the question is very relevant for those in the prediction model field. However, I have some concerns about the methods and paper under review.

My major comment is that this does not extend the existing literature. Where systematically missing confounders (adjustment factors) are not available in some studies, a multivariate meta-analysis approach has already been proposed to borrow strength from partially adjusted results. Please see [1] where this is proposed and also [2] where this is demonstrated. I therefore do not see what this paper is added, apart from looking at the change in C-statistic and Brier score in the example.

Also, I did not find the paper easy to follow. Though the question is well explained early on, I cannot easily see if the methods how the multivariate approach is accounting for studies that do not present all covariates of interest. I also do not follow the rationale for the non-linear modelling of the covariates. Why not just assume normality of the original scale of the alpha and beta coefficients? This is what the authors did in [1], as the regression coefficients are MLE estimates and so should be asymptotically normal I think?

Critically, the multivariate approach presented does not appear to be a random effects model. I doubt whether the fixed effect approach is realistic.

The authors repeatedly say that the method that ignores the studies without the full set of covariates is biased or wrong. Why? If these are a random sample, then why would they not give unbiased regression coefficients and thus the right prediction model (on average)?

The sandwich estimator approach is interesting for dealing with mis-specified within-study correlations, but again how does this build on the existing literature such as [3]?

The simulation studies show minimal improvement in the performance of the prediction model. This is expected, given that I think the regression coefficients should be unbiased for all approaches. The authors note this in their discussion, but say that there is considerable gain in precision. This is interesting – but again a finding that is already well known in this field, shown in many papers that the authors do not seem to be aware of, such as [4-6] (sorry to be referencing a
number of my own papers here, but they do seem very relevant)

Finally, the prediction model’s performance considers discrimination but not calibration explicitly (e.g. calibration slope or calibration in the large), so this is not a complete evaluation of the model performance statistically [7]

Thus – though the authors are covering an important topic and have clearly worked hard to get this far – I am rather left confused as to what the article adds to the current literature and existing knowledge. I hope my comments help the authors going forward.


Level of interest: An article whose findings are important to those with closely related research interests

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 regularly publish methodology work in multivariate meta-analysis and receive funding to my Institution from the MRC for developing multivariate meta-analysis methods.