Reviewer’s report

Title: The calibrated model-based concordance improved assessment of discriminative ability in patient clusters of limited sample size

Version: 0 Date: 18 Dec 2018

Reviewer: Georg Heinze

Reviewer’s report:

The authors investigate the problem of validation of a risk prediction model using multicenter data, using discrimination indices (concordance index, c-index) computed in each center and then probably meta-analyzed. They argue that instead of estimating a c-index in each center separately, it may advantageous to borrow strength across centers and use a multilevel logistic regression model to recalibrate the predictions. The recalibrated predictions, in fact the linear predictors, can then be used to estimate a calibrated version of the so-called model-based concordance index (c-mbc). The c-mbc is shown to have smaller variance than the cluster-specific c-index, while it is biased towards to overall mean c-index. In total, the smaller variance pays off in terms of mean squared error.

Title, abstract: you should state upfront if this article is about a continuous, a binary or a time-to-event outcome (or general).

P. 3, line 64: what do you mean by the assumption that ’regression coefficients are correct’? Unbiased? Correct model specification? Please clarify.

p. 4, line 81: likewise, what do you mean by ’... that the random effect estimates of the calibration intercept and slope are true’? Please clarify.

Simulation study:

The simulation was performed under the assumption that both the model and the recalibration are correctly specified. Unless the simulation is extended, we cannot draw conclusions on the method if one or both models are not correct. The cluster-specific non-parametric c-index is unbiased irrespective of correct model specification. That should be kept in mind if the model specification is in doubt.

==> I would suggest to include in the simulation at least one scenario where the model assumptions are violated.

Were the cluster intercept and slope drawn independently? It is likely that they have some negative correlation in practice. Also, in the example, how about the correlation of calibration intercept and calibration slope?
Just to be sure, were the cluster-specific intercepts and slopes drawn once and then kept fix, or freshly drawn in each of the 2000 replications of the simulation study? (I assume the former, but it is not explicitly stated.)

How about non-normal distributions of the calibration slope and intercept? It may happen that there are two or three types of clusters (e.g., clusters where the prediction formula is more or less misspecified). How does the method perform if the distributional assumption is not fulfilled (but nevertheless a normal distribution assumed)?

To estimate the cluster-specific c-index, did you use the mbc version or the plain c-index (which should be unaffected of recalibration)?

Please check reference 18 in the reference list.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.
I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal