Reviewer's report

Title: The uncertainty with using risk prediction models for individual decision making: an exemplar cohort study examining the prediction of cardiovascular disease in English primary care

Version: 0 Date: 25 Apr 2019

Reviewer: Huining Kang

Reviewer's report:

This is an interesting manuscript that presents an analysis of the uncertainty in predicting the individual CVD risks due to the different modelling strategies. The manuscript has provided sufficient statistical evidence that could heighten the reader's awareness of uncertainty in the prediction of individual risk scores, which is important in clinical decision making. Although the methodology in general is good, I have the following concerns:

1. In the Validation of the models section, the authors used so-called split-sample (or training/test) approach to generate calibration plot (Figure 1), which is appropriate because it gives an unbiased evaluation of the performance. However it is not clear whether all the performance metrics in Table 3 were calculated using split-sample. If so this should be indicated in that section. If the performance metrics were calculated using so called "apparent validation" (also known as "resubstitute validation"), then the estimates in the table are all biased and I would suggest re-do the table using split-sample approach.

2. Tables 4 and 5 were calculated based on the predicted risk scores of the samples that are actually used to fit the models. But what we are more interested in is the uncertainty when the models are used to predict the risk scores of new samples. Should be there a difference in the extent of uncertainty between predicting the risk score of a sample that was used in fitting the model and predicting that of a new sample? Since the sample size is huge in this investigation, it makes more sense to use split-sample (training/test set) approach, again. That is, to use the training set to fit the models and then use the models to predict the risk scores for only individuals in the test set. Perform the uncertainty analysis using the predicted scores of only the samples of the test set.

3. The manuscript is exploring the uncertainty around the risk score beyond the confidence interval. But I think the author would agree that the uncertainty shown in Tables 4 and 5 can be partly explained by the "confidence interval" of model prediction. But to what extent do the confidence intervals contribute the total uncertainty? I would suggest to provide a 95% confidence interval for the predicated risk score of Model A for each of the risk groups of width 1% in Table 4 if it is possible?
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Not applicable

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

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 declare that I have no competing interests

**Statement on potential review bias**

Please complete a statement on potential review bias, considering the following questions:

1. Did you co-author any publication with an author of this manuscript in the last 5 years?

2. Are you currently or recently affiliated at the same institution as an author of this manuscript?

If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria'. If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria

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/](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