Reviewer’s report

Title: External validation, update and development of prediction models for pre-eclampsia using an Individual Participant Data (IPD) meta-analysis: the International Prediction of Pregnancy Complication Network (IPPIC Pre-eclampsia) protocol

Version: 0 Date: 17 Jul 2017

Reviewer: Ben Van Calster

Reviewer's report:

This is a nice and elaborate protocol, that addresses and describes the study plans in appropriate fashion most of the times. I have the following comments, many only ask for clarification but some go a bit further.

P6, I do not understand 'predefine cut-off values for clinical parameters'.

Same page, reference 16 (TRIPOD) is not on IPD MA?

P7, first primary objective: is this meant to involve univariable analysis of each marker?

P10: I assume that the collaborative network already has some IPD datasets through its members, perhaps this can be described?

Same page: what do you mean with secure cyberspace?

Same page, 'the original variable will be replaced with a proxy' sound weird. The advantage of IPD was stated as being able to standardize the definitions of predictors, but this appears to be in contradiction with this sentence? Can you clarify?

P13, improving normality is not necessary by default?

Same page, why will complete case analysis be employed here?

P15, you will impute missing outcome values, but will you use them in the analysis (MI vs MID)?

MoMs: This is an unfortunate issue. I never really understood the relevance of MoMs. Anyway, you seem to suggest that this may render some predictors useless due to different adjustment/scales. Do you expect this to be a major problem affecting many variables?

P16, how will you adjust for the biomarker assay and platform?
P18, the meaning of calibration in-the-large and calibration slope is incorrectly described. E.g. the calibration slope indicates 'under- or over-prediction' is more a description of calibration in the large. Perhaps the wording was unfortunate?

P20, with which instrument will risk of bias be assessed?

Meta-regression: I think this can only be planned as a descriptive analysis.

Why wasn't DCA also mentioned in the section on external validation?

I do not understand how using study-specific intercepts will help to update a model with good overall discrimination and calibration performance but with heterogeneity in calibration. This will mainly be helpful for included studies/centers, but not for other centers?

Why will you not consider recalibration of models with good discrimination but poor overall calibration?

You should make a statement about events per variable (EPV) in the section on the development of new models.

New models will be developed using a separate intercept for each study. Do you mean that a random intercept model will be used?

New models will also check for heterogeneity in predictor effects: Will you check this using random slopes? This would impact on EPV because you need to estimate another parameter. Further, how will this be reconciled with MFP? Will you first drop predictors with too heterogeneous effects, and then perform MFP?

You frequently discuss study-specific intercepts, but what if one study involves data on multiple centres? Isn't it better to use centre-specific intercepts?

IECV: if you use a random intercept model, and then apply that model to one study that was left out, you probably have to use the average (fixed) intercept. This is likely to have an impact on calibration, so is this method fully valid?

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 have a co-authorship with one co-author of this study.

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