Development and external validation of prediction models for pre-eclampsia: An Individual Participant Data (IPD) Meta-analysis

John Allotey[1], Kym Snell[2], Richard Riley[2], Shakila Thangaratinam[1, 3] for the IPPIC Collaborative Network*

[1]Barts Research Centre for Women's Health (BARC), Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[2]Centre for Prognosis Research, School of Primary, Community and Social Care, Keele University, Keele, UK
[3]Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

Background
Pre-eclampsia is the most predicted obstetric outcome, with more than 130 prognostic models developed.[1] A quarter of these have been externally validated, and showed only modest predictive performance, characterised by methodological shortcomings in development including overfitting of models, small event numbers in development datasets and predictors not varied enough to adequately capture the differences between women.

Access to IPD from multiple studies will provide increased sample size with more outcomes to evaluate several candidate prognostic factors, beyond what would have been possible in a single study, to subsequently develop clinically relevant and robust models. It will also enable the evaluation of any prediction model developed across different settings and population case-mix.

Aims
To develop and validate pre-eclampsia prediction models using IPD from multiple studies.

Methods
Logistic regression with a random intercept to account for clustering by study for model development. Internal-external cross-validation using random-effects meta-analysis to summarise performance measures across studies.

Results
The International Prediction of Pregnancy Complications (IPPIC) network[2] is a group of 125 researchers contributing data of 3,674,684 pregnancies from 78 datasets. Twelve prediction models were developed, four each for any, early and late-onset pre-eclampsia.
3-11 datasets were used to develop each model depending on the availability of predictors within datasets. Average models discrimination were good (0.68-0.83), however calibration performance was heterogeneous across datasets. The models showed the highest net-benefit for predicted probability thresholds in nulliparous women at thresholds above 5%.

Conclusions
The IPPIC models on average showed promising predictive performance. However before application in practice, recalibration of model parameters to particular populations and settings may be needed. Additional predictors may improve the predictive performance of the models.

Keywords
Pre-eclampsia, prediction model, individual participant data

References
[1]Allotey J, et al. Accuracy of clinical characteristics…using IPD meta-analysis. *HTA 2020 (in press)*
[2]Allotey J, et al. External validation, update...the IPPIC pre-eclampsia Network protocol. *DAPR* 2017;1:16.