Validation and development of models using clinical, biochemical and ultrasound markers for predicting pre-eclampsia: an individual participant data meta-analysis

John Allotey,[1,2†] Kym IE Snell,[3*†] Melanie Smuk,[2] Richard Hooper,[2] Claire L Chan,[2] Asif Ahmed,[4] Lucy C Chappell,[5] Peter von Dadelszen,[5] Julie Dodds,[1,2] Marcus Green,[6] Louise Kenny,[7] Asma Khalil,[8] Khalid S Khan,[1,2] Ben W Mol,[9] Jenny Myers,[10] Lucilla Poston,[5] Basky Thilaganathan,[8] Anne C Staff,[11,12] Gordon CS Smith,[13] Wessel Ganzevoort,[14] *et al.*[‡] on behalf of 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]Pragmatic Clinical Trials Unit, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[3]Centre for Prognosis Research, School of Primary, Community and Social Care, Keele University, Keele, UK
[4]Aston Medical Research Institute, Aston Medical School, Aston University, Birmingham, UK
[5]Department of Women & Children's Health, School of Life Course Sciences, Faculty of Life Sciences & Medicine, King's College London, London, UK
[6]Action on Pre-eclampsia (APEC), Evesham, UK
[7]Vice Chancellor's Office, Faculty of Health & Life Sciences, University of Liverpool, Liverpool, UK
[8]Fetal Medicine Unit, St George's University Hospitals NHS Foundation Trust and Molecular and Clinical Sciences Research Institute, St George's University of London, London, UK
[9]Department of Obstetrics and Gynaecology, Monash University, Monash Medical Centre, Clayton, VIC, Australia
[10]Maternal and Fetal Health Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester University NHS Foundation Trust, Manchester, UK
[11]Division of Obstetrics and Gynaecology, Oslo University Hospital, Oslo, Norway
[12]Faculty of Medicine, University of Oslo, Oslo, Norway

[13]Department of Obstetrics and Gynaecology, NIHR Biomedical Research Centre, University of Cambridge, Cambridge, UK
[14]Department of Obstetrics, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands

*Corresponding author k.snell@keele.ac.uk
†Joint first authors (both contributed equally)
‡The full list of authors can be found in *Appendix 1*.
§The full list of partners in the IPPIC Collaborative Network can be found in *Acknowledgements*.

Declared competing interests of authors: Gordon CS Smith has received research support from Roche Holding AG (Basel, Switzerland) (supply of equipment and reagents for biomarker studies of ≈£600,000 in value) and Sera Prognostics (Salt Lake City, UT, USA) (≈£100,000), and has been paid by Roche to attend an advisory board and to present at a meeting. He is a named inventor on a patent filed by Cambridge Enterprise (UK Patent Application Number 1808489.7, 'Novel Biomarkers') for the prediction of pre-eclampsia and fetal growth restriction. Ignacio Herraiz reports personal fees from Roche Diagnostics and Thermo Fisher Scientific (Waltham, MA, USA). John Kingdom reports personal fees from Roche Canada (Mississauga, ON, Canada). Lucy C Chappell is chairperson of the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) CET Committee (January 2019 to present). Asma Khalil is a member of the NIHR HTA Board (November 2018 to present). Jane E Norman is a member of the NIHR HTA MNCH Panel, and she reports grants from NIHR and Chief Scientist Office Scotland, as well as consultancy fees from and participation in data monitoring committees for Dilafor AB (Solna, Sweden) and GlaxoSmithKline (Brentford, UK). Kajantie Eero reports grants from the Academy of Finland, the Foundation for Paediatric Research, the Signe and Ane Gyllenberg Foundation (Helsinki, Finland), the Sigrid Jusélius Foundation (Helsinki, Finland), the Juho Vainio Foundation (Helsinki, Finland), the European Commission, the NORFACE DIAL Programme, the Novo Nordisk Foundation (Hellerup, Denmark), the Yrjö Jahnsson Foundation (Helsinki, Finland), Foundation for Cardiovascular Research (Zürich, Switzerland) and the Diabetes Research Foundation. Ben W Mol reports fellowship from the National Health and Medical Research Council (Canberra, ACT, Australia), personal fees from ObsEva (Plan-les-Ouates, Switzerland), personal fees and consultancy fees from Merck Sharp & Dohme (Kenilworth, NJ, USA), personal fees from Guerbet (Villepinte, France), travel funds from Guerbet and grants from Merck Sharp & Dohme. Richard D Riley reports personal fees from the British Medical Journal for statistical reviews, and from Roche and the universities of Leeds, Edinburgh and Exeter for training on individual participant data meta-analysis methods. Jacques Massé reports grants from National Health Research and Development Program, Health and Welfare Canada, during the conduct of the study. Paul T Seed is partly funded by King's Health Partners Institute of Women and Children's Health, Tommy's (registered charity number 1060508) and ARC South London (NIHR). The views expressed are not necessarily those of KHP, Tommy's, the NHS, the NIHR or the Department of Health.

Published December 2020
DOI: 10.3310/hta24720

Scientific summary

IPD meta-analysis of pre-eclampsia markers

Health Technology Assessment 2020; Vol. 24: No. 72
DOI: 10.3310/hta24720

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Pre-eclampsia is a leading cause of maternal and perinatal mortality and morbidity. Current methods of risk assessment for pre-eclampsia are based mainly on clinical history alone and have limited accuracy. Prediction models that incorporate additional information on biochemical and ultrasound markers could improve the predictive performance. Numerous multivariable pre-eclampsia models have been developed to date, but only a few have been externally validated, and none is recommended for use in routine clinical practice. Robust data are needed to externally validate existing models to determine their transportability across new populations and their clinical utility.

Objectives

Primary

The primary objectives were to use individual participant data meta-analysis:

- to validate (across multiple populations and settings) existing models for predicting early-onset, late-onset and any-onset pre-eclampsia based on clinical characteristics only, clinical and biochemical markers, clinical and ultrasound markers, and clinical, biochemical and ultrasound markers
- to develop and validate (across multiple populations and settings) multivariable prediction models for early-onset, late-onset and any-onset pre-eclampsia where existing prediction models have limited performance, or where no such models exist for the relevant pre-eclampsia outcomes
- to estimate the net benefit (clinical utility) of existing and new models to inform clinical decision making based on thresholds of predicted risk
- to estimate the prognostic value of individual clinical, biochemical and ultrasound markers for predicting pre-eclampsia.

Secondary

- To assess the differential performance of the existing models in various predefined subgroups based on population characteristics (unselected; selected) and timing of model use (first trimester; second trimester).
- To study the added accuracy when novel metabolic and microRNA-based biochemical markers are added to the developed model based on clinical, ultrasound and biochemical markers.

Methods

We undertook an individual participant data meta-analysis in line with existing recommendations on prognostic research model development and validation and complied with reporting guidelines for prediction models and individual participant data meta-analysis. We undertook relevant systematic reviews to identify systematic reviews on clinical characteristics, biochemical and ultrasound markers for prediction of pre-eclampsia; prediction models for pre-eclampsia; and relevant studies, birth cohorts or data sets. Primary studies and large birth and population-based cohorts that provided relevant information for assessing the accuracy of clinical, biochemical and ultrasound predictors of pre-eclampsia were included. The primary outcomes were early-onset (delivery at < 34 weeks' gestation), late-onset (delivery at ≥ 34 weeks' gestation) and any-onset pre-eclampsia. We established the International Prediction of Pregnancy Complications collaborative network, and researchers from this group shared their primary data, which required extensive cleaning, standardisation and quality checking.

We externally validated published pre-eclampsia prediction models that reported the full model equation in International Prediction of Pregnancy Complications UK data sets. Partially missing predictors or outcome values missing for < 95% of individuals in a data set were multiply imputed under the missing at random assumption using multiple imputation by chained equations. Imputation was carried out separately in each dataset to account for the clustering of individuals within a data set. The predictive performance of each model was examined using measures of discrimination (*C*-statistics; no discrimination 0.5 to perfect discrimination 1, with values of ≥ 0.7 deemed most promising) and calibration of predicted to observed risks (calibration slope, with an ideal value of 1; and calibration-in-the-large, with an ideal value of 0) first in the individual participant data for each available data set and then across data sets at the meta-analysis level. We also compared the clinical utility (net benefit) of validated prediction models for each pre-eclampsia outcome using a decision curve analysis.

We then developed and validated new prediction models for early-onset, late-onset and any-onset pre-eclampsia based on clinical characteristic variables alone, clinical characteristics and biochemical markers, and clinical characteristics and ultrasound markers. For each model developed, we summarised the data set-specific performance (*C*-statistic, calibration slope and calibration-in-the-large), using a random-effects meta-analysis, in terms of the average performance and (to examine potential generalisability across settings) the heterogeneity in performance. We also assessed the clinical utility of developed models using a decision curve analysis.

Outside model development, we also used the full International Prediction of Pregnancy Complications data set to obtain summary-unadjusted estimates of the prognostic effects of prioritised candidate predictors for early-onset, late-onset and any-onset pre-eclampsia, along with 95% confidence intervals and 95% prediction intervals, using a two-stage individual participant data meta-analysis of complete cases of singleton pregnancies. The two-stage approach involves first fitting a logistic regression model for each study and then pooling the log-odds ratios using a conventional random-effects meta-analysis. Clustering of participants within data sets was accounted for by analysing each data set separately in the first stage.

Results

One hundred and twenty-five researchers from 73 teams in 25 countries joined the International Prediction of Pregnancy Complications network (by October 2017) and provided access to anonymised individual data of 3,674,684 pregnancies (78 data sets). More than half of the data sets (58%, 45/78) were prospective cohort studies, 15% (12/78) were randomised controlled trials and 17% (13/78) were large prospective registry data sets or birth cohorts. One data set included individual participant data from 31 randomised controlled trials.

External validation of existing pre-eclampsia prediction models

Of the 131 models identified, 24 could be validated in one or more of the 11 International Prediction of Pregnancy Complications UK data sets. Eight models predicted any-onset pre-eclampsia (three on clinical characteristics only, three with additional biochemical markers and two with additional ultrasound markers), nine predicted early-onset pre-eclampsia (seven included clinical characteristics only, and one each included additional biochemical or ultrasound markers), and seven predicted late-onset pre-eclampsia (five included clinical characteristics only, and one each included additional biochemical and ultrasound markers). Discrimination performance of the models was modest, with summary *C*-statistics of around 0.6–0.7 for most models. Calibration was generally poor across the data sets, with large heterogeneity in performance across different International Prediction of Pregnancy Complications data sets, with most of the models demonstrating signs of overfitting (summary calibration slope of < 1) and predictions that were systematically too high or too low (calibration-in-the-large $\neq 0$, suggesting poor prediction of overall risk across populations). In most of the data sets, the net benefit of using the models was only slightly greater than the strategy of considering all women to have pre-eclampsia.

Development and validation of International Prediction of Pregnancy Complications pre-eclampsia prediction models

Twelve International Prediction of Pregnancy Complications pre-eclampsia models were developed: four each to predict any-onset, early-onset and late-onset pre-eclampsia (two models each in the first and second trimesters using clinical characteristics, and with additional biochemical markers). We developed each model by meta-analysing 3–11 International Prediction of Pregnancy Complications data sets. The clinical characteristics only models comprised maternal age, body mass index, parity, history of pre-eclampsia, hypertension, diabetes or autoimmune disease and systolic or diastolic blood pressure. In addition to the clinical characteristic predictors, the biochemical marker models included soluble fms-like tyrosine kinase-1, pregnancy-associated plasma protein A and placental growth factor.

For predicting any pre-eclampsia, all second-trimester models (clinical only, clinical and biochemical predictors) showed promising discrimination (average *C*-statistics of ≥ 0.7); first trimester clinical only, and clinical and biochemical models had summary *C*-statistics of 0.68 and 0.70, respectively. All models to predict early-onset pre-eclampsia had promising discrimination; the first trimester (clinical only, clinical and biochemical) models had summary *C*-statistics of 0.72 (95% confidence interval 0.59 to 0.82) and 0.76 (95% confidence interval 0.58 to 0.88) respectively; the corresponding values for second-trimester clinical only and clinical and biochemical models were 0.72 (95% confidence interval 0.60 to 0.82) and 0.83 (95% confidence interval 0.63 to 0.93). For predicting late-onset pre-eclampsia, the second-trimester models (clinical only, clinical and biochemical predictors) showed promising discrimination (average *C*-statistics ≥ 0.7); the first-trimester models' *C*-statistics ranged from 0.68 to 0.69. Summary calibration measures often had wide confidence intervals, and there was often large between-study heterogeneity in the calibration performance, particularly for clinical and biochemical marker models. The net benefit of the models varied across individual data sets, ranging from harm to very little benefit to no benefit.

When validated in individual cohorts with over 100 pre-eclampsia events, the first-trimester clinical model for any pre-eclampsia was well calibrated in the Baschat study (any pregnant women in the USA) (Baschat AA, Magder LS, Doyle LE, Atlas RO, Jenkins CB, Blitzer MG. Prediction of preeclampsia utilizing the first trimester screening examination. *Am J Obstet Gynecol* 2014;**211**:514.e1–7); the predictions were too high for individuals in the World Health Organization study (women with risk factors for pre-eclampsia from low-, middle- and high-income countries) (Widmer M, Cuesta C, Khan KS, Conde-Agudelo A, Carroli G, Fusey S, *et al.* Accuracy of angiogenic biomarkers at 20 weeks' gestation in predicting the risk of pre-eclampsia: a WHO multicentre study. *Pregnancy Hypertens* 2015;**5**:330–8) and low for those at high risk in the Pregnancy Outcome Prediction (POP) (nulliparous, singleton pregnancies in the UK) [Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. *Lancet* 2015;**386**:2089–97]. We observed a consistent net benefit for all International Prediction of Pregnancy Complications models when validated in the POP cohort for probability thresholds of $\geq 5\%$. Very little or no net benefit was observed in other data sets.

Summarising the unadjusted prognostic effect of individual predictors of pre-eclampsia

Any-onset pre-eclampsia

We observed a strong unadjusted association between any-onset pre-eclampsia and history of hypertension (odds ratio 4.76, 95% confidence interval 3.56 to 6.35; $I^2 = 98.39\%$), multiparity (odds ratio 0.88, 95% confidence interval 0.79 to 0.99; $I^2 = 96.6\%$), smoking during pregnancy (odds ratio 0.84, 95% confidence interval 0.76 to 0.93; $I^2 = 86.46\%$) and spontaneous mode of conception (odds ratio 0.73, 95% confidence interval 0.64 to 0.84; $I^2 = 58.67\%$), and increasing placental growth factor in the first (odds ratio 0.22, 95% confidence interval 0.09 to 0.50, $I^2 = 85.44$), second (odds ratio 0.66, 95% confidence interval 0.53 to 0.83; $I^2 = 87.27\%$) or third trimester (odds ratio 0.59, 95% confidence interval 0.45 to 0.77; $I^2 = 96.78\%$) showed a reduction in the odds of any-onset pre-eclampsia.

Early-onset pre-eclampsia

Increasing second-trimester measurement of uterine artery pulsatility index values had the strongest association with early-onset pre-eclampsia (odds ratio 14.73, 95% confidence interval 8.12 to 26.72; $I^2 = 60.11\%$). All statistically significant predictors had evidence of an increase in the odds of early-onset pre-eclampsia with increasing values, except placental growth factor measured in the first (odds ratio 0.08, 95% confidence interval 0.02 to 0.35; $I^2 = 55.69\%$) or second trimester (odds ratio 0.07; 95% confidence interval 0.01 to 0.43; $I^2 = 97.18\%$), which showed a decrease in odds with increasing values.

Late-onset pre-eclampsia

The strongest association with late-onset pre-eclampsia was observed for increasing uterine artery pulsatility index values measured in the second trimester (odds ratio 2.95, 95% confidence interval 2.31 to 3.76; $I^2 = 20.77\%$). Multiparity (odds ratio 0.87, 95% confidence interval 0.78 to 0.97; $I^2 = 95.16\%$) and increasing values of first (odds ratio 0.33, 95% confidence interval 0.16 to 0.68; $I^2 = 82.67\%$), second (odds ratio 0.81, 95% confidence interval 0.69 to 0.94; $I^2 = 76.39\%$) or third (odds ratio 0.68, 95% confidence interval 0.57 to 0.81; $I^2 = 93.60\%$) trimester measurement of placental growth factor and first-trimester soluble fms-like tyrosine kinase-1 (odds ratio 0.98, 95% confidence interval 0.97 to 0.99; $I^2 = 37.07\%$) showed a decrease in the odds of late-onset pre-eclampsia.

There was considerable heterogeneity for most prognostic effects, with wide 95% prediction intervals for the potential prognostic effect of factors in new populations.

Conclusions

Among the 24 existing prediction models that could be validated in individual participant data meta-analysis, their predictive performance was generally poor across data sets (both on average and in terms of heterogeneity in calibration of predicted risks with observed risks), with very limited evidence of clinical utility. Some of the heterogeneity in predictive performance of the models is likely due to different methods and timing of measurement, for example in blood pressure and biochemical marker values. Although the International Prediction of Pregnancy Complications models show promising predictive performance on average across data sets, heterogeneity across settings is likely in calibration performance. Ultrasound markers did not improve the predictive performance of the developed International Prediction of Pregnancy Complications clinical characteristic only models. The International Prediction of Pregnancy Complications pre-eclampsia models show consistent net benefit when applied to a cohort of singleton, nulliparous women in the UK. Before application in practice, calibration performance may need to be improved by recalibrating model parameters, such as the intercept, to particular populations and settings.

Recommendations for further research

Going forward, standardisation of measurement methods, for example across laboratories and hospitals, might reduce heterogeneity in calibration performance. A related point is that prediction models in this field need to be clearer with regard to how included predictors should be measured and exactly when this should occur. Validation, including examination of calibration heterogeneity, is still required for the models that we could not validate. The transportability of these and the International Prediction of Pregnancy Complications models needs to be assessed in multiple large data sets across different settings and populations, as does their acceptability to both women and health-care professionals. The impact of using the models in clinical practice needs to be evaluated beyond pre-eclampsia prediction to include the identification of women most at risk of other severe pregnancy complications. Updated models may be needed in local populations, using recalibration of the International Prediction of Pregnancy Complications models in local data sets, to improve calibration performance. Furthermore, additional strong predictors need to be identified to improve model

performance and consistency. New cohorts must standardise the predictors and outcomes measured, including their timing and measurement methods, to enable more homogenous data sets to be combined in individual participant data meta-analyses. In terms of the prognostic ability of particular factors, further analysis of the International Prediction of Pregnancy Complications data using multilevel multiple imputation for missing data and adjusting for confounders would provide a better evaluation of prognostic association.

Study registration

This study is registered as PROSPERO CRD42015029349.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 72. See the NIHR Journals Library website for further project information.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 14/158/02. The contractual start date was in December 2015. The draft report began editorial review in March 2019 and was accepted for publication in March 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).