European Journal of Obstetrics & Gynecology and Reproductive Biology 217 (2017) 119–125

Contents lists available at ScienceDirect

European Journal of Obstetrics & Gynecology and Reproductive Biology

journal homepage: www.elsevier.com/locate/ejogrb

Full length article

External validation of preexisting first trimester preeclampsia prediction models

Rebecca E. Allen[a,*], Javier Zamora[b], David Arroyo-Manzano[b], Luxmilar Velauthar[d], John Allotey[c], Shakila Thangaratinam[c], Joseph Aquilina[a]

[a] Barts Health NHS Trust, Royal London Hospital, Whitechapel, London, E1 1BB, United Kingdom
[b] Clinical Biostatistics Unit, Hospital Ramon y Cajal, (IRYCIS) Madrid, Spain and CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain
[c] Women's Health Research Unit, Multidisciplinary Evidence Synthesis Hub (mEsh), Centre for Primary Care and Public Health, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, United Kingdom
[d] Barts Health NHS Trust, Newham University Hospital, Plaistow, London, E13 8SL, United Kingdom

ARTICLE INFO

Article history:
Received 14 March 2017
Received in revised form 14 July 2017
Accepted 23 August 2017

Keywords:
Preeclampsia
Prediction models
Screening
Validation

ABSTRACT

Objective: To validate the increasing number of prognostic models being developed for preeclampsia using our own prospective study.
Study design: A systematic review of literature that assessed biomarkers, uterine artery Doppler and maternal characteristics in the first trimester for the prediction of preeclampsia was performed and models selected based on predefined criteria. Validation was performed by applying the regression coefficients that were published in the different derivation studies to our cohort. We assessed the models discrimination ability and calibration.
Results: Twenty models were identified for validation. The discrimination ability observed in derivation studies (Area Under the Curves) ranged from 0.70 to 0.96 when these models were validated against the validation cohort, these AUC varied importantly, ranging from 0.504 to 0.833. Comparing Area Under the Curves obtained in the derivation study to those in the validation cohort we found statistically significant differences in several studies.
Conclusion: There currently isn't a definitive prediction model with adequate ability to discriminate for preeclampsia, which performs as well when applied to a different population and can differentiate well between the highest and lowest risk groups within the tested population. The pre-existing large number of models limits the value of further model development and future research should be focussed on further attempts to validate existing models and assessing whether implementation of these improves patient care.

Introduction

Preeclampsia is a major cause of maternal and neonatal morbidity and mortality. There has been a rise in the number of prognostic models being developed in obstetrics, particularly for preeclampsia, however few have been widely implemented. Currently risk assessment is performed at the booking appointment based on maternal characteristics alone. This approach is thought to falsely classify two thirds of women as being high risk and in need of intensive monitoring and prophylactic aspirin therapy, highlighting the need for a better screening test [1].

Modern medicine is increasingly focusing on preventing disease. As a result it is important to identify patients at increased risk of illness. This could be based on a single risk factor or a combination of multiple predictors. Combining predictors into a prognostic model is likely to allow better risk assessment than the use of single risk factors.

In recent years there has been a massive amount of studies published looking at various biophysical and biochemical markers alone or in combination for the prediction of preeclampsia. A recent systematic review by Kleinrouweler et al. identified 69 prediction models for preeclampsia, only 5 of which had been externally validated and model performance was found to be lower when externally rather than internally validated. They recommended that systematic reviews should be performed to identify and validate existing models to decide whether a new model should be developed or an existing model updated. Following this

* Corresponding author at: Fetal Medicine Centre, 8th Floor, South Tower, Royal London Hospital, Whitechapel, London, E1 1BB, United Kingdom.
E-mail address: rebecca_e_allen@yahoo.co.uk (R.E. Allen).

http://dx.doi.org/10.1016/j.ejogrb.2017.08.031
0301-2115/

the better models should be applied in clinical practice and their influence on patient outcomes evaluated [2].

As a result of this review we have decided to externally validate pre-existing first trimester preeclampsia models using data obtained from our own prospective cohort to predict preeclampsia risk. We focussed on first trimester prediction models as interventions to reduce the risk of preeclampsia, mainly aspirin, have been shown to be more effective if commenced prior to sixteen weeks gestation. [3]

Materials and methods

Objectives

To validate clinical prediction rule for preeclampsia (PET) from models reported in obstetric literature [4–14].

Validation study population

Pregnant women attending for their dating scan at the Royal London Hospital were recruited to the validation study if they were between 11 and 14 weeks gestation. Women with multiple pregnancies and fetal anomalies were excluded. Ethics approval was obtained from the East of England research ethics committee. Information was collected on age, ethnicity, method of conception, parity, smoking, alcohol and drug use, past medical and obstetric history, family history and drug history. Maternal weight and height were measured and BMI calculated as was blood pressure and MAP. Uterine artery Doppler measurements were performed following the Fetal Medicine Foundation recommendations [15]. Maternal serum was taken and the biomarkers placental associated plasma protein-A (PAPP-A), placental growth factor (PlGF), beta-human chorionic gonadotrophin (β-hCG) and alphafetoprotein (AFP) measured.

Prognostic models and study selection

We searched Medline from inception until November 2016 without language restrictions. Full details of the search strategy are given in our previous paper [2]. Models were selected for validation if they assessed pregnant women in the first trimester. Predictors in the prognostic models included were maternal clinical characteristics, previous obstetric history, early pregnancy bleeding, systolic and diastolic blood pressure, first trimester biochemical markers (PAPP-A, PlGF, AFP, hCG, kisspeptin), uterine artery Doppler (resistance index, pulsatility index, unilateral or bilateral notching). Outcomes assessed by the models were early and late preeclampsia (the definition of which was determined by the authors of the papers, however all models that met our criteria for inclusion used 34 weeks gestation as a cut off). Models were excluded if they did not include 1 or more of the predictors and if they did not have a regression coefficient.

Quality assessment of prognostic models

The Quality in Prognosis studies (QUIPS) tool was used assess the risk of bias in the included prognostic studies [16].

Statistical analyses

We describe baseline characteristics of the validation cohort using median and interquartile ranges (IQR) and absolute and relative frequencies. We compared baseline clinical and demographic characteristics of the women recruited in the two cohorts used in the validation process. We used Mann-Whitney *U* tests and Chi-squared tests or Fisher exact tests when needed for these comparisons. The comparison of both cohorts was further extended to evaluate the predictive performance of the models within the two groups of women included in the validation cohort.

Validation of the models included in this review was performed by applying the regression coefficients that were published in the different derivation studies to our cohort. We obtained predicted probabilities of any PE for all models. Since the prevalence of PE on the derivation cohorts were slightly different to the validation one, we recalibrated the regression coefficients of the original models by computing a shrinkage correction coefficient. This allows adjustment of predicted probabilities to the prevalence of PE in the validation cohort [17]. We based all analyses on the shrunk probabilities.

For every model, we assessed its discrimination ability and its calibration. We assessed discrimination ability computing the Area under the ROC curve (AUC). We evaluated model calibration by the Hosmer-Lemeshow goodness-of-fit test.

All the analyses were carried out with Stata v13.

Results

Sample description

Women were obtained from 2 cohorts recruited over a 4 year period (2010–2014) at the Royal London Hospital by two separate researchers. We analyzed 2186 women (~12.5% of the population delivering over this time frame). A total of 2500 women were recruited. 314 were excluded due to incomplete outcome data, late miscarriages and fetal anomalies. A higher proportion of the pregnant population could not be recruited due to the limitations of only one researcher at a time recruiting. The incidence of all preeclampsia was 2.4% (n = 52). We could not separate our analysis between early and late onset disease due to the extremely low incidence of early onset preeclampsia in the validation cohort. The median age (IQR) was 29 years (26; 33), body-mass index 23.7 (21.1; 27.0) kg/m^2 and 802 (36.8%) had a BMI $\geq 25\,kg/m^2$. 824 (37.7%) women were Caucasian, 233 (10.7%) Black, 117 (5.3%) Chinese, 942 (43.3%) Indian or Pakistani, 64 (2.92%) mixed and 0.18% unknown. The population attending the Royal London is ethnically diverse with a large number of South Asian women. The high proportion of women recruited who were of a non-Caucasian ethnicity demonstrates that we managed to capture a large cross sample of our population which has hopefully avoided the introduction of bias (Table 1).

Among all analyzed women we observed 17 (0.8%) who had previous preeclampsia, 25 (1.1%) had pre-existing diabetes mellitus, 21 (0.9%) history of hypertension, 426 (19.5%) had a family history (mother or sister) of PE and 1006 (46.0%) were multiparous. Other variables are described in Table 2.

Study selection

Fig. 1 demonstrates selection of the models included for validation.

Quality assessment

Quality assessment was performed by two reviewers (JA) and (RA). There was a moderate risk of bias in the study participation domain for Akolekar but a low risk of bias for all the other domains and in all domains for the other studies [5]. See Table 3.

Validation

Models being validated are listed in Table 4. Sample sizes of these models ranged from 359 and 8189 with prevalence of PE