Performance of the first-trimester combined screening of preterm pre-eclampsia: results from cohort of 10 110 pregnancies in Spain

D. Cuenca Gómez[1,2], C. de Paco Matallana[3,4], V. Rolle[1,5], N. Valiño[6], R. Revello[7], B. Adiego[8], M. Mendoza[9], F. S. Molina[10], M. P. Carrillo[11], J. L. Delgado[3], A. Wright[12], B. Santacruz[1,2] and M. M. Gil[1,2]

1. Department of Obstetrics and Gynecology, Hospital Universitario de Torrejón, Torrejón de Ardoz, Madrid, Spain
2. Faculty of Medicine, Universidad Francisco de Vitoria, Pozuelo de Alarcón, Madrid, Spain
3. Department of Obstetrics and Gynecology, Hospital Clínico Universitario Virgen de la Arrixaca, El Palmar, Murcia, Spain
4. Institute for Biomedical Research of Murcia, IMIB-Arrixaca, El Palmar, Murcia, Spain
5. Biostatistics and Epidemiology Platform at Instituto de Investigación Sanitaria del Principado de Asturias, Oviedo, Spain
6. Department of Obstetrics and Gynecology, Complejo Hospitalario Universitario A Coruña, A Coruña, Galicia, Spai.
7. Department of Obstetrics and Gynecology, Hospital Universitario Quirón, Pozuelo de Alarcón, Madrid, Spain
8. Department of Obstetrics and Gynecology, Hospital Universitario Fundación de Alcorcón, Alcorcón, Madrid, Spain
9. Department of Obstetrics and Gynecology, Hospital Universitari Vall d'Hebrón, Barcelona, Catalonia, Spain
10. Department of Obstetrics and Gynecology, Hospital Universitario San Cecilio, Granada, Spain and Instituto de Investigación Biosanitaria Ibs., Granada, Spain
11. Department of Obstetrics and Gynecology, Hospital Universitario Virgen de las Nieves, Granada, Spain
12. University of Exeter, Exeter, UK

Corresponding authors:

Dr M. M. Gil

Vicerrectorado de Investigación, Facultad de Medicina, Universidad Francisco de Vitoria, Carretera Pozuelo a Majadahonda, Km 1.800, 28223 Pozuelo de Alarcón, Madrid, Spain

E-mail: mariadelmar.gil@ufv.es

Dr B. Santacruz Martín

Dirección Médica, Hospital Universitario de Torrejón, Calle Mateo Inurria S/N, 28850 Torrejón de Ardoz, Madrid, Spain
E-mail: bsantacruz@torrejonsalud.es

Short title: Performance of FMF first-trimester screening of pre-eclampsia in Spain

Keywords: first-trimester, screening, Fetal Medicine Foundation, pre-eclampsia, competing risk model, biomarkers

CONTRIBUTION

What are the novel findings of this work?

The Fetal Medicine Foundation model for the prediction of preterm pre-eclampsia (PE) by a combination of maternal characteristics and mean arterial pressure, uterine artery pulsatility index, and placental growth factor at 11-13 weeks is effective in predicting preterm PE in the Spanish population.

What are the clinical implications of this work?

The Fetal Medicine Foundation algorithm can be used for identifying the high-risk group for preterm PE that can benefit from the prophylactic use of aspirin.

ABSTRACT

Objective: To evaluate the diagnostic accuracy of the Fetal Medicine Foundation (FMF) competing risk model (the triple test) for the prediction of preterm pre-eclampsia (PE) in a Spanish population.

Methods: This was a prospective cohort study performed in eight fetal-medicine units in five different regions of Spain between September 2017 and December 2019. All pregnant women with singleton pregnancies and non-malformed live fetuses attending their routine ultrasound examination at 11^{+0}-13^{+6} weeks' gestation were invited to participate in the study. We recorded maternal demographic characteristics and medical history and measured MAP, UtA-PI, and serum PlGF and PAPP-A following standardized protocols. We also recorded whether the women were treated with aspirin during pregnancy. The raw values of the biomarkers were converted into multiples of the median (MoM), and audits were periodically performed for the operators and laboratories to receive continuous feedback. Risks for term and preterm PE were calculated according to the FMF competing risks model blinded to outcome. The performance of screening for PE, taking account of aspirin, was assessed by calculating the areas under the receiver-operating-characteristics curve (AUROC) and detection rates (DRs) with 95% confidence intervals (CI) at different fixed screen-positive rates (SPRs). Risk calibration was also assessed.

Results: The study population comprised 10,110 singleton pregnancies, including 72 (0.7%) that developed preterm PE. Compared to those without PE, the median MAP and UtA-PI were significantly higher in the preterm PE group, and the median serum PlGF and pregnancy-associated plasma protein A (PAPP-A) were significantly lower. In the PE group, the deviation in biomarkers from normal was inversely related to the gestational age at delivery. In screening by a combination of maternal characteristics and medical history with MAP, UtA-PI, and PlGF, at an SPR of 10%, the DR of preterm PE was 72.7 (95% CI, 62.9–82.6). An alternative strategy of replacing PlGF with PAPP-A in the triple test was associated with poorer screening

performance; the DR was 66.5% (95% CI, 55.8–77.2). Calibration plots showed good agreement between predicted and observed cases of preterm PE, with a slope of 0.983 (0.846–1.120) and an intercept of 0.154 (-0.091 to 0.397). Our DR of preterm PE at 10% SPR by the triple test was lower than that reported by the FMF (72.7% vs. 74.8%).

Conclusions: The FMF model is effective in predicting preterm PE in the Spanish population. This method of screening is feasible and easy to implement in routine clinical practice, but it must be accompanied by a good audit and monitoring system, which helps ensure the quality of the screening.

INTRODUCTION

Pre-eclampsia (PE), which occurs in 2 to 5% of pregnancies, is a leading cause of maternal and perinatal mortality and morbidity[1]. The risk of adverse perinatal outcome is worse in those cases when the disease is severe and of early onset, requiring delivery before 37 weeks of gestation (preterm PE)[2]. The traditional way to identify women at high risk of developing PE is risk scoring based on a series of factors from maternal demographic characteristics and medical history[3,4]. However, the performance of this method is poor[5]. An alternative approach consists of applying mathematical models that establish the personalized risk of PE by combining maternal characteristics and medical history with biochemical and biophysical biomarkers[6,7]. The most extensively examined and validated algorithm is the competing risks model developed by the Fetal Medicine Foundation (FMF)[8]. This method uses Bayes' theorem to combine the prior distribution of gestational age at delivery with PE from maternal characteristics and medical history with the results of various combinations of biomarkers, including mean arterial pressure (MAP), uterine artery pulsatility index (UtA-PI) and serum placental growth factor (PlGF)[9,10,11,12], which is referred to as the triple test. For a 10% screen positive rate (SPR), this model predicts about 75% of cases of preterm PE[12].

Recently, the International Federation of Gynecology and Obstetrics (FIGO) has recommended that all pregnant women should have an assessment of their risk of PE by the first-trimester triple test, and the high-risk group should receive aspirin prophylaxis at a daily dose of 150 mg from 11^{+0}–13^{+6} weeks of gestation until 36 weeks[1]. The rationale for this recommendation is based on the results of the ASPRE trial, which demonstrated that aspirin administration in a high-risk population is associated with a reduction in the incidence of preterm PE by about 60% and about 90% reduction in the incidence of PE that requires delivery before 32 weeks of gestation[13]. Additionally, closer surveillance of the high-risk group during gestation may help reduce the rate of complications in case pre-eclampsia starts later in pregnancy. Unfortunately, there is no national screening strategy in Spain, and each center decides which approach to use, if any.

Before any mathematical model can be implemented in clinical practice, its performance must be assessed in different populations from those used to construct the algorithm[14]. The FMF competing risks model was recently validated in a large cohort in Asia[15,16]. However, this process of external validation has not been undertaken in other Caucasian populations. The First-trimester pre-eclampsia validation (PREVAL) study is a non-interventional prospective cohort study aiming to implement first-trimester screening in Spain. The objective of this study is to evaluate the performance of the selected algorithm[8] in the prediction of preterm-PE in our population.

METHODS

Study design and population

This is a prospective cohort study performed in eight fetal-medicine units in five different regions of Spain (Hospital Universitario de Torrejón, Hospital Universitario Quirón and Hospital Universitario Fundación de Alcorcón in Madrid; Hospital Clínico Universitario Virgen de la Arrixaca in Murcia; Complejo Hospitalario Universitario A Coruña, in Galicia; Hospital Clínico Universitario San Cecilio and Hospital Universitario Virgen de las Nieves in Granada and Hospital Universitario Vall d'Hebrón in Catalonia) between September 2017 and December 2019. In the participating centers, all pregnant women attending their routine ultrasound examination at 11^{+0}-13^{+6} weeks' gestation (fetal crown-rump length measurement of 45 to 84 mm) were invited to participate. We included women with singleton pregnancies and non-malformed live fetuses at 11^{+0}-13^{+6} weeks' gestation. We excluded pregnancies resulting in miscarriage (n=93) or termination before 20 weeks of gestation (n=64), those with the birth of a malformed neonate (n=3), and those with incomplete pregnancy outcome (n=452). Women gave written informed consent to participate in the study, which was approved by the local Research Ethics Committee at each participating hospital.

Study procedures

During the 11^{+0}-13^{+6} weeks hospital visit, patient characteristics and medical history were recorded in a clinical database (ViewPoint® software, GE Healthcare; Munich, Germany). Patient characteristics included: maternal age; race (White, Black, South Asian, East Asian, or Mixed); method of conception, as natural or using assisted reproductive technology (ART, defined as *in vitro* fertilization or use of ovulation drugs); cigarette smoking during pregnancy; medical history of chronic hypertension, diabetes mellitus, systemic lupus erythematosus (SLE) or antiphospholipid syndrome (APS); family history of PE in the woman's mother or sister; and obstetric history that included parity (parous or nulliparous if no previous pregnancies at ≥24 weeks of gestation), and for parous women, previous PE, interpregnancy

interval, and gestational age at delivery of previous baby. We also recorded whether the women were treated with aspirin during pregnancy by asking the patient directly and examining the patient's electronic prescription records. In the participating centers, screening for pre-eclampsia was not routinely performed, and aspirin prescription was mainly offered for other indications.

In this visit, we also first, carried out an ultrasound scan to measure fetal crown-rump length and nuchal translucency thickness and to diagnose major fetal defects; second, measured the MAP with automated and validated devices (BP3AQ1 Microlife, Taipei, Taiwan) according to a standardized protocol[17]; third, measured UtA-PI by transabdominal color Doppler ultrasound imaging, according to a standardized protocol[18]; and fourth, measured serum PAPP-A and PlGF with an automated device (BRAHMS KRYPTOR™ analyzer, Thermo Fisher Scientific, Hennigsdorf, Germany); the maternal blood sample was collected during the same visit or a few days earlier, but always within the 11^{+0}-13^{+6} weeks window.

All researchers received appropriate training for the measurement of MAP and UtA-PI before starting the study and certification of their competence from the FMF. Audits of UtA-PI, MAP, PlGF, and PAPP-A were periodically performed, and feedback and/or retraining were provided to all staff involved in the study.

Results from biophysical and biochemical testing were recorded in the clinical database and sent to the study statistician every month to audit the results and recommend necessary adjustments in calculating multiple of the median (MoM) values. At the end of the study period and blinded to pregnancy outcome, individual patient risks for preterm-PE were calculated using the FMF competing risk models[10]. The risks were not available to the participants or their clinicians at the time of the first-trimester hospital visit, and any decision concerning aspirin administration was made by the attending clinicians according to the routine protocols of each site.

Diagnosis of pre-eclampsia and ascertainment of pregnancy outcome

Participants were followed up according to the clinical protocols of each center, and any pregnancy complication, as well as delivery data, were recorded by reviewing hospital/regional records or contacting delivering hospitals or the women's general medical practitioners/midwives.

Pre-eclampsia was diagnosed according to the American College of Obstetricians and Gynecologists[19]. This is defined as chronic hypertension or gestational hypertension (systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mmHg, on at least two occasions, four hours apart, and developing after 20 weeks of gestation in previously normotensive women) and at least one of the following: proteinuria (≥300 mg/24h, protein-to-creatinine ratio ≥30 mg/mmoL, or urinary dipstick testing ≥2+), renal insufficiency with serum creatinine >97 µmol/L in the absence of underlying renal disease, hepatic dysfunction with blood concentration of transaminases more than twice the upper limit of normal (≥65 IU/L for our laboratory), thrombocytopenia (platelet count <100,000/µL), neurological complications (e.g., cerebral or visual symptoms), or pulmonary edema[19].

Sample size

The screening performance of the FMF first-trimester combined test in predicting preterm PE was previously reported as having an area under the receiver operating characteristic curve (AUROC) of 0.90[6]. Assuming a standard deviation (SD) of the AUROC of 0.05, to achieve an AUROC of 0.90, a minimum of 60 cases of preterm PE would be required for an alpha of 5% and a power of 80%. For our study, preterm and term PE incidence was estimated at 0.7% and 2.1%, respectively (12). We proposed to conduct a study involving 10,000 pregnancies at 11-13 weeks of gestation. On this basis, it was anticipated that a population of 10,000 pregnancies in Spain would contain 70 cases of preterm PE and 210 cases of term PE.

Statistical analysis

Descriptive data were expressed as the median and interquartile range (IQR) and in proportions (absolute and relative frequencies). Comparisons between PE and non-PE groups were performed by Mann-Whitney U or $\chi 2$-test as appropriate. The level of significance was set at 0.05.

The raw values of MAP, UtA-PI, PlGF, and PAPP-A were converted into MoMs to adjust for maternal and pregnancy characteristics[20]. Patient-specific risks for preterm PE and PE at any gestational age (all PE) were calculated using the FMF competing risks model[8,10,11,12]. The screening performance for preterm PE and all PE was assessed by calculating the AUROC and DRs with their 95% CIs at different fixed screen-positive rates (SPRs). Risk calibration was assessed visually through a figure showing the observed incidence of PE against the incidence predicted from risk calculation. The plot was produced by grouping the data into bins according to the risk, and then, the observed incidence in each group was plotted against the incidence predicted by the model. For perfect calibration (i.e., observed incidence equals the predicted risk of PE), the slope should be 1.0, and the intercept should be 0.

Some of the subjects were prescribed aspirin (100-150 mg/day) according to existing local protocols. Aspirin treatment in a high-risk group converts some outcomes that would, without aspirin, be true positives into false positives. This, therefore, biases the assessment of screening performance. To counter this, the following steps were taken as previously described[16,20]. Ten data sets were generated in which cases of PE that were prevented by aspirin were replaced by cases. These were produced by simulating outcomes for women who received aspirin in the original data set and delivered without PE. For those who were treated with aspirin and did not have PE, the without-aspirin outcome, either PE or not, was simulated from a model using the risk of PE to determine the outcome probability. This imputation process was implemented using Markov chain Monte Carlo methods using a model in which the incidence of PE that would have occurred had it not been for the effect of

treatment was determined from a logistic regression model dependent on the logit transformation of risk. The imputation model assumed that aspirin reduced the incidence of preterm-PE by a prespecified probability of 0.62, as found in ASPRE. Estimates from the ten without-aspirin data sets were pooled.

Markov chain Monte Carlo was implemented using the WinBUGS software[21]. The statistical software package R was used for data analyses[22]. The package pROC was used for the receiver-operating characteristic curve analysis[23]; the package Car[24] was used to clean and manage the data, the package table1[25] was used for descriptive tables, and the package PropCIs[26] to calculate confidence intervals.

Quality control and monitoring

Each center was requested to send raw data on PlGF, PAPP-A, UtA-PI, and MAP and maternal characteristics to the study statistician at 1 to 2 months intervals. The statistician converted the raw values into MoMs according to the FMF reference distributions and examined the distribution of each biomarker by gestational age, maternal weight and in each hospital and for each operator. The hospital-based audit consisted of first, CUSUM plots, where deviations of each measurement from the mean are added and plotted along with randomly generated lines of what would be acceptable behavior and, second, month-to-month plots of the mean MoMs (and 95% CI) over the acceptable variation limits for each biomarker. The individual-based audits represented the mean (and 95% CI) for each biomarker and operator plotted over their acceptable variation limits.

When an increasing cumulative error was detected in CUSUM plots, the research team was notified. In the case of MAP and UtA-PI, individual doctors received further training. In the case of PlGF and PAPP-A, the laboratories were requested to review the handling of the sample from the blood draw and preanalytical and analytical requirements.

RESULTS

Study population

The study population comprised 10,110 singleton pregnancies, including 72 (0.7%) cases of preterm PE and 158 (1.6%) cases of term PE. The characteristics of the population are summarized in Table 1. In total, 9,817 (97.1%) of the population were of White racial origin, and 574 (5.7%) women were treated with aspirin. In the preterm PE, compared to the no PE group, there was a higher mean body mass index, a higher proportion of artificial conception, parous women with previous PE and nulliparous women, diabetes mellitus type 1, chronic hypertension aspirin intake, and family history of PE. In the preterm PE group, the median MAP MoM and UtA-PI MoM were higher, and the median PlGF MoM and PAPP-A MoM were lower. In the term PE, compared to the no PE group, there were more women over 40 years old, there was a higher mean body mass index, a higher proportion of artificial conception, parous women with previous PE and nulliparous women, women of Black race, with chronic hypertension, aspirin intake and family history of PE. The median MAP MoM and UtA-PI MoM were higher, and the median PlGF MoM was lower in both PE groups than in the no-PE group.

In the PE group, the deviation in biomarkers from normal was inversely related to the gestational age at delivery (Figure 1).

Quality control of screening

An example of quality assessment plots on the performance of MAP MoM in one of the participating centers and the consequence of corrective measures in response to the increasing cumulative error is shown in Figure 2. The distribution of UtA-PI MoMs of five operators before and after the audit of their results is shown in Figure 3.

Performance of screening

The performance of screening for preterm PE by various combinations of maternal characteristics and medical history with MAP, UtA-PI, and PlGF, adjusted and unadjusted by the effect of aspirin and expressed as DR at 10% SPR and AUROC are summarized in Table 2, Table S1 and Figure 4. The maternal history-only model had an AUROC of 0.71 (95% CI, 0.65–0.77) and DR of 37.5% (95% CI, 26.8–48.2), at 10% SPR for the screening of preterm PE. The triple test had an AUROC of 0.92 (95% CI, 0.89–0.95) and a DR of 72.7 (95% CI, 62.9–82.6) at 10% SPR, corresponding to a risk cut-off of 1 in 100, for the screening of preterm PE. The performance of screening for PE >34 weeks is provided in table S2.

An alternative strategy of replacing PlGF with PAPP-A in the triple test was associated with poorer screening performance; the DR of preterm PE at 10% SPR was 66.5% (95% CI, 55.8–77.2), respectively.

Calibration plots (Figure 5) showed a good correlation between predicted and observed cases of preterm PE, with a slope of 0.983 (95% CI, 0.846–1.120) and an intercept of 0.154 (95% CI, -0.091 to 0.397).

DISCUSSION

Main findings

The main findings of the study are first, the incidence of preterm and total PE in our Spanish population was 0.7% and 2.3%, which is similar to that in other European populations as reported by the FMF; second, UtA-PI and MAP were higher, and serum PlGF was lower in women who subsequently developed preterm PE as compared to those who did not develop PE and the deviation from normal strongly correlated with the gestational age at delivery; third, the performance of screening for preterm PE showed a DR of 72.7 (95% CI, 62.9–82.6) at SPR of 10% with an excellent agreement between the predicted risks and the observed incidence of PE when continuous quality surveillance is carried out.

Comparison with results of previous studies

Our results are consistent with previous studies carried out in populations of similar characteristics. The FMF group externally validated the algorithm by combining data from three prospective studies, including 61,174 singleton pregnancies from the UK, Italy, Spain, Greece, and Belgium[6,12,21,27]. The authors reported an overall incidence of PE of 2.9%, including 0.8% incidence of preterm PE and 2.1% of term PE, which was similar to that found in our study. Additionally, we found no differences in the DR (72.7% (95% CI, 62.9–82.6)) of preterm PE at 10% SPR by the triple test compared to that reported by the FMF (74.8% (95% CI, 70–80))[6], despite the differences in our population. For example, the proportion of Black women, who are known to have a higher incidence of PE than White women, was substantially lower in our study than in the FMF studies (0.7% vs.16.5%). In the FMF studies of White women, the DR of preterm PE at 10% SPR was 69.1 (95% CI, 63.2–74.5), which is similar to the performance found in our study. The FMF triple test has also been validated in other non-European populations such as East Asia, America, Australia, and Brazil, and minimal adjustments were usually required to ensure good performance[16,28-30].

Strengths and limitations

The main strength of this study relates to the large sample size, which allowed us to study a rare condition, preterm PE, and to assess each of the biomarkers used for screening individually. Participants were prospectively enrolled across different regions of the country, likely representing the global characteristics of the pregnant population in Spain. Therefore, this study provides an objective measure of the distribution of maternal demographic characteristics and medical history in pregnant women in Spain. On the other hand, all researchers were appropriately trained and certified for their competence before the start of the study and received periodical feedback about their audits and continuous monitoring. This is likely a significant contributor to the good screening performance. Finally, there was a good calibration of the model in our population, suggesting that our good results are due to good screening performance rather than overfitting of the model.

The main limitation of our study relates to the low incidence of preterm delivery with PE, with the inevitable uncertainty in estimating performance. However, the screening performance values are consistent with those previously reported in other studies. Another limitation of the study was that 6% of our patients were taking aspirin, which would have influenced the screening performance; however, we have taken this into account and made the necessary adjustment.

Implications for clinical practice

The ASPRE trial demonstrated that the use of aspirin (150 mg/day) in high-risk women starting from 11–14 weeks' reduces the risk of preterm PE by about 60%, and a secondary analysis of the trial reported that the reduction was even greater (76%) if the compliance was ≥90%[13,31]. A meta-analysis, which included the data from ASPRE, reported that aspirin reduces the risk of preterm PE by 67%, provided that the daily dose of the drug is ≥100 mg and the gestational age at onset of therapy is <16 weeks[32]. Therefore, there is a good reason to implement a

screening strategy that will allow the detection of such high-risk cases at the first hospital visit and early administration of aspirin prophylaxis.

There is no national strategy to screen for PE in Spain, so results from this study may improve our knowledge about the inner characteristics of our population, which will help develop clinical guidelines and recommendations and better fit the model to our population. Additionally, we have demonstrated the importance of continuous audits and monitoring for the measurement of the biomarkers, operators' feedback, and corrective measures. Our monitoring system may serve as a model for implementing in parallel not only the screening procedures but also the monitoring measures required to ensure a good quality screening.

Finally, our results are encouraging for other populations that present a higher incidence of pre-eclampsia since, if any difference, the algorithm is likely to perform better.

Conclusions

This study has demonstrated that the triple test effectively predicts preterm PE in the Spanish population. This screening method is feasible and easy to implement in routine clinical practice when there is an existing 11^{+0} to 13^{+6} weeks hospital visit with ultrasound and serum biochemistry assessments, but it must be accompanied by a good audit and monitoring system, which helps ensure the quality of the screening.

ACKNOWLEDGMENTS

Gil MM was awarded a *Leonardo* grant from the *BBVA Foundation* to conduct this study. This study was supported by a grant from Instituto de Salud Carlos III (ISCIII) PI18/01749. *iMaterna Foundation* (Registry No: 2148) provided smaller grants to support the study, Thermo Fisher Scientific (Hennigsdorf, Germany) provided the reagents and instruments to perform the analysis of serum biomarkers and General Electrics (GE Medical Systems, Zipf, Austria) via Health Net Connections (HNC, Spain) provided the software (ViewPoint®) to collect the data. None of these bodies had any involvement in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. This study is part of the Ph.D. Thesis of Diana Cuenca at Universidad de Alcalá, Madrid.

We are particularly grateful for the generous contribution of the patients, the laboratory personnel involved in the analysis of samples and the nurses / auxiliary nurses involved in pregnancy care and the collaboration of Biobank Network of the Region of Murcia, BIOBANC-MUR (reg. number B.0000859), supported by the "Instituto de Salud Carlos III (proyecto PT20/00109), by "Instituto Murciano de Investigación Biosanitaria Virgen de la Arrixaca, IMIB" and by "Consejeria de Salud de la Comunidad Autónoma de la Región de Murcia" and of BIOBANCO CORUÑA.

Finally, the authors are grateful to Professors Kypros H Nicolaides and David Wright for helping with the study design.

REFERENCES

1. Poon L, Shennan A, Hyett J, Dapur A, Hadar E, Divakar H, McAuliffe F, da Silva F, con Dadelszen P, McIntyre HD, Kihara AB, Di Renzo GC, Romero R, D'alton M, Berghella V, Nicolaides K, Hod Me. The International Federation of Gynecology and Obstetrics (FIGO) initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int J Gynecol Obstet* 2019; **145**: 1-33.
2. von Dadelszen P, Magee LA, Roberts JM. Subclassification of preeclampsia. *Hypertens Pregnancy* 2003; **22**: 143-148.
3. NICE. Hypertension in Pregnancy: The Management of Hypertensive Disorders During Pregnancy. London: RCOG Press, 2010.
4. American College of Obstetricians and Gynecologists (ACOG). First-trimestre risk assess for early-onset preeclampsia. Committee Opinión No. 638. *Obstet Gynecol* 2015; **126**: e25-27.
5. Tan MY, Wright D, Syngelaki A, Akolekar R, Cicero S, Janga D, Singh M, Greco E, Wright A, Maclagan K, Poon LC and Nicolaides KH. Comparison of diagnostic accuracy of early screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers; results of SPREE. *Ultrasound Obstet Gynecol* 2018; **51**: 743-750.
6. O´Gorman N, Wright D, Syngelaki A, Akolekar R, Wright A, Poon LC, Nicolaides KH. Competing Ricks model in screening for preeclampsia by maternal factors and biomarkers at 11-13 weeks gestation. *Am J Obstet Gynecol* 2016; **2141**: 03.e-103.e12.
7. Scazzocchio E, Crovetto F, Triunfo S, Gratacós E, Figueras F. Validación of a first-trimester Screening model for pre-eclampsia in an unselected population. *Ultrasound Obstet Gynecol* 2017; **49**: 188-193.
8. Wright D, Wright A, Nicolaides, KH. The competing risk approach for prediction of preeclampsia. *Am J Obstet Gynecol* 2020; **223**: 12-23.
9. Wright D, Akolekar R, Syngelaki A, Poon L, Nicolaides KH. A competing risks model in early screening for preeclampsia. *Fetal Diagn Ther* 2012; **32**: 171–178.
10. Akolekar R, Syngelaki A, Poon L, Wright D, Nicolaides KH. Competing risks model in early screening for preeclampsia by biophysical and biochemical markers. *Fetal Diagn Ther* 2013; **33**: 8–15.
11. Wright D, Syngelaki A, Akolekar R, Poon L, Nicolaides KH. Competing risks model in screening for preeclampsia by maternal characteristics and medical history. *Am J Obstet Gynecol* 2015; **213**: 62.e1–10.
12. O'Gorman N, Wright D, Poon LC, Rolnik DL, Syngelaki A, Wright A, Akolekar R, Cicero S, Janga D, Jani J, Molina FS, de Paco Matallana C, Papantoniou N, Persico N, Plasencia W, Singh M, Nicolaides KH. Accuracy of competing risks model in screening for pre-

eclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. *Ultrasound Obstet Gynecol* 2017; **49**: 751-755

13. Rolnik DL, Wright D, Poon LC, O'Gorman N, Syngelaki A, de Paco Matallana C, Akolekar R, Cicero S, Janga D, Singh M, Molina FS, Persico N, Jani JC,Plasencia W, Papaioannou G, Tenenbaum-Gavish K, Meiri H, Gizurarson S,Maclagan K, Nicolaides KH. Aspirin versus placebo in Pregnancies at Hight Risk for Preterm Preeclampsia. *N Engl J Med* 2017; **377**: 613-622.
14. Collins, G.S., de Groot, J.A., Dutton, S. et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol* 2014; **14**: 40.
15. Chaemsaithong P, Sahota D, Pooh RK, Zheng M, Ma R, Chaiyasit N, Koide K, Shaw SW, Seshadri S, Choolani M, Panchalee T, Yapan P, Sim WS, Sekizawa A, Hu Y, Shiozaki A, Saito S, Leung TY, Poon LC. First-trimester pre-eclampsia biomarker profiles in Asian population: multicenter cohort study. *Ultrasound Obstet Gynecol* 2020; **56**: 206-214.
16. Chaemsaithong P, Pooh RK, Zheng M, Ma R, Chaiyasit N, Tokunaka M, Shaw SW, Seshadri S, Choolani M, Wataganara T, Yeo GSH, Wright A, Leung WC, Sekizawa A, Hu Y, Naruse K, Saito S, Sahota D, Leung TY, Poon LC. Prospective evaluation of screening performance of first-trimester prediction models for preterm preeclampsia in an Asian population. *Am J Obstet Gynecol* 2019; **221**: 650.e1-650.e16)
17. Poon LC, Zymeri NA, Zamprakou A, Syngelaki A, Nicolaides KH. Protocol for measurement of mean arterial pressure at 11- 13 weeks' gestation. *Fetal Diagn Ther* 2012; **31**: 42–48.
18. Plasencia W, Maiz N, Bonino S, Kaihura C, Nicolaides KH. Uterine artery Doppler at 11 + 0 to 13 + 6 weeks in the prediction of pre-eclampsia. *Ultrasound Obstet Gynecol* 2007; **30**: 742-9.
19. American College of Obstetricians and Gynecologists, and the Task Force on Hypertension in Pregnancy. Hypertension in Pregnancy. *Obstet Gynecol* 2013; **122**: 1122–1131.
20. Tan MY, Syngelaki A, Poon LC, Rolnik DL, O'Gorman N, Delgado JL, Akolekar R, Konstantinidou L, Tsavdaridou M, Galeva S, Ajdacka U, Molina FS, Persico N, Jani JC, Plasencia W, Greco E, Papaioannou G, Wright A, Wright D, Nicolaides KH. Screening for preeclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. *Ultrasound Obstet Gynecol* 2018; **52**: 186.-195.
21. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000; **10**: 325–337.
22. Tan MY, Wright D, Syngelaki A, Akolekar R, Cicero S, Janga D, Singh M, Greco E, Wright A, Maclagan K, Poon LC, Nicolaides KH. Comparison of diagnostic accuracy of early

screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers: results of SPREE. *Ultrasound Obstet Gynecol* 2018; **51**: 743-750.

23. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2011.
24. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**:77.
25. Fox J, Sandford W. An R Companion to Applied Regression. Third Edit. Thousand Oaks; 2019. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.
26. Benjamin R. (2020). table1: Tables of Descriptive Statistics in HTML. R package version 1.2. https://CRAN.R-project.org/package=table1
27. Ralph S. (2018). PropCIs: Various Confidence Interval Methods for Proportions. R package version 0.3-0. https://CRAN.R-project.org/package=PropCIs
28. Tan MY, Wright D, Syngelaki A, Akolekar R, Cicero S, Janga D, Singh M, Greco E, Wright A, Maclagan K, Poon LC, Nicolaides KH. Comparison of diagnostic accuracy of early screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers: results of SPREE. *Ultrasound Obstet Gynecol* 2018; **51**: 743-750.
29. Sonek J, Krantz D, Carmichael J, Downing C, Jessup K, Haidar Z, Ho S, Hallahan T, Kliman H, Mckenna D. First-trimester screening for early and late preeclampsia using maternal characteristics, biomarkers, and estimated placental volume. *Am J Obstet Gynecol* 2018; **218**: 126.e1-.e13.
30. Park FJ, Leung CH, Poon LC, Williams PF, Rothwell SJ, Hyett JA. Clinical evaluation of a first trimester algorithm predicting the risk of hypertensive disease of pregnancy. *Aust N Z J Obstet Gynaecol* 2013; **53**: 532-539.
31. Lobo GAR, Nowak PM, Panigassi AP, Lima AIF, Araujo Junior E, Nardozza LMM, Baptista D. Validation of Fetal Medicine Foundation algorithm for prediction of pre-eclampsia in the first trimester in an unselected Brazilian population. *J Matern Fetal Neonatal Med* 2019; **32**: 286-292.
32. Wright D, Poon LC, Rolnik DL, Syngelaki A, Delgado JL, Vojtassakova D, de Alvarado M, Kapeti E, Rehal A, Pazos A, Carbone IF, Dutemeyer V, Plasencia W, Papantoniou N, Nicolaides KH. Aspirin for evidence-based preeclampsia prevention trial: influence of compliance on beneficial effect of aspirin in prevention of preterm preeclampsia. *Am J Obstet Gynecol* 2017; **217**: 685.e1–5.
33. Roberge S, Bujold E, Nicolaides KH. Aspirin for the prevention of preterm and term preeclampsia: systematic review and metaanalysis. *Am J Obstet Gynecol* 2018; **218**: 287-293.

FIGURE LEGENDS

Figure 1. Scatter diagrams and regression lines of biomarker distributions according to gestational age at delivery with pre-eclampsia. A) Mean arterial pressure; B) Mean uterine artery pulsatility index; C) Serum placental growth factor and pregnancy-associated plasma protein-A. The solid lines represent the reference as the median in unaffected pregnancies. The dotted lines represent the regression line in pregnancies complicated with pre-eclampsia. MoM: multiples of the median.

Figure 2. Quality assessment plots on the performance of mean arterial pressure in one of the participating centers. Blue lines represent individual operator measurements, and the black line represents the median of the participating center. An increasing cumulative error was detected, and corrective measures were applied in August (red arrow), as reflected as a straight line (panel A) or centered MoMs from September (panel B).

Figure 3. Distribution of measurements of mean uterine artery pulsatility index for each one of the operators participating in the study. Dashed lines represent the measurements before the audit and solid lines those after the audit. The numbers show the number of examinations per operator and respective audit. The darkest grey area represents the good quality zone, and the lighter grey area is the warning zone. Measurements outside these areas are not acceptable. A) Mean arterial pressure; B) Mean uterine artery pulsatility index (PI); C) Serum placental growth factor (PlGF). MoM: multiples of the median.

Figure 4. Receiver operating characteristics curves of the different models to predict pre-eclampsia < 34 weeks (left), <37 weeks (middle), and all pre-eclampsia (right). Prediction models represented are: maternal factors alone (black lines); maternal factors plus mean arterial pressure (blue); maternal factors plus mean arterial pressure plus mean uterine artery pulsatility index (red); and maternal factors plus mean arterial pressure plus mean uterine artery pulsatility index plus placental growth factor (green).

Figure 5. Calibration of the aspirin-unadjusted model. The y-axis represents the observed incidence of preterm pre-eclampsia, and the x-axis represents the predicted incidence of preterm pre-eclampsia (screen-positive cases). Grey numbers on top of each group are the observed number of PE cases delivered preterm and black numbers are the observed number of cases delivered without pre-eclampsia for each bin of predicted risk.

SUPPORTING INFORMATION ON THE INTERNET

Table S1: Detection rate for term pre-eclampsia, at 10% screen positive rate, in screening by various combinations of biomarkers with and without adjustment for use of aspirin

Table S2: Detection rate for pre-eclampsia < 34 weeks, at 10% screen positive rate, in screening by various combinations of biomarkers with and without adjustment for use of aspirin

Table 1. Characteristics of the study population.

Maternal characteristics	**Preterm PE (n=72)**	**Term PE (n=158)**	**No PE (n=9880)**
Maternal age (years)	34.2 [31.7, 38.6]	34.7 [31.0, 38.3] *	33.9 [30.1, 36.9]
< 35	37 (51.4)	84 (53.2)	5862 (59.3)
≥ 35	35 (48.6)	74 (46.8)	4018 (40.7)
< 40	62 (86.1)	135 (85.4)	9052 (91.6)
≥ 40	10 (13.9)	23 (14.6) *	828 (8.4)
Body mass index (kg/m^2)	26.5 [23.5, 31.6] *	27.6 [23.9, 31.8] *	24.0 [21.7, 27.2]
Underweight (<18.5)	2 (2.8)	1 (0.633)	257 (2.6)
Normal weight (18.5 to 24.9)	25 (34.7) *	49 (31.0) *	5625 (56.9)
Overweight (25 to 29.9)	22 (30.6)	52 (32.9)	2657 (26.9)
Obese or more (≥ 30)	23 (31.9) *	56 (35.4) *	1341 (13.6)
Conception	*	*	
Spontaneous	58 (80.6) *	136 (86.1) *	9091 (92.0)
Assisted	14 (19.4) *	22 (13.9) *	789 (7.99)
Cigarette smoker	7 (9.7)	16 (10.1)	1137 (11.5)
Obstetric history	*	*	
Parous with previous PE	7 (9.7) *	20 (12.7) *	143 (1.5)
Parous without previous PE	21 (29.2) *	40 (25.3) *	4782 (48.4)
Nulliparous	44 (61.1)	98 (62.0)	4955 (50.2)
Racial origin			

White	70 (97.2)	143 (90.5) *	9604 (97.2)
Mixed	1 (1.4)	7 (4.4) *	178 (1.8)
Black	1 (1.4)	7 (4.4) *	70 (0.7)
East Asian	0 (0.0)	1 (0.6)	21 (0.2)
South Asian	0 (0.0)	0 (0)	7 (0.1)
Medical history			
Diabetes mellitus type 1	4 (5.6) *	2 (1.3)	45 (0.5)
Diabetes mellitus type 2	1 (1.4)	2 (1.3)	20 (0.2)
SLE or APS	1 (1.4) *	0 (0) *	52 (0.5)
Chronic hypertension	8 (11.1) *	16 (10.1) *	82 (0.8)
Aspirin intake	9 (12.5) *	21 (13.3) *	534 (5.4)
Family history of PE	6 (8.3) *	10 (6.3) *	179 (1.8)
Biomarkers			
Mean arterial pressure MoM	1.1 [1.0, 1.1] *	1.1 [1.0, 1.1] *	1.0 [0.9, 1.1]
Uterine artery pulsatility index MoM	1.2 [0.9, 1.3] *	1.2 [0.9, 1.3] *	1.0 [0.8, 1.2]
Placental growth factor MoM	0.9 [0.6, 1.1] *	0.9 [0.6, 1.1] *	1.0 [0.8, 1.3]
Pregnancy associated plasma protein A MoM	0.9 [0.6, 1.4] *	0.9 [0.6, 1.4] *	1.0 [0.6, 1.6]

Data are given as median (interquartile range) or n (%). PE: pre-eclampsia; SLE: systemic lupus erythematosus; APS: antiphospholipid syndrome; MoM: multiple of the median. *Statistically significantly different from values in the no-PE group.

Table 2. Detection rate for preterm pre-eclampsia, at 10% screen positive rate, in screening by various combinations of biomarkers with and without adjustment for use of aspirin.

Method of screening	Unadjusted for aspirin		Adjusted for aspirin	
	DR at 10% SPR	AUROC	DR at 10% SPR	AUROC
Maternal characteristics and medical history	31.9 (21.4; 44)	0.68 (0.62; 0.75)	37.5 (26.8; 48.2)	0.71 (0.65; 0.77)
+ MAP	38.9 (27.6; 51.1)	0.79 (0.73; 0.84)	44.1 (33.0; 55.1)	0.80 (0.76; 0.85)
+ UtA-PI	47.2 (35.3; 59.3)	0.80 (0.75; 0.86)	51.8 (40.7; 62.8)	0.82 (0.77; 0.87)
+ PlGF	55.6 (43.4; 67.3)	0.85 (0.80; 0.89)	59.0 (48.0; 69.9)	0.86 (0.82; 0.90)
+ PAPP-A	36.1 (25.1; 48.3)	0.70 (0.63; 0.76)	40.4 (29.4; 51.5)	0.72 (0.66; 0.78)
+ MAP + UtA-PI	61.1 (48.9; 72.4)	0.86 (0.82; 0.91)	64.2 (53.5; 74.9)	0.87 (0.83; 0.91)
+ MAP + PlGF	62.5 (50.3; 73.6)	0.88 (0.85; 0.92)	65.3 (54.7; 75.9)	0.89 (0.86; 0.92)
+ MAP + PAPP-A	44.4 (32.7; 56.6)	0.78 (0.73; 0.84)	49.0 (37.9; 60.0)	0.80 (0.75; 0.85)
+ UtA-PI + PlGF	63.9 (51.7; 74.9)	0.88 (0.84; 0.92)	66.6 (55.8; 77.4)	0.89 (0.85; 0.93)
+ Ut-PI + PAPP-A	45.8 (34.0; 58.0)	0.80 (0.74; 0.85)	50.3 (39.2; 61.4)	0.81 (0.76; 0.87)
+ PlGF + PAPP-A	56.9 (44.7; 68.6)	0.84 (0.79; 0.88)	59.7 (48.7; 70.6)	0.85 (0.81; 0.89)
+ MAP + UtA-PI + PlGF	70.8 (58.9; 81.0)	0.91 (0.88; 0.94)	72.7 (62.9; 82.6)	0.92 (0.89; 0.95)
+ MAP + UtA-PI + PAPP-A	63.9 (51.7; 74.9)	0.86 (0.81; 0.90)	66.5 (55.8; 77.2)	0.87 (0.83; 0.91)
+ MAP + PlGF + PAPP-A	63.9 (51.7; 74.9)	0.88 (0.84; 0.91)	66.6 (56.1; 77.2)	0.89 (0.85; 0.92)
+ MAP + UtA-PI + PlGF + PAPP-A	69.4 (57.5; 79.8)	0.91 (0.88; 0.94)	71.6 (61.7; 81.5)	0.92 (0.88; 0.95)

Values in parentheses are 95% CI. DR: detection rate; SPR: screen positive rate; AUROC: area under the curve; MAP: mean arterial pressure; UtA-PI: uterine artery pulsatility index, PlGF: placental growth factor; PAPP-A: pregnancy-associated plasma protein A.