OBSTETRICS

Predictive performance of the competing risk model in screening for preeclampsia

David Wright, PhD; Min Yi Tan, MD; Neil O'Gorman, MD; Liona C. Poon, MD; Argyro Syngelaki, PhD; Alan Wright, PhD; Kypros H. Nicolaides, MD

BACKGROUND: The established method of screening for preeclampsia is to identify risk factors from maternal demographic characteristics and medical history; in the presence of such factors the patient is classified as high risk and in their absence as low risk. However, the performance of such an approach is poor. We developed a competing risks model, which allows combination of maternal factors (age, weight, height, race, parity, personal and family history of preeclampsia, chronic hypertension, diabetes mellitus, systemic lupus erythematosus or antiphospholipid syndrome, method of conception and interpregnancy interval), with biomarkers to estimate the individual patient-specific risks of preeclampsia requiring delivery before any specified gestation. The performance of this approach is by far superior to that of the risk scoring systems.

OBJECTIVE: The objective of the study was to examine the predictive performance of the competing risks model in screening for preeclampsia by a combination of maternal factors, mean arterial pressure, uterine artery pulsatility index, and serum placental growth factor, referred to as the triple test, in a training data set for the development of the model and 2 validation studies.

STUDY DESIGN: The data for this study were derived from 3 previously reported prospective, nonintervention, multicenter screening studies for preeclampsia in singleton pregnancies at 11^{+0} to 13^{+6} weeks' gestation. In all 3 studies, there was recording of maternal factors and biomarkers and ascertainment of outcome by appropriately trained personnel. The first study of 35,948 women, which was carried out between February 2010 and July 2014, was used to develop the competing risks model for prediction of preeclampsia and is therefore considered to be the training set. The 2 validation studies were comprised of 8775 and 16,451 women, respectively, and they were carried out between February and September 2015 and between April and December 2016, respectively. Patient-specific risks of delivery with preeclampsia at <34, <37, and $<41^{+3}$ weeks' gestation were calculated using the competing risks model and the performance of screening for preeclampsia by maternal factors alone and the triple test in each of the 3 data sets was assessed. We examined the predictive performance of the model by first, the ability of the model to discriminate between the preeclampsia and no-preeclampsia groups using the area under the receiver operating characteristic curve and the detection rate at fixed screen-positive rate of 10%, and second, calibration by measurements of calibration slope and calibration in the large.

RESULTS: The detection rate at the screen-positive rate of 10% of early-preeclampsia, preterm-preeclampsia, and all-preeclampsia was about 90%, 75%, and 50%, respectively, and the results were consistent between the training and 2 validation data sets. The area under the receiver operating characteristic curve was >0.95, >0.90, and >0.80, respectively, demonstrating a very high discrimination between affected and unaffected pregnancies. Similarly, the calibration slopes were very close to 1.0, demonstrating a good agreement between the predicted risks and observed incidence of preeclampsia. In the prediction of early-preeclampsia and preterm-preeclampsia, the observed incidence in the training set and 1 of the validation data sets was consistent with the predicted one. In the other validation data set, which was specifically designed for evaluation of the model, the incidence was higher than predicted, presumably because of better ascertainment of outcome. The incidence of all-preeclampsia was lower than predicted in all 3 data sets because at term many pregnancies deliver for reasons other than preeclampsia, and therefore, pregnancies considered to be at high risk for preeclampsia that deliver for other reasons before they develop preeclampsia can be wrongly considered to be false positives.

CONCLUSION: The competing risks model provides an effective and reproducible method for first-trimester prediction of early preeclampsia and preterm preeclampsia as long as the various components of screening are carried out by appropriately trained and audited practitioners. Early prediction of preterm preeclampsia is beneficial because treatment of the high-risk group with aspirin is highly effective in the prevention of the disease.

Key words: aspirin, calibration, Combined Multimarker Screening and Randomized Patient Treatment with Aspirin for Evidence-Based Preeclampsia Prevention trial, competing risks model, discrimination, first-trimester screening, mean arterial blood pressure, performance of screening, placental growth factor, preeclampsia, survival model, uterine artery Doppler

EDITORS' CHOICE

The established method of screening for preeclampsia (PE) is to identify risk factors from maternal demographic characteristics and medical history; in the presence of such factors, the patient is classified as high risk and in their absence as low risk.[1,2] The performance of this approach is poor[3–5] and, although it is simple, it does not quantify individual patient-specific risks.

An alternative way of screening is to use logistic regression models fitted to maternal characteristics and medical history alone or in combination with biomarkers to predict early, late, or all PE.[6–10] Such models are useful in quantifying the individual patient specific risk for PE, rather than just classifying women into high- and low-risk groups. However, they do not allow the flexibility of selecting different gestational age cutoffs for categorizing the severity of PE, they do not take into

Cite this article as: Wright D, Tan MY, O'Gorman N, et al. Predictive performance of the competing risk model in screening for preeclampsia. Am J Obstet Gynecol 2019;220:199.e1-13.

0002-9378/free

https://doi.org/10.1016/j.ajog.2018.11.1087

AJOG at a Glance

Why was this study conducted?

To assess the predictive performance of the competing risks model for preeclampsia using the first-trimester triple test that combines maternal factors, mean arterial pressure, uterine artery pulsatility index, and serum placental growth factor.

Key findings

Results from 2 prospective multicenter validation data sets show that, with appropriately trained staff and quality control of measurement, preeclampsia, especially that leading to early delivery, can be predicted effectively using the triple test. These results are consistent with those obtained from the training data set.

What does this add to what is known?

The competing risks model provides an effective and reproducible method for first-trimester prediction of preeclampsia.

account the fact that the deviation in biomarker levels from normal depends on the severity of the disease, and they cannot be easily expanded to include additional biomarkers measured at different stages in pregnancy.

We have developed a competing risks approach that allows combination of maternal factors with biomarkers to estimate the individual patient-specific risks of PE requiring delivery before any specified gestation.[11,12] This is based on a survival-time model for the gestational age at delivery with PE, and it is assumed that if the pregnancy was to continue indefinitely, all women would develop PE and whether they do so or not before a specified gestational age depends on competition between delivery before or after the development of PE.

Based on model fit and ease of interpretation, a Gaussian model for gestational age at delivery was chosen. The effects of variables from maternal factors and biomarkers is to modify the distribution of gestational age at delivery with PE so that in pregnancies at low risk for PE, the gestational age distribution is shifted to the right with the implication that in most pregnancies delivery will actually occur before the development of PE.

In high-risk pregnancies, the distribution is shifted to the left, and the smaller the mean gestational age, the higher is the risk for PE. In 1 previous study of 120,492 singleton pregnancies undergoing screening at 11−13 weeks' gestation, we reported the development of the competing risks model based on maternal characteristics and medical history, including age, weight, height, race, parity, personal and family history of PE, chronic hypertension, diabetes mellitus, systemic lupus erythematosus or antiphospholipid syndrome, method of conception and interpregnancy interval.[3]

In another study of 35,948 singleton pregnancies, we reported effective screening for preterm PE, with delivery at <37 weeks' gestation, by a combination of maternal factors with mean arterial pressure (MAP), uterine artery pulsatility index (PI), and serum placental growth factor (PLGF).[13] A limitation of the study was that the performance of screening by a model derived and tested using the same data set is overestimated. We used cross-validation to reduce this effect but suggested the necessity for external validation on independent data from different sources.

The objective of this study is to examine the predictive performance of the competing risks model in screening for PE with delivery <34 weeks (early-PE), <37 weeks (preterm-PE), and delivery at any gestation (all-PE) by maternal factors alone and a combination of maternal factors, MAP, UtA-PI and PLGF (triple test) in the training data set[13] for development of the model and 2 validation studies.

Materials and Methods

Study populations

The data for this study were derived from 3 previously reported prospective nonintervention screening studies at 11^{+0} to 13^{+6} weeks' gestation with a combined total of 61,174 singleton pregnancies, including 1770 (2.9%) that developed PE.[4,13,14]

The first study comprised of 35,948 women attending for their routine first hospital visit in pregnancy at King's College Hospital and Medway Maritime Hospital (United Kingdom) between February 2010 and July 2014.[13] This data set was used to develop the competing risks model for the prediction of PE and is therefore considered to be the training set.

The second study, referred to as the screening quality study (SQS), comprised of 8775 singleton pregnancies undergoing first-trimester screening for PE, using the competing risks model developed in the first study,[13] in 12 maternity hospitals in England, Spain, Belgium, Italy, and Greece between February and September 2015.[14] This study was carried out before the Combined Multimarker Screening and Randomized Patient Treatment with Aspirin for Evidence-Based Preeclampsia Prevention (ASPRE) trial[15] and was primarily designed to examine the feasibility of multicenter screening and establish methods for quality assurance of the biomarkers, and the results from screening were not made available to the patients or their obstetricians.

The third study, referred to as Superior Province Rifting EarthScope Experiment (SPREE), was a multicenter cohort study in 16,451 women carried out in 7 National Health Service maternity hospitals in England between April and December 2016.[4] This study was specifically designed to examine the performance of screening by the algorithm established in the first study[13] in comparison with that of the method advocated by the National Institute for Health and Care Excellence (NICE); the results from screening by the competing

TABLE 1
Maternal and pregnancy characteristics in the 3 populations

Variables	Training set (n = 35,948)	SQC (n = 8775)	SPREE (n = 16,451)
Maternal age, y, median (IQR)	31.3 (26.8, 35.0)	31.5 (27.3, 35.0)[a]	31.5 (27.4, 35.1)[a]
Maternal weight, kg, median (IQR)	66.7 (59.0, 77.2)	66.5 (59.0, 77.0)[b]	67.0 (59.2, 78.0)[a]
Maternal height, cm, median (IQR)	164.5 (160.0, 169.0)	164.5 (160.0, 169.0)[b]	165.0 (160.0, 169.0)[a]
Body mass index, kg/m^2, median (IQR)	24.5 (22.0, 28.4)	24.5 (21.9, 28.4)[b]	24.7 (22.0, 28.7)[a]
Gestational age, wks, median (IQR)	12.7 (12.3, 13.1)	12.7 (12.3, 13.1)[a,b]	12.9 (12.4, 13.3)[a]
Racial origin		[a,b]	[a]
White, n (%)	25,879 (71.99)	6,883 (78.44)	11,922 (72.47)
Black, n (%)	6681 (18.59)	1,090 (12.42)	2,337 (14.21)
South Asian, n (%)	1623 (4.51)	462 (5.26)	1,361 (8.27)
East Asian, n (%)	846 (2.35)	154 (1.75)	407 (2.47)
Mixed, n (%)	919 (2.56)	186 (2.12)	424 (2.58)
Conception		[a,b]	[a]
Natural	34,743 (96.65)	8,483 (96.67)	15,765 (95.83)
Assisted by use of ovulation drugs	349 (0.97)	64 (0.73)	125 (0.76)
In vitro fertilization	856 (2.38)	227 (2.59)	561 (3.41)
Medical history			
Chronic hypertension	561 (1.56)	100 (1.14)[b]	137 (0.83)[a]
Diabetes mellitus type 1	137 (0.38)	31 (0.35)[b]	46 (0.28)[a]
Diabetes mellitus type 2	188 (0.52)	37 (0.42)[b]	71 (0.43)[a]
SLE/APS	53 (0.15)	19 (0.22)	39 (0.24)[a]
Cigarette smokers, n (%)	3,263 (9.08)	732 (8.34)[b]	1,105 (6.72)[a]
Family history of preeclampsia, (n, %)	1,518 (4.22)	339 (3.86)[a]	535 (3.25)[a]
Parity		[a,b]	[a]
Nulliparous, n (%)	17,361 (48.29)	4,127 (47.03)	7,587 (46.12)
Parous with no previous PE, n (%)	17,311 (48.16)	4,459 (50.81)	8,483 (51.57)
Parous with previous PE, n (%)	1,276 (3.55)	189 (2.15)	381 (2.32)
Preeclampsia			
Total, n (%)	1,058 (2.94)	239 (2.72)	439 (2.67)
Delivery <37 wks, n (%)	292 (0.81)	59 (0.67)	135 (0.82)
Delivery <34 wks, n (%)	128 (0.36)	27 (0.31)	58 (0.35)

Comparisons between outcome groups were by χ^2 or Fisher exact test for categorical variables and a Mann-Whitney *U* test for continuous variables.

APS, antiphospholipid syndrome; *IQR*, interquartile range; *PE*, preeclampsia; *SLE*, systemic lupus erythematosus; *SPREE*, Superior Province Rifting EarthScope Experiment; *SQS*, screening quality study.

[a] Significance value of $P < .05$ in comparison of SQS and SPREE with the training set; [b] Significance value of $P < .05$ in comparison of SQS and SPREE.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

risks model were not made available to the patients or their obstetricians.

In all 3 studies, women with singleton pregnancies in the participating hospitals had a routine examination at 11^{+0} to 13^{+6} weeks' gestation. This visit included first, recording of maternal characteristics and medical history,[3] second, measurement of the left and right UtA-PI by transabdominal color Doppler ultrasound and calculation of the mean PI,[16] third, measurement of MAP by validated automated devices and standardized protocol,[17] and fourth, measurement of serum concentration of PLGF (Delfia Xpress system; PerkinElmer Life and Analytical Sciences, Waltham, MA, or Brahms Kryptor analyzer; Thermo Fisher Scientific, Hennigsdorf, Germany).

TABLE 2
Performance of screening, with 95% confidence interval, for early-PE, preterm-PE, and all-PE by the triple test in the 3 data sets

	Discrimination		Calibration	
Method of screening	AUROC curve	DR for 10% SPR	Slope	Intercept
Early-PE				
Training set	0.95 (0.93, 0.97)	87 (80, 92)	0.92 (0.84, 1.01)	0.05 (-0.14, 0.23)
SQS	0.97 (0.95, 0.99)	93 (76, 99)	0.98 (0.80, 1.17)	0.05 (-0.38, 0.48)
SPREE	0.96 (0.93, 0.98)	90 (78, 96)	0.92 (0.79, 1.04)	0.45 (0.16, 0.73)
Preterm-PE				
Training set	0.91 (0.89, 0.93)	75 (70, 80)	0.95 (0.89, 1.02)	-0.19 (-0.32, -0.07)
SQS	0.93 (0.89, 0.96)	75 (62, 85)	1.00 (0.85, 1.15)	−0.19 (−0.47, 0.09)
SPREE	0.93 (0.92, 0.95)	83 (76, 89)	1.05 (0.95, 1.15)	0.17 (−0.01, 0.35)
All-PE				
Training set	0.83 (0.81, 0.84)	52 (49, 55)	1.07 (1.02, 1.12)	−0.57 (−0.64, −0.50)
SQS	0.82 (0.80, 0.85)	49 (43, 56)	1.06 (0.94, 1.17)	−0.44 (−0.58, −0.29)
SPREE	0.85 (0.83, 0.87)	53 (49, 58)	1.17 (1.08, 1.26)	−0.41 (−0.52, −0.31)

AUROC, area under the receiver-operating characteristic; *DR*, detection rate; *SLE*, systemic lupus erythematosus; *SPREE*, Superior Province Rifting EarthScope Experiment; *SQS*, screening quality study.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

The measurements of MAP were carried out by health care assistants or sonographers who had received specific training for this purpose, and measurements of UtA-PI were performed by doctors or sonographers who had obtained the Fetal Medicine Foundation Certificate of Competence in Doppler ultrasound. In both validation studies, quality control was applied on a monthly basis to achieve consistency of measurement of biomarkers across different hospitals throughout the duration of the study.

The distribution of measurements of MAP and UtA-PI were reported to the coordinator who provided feedback and if necessary retraining of the personnel with large deviations from the expected values. Similarly, the laboratories were provided with diagnostics for PLGF measurements so that appropriate corrective actions could be undertaken. Gestational age was determined from the fetal crown-rump length.[18] The women gave written informed consent to participate in the studies, which were approved by the relevant research ethics committee in each participating hospital.

The inclusion criteria were singleton pregnancy undergoing first-trimester combined screening for PE and subsequently delivering a morphologically normal live birth or stillbirth at ≥24 weeks' gestation. We excluded pregnancies with aneuploidies and major fetal abnormalities and those ending in termination, miscarriage, or fetal death before 24 weeks' gestation.

Outcome measures were early-PE, preterm-PE, and all-PE. Data on pregnancy outcome were collected from the hospital maternity records or the general medical practitioners of the women. The obstetric records of all women with preexisting or pregnancy-associated hypertension were examined to determine whether the condition was PE, as defined by the International Society for the Study of Hypertension in Pregnancy.[19] This includes the finding of hypertension (systolic blood pressure of ≥140 mm Hg or diastolic blood pressure of ≥90 mm Hg on at least 2 occasions 4 hours apart developing after 20 weeks' gestation in previously normotensive women) and proteinuria (≥300 mg per 24 hours or protein to creatinine ratio ≥30 mg/mmol or ≥2 + on dipstick testing).

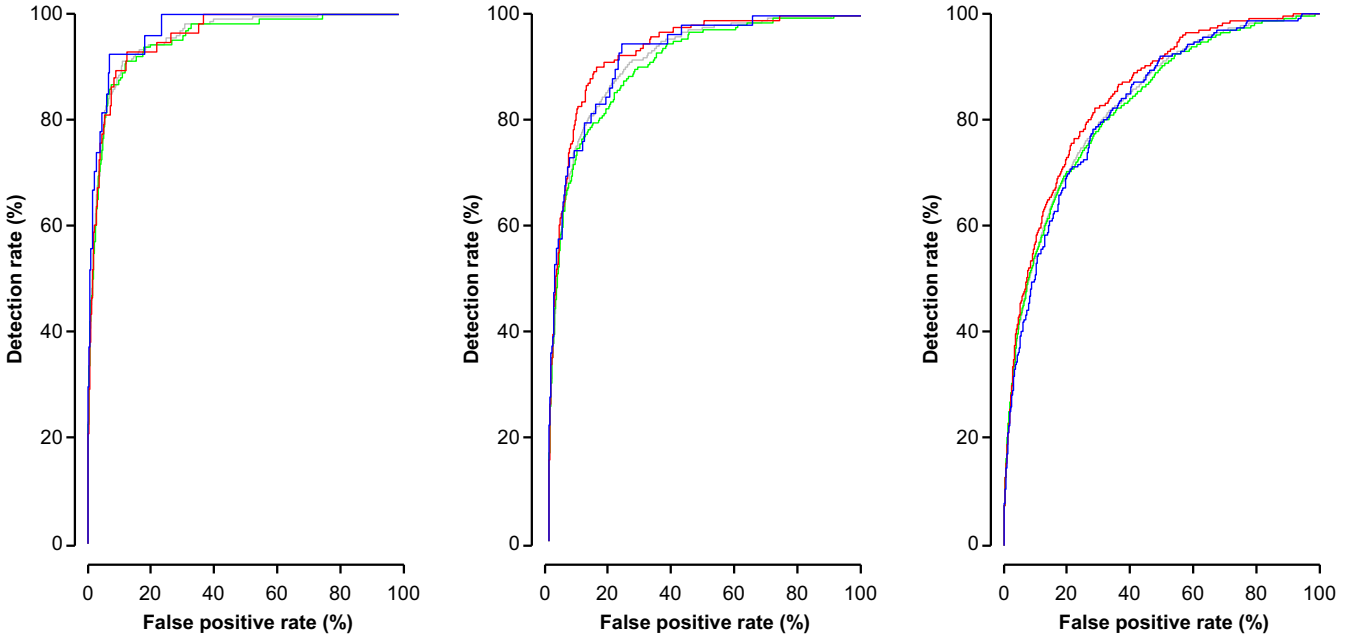
Statistical analysis

Patient-specific risks of delivery with PE at <34, <37, and $<41^{+3}$ weeks' gestation were calculated using the competing risks model to combine the prior distribution of the gestational age at delivery with PE, obtained from maternal characteristics and medical history, with the multiple of the median values of MAP, UtA-PI, and PLGF.[3,13] The performance of the screening for early-PE, preterm-PE, and all-PE by the triple test in each of the 3 data sets was assessed.

We examined the predictive performance of the model by first, the ability of the model to discriminate between the PE and no-PE groups and second, calibration, which assesses agreement between predicted risks and outcomes (for a well-calibrated model, among those women with a risk of 1 in n, the incidence should be 1 in n).

Discrimination was assessed by the area under the receiver-operating characteristic (AUROC) curve (this indicates perfect discrimination if the value is 1 and no discrimination beyond chance if the value is 0.5) and the detection rate (DR) at a fixed screen-positive rate (SPR) of 10%. Calibration was assessed visually

FIGURE 1
Receiver-operating characteristic plots of screening for the 3 groups



Receiver-operating characteristic plots of screening for early-PE, preterm-PE, and all-PE by the triple test in the training set (*green line*), SQS (*blue line*), SPREE (*red line*), and the combination of the 3 data sets (*gray line*).

PE, preeclampsia; *SPREE*, Superior Province Rifting EarthScope Experiment; *SQS*, screening quality study.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

through a series of figures showing the observed incidence against that predicted from risk for PE <32, <34, <37, and $<41^{+3}$ weeks' gestation by maternal factors and the triple test.

The plots were produced by grouping the data into bins according to risk. The observed incidence in each group was then plotted against the incidence predicted by the model (ie, the mean risks within each group). Quantitative assessment of calibration was by recording of measurements of calibration in the large and calibration slope.

Calibration in the large is a measure of whether generally the risks are too high or too low. This is quantified by the estimated intercept from a logistic regression of incidence on the logit of risk with the slope fixed at 1. The intercept is a measure of the deviation of the observed incidence from the predicted. For perfectly calibrated risks, the intercept should be zero. If there is a general tendency for underestimation so that the observed incidence is larger than that predicted, the intercept will be positive. Conversely, for overestimation, the intercept will be negative.

The calibration slope assesses the calibration across the range of risks and is the slope of the regression line of the logistic regression of incidence on the logit of risk. If the risk is well calibrated, then the slope should be 1.0. A slope less than 1 means that the relationship between risk and incidence is flatter than it should be. A calibration slope greater than 1 means the relationship is steeper than it should be.

The risks produced from our competing risks model are for delivery with PE before a specific gestation assuming no other cause for delivery. Because other cause deliveries are effectively censored observations, the actual incidence of PE would be expected to be lower than predicted. For early gestations, when there are few other cause deliveries, the effects would be small. At later gestations, with many other cause deliveries, the effect of censoring may be substantial. Consequently, we applied survival analysis (Kaplan Meier) to estimate the incidence of delivery with PE treating the deliveries from other causes as censored observations.

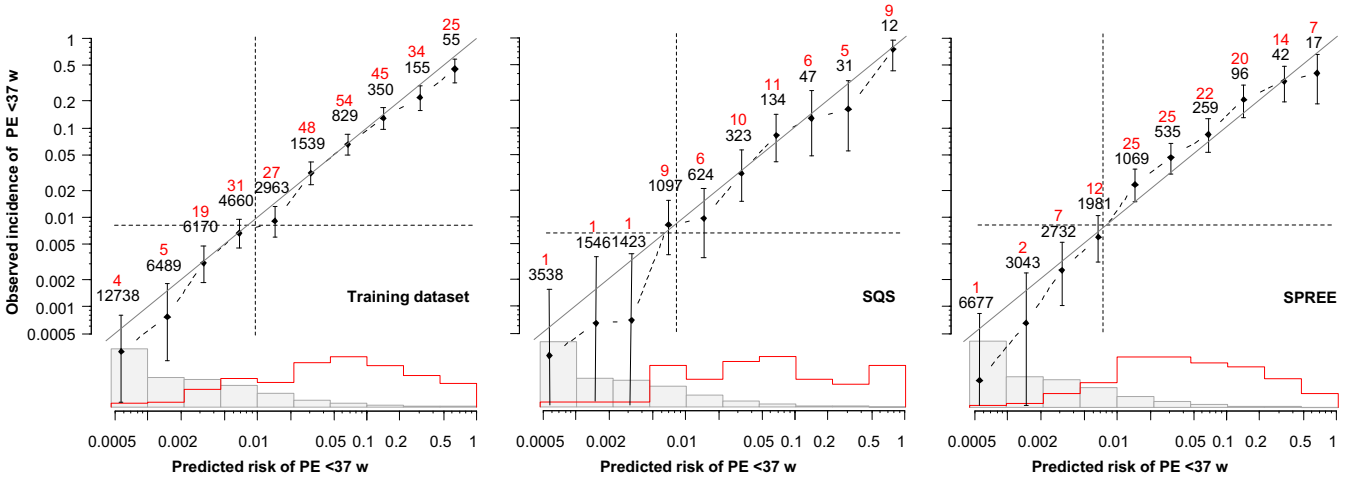
The statistical software package R was used for data analyses.[20] The package pROC was used for the receiver-operating characteristic curve analysis, and the package survival was used for survival analysis.[21–23]

Results

Maternal and pregnancy characteristics in the training set, SQS, and SPREE populations are provided and compared in Table 1.

Performance of screening for early-PE, preterm-PE, and all-PE is given in Table 2. Receiver-operating characteristics curves for the performance of screening for early-PE, preterm-PE, and all-PE in the 3 data sets and their combination by the triple test are shown in Figures 1, Supplemental Figure 1, Supplemental Figure 2 and Supplemental Figure 3. Calibration plots of the predictive performance of the competing risks model for early-PE, preterm-PE, and

FIGURE 2
Calibration plots for screening using the triple test



Calibration plots for screening using the competing risk model for prediction of preterm-PE by the triple test in the 3 data sets. The *diagonal gray line* is the line of perfect agreement. The overall mean risk is shown by the *vertical interrupted line* and the overall incidence by the *horizontal interrupted line*. The histograms show the distribution of risks in pregnancies with preterm-PE (*red*) and those without preterm-PE (*gray*).

PE, preeclampsia.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

all-PE using the triple test in the 3 data sets are shown in Figures 2, Supplemental Figure 4, Supplemental Figure 5, and Supplemental Figure 6.

The AUROC curve and DR at SPR of 10% of early-PE, preterm-PE, and all-PE in the 2 testing data sets were very similar to that in the training set. In the prediction of early-PE, preterm-PE, and all-PE by the triple test, the AUROC curve was >0.95, >0.90, and >0.80, respectively, demonstrating a very high discrimination between affected and unaffected pregnancies. Similarly, the calibration slopes were very close to 1.0, demonstrating a good agreement between the predicted risks and observed incidence of PE.

In the prediction of early-PE and preterm-PE, the observed incidence in the training and SQS data sets was consistent with the predicted one, but for SPREE the incidence was higher than predicted; this is likely to be due to better ascertainment of outcome in SPREE. The incidence of all-PE was lower than predicted in all 3 data sets (Supplemental Figure 5). After adjustment for the effect of censoring because of births from causes other than PE, the incidence was consistent with the predicted one (Supplemental Figure 6).

Comment

Main findings of the study

This study on the predictive performance of the competing risks model for PE demonstrate that the results from 2 validation data sets, derived from prospectively collected data from multicenter studies, are consistent with those of the training set used for the development of the model.

The triple test provided very high discrimination between affected and unaffected pregnancies for early-PE and preterm-PE in each of the 3 data sets with values for the AUROC curve of >0.95 and >0.90, respectively, and DR at 10% SPR of about 90% and 75%, respectively. The performance of screening at 11−13 weeks for term-PE is poor[13] and because about 70% of all cases of PE occur at term, the AUROC curve for all-PE and the DR at 10% SPR were about 0.8 and 50%, respectively.

There are 2 main reasons for poor discrimination for term-PE from screening at 11−13 weeks. First, in pregnancies with PE the deviation from normal in MAP, UtA-PI, and PLGF multiple of the median decreases with increasing gestational age, and especially for UtA-PI, there is little discrimination between term-PE and unaffected pregnancies.[13]

Second, at term many pregnancies deliver for reasons other than PE. Therefore, pregnancies considered to be at high risk for PE that deliver for other reasons are counted as false positives, even though many would have developed PE if the pregnancy had continued. More effective screening for term-PE by the competing risks model can be provided at 35−37 weeks' gestation using a combination of maternal factors, MAP, PLGF, and serum soluble fms-like tyrosine kinase-1.[24,25]

Calibration refers to how well the predictions from the model agree with the observed outcomes. Deviations between the predicted and observed outcome do not only reflect on the accuracy of a given model but could also be the consequence of differences between the studies used for development of the model and those used for validation in terms of first, methodology, and accuracy of recording maternal characteristics and medical history and the measurement of biomarkers and second, definition and ascertainment of the outcome measure.

In all 3 data sets there was prospective collection of data on maternal factors and biomarkers using a standardized protocol, the same definition of PE was used and the approach to ascertainment of outcome was similar. The results of the study demonstrate that in both the training and validation data sets calibration of risks for PE were generally good with the calibration slope very close to 1.0.

In SPREE there was a tendency for the risks to underestimate the incidence of early-PE and preterm-PE. A possible explanation for this finding is that in the training set, there was general screening for many pregnancy complications, the data were collected over many years, and many doctors were involved in the ascertainment of outcome. In contrast, SQS and SPREE were specifically designed for prediction of PE, recruitment was completed within a few months, and only 1 doctor was overall responsible for ascertainment of outcome. Indeed, the ascertainment is SPREE is likely to have been higher than in SQS because the latter focused more on the quality assurance of the biomarkers rather than the performance of the screening.

In all 3 data sets, the observed incidence of all-PE was lower than the predicted one. The likely explanation for this finding is the same as for the poorer performance of the competing risks model for term-PE because many pregnancies with estimated high risk for PE would deliver earlier than the expected event for reasons other than PE.

After adjustment for the effect of censoring because of births from causes other than PE, the observed incidence in the training set and SQS was closer to the predicted one, but in the case of SPREE, the observed incidence became higher than the predicted one; this finding could be a reflection of the higher ascertainment of cases of PE in SPREE.

Strengths and limitations

The strengths of this study include the following: first, use of a large data set of prospectively collected data on maternal factors and biomarkers to develop the model; second, prospective evaluation of discrimination and calibration of the prespecified model in 2 independent multicenter studies,[4,12] which were overseen by an independent clinical trials unit; and third, assessment of calibration allowing for the effect of censoring because of births from causes other than PE. The combined data from the studies are now being used to refine the competing risk model.

The results of the study have confirmed the accuracy of the competing risks model. However, application of the model in clinical practice necessitates the appropriate infrastructure for accurate recording of maternal characteristics and medical history, appropriate training of personnel undertaking the measurement of biomarkers and regular audit of their results, standardization of biomarkers that may vary in different populations and with different assays, use of the same outcome measures, and good ascertainment for such outcome.

Results of previous studies

A previous study examined our competing risks model in 541 nulliparous women at 11−13 weeks' gestation and reported that the DR of preterm-PE and all-PE, at a false-positive rate (FPR) of 10%, was 80% and 40%, respectively.[26] The number of cases examined is very small but the results are consistent with our findings.

A prospective study in 3066 women evaluated a previously published first-trimester algorithm for prediction of early-PE that was derived by logistic regression using maternal factors and biomarkers and reported that the DR, at 10% FPR, of early-PE was 92%, which was similar to the 95% reported in the original model.[27]

Another prospective study evaluated previously published first-trimester algorithms for prediction of PE that were derived by logistic regression using maternal factors and biomarkers.[28] The validation data set consisted of between 871 and 2962 women, depending on the variables required in the published algorithms. The DR, at an FPR of 10%, in 6 algorithms for early-PE varied from 29% to 80%, and in 2 algorithms for late-PE (≥34 weeks), it was 18% and 53%.

Implications for clinical practice

NICE and the American College of Obstetricians and Gynecologists recommend screening for PE by maternal factors and treatment of the screen-positive group with aspirin at a daily dose of 75 mg and 81 mg, respectively.[1,2] However, recent evidence suggests such an approach to the prediction and prevention of PE is likely to be ineffective because the performance of the screening method is poor and the recommended dose of aspirin is inadequate.

As demonstrated by this study, the competing risks model using the triple test can predict about 90% of early-PE and 75% of preterm-PE at an SPR of 10%; at the same SPR, the DR achieved by the methods recommended by NICE and the American College of Obstetricians and Gynecologists is half as much.[4] Treatment of the group identified by the competing risks model as being at high risk for preterm-PE with aspirin (150 mg/d from 11−14 weeks' gestation to 36 weeks) reduces the rate of preterm-PE by about 60%, early-PE by about 80%, and very early-PE by about 90%, but there is little evidence of a reduction in incidence of PE with delivery at term.[15]

Screening and the prevention of PE is also associated with a reduction in the length of stay in the neonatal intensive care unit by about 70% because about 85% of such a length of stay is due to births at <32 weeks, which are substantially reduced.[29] A secondary analysis of the ASPRE trial demonstrated that the beneficial effect of aspirin depends on adherence and the reduction in incidence of preterm-PE may be about 75% in those with adherence of ≥90% and only 40% in those with adherence of <90%.[30]

A subgroup analysis of the ASPRE trial demonstrated that there was no evidence

of heterogeneity in the beneficial effect of aspirin in reducing the incidence of preterm-PE in subgroups defined according to maternal age, body mass index, racial origin, method of conception, smoking, family history of PE, obstetrical history, and history of preexisting medical conditions, except for chronic hypertension; in chronic hypertension prophylactic use of aspirin may not be useful in the prevention of preterm-PE.[31]

Meta-analyses of trials on the use of aspirin in women at high risk for PE have reported that first, use of aspirin at a daily dose of <100 mg or onset at >16 weeks' gestation did not prevent PE, second, aspirin at ≥100 mg/d started at <16 weeks reduced the risk of preterm-PE by 67% but has no significant effect on the incidence of term-PE, and third, aspirin at ≥100 mg/d started at >16 weeks may increase the risk of placental abruption and antepartum hemorrhage.[32,33]

Conclusion

The competing risks model provides an effective and reproducible method for first-trimester prediction of preterm-PE, provided the various components of screening are carried out by appropriately trained and audited practitioners. Early prediction of preterm-PE is beneficial because treatment of the high-risk group with aspirin at a daily dose of ≥100 mg is highly effective in the prevention of the disease. ■

References

1. National Collaborating Centre for Women's and Children's Health (United Kingdom). Hypertension in pregnancy: the management of hypertensive disorders during pregnancy. London: RCOG Press; 2010.
2. American College of Obstetricians and Gynecologists. Summary: low-dose aspirin use during pregnancy. ACOG Committee opinion no. 743. Obstet Gynecol 2018;132:254–6.
3. Wright D, Syngelaki A, Akolekar R, Poon LC, Nicolaides KH. Competing risks model in screening for preeclampsia by maternal characteristics and medical history. Am J Obstet Gynecol 2015;213:62.e1–10.
4. Tan MY, Wright D, Syngelaki A, et al. 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;51:743–50.
5. Gallo DM, Wright D, Casanova C, Campanero M, Nicolaides KH. Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 19-24 weeks' gestation. Am J Obstet Gynecol 2016;214:619.e1–17.
6. Poon LC, Kametas NA, Chelemen T, Leal A, Nicolaides KH. Maternal risk factors for hypertensive disorders in pregnancy: a multivariate approach. J Hum Hypertens 2010;24:104–10.
7. Poon LC, Akolekar R, Lachmann R, Beta J, Nicolaides KH. Hypertensive disorders in pregnancy: screening by biophysical and biochemical markers at 11–13 weeks. Ultrasound Obstet Gynecol 2010;35:662–70.
8. Akolekar R, Syngelaki A, Sarquis R, Zvanca M, Nicolaides KH. Prediction of early, intermediate and late pre-eclampsia from maternal factors, biophysical and biochemical markers at 11-13 weeks. Prenat Diagn 2011;31:66–74.
9. Scazzocchio E, Figueras F, Crispi F, et al. Performance of a first-trimester screening of preeclampsia in a routine care low-risk setting. Am J Obstet Gynecol 2013;208:203.e1–10.
10. 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.
11. Wright D, Akolekar R, Syngelaki A, Poon L, Nicolaides KH. A competing risks model in early screening for preeclampsia. Fetal Diagn Ther 2012;32:171–8.
12. 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.
13. O'Gorman N, Wright D, Syngelaki A, et al. Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. Am J Obstet Gynecol 2016;214:103.e1–12.
14. O'Gorman N, Wright D, Poon LC, et al. Accuracy of competing-risks model in screening for pre-eclampsia by maternal factors and biomarkers at 11–13 weeks' gestation. Ultrasound Obstet Gynecol 2017;49:751–5.
15. Rolnik DL, Wright D, Poon LC, et al. Aspirin versus placebo in pregnancies at high risk for preterm preeclampsia. N Engl J Med 2017;377:613–22.
16. 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.
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;31:42–8.
18. Robinson HP, Fleming JE. A critical evaluation of sonar crown rump length measurements. Br J Obstet Gynaecol 1975;82:702–10.
19. Brown MA, Lindheimer MD, de Swiet M, Van Assche A, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension In Pregnancy (ISSHP). Hypertens Pregnancy 2001;20:IX–XIV.
20. Geisser S. Predictive Inference. New York: Chapman and Hall; 1993.
21. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.
22. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
23. Therneau T. A package for survival analysis in S. R package version 2.37-7 2014. Available at: http://CRAN.R-project.org/package=survival. Accessed October 5, 2018.
24. Andrietti S, Silva M, Wright A, Wright D, Nicolaides KH. Competing-risks model in screening for pre-eclampsia by maternal factors and biomarkers at 35–37 weeks' gestation. Ultrasound Obstet Gynecol 2016;48:72–9.
25. Panaitescu AM, Wright D, Militello A, Akolekar R, Nicolaides KH. Proposed clinical management of pregnancies after combined screening for pre-eclampsia at 35–37 weeks' gestation. Ultrasound Obstet Gynecol 2017;50:383–7.
26. Skråstad RB, Hov GG, Blaas HG, Romundstad PR, Salvesen KÅ. Risk assessment for preeclampsia in nulliparous women at 11–13 weeks gestational age: prospective evaluation of two algorithms. BJOG 2015;122:1781–8.
27. 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;53:532–9.
28. Oliveira N, Magder LS, Blitzer MG, Bashat A. First-trimester prediction of pre-eclampsia: external validity of algorithms in a prospectively enrolled cohort. Ultrasound Obstet Gynecol 2014;44:279–85.
29. Wright D, Rolnik DL, Syngelaki A, et al. Aspirin for Evidence-Based Preeclampsia Prevention trial: effect of aspirin on length of stay in the neonatal intensive care unit. Am J Obstet Gynecol 2018;218:612.e1–6.
30. Wright D, Poon LC, Rolnik DL, et al. 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.
31. Poon LC, Wright D, Rolnik DL, et al. Aspirin for Evidence-Based Preeclampsia Prevention trial: effect of aspirin in prevention of preterm preeclampsia in subgroups of women according

to their characteristics and medical and obstetrical history. Am J Obstet Gynecol 2017;217:585.e1–5.

32. Roberge S, Bujold E, Nicolaides KH. Aspirin for the prevention of preterm and term preeclampsia: systematic review and meta-analysis. Am J Obstet Gynecol 2018;218:287–93.e1.

33. Roberge S, Bujold E, Nicolaides KH. Meta-analysis on the effect of aspirin use for prevention of preeclampsia on placental abruption and antepartum hemorrhage. Am J Obstet Gynecol 2018;218:483–9.

Author and article information

From the Institute of Health Research, University of Exeter, Exeter (Drs D. Wright and A. Wright), and the Harris Birthright Research Centre for Fetal Medicine, King's College, London (Drs Tan, O'Gorman, Poon, Syngelaki, and Nicolaides), United Kingdom.

Received Sept. 11, 2018; revised Oct. 28, 2018; accepted Nov. 8, 2018.

The supporters of this work had no involvement in the study design; the collection, analysis, and interpretation of the data; the writing of the report; and the decision to submit the article for publication.

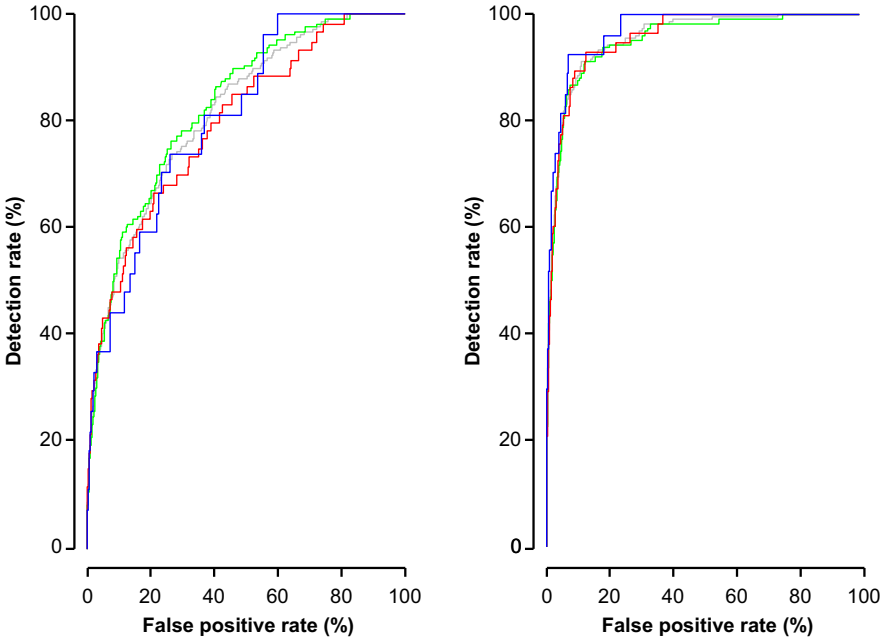
This study was supported by a grant from the Fetal Medicine Foundation (UK Charity number 1037116).

The authors report no conflict of interest.

Corresponding author: Kypros H. Nicolaides, MD. kypros@fetalmedicine.com

SUPPLEMENTAL FIGURE 1

Receiver-operating characteristic plots of screening for early-PE

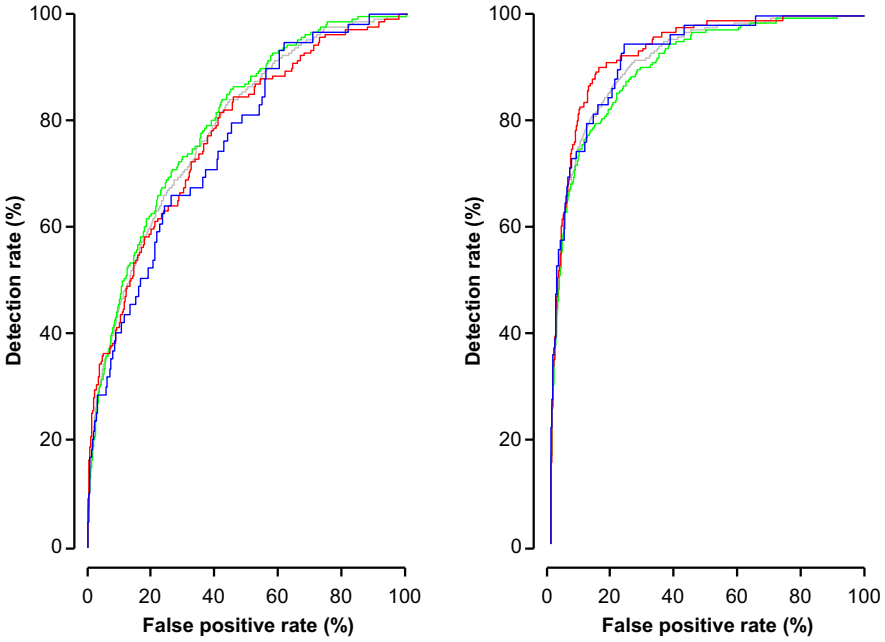


Receiver-operating characteristic plots of screening for early-PE in the 3 data sets and their combination by maternal factors (left) and the triple test (right). Training set, *green line*; SPREE, *red line*; SQS, *blue line*; combination of the 3 data sets, *gray line*.

PE, preeclampsia; *SPREE*, Superior Province Rifting EarthScope Experiment; *SQS*, screening quality study.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

SUPPLEMENTAL FIGURE 2
Receiver operating characteristic plots of screening for preterm-PE

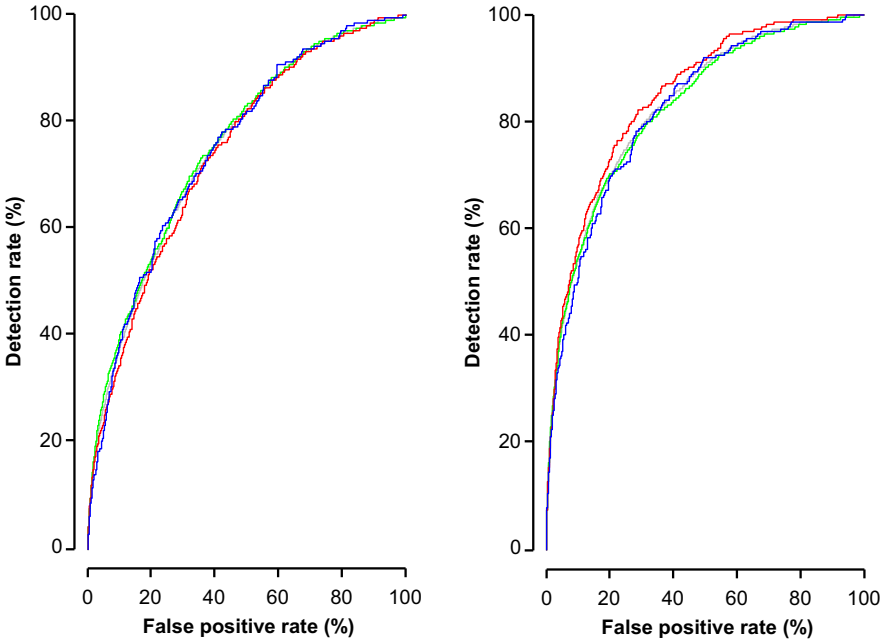


Receiver operating characteristic plots of screening for preterm-PE in the 3 data sets and their combination by maternal factors (left) and the triple test (right). Training set, *green line*; SPREE, *red line*; SQS, *blue line*; combination of the 3 data sets, *gray line*.

PE, preeclampsia; *SPREE*, Superior Province Rifting EarthScope Experiment; *SQS*, screening quality study.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

SUPPLEMENTAL FIGURE 3
Receiver operating characteristic plots of screening for all-PE

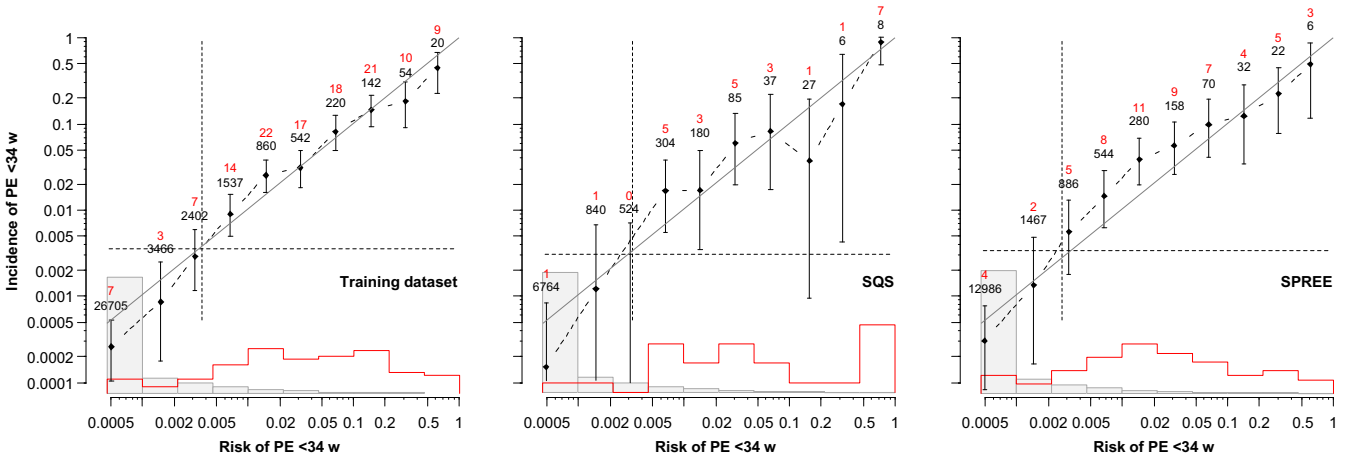


Receiver operating characteristic plots of screening for all-PE in the 3 data sets and their combination by maternal factors (left) and the triple test (right). Training set, *green line*; SPREE, *red line*; SQS, *blue line*; combination of the 3 data sets, *gray line*.

PE, preeclampsia; *SPREE*, Superior Province Rifting EarthScope Experiment; *SQS*, screening quality study.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

SUPPLEMENTAL FIGURE 4
Calibration plots for early-PE for screening using the triple test

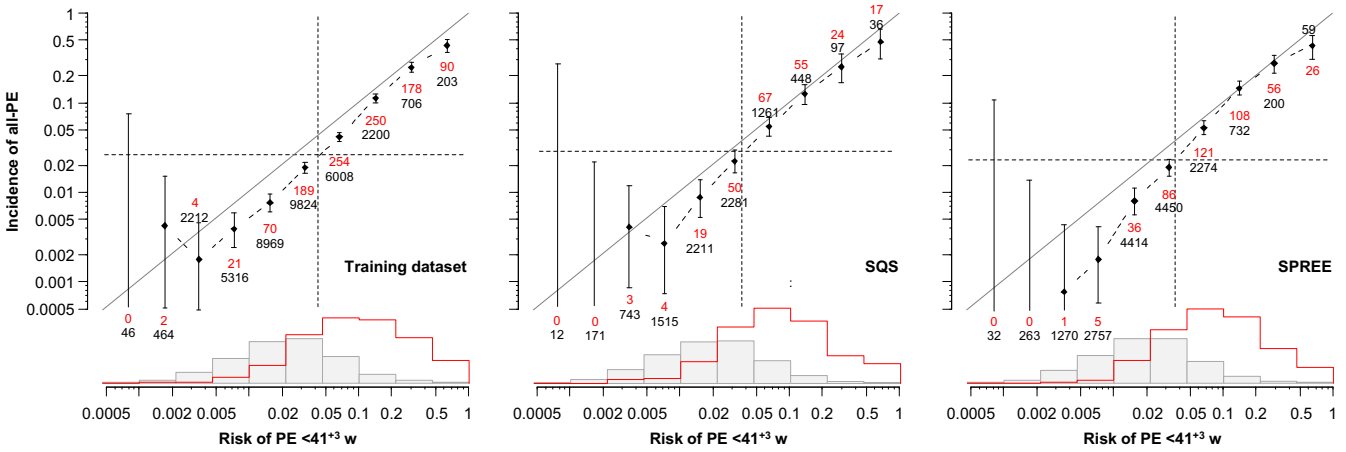


Calibration plots for screening using the competing risk model for prediction of early-PE by the triple test in the 3 data sets. The *diagonal gray line* is the line of perfect agreement. The overall mean risk is shown by the *vertical interrupted line* and the overall incidence by the *horizontal interrupted line*. The histograms show the distribution of risks in pregnancies with preterm-PE (*red*) and those without preterm-PE (*gray*).

PE, preeclampsia.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

SUPPLEMENTAL FIGURE 5
Calibration plots for all-PE for screening using the triple test

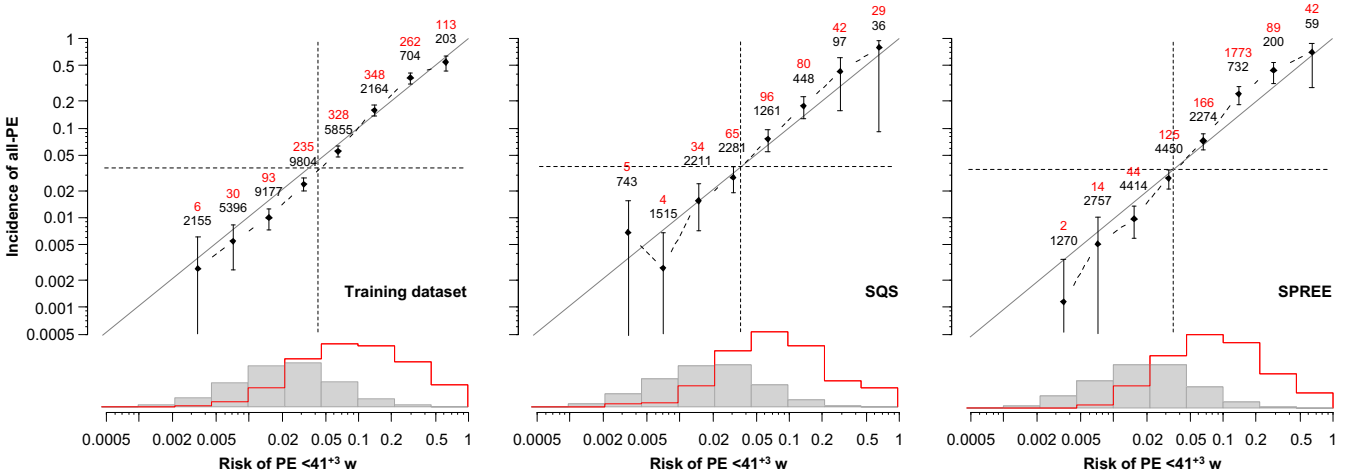


Calibration plots for screening using the competing risk model for prediction of all-PE by the triple test in the 3 data sets. The *diagonal gray line* is the line of perfect agreement. The overall mean risk is shown by the *vertical interrupted line* and the overall incidence by the *horizontal interrupted line*. The histograms show the distribution of risks in pregnancies with preterm-PE (*red*) and those without preterm-PE (*gray*).

PE, preeclampsia.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.

SUPPLEMENTAL FIGURE 6
Calibration plots for all-PE for screening using the triple test adjusted for censoring



Calibration plots for screening using the competing risk model for prediction of all-PE by the triple test in the 3 data sets. The *diagonal gray line* is the line of perfect agreement. The overall mean risk is shown by the *vertical interrupted line* and the overall incidence by the *horizontal interrupted line*. The histograms show the distribution of risks in pregnancies with all-PE (*red*) and those without PE (*gray*). The incidence counts were adjusted for the effect of censoring by multiplying the estimated incidence by the number of observations in each bin.

PE, preeclampsia.

Wright et al. Competing risks model in screening for preeclampsia. Am J Obstet Gynecol 2019.