Ultrasound Obstet Gynecol 2019; **53**: 649–654
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.20265

Validation of competing-risks model in screening for pre-eclampsia in twin pregnancy by maternal factors

Z. BENKŐ[1], P. CHAVEEVA[2], C. DE PACO MATALLANA[3], E. ZINGLER[1], A. WRIGHT[4], D. WRIGHT[4] and K. H. NICOLAIDES[1]

[1]*Harris Birthright Research Centre for Fetal Medicine, King's College Hospital, London, UK;* [2]*Dr. Shterev Hospital, Sofia, Bulgaria;* [3]*Hospital Clínico Universitario Virgen de la Arrixaca, Murcia, Spain;* [4]*Institute of Health Research, University of Exeter, Exeter, UK*

KEYWORDS: calibration; competing-risks model; discrimination; first-trimester screening; performance of screening; pre-eclampsia; survival model; twin pregnancy

ABSTRACT

Objective *To examine the predictive performance of the competing-risks model in screening for pre-eclampsia (PE) by maternal demographic characteristics and medical history in twin pregnancy, in a training dataset used for development of the model and a validation dataset.*

Methods *The data for this study were derived from two prospective non-intervention multicenter screening studies for PE in twin pregnancies at 11 + 0 to 13 + 6 weeks' gestation. The first study of 2219 women, which was reported previously, was used to develop the competing-risks model for prediction of PE and is therefore considered to be the training set. The validation study comprised 2999 women. Patient-specific risks of delivery with PE at < 34 (early), < 37 (preterm) and < 41 + 3 (all) weeks' gestation were calculated using the competing-risks model and the performance of screening for PE in the training and validation datasets was assessed. We examined the predictive performance of the model by, first, its ability to discriminate between the PE and no-PE groups using the area under the receiver–operating characteristics curve (AUC) and, second, calibration, which assesses agreement between the predicted risk and observed incidence of PE.*

Results *The incidence of early PE, preterm PE and all PE in the training and validation datasets was similar (1.8%* vs *1.4%, 5.6%* vs *5.6% and 7.7%* vs *7.2%, respectively) and this was substantially higher than in our previous studies in singleton pregnancies. The training and validation datasets had similar AUCs for early PE (0.670 (95% CI, 0.593–0.747)* vs *0.677 (95% CI, 0.594–0.760)), preterm PE (0.666 (95% CI, (0.617–0.715)* vs *0.652 (95% CI, 0.609–0.694)) and all PE (0.656 (95% CI, 0.615–0.697)* vs *0.644 (95% CI, 0.606–0.682)). Calibration plots of the predictive performance of the competing-risks model demonstrated that, in both the training and validation datasets, the observed incidence of PE was lower than the predicted one and such overestimation of risk was particularly marked for early PE.*

Conclusions *Discrimination and calibration of the competing-risks model for PE in a validation dataset are consistent with those in the training dataset. However, the model needs to be adjusted to correct the observed overestimation of risk for early PE. Copyright © 2019 ISUOG. Published by John Wiley & Sons Ltd.*

INTRODUCTION

In twin pregnancy, the incidence of pre-eclampsia (PE) is about 9%[1–11], which is three-times higher than in singleton pregnancies. However, twins are delivered at an earlier gestational age than singletons and, consequently, comparison of the overall rates of PE between twin and singleton pregnancies underestimates the relative risk of preterm PE in twins, which is nine-times higher[11]. In screening for PE in singleton pregnancies, we proposed the competing-risks approach, which is based on a survival-time model for the gestational age at delivery with PE[12–14]. Each woman has a personalized distribution of gestational age at delivery with PE, and the risk of delivery with PE before a specified gestational age, assuming no other cause of delivery, is given by the area under the probability density curve. In this approach, 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 and after development of PE. The effect of variables from maternal factors and biomarkers

Correspondence to: Prof. K. H. Nicolaides, Fetal Medicine Research Institute, King's College Hospital, 16–20 Windsor Walk, Denmark Hill, London SE5 8BB, UK (e-mail: kypros@fetalmedicine.com)

Accepted: 11 March 2019

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 development of PE. In high-risk pregnancies, the distribution is shifted to the left and the smaller the mean gestational age the higher the risk for PE.

We have examined previously 2219 twin pregnancies and proposed that the same competing-risks model developed in singleton pregnancies can be adapted for use in twins[15]. In this model, the mean gestational age at delivery with PE was 55 weeks in a reference population (white race, weight 69 kg, height 164 cm, nulliparous, spontaneous conception, no family history of PE and no history of diabetes mellitus, systemic lupus erythematosus or antiphospholipid syndrome). In dichorionic (DC) and monochorionic (MC) twin pregnancies with the same characteristics as singleton pregnancies, the distribution of gestational age at delivery with PE was shifted to the left by 8 and 10 weeks, respectively. The estimated risk of PE at < 37 weeks' gestation was 0.6% for singletons, 9.0% for DC twins and 14.2% for MC twins; the respective values for PE at < 42 weeks were 3.6%, 27.0% and 36.5%. A limitation of the study was that the performance of screening by a model derived and tested using the same dataset is overestimated and we suggested the necessity for external validation using independent data from different sources.

The objective of this study was 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 at any gestational age (all PE) in twins in the training dataset[15] used for development of the model and in a validation dataset.

METHODS

Study population

The data for this study were derived from prospective screening for adverse obstetric outcome in women attending for their routine hospital visit at 11 + 0 to 13 + 6 weeks' gestation. At this visit, we recorded maternal demographic characteristics and medical history, measured maternal weight and height and performed an ultrasound scan to determine if both fetuses were alive and had any major abnormalities, estimate gestational age from the measurement of fetal crown–rump length[16] of the larger twin, and determine chorionicity by examining the intertwin membrane at its junction with the placenta[17].

The training dataset was derived from pregnancies examined at King's College Hospital, London and Medway Maritime Hospital, Gillingham, UK, between January 2006 and December 2015[15].

The validation dataset was derived from pregnancies examined at five hospitals in the UK (King's College Hospital and Medway Maritime Hospital, between December 2015 and April 2018; Homerton University Hospital, London, between January 2014 and April 2018; North Middlesex University Hospital, London, between May 2015 and April 2018; and Southend University Hospital, Essex, between June 2015 and April 2018), one hospital in Bulgaria (Dr. Shterev Hospital, Sofia, between January 2013 and April 2018) and one hospital in Spain (Hospital Clínico Universitario Virgen de la Arrixaca, Murcia, between March 2009 and April 2018). The study was approved by the NHS Research Ethics Committee in England and the Hospital Ethics Committees of the participating hospitals in Bulgaria and Spain.

Patient characteristics included maternal age and racial origin (white, black, South Asian, East Asian and mixed), method of conception (spontaneous or assisted requiring *in-vitro* fertilization or the use of ovulation drugs), cigarette smoking during pregnancy, history of chronic hypertension, diabetes mellitus, systemic lupus erythematosus or antiphospholipid syndrome, family history of PE in the mother of the patient and obstetric history including parity (parous or nulliparous if no previous pregnancy ≥ 24 weeks), previous pregnancy with PE, gestational age at delivery and birth weight of the neonate in the last pregnancy, and interval in years between delivery of the previous pregnancy and estimated date of conception of the current pregnancy.

The inclusion criteria for this study on screening for PE were twin pregnancy with delivery of a phenotypically normal liveborn or stillborn neonate at ≥ 24 weeks' gestation. We excluded pregnancies with aneuploidy or major fetal abnormality, those ending in termination, miscarriage or fetal death before 24 weeks and those with an interval of more than 3 days between death of one fetus and live birth of the second twin.

Outcome measures

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 pre-existing or pregnancy-associated hypertension were examined to determine if the condition was PE, as defined by the International Society for the Study of Hypertension in Pregnancy[18].

Statistical analysis

Patient-specific risks of delivery with PE at < 34, < 37 and < 41 + 3 weeks' gestation were calculated using the competing-risks model based on maternal characteristics and medical history[13]. The performance of screening for early PE, preterm PE and all PE in the training and validation datasets was assessed. The number of affected cases was too small to provide separate results for DC and MC twins.

We examined the predictive performance of the model by, first, its ability to discriminate between the PE and

no-PE groups using the area under the receiver–operating characteristics (ROC) curve (AUC) (a value of 1 indicates perfect discrimination and 0.5 indicates no discrimination beyond chance) and, second, calibration, which assesses agreement between predicted risk and outcome. Calibration was assessed visually through a series of figures showing the observed incidence against that predicted from risk for PE < 34, < 37 and < 41 + 3 weeks' gestation. 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 (i.e. the mean risk within each group).

The risks produced from our competing-risks model are for delivery with PE before a specific gestational age assuming no other cause for delivery. Because other causes of delivery are effectively censored observations, the actual incidence of PE would be expected to be lower than predicted. Consequently, we applied survival analysis (Kaplan–Meier) to estimate the incidence of delivery with PE, treating deliveries due to other causes as censored observations.

The statistical software package R was used for data analyses[19]. The package pROC was used for the ROC curve analysis and the package 'survival' was used for survival analysis[20–22].

RESULTS

Maternal and pregnancy characteristics in the training and validation datasets are provided and compared in Table 1. The incidence of early PE, preterm PE and all PE in the two datasets was similar.

The ROC curves for the performance of screening for early PE, preterm PE and all PE in the two datasets and their combination are shown in Figure 1. The two datasets had similar AUCs for early PE (training dataset 0.670 (95% CI, 0.593–0.747); validation dataset 0.677 (95% CI, 0.594–0.760)), preterm PE (training dataset 0.666 (95% CI, 0.617–0.715); validation dataset 0.652 (95% CI, 0.609–0.694)) and all PE (training dataset 0.656 (95% CI, 0.615–0.697); validation dataset 0.644 (95% CI, 0.606–0.682)). Calibration plots of the predictive performance of the competing-risks model for early PE,

Table 1 Maternal and pregnancy characteristics in women with twin pregnancy included in training and validation datasets for pre-eclampsia (PE) screening model

Characteristic	*Training set* (n = *2219*)	*Validation set* (n = *2999*)	P
Maternal age (years)	32.9 (28.7–36.3)	33.7 (30.1–36.9)	< 0.00001
Maternal weight (kg)	68.0 (60.0–79.0)	66.0 (58.8–76.0)	< 0.00001
Maternal height (cm)	165 (160–170)	165 (161–170)	0.739
Maternal body mass index (kg/m^2)	24.9 (22.3–28.6)	23.9 (21.6–27.7)	< 0.00001
Gestational age (weeks)	12.9 (12.5–13.3)	12.6 (12.1–13.1)	< 0.00001
Racial origin			< 0.00001
White	1710 (77.1)	2627 (87.6)	
Black	353 (15.9)	240 (8.0)	
South Asian	80 (3.6)	78 (2.6)	
East Asian	33 (1.5)	20 (0.7)	
Mixed	43 (1.9)	34 (1.1)	
Conception			< 0.00001
Natural	1547 (69.7)	1619 (54.0)	
Assisted by use of ovulation drugs	55 (2.5)	63 (2.1)	
In-vitro fertilization	617 (27.8)	1317 (43.9)	
Medical history			
Chronic hypertension	30 (1.4)	57 (1.9)	< 0.00001
Diabetes mellitus	23 (1.0)	17 (0.6)	< 0.00001
SLE/APS	4 (0.2)	12 (0.4)	0.243
Cigarette smoker	203 (9.1)	190 (6.3)	< 0.001
Family history of PE	97 (4.4)	35 (1.2)	< 0.00001
Parity			< 0.00001
Nulliparous	1184 (53.4)	1877 (62.6)	
Parous with no previous PE	967 (43.6)	1095 (36.5)	
Parous with previous PE	68 (3.1)	27 (0.9)	
Chorionicity			0.103
Dichorionic	1789 (80.6)	2472 (82.4)	
Monochorionic	430 (19.4)	527 (17.6)	
PE			
Total	171 (7.7)	215 (7.2)	0.497
Delivery < 37 weeks	124 (5.6)	167 (5.6)	1
Delivery < 34 weeks	41 (1.8)	43 (1.4)	0.288

Data are given as median (interquartile range) or *n* (%). Comparisons between outcome groups were by chi-square or Fisher's exact test for categorical variables and Mann–Whitney *U*-test for continuous variables. APS, antiphospholipid syndrome; SLE, systemic lupus erythematosus.

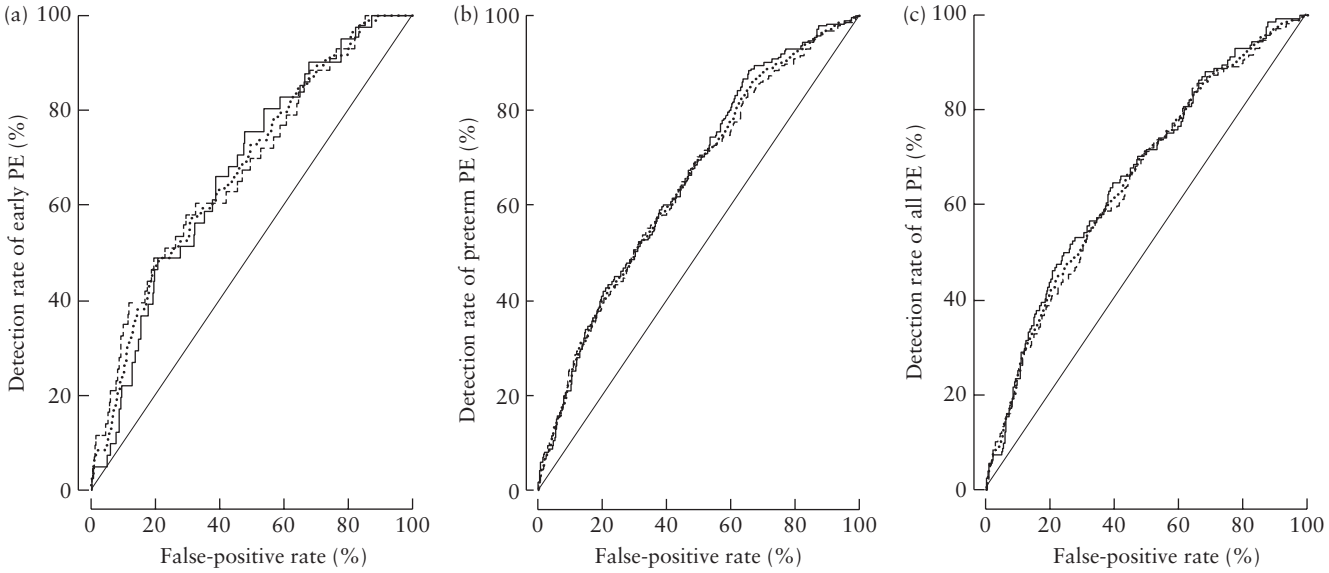

Figure 1 Receiver–operating characteristics curves in screening for early (< 34 weeks) pre-eclampsia (PE) (a), preterm (< 37 weeks) PE (b) and all (< 41 + 3 weeks) PE (c) in training dataset (——), validation dataset (– – –) and combination of the two datasets (·····).

preterm PE and all PE in the two datasets are shown in Figure 2. In both the training and validation datasets, there was a general tendency for overestimation of risk, which was most marked for early PE.

DISCUSSION

Main findings

In both the training and validation datasets, the incidence of early PE and preterm PE in twin pregnancies was substantially higher than in our previous studies in singleton pregnancies[12–14]. The findings on the predictive performance of the competing-risks model for PE in twin pregnancy demonstrate that the results in the validation dataset, derived from prospectively collected data from multicenter studies, are consistent with those in the training set used for development of the model.

The competing-risks model provided moderate discrimination between affected and unaffected pregnancies in both the training and validation datasets, with AUC values of about 0.65. This is not surprising because all twin pregnancies, compared to singletons, are at substantially increased risk of PE.

Calibration refers to how well the predicted risk from the model agrees with the observed incidence of PE. The results of the study demonstrate that, in both the training and validation datasets, the observed incidence of PE was lower than the predicted one and such overestimation of risk was particularly marked for early PE.

Strengths and limitations

The strengths of this study include, first, prospective evaluation of discrimination and calibration of the prespecified model in an independent multicenter study and, second, assessment of calibration allowing for the effect of censoring due to births from causes other than PE. A limitation of the study is that the number of twin pregnancies was too small to be divided according to chorionicity.

Comparison with previous studies

In previous studies, we established a competing-risks approach for the prediction of PE in singleton pregnancies based on maternal factors and extended this model to include twin pregnancies[13,15]. Other studies in twin pregnancies merely reported that the rate of PE is about three-times higher than in singleton pregnancies[1–11]. In a previous study, we evaluated the predictive performance of the competing-risks model in singleton pregnancies using two validation datasets and demonstrated very high discrimination between affected and unaffected pregnancies and very good agreement between the predicted risk and observed incidence of PE[15,23–25]. In this study, we compared the predictive performance of the model developed for twin pregnancies[15].

Implications for further research

In the initial development of the competing-risks model of PE in twin pregnancies, we adopted the simple approach of adjusting the model for singletons; in DC and MC twin pregnancies with the same characteristics as singleton pregnancies, the distribution of gestational age at delivery with PE was shifted to the left by 8 and 10 weeks, respectively[15]. This study has demonstrated that such an approach did not adequately address the effect of twin pregnancy on risk of PE and this was particularly so for early PE. Therefore, a new model needs to be fitted in which the effect of twins in shifting the distribution of risk in singletons to the left should not be the same for all gestational ages but such shift should be less for lower than for higher gestational ages.

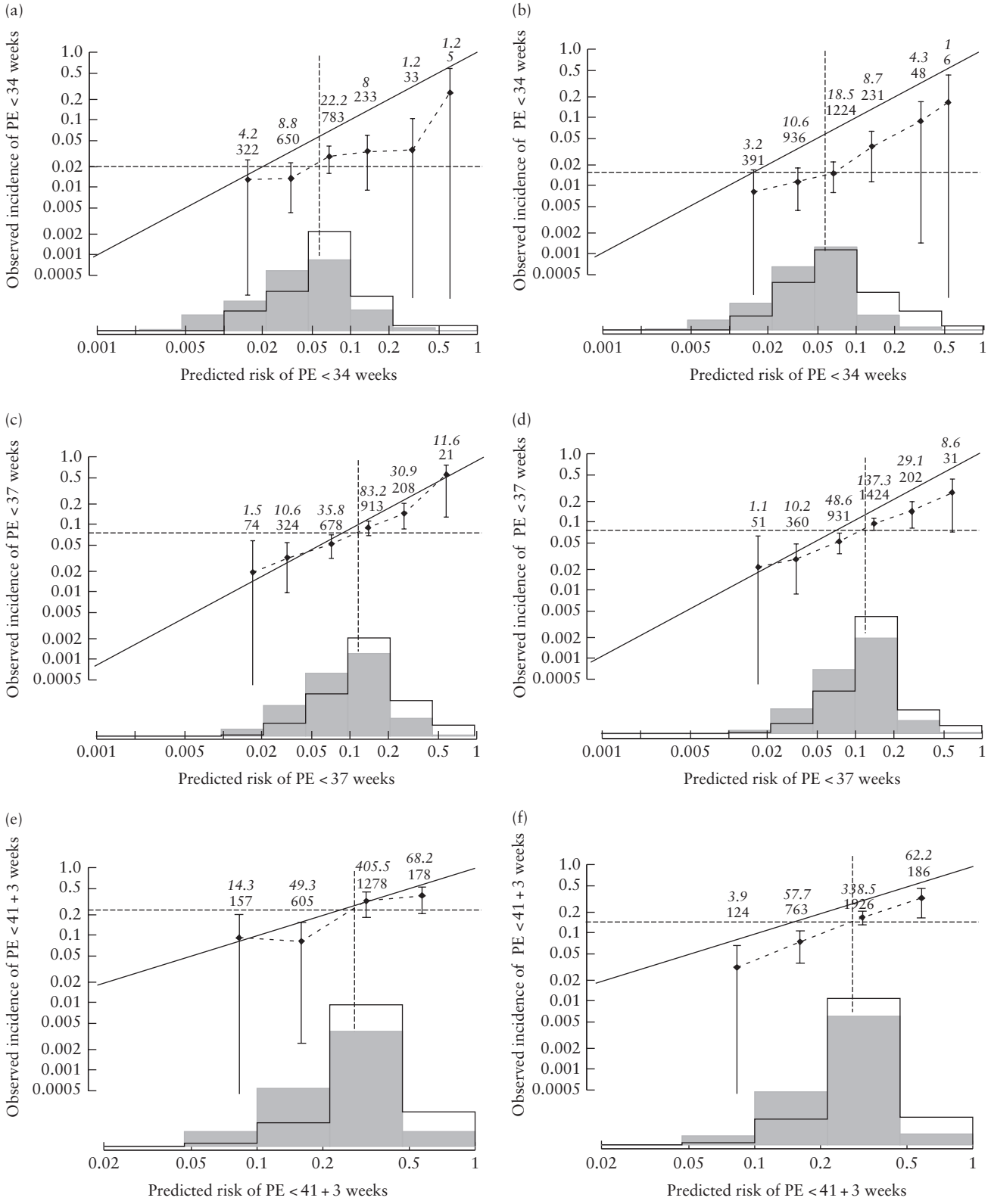

Figure 2 Calibration plots for screening using competing-risks model for prediction of early (< 34 weeks) pre-eclampsia (PE) (a,b), preterm (< 37 weeks) PE (c,d) and all (< 41 + 3 weeks) PE (e,f) in training (a,c,e) and validation (b,d,f) datasets after adjustment for effect of censoring due to births from causes other than PE. Diagonal line is line of perfect agreement. Overall mean risk is shown by vertical dashed line and overall incidence by horizontal dashed line. Vertical solid lines are confidence intervals. Numbers of women with PE are shown in italics above total number in that predicted-risk group. Histograms show distribution of risk in affected (□) and unaffected (■) pregnancies.

ACKNOWLEDGMENT

This study was supported by grants from The Fetal Medicine Foundation (Charity No: 1037116).

REFERENCES

1. Geipel A, Berg C, Germer U, Katalinic A, Krapp M, Smrcek J, Gembruch U. Doppler assessment of the uterine circulation in the second trimester in twin pregnancies: prediction of pre-eclampsia, fetal growth restriction and birth weight discordance. *Ultrasound Obstet Gynecol* 2002; **20**: 541–545.
2. Savvidou MD, Karanastasi E, Skentou C, Geerts L, Nicolaides KH. Twin chorionicity and preeclampsia. *Ultrasound Obstet Gynecol* 2001; **18**: 228–231.
3. Yu CKH, Papageorghiou AT, Boli A, Cacho AM, Nicolaides KH. Screening for pre-eclampsia and fetal growth restriction in twin pregnancies at 23 weeks of gestation by transvaginal uterine artery Doppler. *Ultrasound Obstet Gynecol* 2002; **20**: 535–540.
4. Klein K, Mailath-Pokorny M, Elhenicky M, Schmid M, Zeisler H, Worda C. Mean, lowest, and highest pulsatility index of the uterine artery and adverse pregnancy outcome in twin pregnancies. *Am J Obstet Gynecol* 2011; **205**: 549.e1–7.
5. Sparks TN, Cheng YW, Phan N, Caughey AB Sparks J. Does risk of preeclampsia differ by twin chorionicity? *Matern Fetal Neonatal Med* 2013; **26**: 1273–1277.
6. Rizzo G, Pietrolucci ME, Aiello E, Capponi A, Arduini D. Uterine artery Doppler evaluation in twin pregnancies at 11 + 0 to 13 + 6 weeks of gestation. *Ultrasound Obstet Gynecol* 2014; **44**: 557–561.
7. Bensdorp AJ, Hukkelhoven CW, van der Veen F, Mol BW, Lambalk CB, van Wely M. Dizygotic twin pregnancies after medically assisted reproduction and after natural conception: maternal and perinatal outcomes. *Fertil Steril* 2016; **106**: 371–377.e2.
8. Lučovnik M, Blickstein I, Lasič M, Fabjan-Vodušek V, Bržan-Simenc G, Verdenik I, Tul N. Hypertensive disorders during monozygotic and dizygotic twin gestations: A population-based study. *Hypertens Pregnancy* 2016; **35**: 542–547.
9. Wang YA, Chughtai AA, Farquhar CM, Pollock W, Lui K, Sullivan EA. Increased incidence of gestational hypertension and preeclampsia after assisted reproductive technology treatment. *Fertil Steril* 2016; **105**: 920–926.e2.
10. Barda G, Gluck O, Mizrachi Y, Bar J. A comparison of maternal and perinatal outcome between in vitro fertilization and spontaneous dichorionic-diamniotic twin pregnancies. *J Matern Fetal Neonatal Med* 2017; **30**: 2974–2977.
11. Francisco C, Wright D, Benkő Z, Syngelaki A, Nicolaides KH. Hidden high rate of pre-eclampsia in twin compared with singleton pregnancy. *Ultrasound Obstet Gynecol* 2017; **50**: 88–92.
12. Wright D, Akolekar R, Syngelaki A, Poon LC, Nicolaides KH. A competing risks model in early screening for preeclampsia. *Fetal Diagn Ther* 2012; **32**: 171–178.
13. 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.
14. O'Gorman N, Wright D, Syngelaki A, Akolekar R, Wright A, Poon LC, Nicolaides KH. 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.
15. Francisco C, Wright D, Benkő Z, Syngelaki A, Nicolaides KH. Competing-risks model in screening for pre-eclampsia in twin pregnancy by maternal characteristics and medical history. *Ultrasound Obstet Gynecol* 2017; **50**: 501–506.
16. Robinson HP, Fleming JE. A critical evaluation of sonar crown rump length measurements. *Br J Obstet Gynaecol* 1975; **82**: 702–710.
17. Sepulveda W, Sebire NJ, Hughes K, Odibo A, Nicolaides KH. The lambda sign at 10-14 weeks of gestation as a predictor of chorionicity in twin pregnancies. *Ultrasound Obstet Gynecol* 1996; **7**: 421–423.
18. 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.
19. Geisser S. *Predictive Inference*. Chapman and Hall: NY, New York, 1993.
20. R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. https://www.r-project.org/
21. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**: 77.
22. Therneau TM. *A Package for Survival Analysis in S. R package version* 2.37-7, 2014. http://CRAN.R-project.org/package=survival.
23. O'Gorman N, Wright D, Poon LC, Rolnik DL, Syngelaki A, Wright A, Akolekar R, Cicero S, Janga D, Jani J, Molina FS. 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.
24. Tan MY, Wright D, Syngelaki A, Akolekar R, Cicero S, Janga D, Singh M, Greco E, Wright A, Maclagan K, Poon LC. 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.
25. Wright D, Tan MY, O'Gorman N, Poon LC, Syngelaki A, Wright A, Nicolaides KH. Predictive performance of the competing risk model in screening for preeclampsia. *Am J Obstet Gynecol* 2019; **220**: 199.e1–13.

