Ultrasound Obstet Gynecol 2020; **56**: 196–205
Published online 10 July 2020 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.22129

Competing-risks model for prediction of small-for-gestational-age neonate from maternal characteristics and medical history

I. PAPASTEFANOU[1], D. WRIGHT[2] and K. H. NICOLAIDES[1]

[1]*Fetal Medicine Research Institute, King's College Hospital, London, UK;* [2]*Institute of Health Research, University of Exeter, Exeter, UK*

KEYWORDS: Bayes' theorem; fetal growth restriction; first-trimester screening; SGA; survival model

CONTRIBUTION

What are the novel findings of this work?
This study presents a new model for prediction of a small-for-gestational-age (SGA) neonate, in which gestational age (GA) at delivery and birth-weight *Z*-score are treated as continuous variables. In pregnancies at low risk for SGA, the joint distribution of GA at delivery and birth-weight *Z*-score is shifted to higher GA at delivery and birth-weight *Z*-score values, and in high-risk pregnancies, the model shifts the joint distribution towards lower values.

What are the clinical implications of this work?
Prediction of SGA generally involves dichotomization of both GA at delivery and birth-weight *Z*-score. A continuous model has been developed, in which any specific cut-off of birth-weight *Z*-score and GA at delivery can be applied to define a risk. Therefore, a single model can be used for any choice of cut-offs for birth-weight *Z*-score and GA at delivery. This model will form the basis for a Bayesian update by adding biomarkers.

ABSTRACT

Background *The established method of identifying a group of women at high risk of delivering a small-for-gestational-age (SGA) neonate, requiring increased surveillance, is use of risk scoring systems based on maternal demographic characteristics and medical history. Although this approach is relatively simple to perform, it does not provide patient-specific risks and has an uncertain performance in predicting SGA. Another approach to predict delivery of a SGA neonate is to use logistic regression models that combine maternal factors with first-trimester biomarkers. These models provide patient-specific risks for different prespecified cut-offs of birth-weight percentile and gestational age (GA) at delivery.*

Objectives *First, to develop a competing-risks model for prediction of SGA based on maternal demographic characteristics and medical history, in which GA at the time of delivery and birth-weight Z-score are treated as continuous variables. Second, to compare the predictive performance of the new model for SGA neonates to that of previous methods.*

Methods *This was a prospective observational study in 124 443 women with singleton pregnancy undergoing routine ultrasound examination at 11 + 0 to 13 + 6 weeks' gestation. The dataset was divided randomly into a training and a test dataset. The training dataset was used to develop a model for the joint distribution of GA at delivery and birth-weight Z-score from variables of maternal characteristics and medical history. This patient-specific joint Gaussian distribution of GA at delivery and birth-weight Z-score allows risk calculation for SGA defined in terms of different birth-weight percentiles and GA. The new model was then validated in the test dataset to assess performance of screening and we compared its predictive performance to that of logistic regression models for different SGA definitions.*

Results *In the new model, the joint Gaussian distribution of GA at delivery and birth-weight Z-score is shifted to lower GA at delivery and birth-weight Z-score values, resulting in an increased risk for SGA, by lower maternal weight and height, black, East Asian, South Asian and mixed racial origin, medical history of chronic hypertension, diabetes mellitus and systemic lupus erythematosus and/or antiphospholipid syndrome, conception by* in-vitro *fertilization and smoking. In parous women, variables*

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 June 2020

from the last pregnancy that increased the risk for SGA were history of pre-eclampsia or stillbirth, decreasing birth-weight Z-score and decreasing GA at delivery of the last pregnancy and interpregnancy interval < 0.5 years. In the test dataset, at a false-positive rate of 10%, the new model predicted 30.1%, 32.1%, 32.2% and 37.8% of cases of a SGA neonate with birth weight < 10th percentile delivered at < 42, < 37, < 34 and < 30 weeks' gestation, respectively, which were similar or higher than the respective values achieved by a series of logistic regression models. The calibration study demonstrated good agreement between the predicted risks and the observed incidence of SGA in both the training and test datasets.

Conclusions *A new competing-risks model, based on maternal characteristics and medical history, provides estimation of patient-specific risks for SGA in which GA at delivery and birth-weight Z-score are treated as continuous variables. Such estimation of the* a-priori *risk for SGA is an essential first step in the use of Bayes' theorem to combine maternal factors with biomarkers for the continuing development of more effective methods of screening for SGA.*

INTRODUCTION

Small-for-gestational-age (SGA) neonates are at increased risk of adverse perinatal outcome and development of metabolic and cardiovascular diseases in adult life[1–5]. With the expectation that these risks can be potentially reduced by medical intervention, national societies from many developed countries have issued guidelines on monitoring and criteria for delivery of such pregnancies[6]. However, there remains considerable uncertainty as to how best to identify SGA fetuses[7]. The established method of identifying a group of women at high risk of delivering a SGA neonate, requiring increased surveillance, is use of risk scoring systems; although this approach is relatively simple to perform, it does not provide patient-specific risks and has uncertain performance in predicting a SGA neonate[8]. Another approach to early prediction of delivery of a SGA neonate is to use logistic regression models that combine maternal factors with first-trimester biomarkers[9–12]. These models provide patient-specific risks for different prespecified cut-offs of birth-weight percentile and gestational age (GA) at delivery, which has led to an arbitrary dichotomization of the disease; different models for different SGA definitions are required, and adding new biomarkers requires refitting the whole model.

An alternative approach for prediction of a SGA neonate is to consider SGA as a spectrum disorder, the severity of which is continuously reflected in both GA at delivery and *Z*-score of birth weight for GA. The concept of this approach is similar to that of the competing-risks model in the assessment of risk for pre-eclampsia (PE)[13–15]. In this approach, which is based on a survival-time model, every woman has a personalized distribution of GA at delivery with PE and it is assumed that, if the pregnancy were to continue indefinitely, all women would develop PE; whether PE occurs depends on competition between delivery before or after development of PE. The risk of delivery with PE before a specified GA, assuming no other cause of delivery, is given by the area under the probability density curve. The new model for SGA prediction uses a continuous personalized joint bivariate Gaussian distribution of GA at delivery and *Z*-score of birth weight. The risk for any desired SGA definition is the volume under the surface of the joint probability distribution.

The objectives of this study were, first, to develop a new model for prediction of a SGA neonate, based on maternal characteristics and history, in which GA at the time of delivery and birth-weight *Z*-score are treated as continuous variables, and, second, to compare the predictive performance of the new model for a SGA neonate to that of previous methods.

METHODS

Study population

The data for this study were derived from prospective screening for adverse obstetric outcomes in women attending for their routine first hospital visit in pregnancy at King's College Hospital, London and Medway Maritime Hospital, Gillingham, UK, between March 2006 and December 2016. At this visit, at 11 + 0 to 13 + 6 weeks' gestation, we recorded maternal characteristics and medical history and performed combined screening for aneuploidy[16]. GA was determined by the measurement of fetal crown–rump length[17]. The participants gave written informed consent for the study, which was approved by the UK National Health Service Research Ethics Committee.

The inclusion criteria for the study were singleton pregnancy delivering a non-malformed liveborn or stillborn neonate at ≥ 24 weeks' gestation. We excluded pregnancies with aneuploidy or major fetal abnormality and those ending in termination, miscarriage or fetal death at < 24 weeks' gestation.

Patient characteristics

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

of conception of the current pregnancy. Maternal weight and height were measured, and body mass index was calculated.

Outcome measures

Data on pregnancy outcome were collected from hospital maternity records or the general medical practitioners of the women. The outcome measures of the study were birth of a neonate at or below different thresholds of birth-weight percentile for different cut-offs of GA at delivery. The Fetal Medicine Foundation fetal and neonatal population weight charts were used to convert birth weight to percentiles and *Z*-scores[18].

Statistical analysis

Model development

The new approach uses a single continuous model that provides a personalized joint Gaussian distribution of birth weight expressed as *Z*-scores and GA at delivery (Figure 1). Therefore, any specific cut-off for birth-weight *Z*-score and GA at delivery can be applied to define a risk. Our model for personalized bivariate Gaussian distribution of birth-weight *Z*-score and GA at delivery was specified by the following elements: a regression model for the mean for birth-weight *Z*-score, determined from maternal characteristics; a regression model for the mean for GA at delivery, determined from the mean for birth-weight *Z*-score and maternal characteristics; standard deviations for GA at delivery and birth-weight *Z*-score, which were assumed to be the same for all women and independent from maternal factors; and the correlation coefficient between GA at delivery and birth-weight *Z*-score, which was assumed to be constant for all women and independent from maternal factors. We assumed Gaussian distributions, constant standard deviations and a constant correlation coefficient for simplicity of interpretation. The model was fitted using Markov chain Monte Carlo techniques which enabled all parameters to be estimated within a single analysis. To focus the model fit on preterm SGA, GAs greater than 37 weeks were treated as censored observations at 37 weeks and birth-weight *Z*-scores greater than −1.2816 were censored at −1.2816 (Figure 2). The risk for SGA is given by the volume under the distribution surface for the region defined by the chosen birth-weight *Z*-score and GA cut-offs (Figure 1). Established risk factors, including maternal age in years, weight in kg, height in cm, racial origin, method of conception, chronic hypertension, diabetes mellitus and systemic lupus erythematosus and/or antiphospholipid syndrome were included as covariates. For parous women, interpregnancy interval in years, GA at delivery of the previous pregnancy in weeks, previous

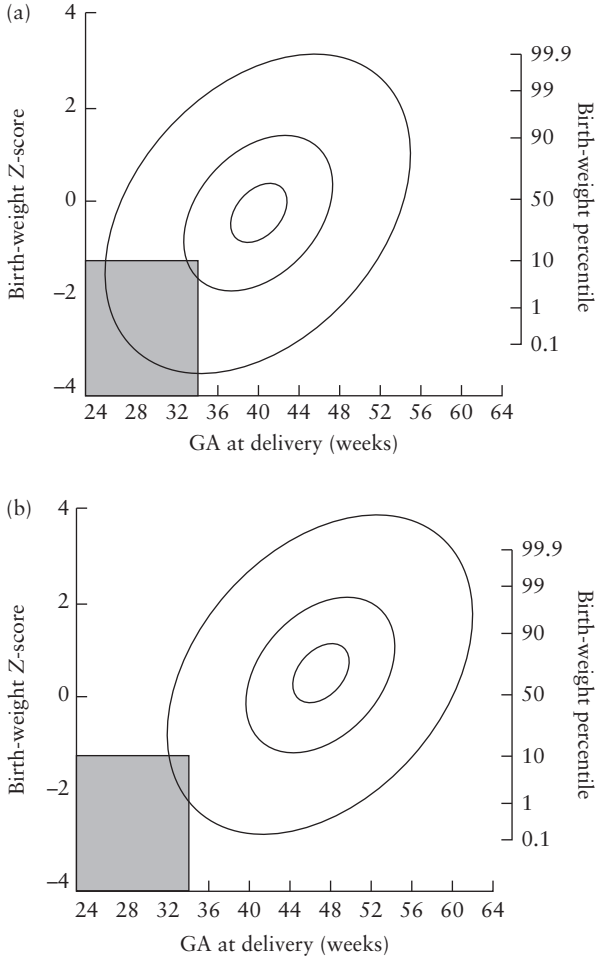

Figure 1 Contour plots of joint Gaussian distribution of birth-weight *Z*-scores and gestational age (GA) at delivery in high-risk pregnancy (a) and low-risk pregnancy (b). Birth weight expressed in percentiles is shown on vertical right axis. 50%, 75% and 95% contours are presented. Shaded area corresponds to risk of delivery before 34 weeks' gestation with small-for-gestational-age neonate with birth weight < 10th percentile.

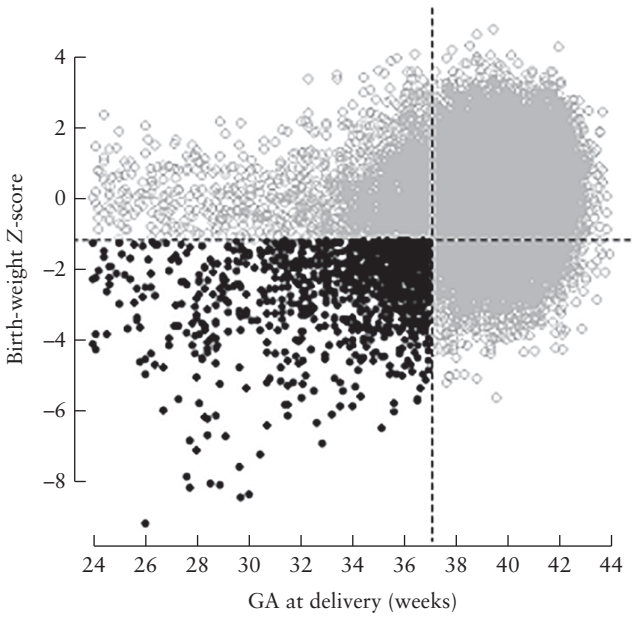

Figure 2 Scatterplot showing birth-weight *Z*-score in 124 443 singleton pregnancies, according to gestational age (GA) at delivery. Cases with birth weight > 10th percentile or GA at delivery > 37 weeks were censored (○).

birth-weight *Z*-score, history of PE and history of stillbirth were included as factors in our analysis. For maternal height, previous pregnancy birth-weight *Z*-score and GA at delivery in the previous pregnancy, a linear association was assumed. For maternal weight, quadratic terms were used, and for interpregnancy interval, fractional polynomials were adopted.

Training and test datasets

Data were partitioned into a training and a test dataset by random assignment of the sample of 124 443 pregnancies into a training dataset of 62 221 pregnancies and a test dataset of 62 222 pregnancies. The training data were used for model development and fitting. The model was then assessed in the test dataset for internal validation purposes.

Predictive performance

The predictive performance of the model was assessed by, first, the ability of the model to discriminate between the SGA and non-SGA groups using the area under the receiver-operating-characteristics curve (AUC) and the detection rate of a SGA neonate of different severities (birth weight $< 10^{th}$ or $< 3^{rd}$ percentile) at different GA cut-offs (< 42, < 37, < 34 and < 30 weeks), at fixed false-positive rates of 5%, 10% and 20%, and, second, calibration by measurements of calibration intercept and slope using logistic regression analysis of outcome incidence against the logit of the respective risks.

Comparison with previous definitions of SGA and logistic regression models

There is an apparent contradiction in the relationship between ultrasonographic estimated fetal weight (EFW) and birth-weight charts. Although the EFW recorded within a few days of birth correlates strongly with birth weight, and for a given GA they have essentially the same median[19], in reported reference ranges, the median birth weight for gestational age for babies born preterm is substantially lower than that of EFW[9,20,21]. This difference is likely to be the consequence of pathological fetal growth in a high proportion of preterm births. Reference ranges of EFW are representative of the whole population, whereas in the construction of reference ranges of birth weight, particularly for GAs at < 37 weeks, there is overrepresentation of pathological pregnancies. One-third of preterm births are iatrogenic, mainly for hypertensive disorders and/or suspected fetal growth restriction, but there is also evidence that, in a substantial proportion of spontaneous preterm births, there is impaired placentation[22–24]. This problem of underestimation of growth restriction in preterm births has been overcome through the construction of a birth-weight chart for the population of all babies at a given GA, including those still *in utero*[18].

We constructed a series of logistic regression models to predict SGA (birth weight $< 10^{th}$ and $< 3^{rd}$ percentiles for GAs at birth < 42, < 37, < 34 and < 30 weeks' gestation), defined by The Fetal Medicine Foundation birth-weight charts[18]. All these models were also validated in the test dataset.

The model was fitted within a Bayesian framework using Markov chain Monte Carlo implemented in WinBUGS[25]. The statistical software package R was also used for data analyses[26].

RESULTS

Characteristics of study population

The study population included 124 443 singleton pregnancies and the maternal and pregnancy characteristics are given in Table 1.

Model for prediction of SGA neonate

The new model provides a personalized joint distribution of birth-weight *Z*-score and GA at delivery. The model for the mean of this joint distribution is specified, first, by a regression model for the mean of birth-weight

Table 1 Maternal and pregnancy characteristics in study population of 124 443 singleton pregnancies

Characteristic	*Value*
Maternal age (years)	31.1 (26.9–35.3)
Maternal weight (kg)	67.0 (57.9–76.1)
Maternal height (cm)	164 (160–169)
BMI (kg/m^2)	24.5 (21.4–27.8)
GA at examination (weeks)	12.7 (12.3–13.1)
Racial origin	
White	93 954 (75.5)
Black	19 699 (15.8)
South Asian	5297 (4.3)
East Asian	2454 (2.0)
Mixed	3039 (2.4)
Conception	
Natural	120 302 (96.7)
Ovulation induction	1492 (1.2)
In-vitro fertilization	2649 (2.1)
Medical history	
Chronic hypertension	1569 (1.3)
Diabetes mellitus	1075 (0.9)
SLE/APS	244 (0.2)
Cigarette smoker	12 572 (10.1)
Family history of PE	5303 (4.3)
Parity	
Nulliparous	58 492 (47.0)
Parous with previous PE or SGA $< 10^{th}$	12 557 (10.1)
Parous with previous SGA $< 10^{th}$	8580 (6.9)
Parous with previous PE and SGA $< 10^{th}$	924 (0.7)
Pregnancy interval (years)	3.0 (1.5–4.5)
GA at delivery of last pregnancy (weeks)	40.0 (39.5–40.5)

Data are given as median (interquartile range) or *n* (%). 10^{th}, 10^{th} percentile; APS, antiphospholipid syndrome; BMI, body mass index; GA, gestational age; PE, pre-eclampsia; SGA, small-for-gestational age; SLE, systemic lupus erythematosus.

Z-score, and, second, by a regression model for the mean GA at delivery given the mean of birth-weight *Z*-score. Therefore, the mean GA at delivery is determined from the mean of birth-weight *Z*-score so that, when the mean of birth-weight *Z*-score is low, babies tend to be born earlier. This is reflected in the coefficient of mean birth-weight *Z*-score in the GA model (Table 2). Given the effect of smallness on GA at delivery, the new model quantifies the simultaneous effect of other variables on GA at delivery (Table 2).

This joint distribution depicted in a two-dimensional plane is a contour plot (Figure 1). The center of this contour plot is defined by the predicted mean birth-weight *Z*-score and the predicted mean GA at delivery (Table 2, Figure 1). Therefore, the coordinates of the contour plot's center are governed by maternal characteristics and medical history. The risk for SGA is given by the volume under the distribution surface for the region defined by the chosen birth-weight *Z*-score and GA cut-offs (Figure 1, Appendix S1). The lower the predicted mean birth-weight *Z*-score and the predicted mean GA, the more the contour plot falls within the chosen region and the higher the risk for SGA (Figure 1).

The factors that decreased birth-weight *Z*-score were black, South Asian, East Asian and mixed racial origin, conception by IVF, smoking, PE in the previous pregnancy, history of stillbirth in the previous pregnancy, chronic hypertension and systemic lupus erythematosus and/or antiphospholipid syndrome, whereas being parous increased birth-weight *Z*-score (Table 2). The effect on birth-weight *Z*-score for categorical variables is shown in Figure 3. The effect of maternal height was linear whereas the effect of maternal weight on birth-weight *Z*-score was positive and quadratic (Figure 4). Application of fractional polynomials revealed that interpregnancy interval had a non-linear effect on birth-weight *Z*-score with a positive peak at 2 years (Figure 4). The lower the birth weight for GA in the last pregnancy and the earlier the GA at delivery, the lower the birth-weight *Z*-score in the index pregnancy (Table 2).

The factors that decreased GA at delivery were conception by IVF, chronic hypertension, systemic lupus erythematosus and/or antiphospholipid syndrome, history of stillbirth and diabetes mellitus, whereas being parous increased predicted GA at delivery (Table 2). The effect on GA at delivery for categorical variables is shown

Table 2 Model for joint distribution of birth-weight (BW) *Z*-score and gestational age at delivery (GA)

Parameter	*Estimate*	*SD*	*LCL*	*UCL*
BW *Z*-score				
Intercept	0.4358	0.0222	0.3923	0.4793
Racial origin				
Black	−0.5436	0.0241	−0.5909	−0.4963
East Asian	−0.0468	0.0603	−0.1650	0.0714
South Asian	−0.4902	0.0390	−0.5667	−0.4137
Mixed	−0.2533	0.0545	−0.3601	−0.1465
Height (in cm) − 165	0.02789	0.00146	0.0250	0.0308
Weight (in kg) − 69	0.01138	0.00079	0.009832	0.012928
(Weight (in kg) − 69)^2	−0.0002005	0.0000216	−0.0002428	−0.0001582
In-vitro fertilization	−0.1838	0.0593	−0.2999	−0.0677
Smoker	−0.6602	0.0271	−0.7133	−0.6071
Chronic hypertension	−0.6267	0.0675	−0.7590	−0.4944
SLE/APS	−0.3309	0.1845	−0.6925	0.0307
Parous	0.05933	0.07173	−0.0813	0.1999
Last GA* (in weeks) − 40	0.06155	0.00563	0.0505	0.0726
Previous BW *Z*-score	0.3665	0.0113	0.3444	0.3886
Interpregnancy interval (in years)$^{-1}$	−0.6062	0.1179	−0.8373	−0.3751
Interpregnancy interval (in years)$^{-0.5}$	1.2990	0.1911	0.9244	1.6736
Previous PE	−0.1499	0.0513	−0.2505	−0.0493
Previous IUD	−0.1589	0.1010	−0.3569	0.0391
SD for BW *Z*-score	1.3850			
GA				
Intercept	46.790	0.1863	46.4249	47.1551
Mean BW *Z*-score	1.680	0.0519	1.5784	1.7816
In-vitro fertilization	−1.469	0.3111	−2.0788	−0.8592
Chronic hypertension	−1.827	0.3361	−2.4858	−1.1682
SLE/APS	−1.929	0.8833	−3.6603	−0.1977
Diabetes mellitus	−4.744	0.3832	−5.4951	−3.9929
Previous IUD	−1.604	0.4373	−2.4611	−0.7469
Parous	0.339	0.1086	0.1261	0.5519
Last GA* (in weeks) − 40	0.538	0.0271	0.4850	0.5912
SD for GA	6.1865			
Correlation	0.3761			

Posterior means, SD and lower (LCL) and upper (UCL) credibility limits are shown. *GA at delivery of last pregnancy. APS, antiphospholipid syndrome; IUD, intrauterine death; PE, pre-eclampsia; SLE, systemic lupus erythematosus.

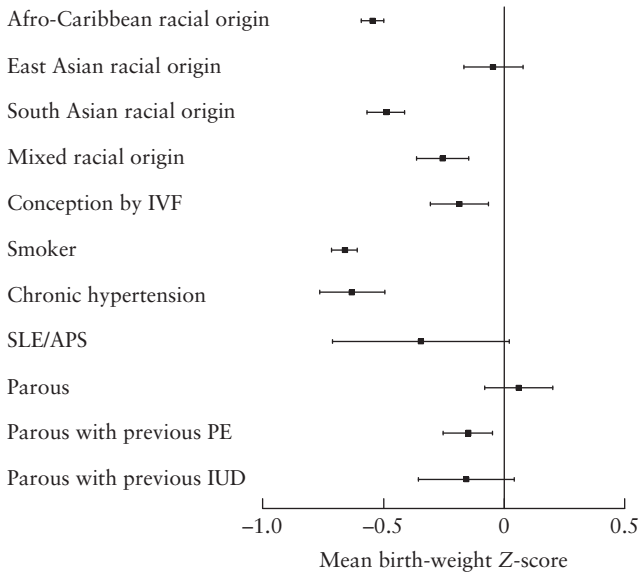

Figure 3 Forest plot showing effect of maternal factors on mean birth-weight *Z*-score. Horizontal bars represent credibility intervals. APS, antiphospholipid syndrome; IUD, intrauterine death; IVF, *in-vitro* fertilization; PE, pre-eclampsia; SLE, systemic lupus erythematosus.

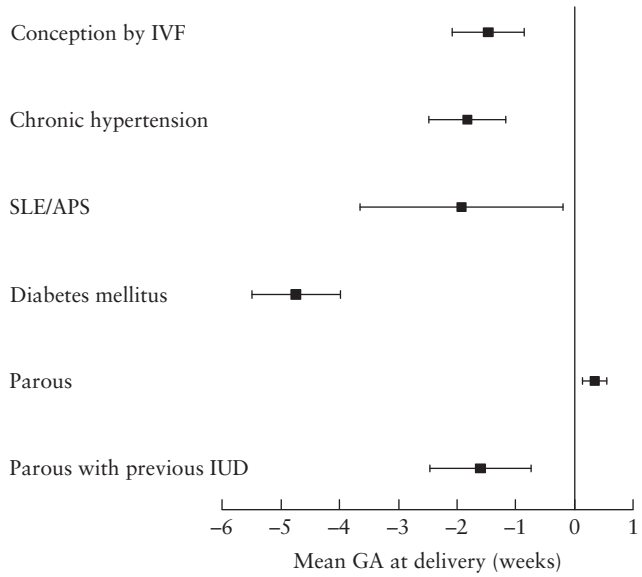

Figure 5 Forest plot showing effect of maternal factors on mean gestational age (GA) at delivery. Horizontal bars represent credibility intervals. APS, antiphospholipid syndrome; IUD, intrauterine death; IVF, *in-vitro* fertilization; SLE, systemic lupus erythematosus.

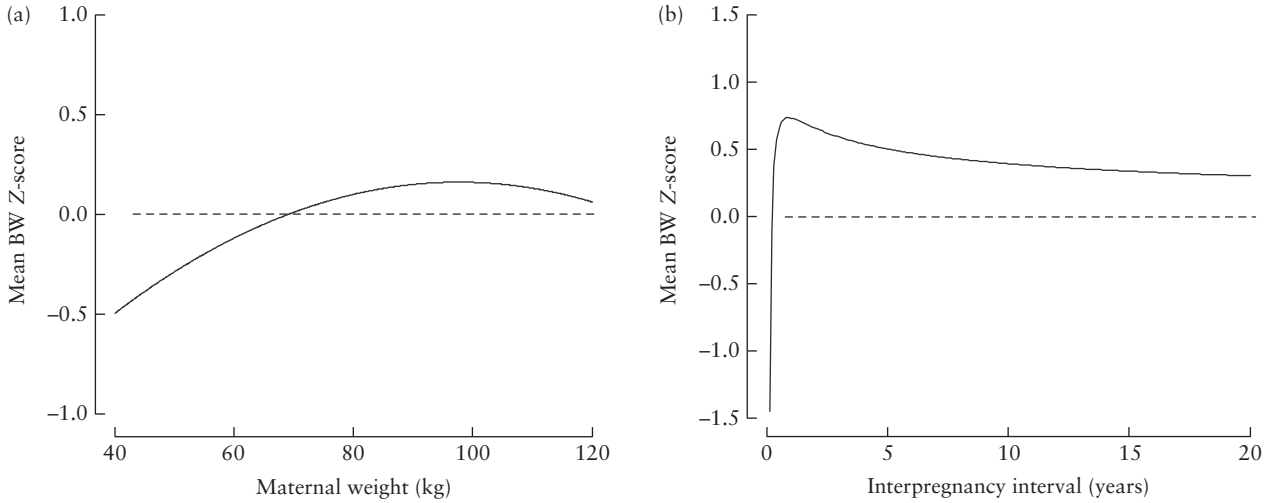

Figure 4 Non-linear effect of maternal weight (a) and interpregnancy interval (b) on mean birth-weight (BW) *Z*-score.

in Figure 5. The earlier the GA at delivery in the last pregnancy, the earlier the predicted GA at delivery.

Model evaluation

The prediction for several SGA definitions and fixed false-positive rates is presented in Table 3. The prediction was progressively better for earlier GAs and increasing severity of SGA and in parous women. The detection rates were, in the majority of cases, lower in the test dataset, as expected.

We assessed the agreement between the predicted risks by the competing-risks model for SGA and the observed incidence for different SGA definitions. The new model had satisfactory calibration for all the outcomes (Table 4, Figure 6).

Comparison of performance of new model with logistic regression models

The predictive performance of the new model for SGA with birth weight $< 10^{th}$ and $< 3^{rd}$ percentiles for GAs at birth < 42, < 37, < 34 and < 30 weeks' gestation, at fixed false-positive rates, was comparable to that of several logistic regression models (Table 5). Internal validation revealed that the new model is more stable with superior performance for preterm SGA (Table 5).

Table 3 Performance of new model in prediction of small-for-gestational-age neonate with birth weight (BW) < 10th or < 3rd percentile, for different gestational-age cut-offs at delivery, in training and test datasets, overall and according to parity

			DR (%) at FPR of:					
	AUC		*5%*		*10%*		*20%*	
Outcome measure	*Training*	*Test*	*Training*	*Test*	*Training*	*Test*	*Training*	*Test*
Delivery < 42 weeks								
BW < 10th percentile								
All pregnancies	0.7185	0.7175	18.4	17.9	30.0	30.1	47.5	48.6
Nulliparous	0.6535	0.6585	13.7	13.8	23.9	24.2	39.4	41.4
Parous	0.7648	0.7559	24.5	23.5	37.5	37.6	57.1	55.6
BW < 3rd percentile								
All pregnancies	0.7442	0.7357	21.4	20.5	34.0	33.1	52.3	51.5
Nulliparous	0.6783	0.6751	15.1	15.6	25.8	25.5	43.7	44.3
Parous	0.7936	0.7792	29.8	27.0	43.1	41.7	62.3	58.9
Delivery < 37 weeks								
BW < 10th percentile								
All pregnancies	0.7324	0.7155	24.0	21.5	33.8	32.1	51.0	48.6
Nulliparous	0.6422	0.6291	13.1	12.8	22.3	21.2	37.9	37.0
Parous	0.7972	0.7717	34.6	30.1	47.1	45.5	64.3	61.0
BW < 3rd percentile								
All pregnancies	0.7462	0.7271	25.2	21.9	35.7	32.2	53.0	50.3
Nulliparous	0.6597	0.6370	13.5	12.8	23.1	21.7	40.4	38.3
Parous	0.8091	0.7905	37.1	33.6	50.1	48.2	67.3	62.1
Delivery < 34 weeks								
BW < 10th percentile								
All pregnancies	0.7398	0.7151	26.1	24.2	36.0	32.2	52.2	49.5
Nulliparous	0.6459	0.6249	12.3	14.0	20.8	20.6	38.8	38.5
Parous	0.8052	0.7742	39.8	35.9	53.6	48.1	65.9	61.2
BW < 3rd percentile								
All pregnancies	0.7480	0.7189	25.7	22.8	36.2	30.8	52.8	48.4
Nulliparous	0.6566	0.6212	11.1	14.0	21.2	21.0	40.2	39.0
Parous	0.8155	0.7870	42.2	35.4	54.6	48.3	64.9	61.2
Delivery < 30 weeks								
BW < 10th percentile								
All pregnancies	0.7325	0.7355	26.2	28.7	35.5	37.8	50.0	53.0
Nulliparous	0.6141	0.6604	9.3	21.0	18.6	26.3	32.6	42.1
Parous	0.8062	0.7827	40.7	40.0	54.7	50.0	67.4	63.3
BW < 3rd percentile								
All pregnancies	0.7467	0.7326	28.8	27.9	39.6	38.6	51.1	50.7
Nulliparous	0.6280	0.6504	10.2	19.2	20.3	27.4	33.3	39.7
Parous	0.8240	0.7835	44.3	37.3	58.6	50.8	68.6	61.2

AUC, area under the receiver-operating-characteristics curve; DR, detection rate; FPR, false-positive rate.

Table 4 Calibration study for new model in prediction of small-for-gestational-age neonate with birth weight (BW) < 10th or < 3rd percentile, for different gestational-age (GA) cut-offs at delivery, in training and test datasets

	BW < 10th percentile		*BW < 3rd percentile*	
GA at delivery	*Slope*	*Intercept*	*Slope*	*Intercept*
< 42 weeks				
Training dataset	0.99573	0.99574	0.97292	0.65287
Test dataset	0.97931	0.98209	0.92914	0.62484
< 37 weeks				
Training dataset	0.96260	0.03388	0.91743	0.10603
Test dataset	0.87160	0.00327	0.83798	0.06913
< 34 weeks				
Training dataset	0.90030	−0.10263	0.8621	0.05647
Test dataset	0.83078	−0.09771	0.78224	0.05558
< 30 weeks				
Training dataset	0.79227	0.39565	0.79904	0.55026
Test dataset	0.7761	0.45864	0.74863	0.54341

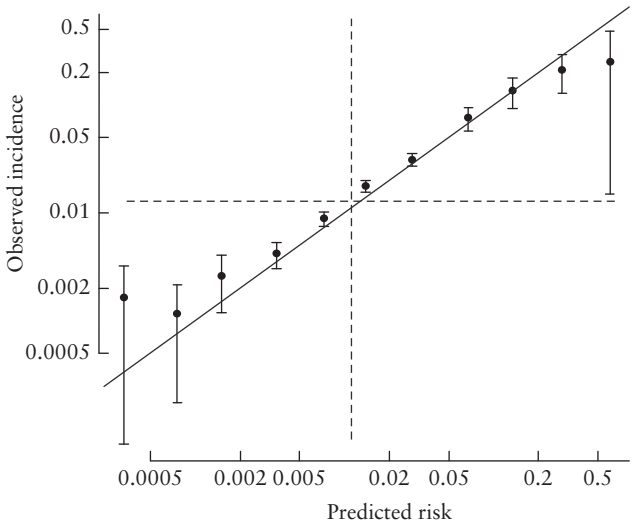

Figure 6 Calibration plot for prediction of small-for-gestational-age neonate with birth weight < 3rd percentile born < 37 weeks' gestation. Data are mean risks with confidence intervals. Horizontal dashed line represents mean observed incidence. Vertical dashed line represents mean risk predicted by new model.

DISCUSSION

Principal findings

We developed a new model based on maternal characteristics and history which provides estimation of patient-specific risks for birth of a SGA neonate, in which GA at the time of delivery and birth-weight *Z*-score are treated as continuous variables. All women have a personalized joint Gaussian distribution of birth-weight *Z*-score and GA at delivery, and maternal risk factors modify the mean of this distribution. The mean of such a joint distribution is comprised of two coordinates. The first coordinate is the predicted mean birth-weight *Z*-score and the second coordinate is the predicted mean of GA at delivery conditional on the predicted mean birth-weight *Z*-score. In pregnancies at low risk for SGA, the distribution is shifted upwards and right towards higher birth-weight *Z*-scores and gestational weeks. In high-risk pregnancies, the model shifts the distribution downwards and to the left towards lower birth-weight *Z*-score and GA values (Figure 1). A single continuous model can be

Table 5 Comparison of screening performance between new model and logistic regression models for small-for-gestational-age neonate with birth weight (BW) < 10th or < 3rd percentile, for different gestational-age cut-offs at delivery, in training and test datasets

			DR (%) at FPR of:					
	AUC		*5%*		*10%*		*20%*	
Outcome measure	*Training*	*Test*	*Training*	*Test*	*Training*	*Test*	*Training*	*Test*
Delivery < 42 weeks								
BW < 10th percentile								
New model	0.7185	0.7175	18.4	17.9	30.0	30.1	47.5	48.6
Logistic regression	0.7101	0.7193	18.9	18.1	30.4	30.6	47.9	48.5
BW < 3rd percentile								
New model	0.7442	0.7357	21.4	20.5	34.0	33.1	52.3	51.5
Logistic regression	0.7423	0.7325	21.1	20.2	33.6	32.5	52.3	50.5
Delivery < 37 weeks								
BW < 10th percentile								
New model	0.7324	0.7155	24.0	21.5	33.8	32.1	51.0	48.6
Logistic regression	0.7299	0.7158	23.2	21.8	33.3	31.4	50.5	48.5
BW < 3rd percentile								
New model	0.7462	0.7271	25.2	21.9	35.7	32.2	53.0	50.3
Logistic regression	0.7497	0.7318	24.6	22.5	36.3	32.7	55.0	51.4
Delivery < 34 weeks								
BW < 10th percentile								
New model	0.7398	0.7151	26.1	24.2	36.0	32.2	52.2	49.5
Logistic regression	0.7521	0.7256	27.4	22.1	38.3	33.6	55.9	50.6
BW < 3rd percentile								
New model	0.7480	0.7189	25.7	22.8	36.2	30.8	52.8	48.4
Logistic regression	0.7512	0.7230	27.2	22.7	41.7	32.3	56.5	49.6
Delivery < 30 weeks								
BW < 10th percentile								
New model	0.7325	0.7355	26.2	28.7	35.5	37.8	51.2	53.0
Logistic regression	0.7534	0.7205	29.0	23.0	41.9	31.4	55.8	47.0
BW < 3rd percentile								
New model	0.7467	0.7326	28.8	27.9	39.6	38.6	51.1	50.7
Logistic regression	0.7677	0.7278	32.3	25.0	45.3	30.2	56.8	45.0

AUC, area under the receiver-operating-characteristics curve; DR, detection rate; FPR, false-positive rate.

used for any choice of birth-weight *Z*-score and GA at delivery cut-offs.

Internal validation, which is actually the clinical application of the model, revealed that a single unified equation has better performance compared to a series of different logistic regression models that were fitted separately for the different SGA definitions (Table 5). The new model is more stable, with similar performance in the training and validation datasets. On the contrary, the logistic regression approach loses discrimination in the validation dataset, especially for the preterm cases. These observations provide support for the argument that SGA should be treated as a spectrum disorder rather than being fragmented arbitrarily by different cut-offs according to birth-weight percentile and GA at delivery. This study has also demonstrated that the calibration of the model is good, and this may improve stratification of pregnancy care based on risk assessment, especially for high-risk cases for preterm SGA.

Comparison with previous studies

Previous first-trimester studies that aimed to predict delivery of a SGA neonate reported similar sensitivities compared to the one achieved by the new model[9–12]. However, the predictive performance of the new approach is actually higher than that of previous models because our definition of SGA was based on the new Fetal Medicine Foundation birth-weight charts; these charts modeled efficiently the over-representation of preterm SGA pregnancies, and this has led to an increasing percentage of SGA for lower GA cut-offs[18]. Thus, we are predicting an outcome that is less extreme, compared to the previous definitions, and consequently more difficult to predict[9,18].

Implications for clinical practice

The building block of the new model for SGA prediction is an individualized joint Gaussian distribution that is defined by maternal characteristics and medical history. An important functionality of the new approach is the ability of a clinician to select any desired cut-off in birth-weight *Z*-score and GA at delivery. The selected cut-off for birth-weight *Z*-score may depend on local resources. The GA cut-off can be changed several times during the pregnancy and this flexibility will probably enhance efficient risk stratification. Such a prior model augmented by the Bayesian incorporation of biophysical and biochemical markers will improve prediction of a SGA fetus and will inevitably lead to improved future research for preventative therapeutic interventions and stratification of intensity of pregnancy monitoring. Ultimately, this may lead to improved perinatal outcome for fetuses that are growth restricted.

Strengths and limitations

The strengths of this study are, first, the large number of prospectively examined pregnancies in which maternal characteristics were recorded and specific questions were asked to obtain the medical history, as a part of an implemented screening program, second, application of a multivariate analysis that best describes the effect of each predictor, third, use of a joint distribution model that allows estimation of patient-specific risks for any desired SGA definition, and, fourth, potential for use of the model to derive the prior distribution in a Bayesian update process at different stages of pregnancy.

We have used internal validation to examine the internal validity of our model. We estimated the discrimination and calibration of the model, if it is to be trained by a dataset and then applied to a new dataset. Therefore, we know what to expect from the model in our population. A limitation of the study is the lack of external validation; we cannot demonstrate the applicability of our results in other populations, and independent data from different sources are required.

Conclusions

Birth-weight deviation and GA are intimately related; SGA is defined by its severity and prematurity. These two important elements can be combined and reflected in a continuous joint distribution. Such a unified approach facilitates the understanding and interpretation of these two important determinants of perinatal outcome. The same coefficients provide effective screening for any SGA definition. The new method of screening supports the hypothesis that SGA is a disease with a continuous severity spectrum, and arbitrary dichotomization of the condition should be avoided.

A new efficient clinical tool, with clinical applicability, has been developed. The new approach provides a framework in which different desired cut-offs of GA at delivery and birth-weight *Z*-score may be used in the context of the same model. This model will form the basis for a Bayesian update by various biomarkers at different stages in pregnancy.

ACKNOWLEDGMENT

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

REFERENCES

1. Pilliod RA, Cheng YW, Snowden JM, Doss AE, Caughey AB. The risk of intrauterine fetal death in the small-for-gestational-age fetus. *Am J Obstet Gynecol* 2012; **207**: 318–416.
2. Trudell AS, Cahill AG, Tuuli MG, Macones GA, Odibo AO. Risk of stillbirth after 37 weeks in pregnancies complicated by small-for-gestational-age fetuses. *Am J Obstet Gynecol* 2013; **208**: 376.e1–7.
3. Bukowski R, Burgett AD, Gei A, Saade GR, Hankins GD. Impairment of fetal growth potential and neonatal encephalopathy. *Am J Obstet Gynecol* 2003; **188**: 1011–1015.
4. McIntyre S, Blair E, Badawi N, Keogh J, Nelson KB. Antecedents of cerebral palsy and perinatal death in term and late preterm singletons. *Obstet Gynecol* 2013; **122**: 869–877.
5. Barker DJ. Adult consequences of fetal growth restriction. *Clin Obstet Gynecol* 2006; **49**: 270–283.

6. McCowan LM, Figueras F, Anderson NH. Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy. *Am J Obstet Gynecol* 2018; **218**: S855–S868.
7. Ciobanu A, Rouvali, A, Syngelaki A, Akolekar R, Nicolaides KH. Prediction of small for gestational age neonates: Screening by maternal factors, fetal biometry and biomarkers at 35–37 weeks' gestation. *Am J Obstet Gynecol* 2019; **220**: 486.e1–11.
8. Royal College of Obstetricians and Gynaecologists. The Investigation and Management of the Small-for-Gestational-Age Fetus. Green-top guideline No. 31. 2014.
9. Poon LC, Karagiannis G, Staboulidou I, Shafiei A, Nicolaides KH. Reference range of birth weight with gestation and first-trimester prediction of small-for-gestation neonates. *Prenat Diagn* 2011; **31**: 58–65.
10. Karagiannis G, Akolekar R, Sarquis R, Wright D, Nicolaides KH. Prediction of small-for-gestation neonates from biophysical and biochemical markers at 11–13 weeks. *Fetal Diagn Ther* 2011; **29**: 148–154.
11. Papastefanou I, Souka AP, Pilalis A, Eleftheriades M, Michalitsi V, Kassanos D. First trimester prediction of small and large for gestation neonates by an integrated model incorporating ultrasound parameters, biochemical indices and maternal characteristics. *Acta Obstet Gynecol Scand* 2012; **91**: 104–111.
12. Crovetto F, Triunfo S, Crispi F, Rodriguez-Sureda V, Dominguez C, Figueras F, Gratacos E. Differential performance of first-trimester screening in predicting small-for-gestational-age neonate or fetal growth restriction. *Ultrasound Obstet Gynecol* 2017; **49**: 349–356.
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. 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.
16. Nicolaides KH. Screening for fetal aneuploidies at 11 to 13 weeks. *Prenat Diagn* 2011; **31**: 7–15.
17. Robinson HP, Fleming JE. A critical evaluation of sonar crown rump length measurements. *Br J Obstet Gynaecol* 1975; **82**: 702–710.
18. Nicolaides KH, Wright D, Syngelaki A, Wright A, Akolekar R. Fetal Medicine Foundation fetal and neonatal population weight charts. *Ultrasound Obstet Gynecol* 2018; **52**: 44–51.
19. Hammami A, Zumaeta AM, Syngelaki A, Akolekar R, Nicolaides KH. Ultrasonographic estimation of fetal weight: development of new model and assessment of performance of previous models. *Ultrasound Obstet Gynecol* 2018; **52**: 35–43.
20. Stirnemann J, Villar J, Salomon LJ, Ohuma E, Ruyan P, Altman DG, Nosten F, Craik R, Munim S, Cheikh Ismail L, Barros FC, Lambert A, Norris S, Carvalho M, Jaffer YA, Noble JA, Bertino E, Gravett MG, Purwar M, Victora CG, Uauy R, Bhutta Z, Kennedy S, Papageorghiou AT; The International Fetal and Newborn Growth Consortium for The 21st Century (Intergrowth-21st). International estimated fetal weight standards of the Intergrowth-21st project. *Ultrasound Obstet Gynecol* 2017; **49**: 478–486.
21. Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Neerup Jensen L, Giordano D, Cecatti JG, Abdel Aleem H, Talegawkar SA, Benachi A, Diemert A, TshefuKitoto A, Thinkhamrop J, Lumbiganon P, Tabor A, Kriplani A, Gonzalez Perez R, Hecher K, Hanson MA, Gülmezoglu AM, Platt LD. The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. *PLoS Med* 2017; **14**: e1002220.
22. Arias F, Rodriguez L, Rayne SC, Kraus FT. Maternal placental vasculopathy and infection: two distinct subgroups among patients with preterm labor and preterm ruptured membranes. *Am J Obstet Gynecol* 1993; **168**: 585–591.
23. Kim YM, Chaiworapongsa T, Gomez R, Bujold E, Yoon BH, Rotmensch S, Thaler HT, Romero R. Failure of physiologic transformation of the spiral arteries in the placental bed in preterm premature rupture of membranes. *Am J Obstet Gynecol* 2002; **187**: 1137–1142.
24. Kim YM, Bujold E, Chaiworapongsa T, Gomez R, Yoon BH, Thaler HT, Rotmensch S, Romero R. Failure of physiologic transformation of the spiral arteries in patients with preterm labor and intact membranes. *Am J Obstet Gynecol* 2003; **189**: 1063–1069.
25. Gilks WR, Thomas A, Spiegelhalter DJ. A language and program for complex Bayesian modelling. *The Statistician* 1994; **43**: 169–177.
26. R Development Core Team. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Appendix S1 Risk calculation

Ultrasound Obstet Gynecol 2020; **56**: 196–205
Published online 10 July 2020 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.22129

Modelo de riesgos en competencia para la predicción de recién nacidos pequeños para la edad gestacional a partir de las características maternas y su historial médico

RESUMEN

Antecedentes El método establecido para identificar a un grupo de mujeres con alto riesgo de dar a luz a un recién nacido de tamaño pequeño para la edad gestacional (PEG), que requiere una mayor vigilancia, es el uso de sistemas de puntuación de riesgos basados en las características demográficas maternas y el historial médico. Aunque este enfoque es relativamente sencillo de aplicar, no detecta riesgos específicos para la paciente y tiene un rendimiento incierto en la predicción de neonatos PEG. Otro enfoque para predecir el parto de neonatos PEG es utilizar modelos de regresión logística que combinan factores maternos con biomarcadores del primer trimestre. Estos modelos detectan riesgos específicos para cada paciente para diferentes límites preestablecidos del percentil de peso al nacer y la edad gestacional (EG) en el momento del parto.

Objetivos En primer lugar, desarrollar un modelo de riesgos en competencia para la predicción de neonatos PEG basado en las características demográficas maternas y el historial médico, en el que la EG en el momento del parto y la puntuación Z del peso al nacer se tratan como variables continuas. En segundo lugar, comparar el comportamiento predictivo del nuevo modelo para neonatos PEG con el de los métodos anteriores.

Métodos Este fue un estudio prospectivo de observación a 124 443 mujeres con embarazos de feto único que se sometieron a una ecografía de rutina entre las 11+0 y las 13+6 semanas de gestación. El conjunto de datos se dividió al azar en un conjunto de datos de entrenamiento y otro para la prueba. El conjunto de datos de entrenamiento se utilizó para elaborar un modelo para la distribución conjunta de la EG en el parto y la puntuación Z del peso al nacer a partir de variables de las características maternas y su historial médico. Esta distribución gaussiana conjunta específica de la paciente de la EG en el parto, junto con la puntuación Z del peso al nacer, permiten el cálculo del riesgo de un neonato PEG definido en términos de diferentes percentiles del peso al nacer y la EG. El nuevo modelo fue validado con el conjunto de datos de prueba para evaluar el desempeño del cribado y se comparó su rendimiento predictivo con el de los modelos de regresión logística para las diferentes definiciones de PEG.

Resultados En el nuevo modelo, la distribución gaussiana conjunta de la EG en el parto y la puntuación Z del peso al nacer se desplaza a valores más bajos de puntuación Z del peso al nacer y de la EG en el parto, lo que da lugar a un mayor riesgo de neonatos PEG cuando el peso y la altura de la madre disminuyen, para las madres de origen negro, de Asia oriental, de Asia meridional y de razas mixtas, con antecedentes médicos de hipertensión crónica, con diabetes mellitus y lupus eritematoso sistémico y/o síndrome antifosfolipídico, con la concepción por fecundación *in vitro* y con el tabaquismo. En las mujeres que ya han tenido algún hijo, las variables del último embarazo que aumentaron el riesgo de un neonato PEG fueron un historial de preeclampsia o el éxitus fetal, la disminución de la puntuación Z del peso al nacer sumada a una menor EG en el parto del último embarazo, y un intervalo entre embarazos de <0,5 años. En el conjunto de datos de la prueba, con una tasa de falsos positivos del 10%, el nuevo modelo predijo el 30,1%, 32,1%, 32,2% y 37,8% de los casos de neonatos PEG con un peso al nacer <10º percentil para partos a <42, <37, <34 y <30 semanas de gestación, respectivamente, que fueron similares o más altos que los valores respectivos logrados por una serie de modelos de regresión logística. El estudio de calibración demostró una buena concordancia entre los riesgos previstos y la incidencia observada de neonatos PEG tanto en el conjunto de datos de entrenamiento como en el de la prueba.

Conclusiones Un nuevo modelo de riesgos en competencia, basado en las características maternas y el historial clínico, proporciona una estimación de los riesgos específicos de cada paciente de un neonato PEG en el que la EG en el momento del parto y la puntuación Z del peso al nacer se tratan como variables continuas. Esa estimación del riesgo a priori para neonatos PEG es un primer paso esencial en el uso del teorema de Bayes para combinar los factores maternos con los biomarcadores a fin de seguir desarrollando métodos más eficaces de detección de neonatos PEG.

从产妇特征和病史来预测小于胎龄新生儿的互竞风险模型

摘要

背景 识别生下小于胎龄（SGA）新生儿的一组高风险孕妇的常见方法需要增加监测，使用以产妇个人背景特征和病史为基础的风险评分系统。尽管这种方法相对简单易行，但是却不提供患者特异风险，而且在预测SGA上也有不确定性。另一种预测生下SGA新生儿的方法是使用回归分析模型，将产妇因素与早期妊娠生物指标相结合。这些模型为出生体重百分位数的不同预设界限和分娩时孕龄（GA）提供患者特异风险。

目标 首先，根据产妇个人背景特征和病史为预测SGA开发一个互竞风险模型，在该模型中将分娩时GA和出生体重Z评分看作连续变量。其次，把SGA新生儿的新模型的预测能力与以前的方法进行比较。

方法 这是一项前瞻性观察研究，有124443名单胎受试孕妇于孕期第11+0周到第13+6周进行常规超声检查。数据集被随机分成培训数据集和测试数据集。培训数据集用于根据产妇特征和病史变量，为分娩时GA的联合分布和出生体重Z评分开发一个模型。该患者特异分娩时GA的联合高斯分布和出生体重Z评分，允许从不同出生体重百分位数和GA的角度来对SGA定义风险计算。然后，将新模型在测试数据集中进行验证以评估筛检表现，而且我们针对不同的SGA定义将其预测能力与回归分析模型进行比较。

结果 在新模型中，将分娩时GA的联合高斯分布和出生体重Z评分转移到分娩时更低的GA和出生体重Z评分值，结果是孕妇体重更低、身高更矮，黑人、东亚人、南亚人和混合族源，有慢性高血压、糖尿病、全身性红斑狼疮和/或抗磷脂综合征的病史，通过体外人工授精怀孕和吸烟，都会增加SGA的风险。在经产妇女中，上一次怀孕时增加SGA风险的变量有先兆子痫病史或死产、出生体重Z评分降低和上一次怀孕的分娩时GA减少，以及怀孕间隔小于半年。在测试数据集中，当假阳性率为10%时，新模型预测了在孕期小于42周、小于37周、小于34周和小于30周时生下出生体重低于第10百分位数的SGA新生儿分别为总受试病例的30.1%、32.1%、32.2%和37.8%，这个结果类似或高于通过一系列回归分析模型得出的各个值。校验研究证明了预测风险和观测到的SGA发病率在培训数据集和测试数据集中有良好的一致性。

结论 一个基于产妇特征和病史的新互竞风险模型，为SGA提供患者特异风险估计，分娩时GA和出生体重Z评分在该模型中被看作连续变量。此类对SGA先天风险的评估，是使用贝叶斯定理结合产妇因素与生物指标的至关重要的第一步，从而为SGA开发更为有效的筛检方法。