BMJ Paediatrics Open

Cross-validated prediction model for severe adverse neonatal outcomes in a term, non-anomalous, singleton cohort

Christopher Flatley,[1] Kristen Gibbons,[2] Cameron Hurst,[3] Vicki Flenady,[1,4] Sailesh Kumar[1]

To cite: Flatley C, Gibbons K, Hurst C, *et al*. Cross-validated prediction model for severe adverse neonatal outcomes in a term, non-anomalous, singleton cohort. *BMJ Paediatrics Open* 2019;**3**:e000424. doi:10.1136/bmjpo-2018-000424

Received 19 December 2018
Revised 22 January 2019
Accepted 23 January 2019

[1]Mater Research, Mater Research Institute/University of Queensland, Brisbane, Queensland, Australia
[2]Mater Medical Research Institute, South Brisbane, Queensland, Australia
[3]QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
[4]Centre for Research Excellence in Stillbirth, Mater Research Institute/University of Queensland, Brisbane, Queensland, Australia

Correspondence to
Dr Sailesh Kumar; sailesh.kumar@mater.uq.edu.au

ABSTRACT

Objective The aim of this study was to develop a predictive model using maternal, intrapartum and ultrasound variables for a composite of severe adverse neonatal outcomes (SANO) in term infants.

Design Prospectively collected observational study. Mixed effects generalised linear models were used for modelling. Internal validation was performed using the K-fold cross-validation technique.

Setting This was a study of women that birthed at the Mater Mother's Hospital in Brisbane, Australia between January 2010 and April 2017.

Patients We included all term, non-anomalous singleton pregnancies that had an ultrasound performed between 36 and 38 weeks gestation and had recordings for the umbilical artery pulsatility index, middle cerebral artery pulsatility index and the estimated fetal weight (EFW).

Main outcome measures The components of the SANO were: severe acidosis arterial, admission to the neonatal intensive care unit, Apgar score of ≤3 at 5 min or perinatal death.

Results There were 5439 women identified during the study period that met the inclusion criteria, with 11.7% of this cohort having SANO. The final generalised linear mixed model consisted of the following variables: maternal ethnicity, socioeconomic score, nulliparity, induction of labour, method of birth and z-scores for EFW and cerebroplacental ratio. The final model had an area under the receiver operating characteristic curve of 0.71.

Conclusions The results of this study demonstrate it is possible to predict infants that are at risk of SANO at term with moderate accuracy using a combination of maternal, intrapartum and ultrasound variables. Cross-validation analysis suggests a high calibration of the model.

What is already known on this topic?

- Both the estimated fetal weight and the cerebroplacental ratio are able to identify separate cohorts of 'at-risk' fetuses.
- Individually, the estimated fetal weight and the cerebroplacental ratio are poor predictors of severe neonatal outcomes at term.
- Predictive model diagnostic accuracies have been improved with the incorporation of maternal and intrapartum variables.

What this study hopes to add?

- A model that combines maternal, intrapartum and ultrasound variables is able to predict serious adverse neonatal composite outcome with moderate accuracy.
- The proposed model supports the incorporation of a late pregnancy ultrasound as part of routine antenatal care.
- Dichotomising risk variables in any predictive model when there are significant linear associations with outcomes may improve model performance, though will miss cases that are close to the cut-off thresholds.

INTRODUCTION

Globally, hypoxia remains a major contributor to stillbirth, hypoxic ischaemic encephalopathy and cerebral palsy. For parents and families, the psychosocial and financial impact of these complications are profound and long-lasting. The majority of these catastrophic events occur despite a lack of obvious risk factors.[1] This problem is significant and pressing, with the Royal College of Obstetricians and Gynaecologists, Gates Foundation, *The Lancet* and WHO urging focused research in this area. Indeed, a recent major 2017 UK report ('Each Baby Counts') of stillbirths, neonatal deaths and perinatal brain injury occurring has set an ambitious 50% reduction target by 2020.[2]

One prerequisite of any strategy to reduce adverse outcomes is the need to identify an at-risk population of fetuses. However, there is often lack of clarity of the population being screened and the perinatal outcomes chosen. Furthermore, clinically plausible and accurate interpretation of the relationship between risk variables and health outcomes is vital to ensure the robustness of any predictive model.[3] The development of risk algorithms and predictive models using

both ultrasound and demographic variables to enable risk stratification and individualised care is an increasing focus of research to reduce stillbirth and other adverse outcomes in high-income country settings.[4] The accuracy of these models depends on careful consideration of the association between risk factors and outcomes, and importantly how these factors interact with and on occasion, confound each other.

The cerebroplacental ratio (CPR) is the ratio of the middle cerebral artery pulsatility index (MCA PI) divided by the umbilical artery pulsatility index (UA PI) and is now shown to be a possible marker of suboptimal fetal growth regardless of gestation.[5–7] A low CPR is associated with a variety of adverse perinatal outcomes including stillbirth, intrapartum fetal compromise and acidosis at birth, a low Apgar score and neonatal unit admission regardless of gestational age or weight.[8–11] The CPR is now increasingly being incorporated into clinical practice despite its relatively poor performance as a screening test for adverse perinatal outcomes.[9 12–14] Previously, we have shown that both the CPR and estimated fetal weight (EFW) identified distinct at-risk cohorts and that a model incorporating both these factors improved the predictive capability for adverse perinatal outcomes.[15] Others[16 17] have used a larger number of variables including the CPR, fetal gender, parity, maternal age, EFW and gestational age at birth to develop models for prediction of adverse pregnancy outcomes.

The aim of this study was to develop a predictive model using a range of maternal, pregnancy, intrapartum and ultrasound variables for a composite of severe adverse neonatal outcomes (SANO) for term infants.

METHODS

This study used information from clinical records of women that birthed at the Mater Mother's Hospital in Brisbane, Australia between January 2010 and April 2017. The predictive model was developed using routine prospectively collected demographic, ultrasound, intrapartum and perinatal data.

We included all term (>37 weeks gestation), non-anomalous singleton pregnancies that had an ultrasound performed between 36 and 38 weeks gestation and had recordings for the UA PI, MCA PI and the EFW. Gestational age was determined using a first trimester ultrasound examination. Fetal biometry and EFW was measured and calculated using the formula by Hadlock *et al*.[18]

The following maternal demographic, pregnancy and birth variables were extracted for the analysis: maternal age, body mass index, ethnicity, parity, smoking status, alcohol consumption, use of illicit drugs, diabetes mellitus (gestational, type 1 or type 2), hypertension (gestational, chronic or pre-eclampsia), assisted reproductive techniques, induction of labour (IOL), fetal gender, mode of birth, gestational age at birth and socioeconomic index for areas (SEIFA) score. The SEIFA score is an Australian measure of an individual's socioeconomic status where the average score is 1000 and a lower score represents relative socioeconomic deprivation.[19]

The components of the SANO were: severe acidosis (cord artery pH<7.0, lactate>6 mmol/L and/or base excess≤−12 mmol/L), admission to the neonatal intensive care unit, Apgar score of ≤3 at 5 min and/or perinatal death. Perinatal death was defined as stillbirth that occurred after >37 weeks gestation or neonatal death within 28 days of birth.

STATISTICAL ANALYSIS

Due to the change in the mean and SD over gestation for the measures of the CPR, UA PI, MCA PI and EFW, z-scores were first calculated for each gestational age when the ultrasound scan was performed, using previously published reference centiles.[20 21]

Data measured on a continuous scale are reported as mean (SD). Proportions are reported as a percentage and number of observations. Mixed effects generalised linear models with a binomial distribution were used to account for the correlation of observations from women having more than one birth within the study period. Univariable analysis was performed and all variables with a p value <0.20 were included in the initial model. This was done in consideration of the prevailing consensus opinion that at least 10 events per variable are required to avoid overfitting the model.[3 22 23]

Model building was performed using the backwards stepwise approach as previously described by Sauerbrei *et al*.[23] Variables were removed based on the highest p value and subsequent model improvement assessed through a decrease in the Akaike information criterion, a widely used criterion to assess model goodness of fit and parsimony.[24] All variables removed were individually reinserted into the model and reassessed for any model improvement.

Receiver operating characteristic (ROC) curves, sensitivity, percentage of cases correctly classified, positive and negative likelihood ratios (PLR and NLR) and positive and negative predictive values (PPV and NPV) were used to evaluate the diagnostic accuracy of the final model.

Internal validation of the model was performed using the K-fold cross-validation technique using 50 folds.[25 26] The number of SANO outcomes were compared with the number of SANO predicted by the model through the use of cross-tabulation of actual and predicted outcomes (aka confusion matrix) for the cross-validation model versus the original predictive model, and comparison of diagnostic accuracies using the original predictive model's optimum threshold from the ROC curves.

Statistical analysis was performed using Stata statistical software, V.14 (StataCorp, College Station, Texas, USA).

RESULTS

There were 5439 women during the study period that met the inclusion criteria, with 11.7% (639/5439) of this cohort having the SANO. Infants with the composite

Table 1 Demographics

		Severe adverse neonatal outcome			
	Total cohort (5439)	No (n=4800)	Yes (n=639)	OR (95% CI)	P value
Age (years)	31.0 (5.5)	31.0 (5.5)	30.3 (5.5)	0.97 (0.96 to 0.99)	0.003
BMI (kg/m^2)	25.0 (6.5)	25.0 (6.5)	25.1 (6.6)	1.00 (0.99 to 1.02)	0.60
Ethnicity					
Caucasian	50.7% (2756/5439)	50.3% (2413/4800)	53.7% (343/639)	1	
Indigenous	3.3% (180/5439)	3.3% (156/4800)	3.8% (24/639)	1.08 (0.66 to 1.77)	0.76
Asian	29.6% (1608/5439)	29.8% (1428/4800)	28.2% (180/639)	0.88 (0.71 to 1.08)	0.22
Other	16.5% (895/5439)	16.7% (803/4800)	14.4% (92/639)	0.79 (0.60 to 1.04)	0.09
SEIFA score	1017 (74)	1018 (73)	1011 (76)	0.999 (0.997 to 0.9999)	0.04
Diabetes mellitus	23.9% (1298/5439)	23.9% (1149/4800)	23.3% (149/639)	0.96 (0.77 to 1.20)	0.73
Hypertension	8.3% (449/5439)	8.3% (400/4800)	7.7% (49/639)	0.91 (0.65 to 1.28)	0.58
ART	4.9% (264/5439)	4.9% (235/4800)	4.5% (29/639)	0.90 (0.58 to 1.41)	0.66
Smokes	13.5% (733/5439)	13.5% (648/4800)	13.3% (85/639)	0.98 (0.75 to 1.29)	0.89
Alcohol	4.8% (259/5439)	4.7% (227/4800)	5.0% (32/639)	1.07 (0.70 to 1.63)	0.76
Illicit drug use	10.6% (578/5439)	10.4% (497/4800)	12.7% (81/639)	1.30 (0.97 to 1.74)	0.08
Nulliparous	45.8% (2489/5439)	43.4% (2081/4800)	63.9% (408/639)	2.50 (1.89 to 3.13)	<0.001
IOL	44.4% (2415/5439)	43.1% (2067/4800)	54.5% (348/639)	1.67 (1.33 to 2.11)	<0.001
Gestation	38.7 (1.1)	38.7 (1.1)	38.7 (1.3)	0.97 (0.89 to 1.05)	0.41
Gender (female)	50.4% (2740/5439)	50.9% (2443/4800)	46.5% (297/639)	0.83 (0.69 to 0.99)	0.04
Method of birth					
SVD	53.2% (2895/5439)	56.0% (2687/4800)	32.6% (208/639)	1	
Instrumental	13.2% (716/5439)	10.8% (520/4800)	30.7% (196/639)	5.97 (3.52 to 10.13)	<0.001
Emergency CS	16.4% (890/5439)	15.1% (726/4800)	25.7% (164/639)	3.28 (2.26 to 4.76)	<0.001
Elective CS	17.3% (938/5439)	18.1% (867/4800)	11.1% (71/639)	1.07 (0.79 to 1.45)	0.68
US gestation	36.6 (0.72)	36.6 (0.7)	36.5 (0.7)	0.85 (0.74 to 0.97)	0.02
Time from ultrasound to birth (days)	15.3 (8.6)	15.3 (8.4)	15.4 (9.5)	1.00 (0.99 to 1.01)	0.69
EFW	2969 (458)	2976 (452)	2911 (503)	0.9997 (0.999 to 0.9999)	0.002
CPR	1.99 (0.51)	2.00 (0.50)	1.93 (0.55)	0.73 (0.61 to 0.89)	0.001
UA PI	0.83 (0.15)	0.83 (0.15)	0.86 (0.16)	3.89 (1.98 to 7.65)	<0.001
MCA	1.61 (0.33)	1.61 (0.32)	1.59 (0.34)	0.83 (0.63 to 1.10)	0.20
EFW z-score	0.43 (1.10)	0.45 (1.08)	0.32 (1.24)	0.89 (0.82 to 0.97)	0.01
CPR z-score	−0.15 (1.02)	−0.13 (1.0)	−0.31 (1.12)	0.83 (0.75 to 0.91)	<0.001
UA PI z-score	0.09 (1.03)	0.07 (1.02)	0.25 (1.13)	1.20 (1.09 to 1.32)	<0.001
MCA PI z-score	−0.15 (0.98)	−0.14 (0.97)	−0.23 (1.03)	0.90 (0.82 to 0.99)	0.04

Data are reported as % (n) for categorical data and mean (SD) for continuous data.
ART, artificial reproductive technologies; BMI, body mass index; CPR, cerebroplacental ratio; CS, caesarean section; EFW, estimated fetal weight; IOL, induction of labour; MCA PI, middle cerebral artery pulsatility index; SEIFA, socioeconomic indexes for areas; SVD, spontaneous vaginal delivery; UA PI, umbilical artery pulsatility index; US, ultrasound.

SANO were more likely to be born to women who were younger (30.3 vs 31.0, p=0.001), nulliparous (63.9% vs 43.4%, p<0.001), had lower SEIFA score (1011 vs 1018, p=0.03) and were less likely to be female (46.5% vs 50.9%, p=0.04). These women were more likely to be induced (54.5% vs 43.1%, p<0.001) and have an operative delivery (instrumental delivery [30.7% vs 10.8%] and emergency caesarean [25.7% vs 15.1%], p<0.001). For the ultrasound variables, fetuses in the SANO cohort had lower mean EFW (2911 vs 2976 g, p<0.001), lower mean CPR (1.93 vs 2.00, p<0.001) and higher mean UA PI (0.86 vs 0.83, p<0.001). There was however no difference in the mean MCA PI (1.59 vs 1.61, p=0.19). After standardisation, z-scores for the EFW (0.32 vs 0.45, p=0.01), CPR (−0.31 vs −0.13, p<0.001) and MCA PI (−0.23 vs −0.14, p=0.03) were all lower in the SANO cohort while the UA

Table 2 Final model—severe adverse neonatal outcome

	OR (95% CI)	P value
Ethnicity		
Caucasian	1	
Indigenous	1.03 (0.60 to 1.79)	0.91
Asian	0.66 (0.51 to 0.86)	0.002
Other	0.73 (0.54 to 1.00)	0.049
SEIFA score	0.998 (0.996 to 0.999)	0.003
Nulliparous	1.50 (1.18 to 1.90)	0.001
IOL	1.34 (1.07 to 1.69)	0.01
Method of birth		
SVD	1	
Instrumental	5.69 (3.41 to 9.49)	<0.001
Emergency CS	3.15 (2.17 to 4.57)	<0.001
Elective CS	1.33 (0.94 to 1.88)	0.11
EFW z-score	0.88 (0.79 to 0.97)	0.01
CPR z-score	0.88 (0.79 to 0.98)	0.02

CS, caesarean section; CPR, cerebroplacental ratio, EFW; estimated fetal weight; IOL, induction of labour; SVD, spontaneous vaginal delivery.o.

PI was higher (0.25 vs 0.07, p<0.001). There was no difference in the time from ultrasound to delivery between the two groups.

After univariable analysis, associations between the SANO and maternal age (OR 0.97, 95% CI 0.96 to 0.99, p=0.003), SEIFA score (OR 0.999, 95% CI 0.997 to 0.999, p=0.04), nulliparity (OR 2.50, 95% CI 1.89 to 3.13, p<0.001), IOL (OR 1.67, 95% CI 1.33 to 2.11, p<0.001) and female gender (OR 0.83, 95% CI 0.69 to 0.99, p=0.04) were identified. The composite outcome was also associated with instrumental birth (OR 5.97, 95% CI 3.52 to 10.13, p<0.001) and emergency caesarean (OR 3.28, 95% CI 2.26 to 4.76, p<0.001) as well as z-scores for EFW (OR 0.89, 95% CI 0.82 to 0.97, p=0.01), CPR (OR 0.83, 95% CI 0.75 to 0.91, p<0.001), UA PI (OR 1.20, 95% CI 1.09 to 1.32, p<0.001) and MCA PI z-score (OR 0.90, 95% CI 0.82 to 0.99, p=0.04) (table 1).

The initial multivariable model consisted of maternal age, ethnicity, SEIFA score, illicit drug use, nulliparity, IOL, gender, method of birth, EFW z-score and CPR z-score. The UA PI and MCA PI z-scores were not included due to the association with the CPR z-score. Model selection was performed as previously described. The final generalised linear mixed model consisted of maternal ethnicity (Caucasian—reference, Indigenous [adjusted OR (aOR) 1.03, 95% CI 0.60 to 1.79, p=0.91], Asian [aOR 0.66, 95% CI 0.51 to 0.86, p=0.002], other [aOR 0.73, 95% CI 0.54 to 1.00, p=0.049]), SEIFA score (aOR 0.998, 95% CI 0.996 to 0.999, p=0.003), nulliparity (aOR 1.50, 95% CI 1.18 to 1.90, p=0.001), IOL (aOR 1.34, 95% CI 1.07 to 1.69, p=0.01), method of birth (spontaneous vaginal delivery [SVD] reference, instrumental [aOR 5.69, 95% CI 3.41 to 9.49, p<0.001], emergency caesarean [aOR 3.15, 95% CI 2.17 to 4.57, p<0.001], elective caesarean [aOR 1.33, 95% CI 0.94 to 1.88, p=0.11]) and z-scores for EFW (aOR 0.88, 95% CI 0.79 to 0.97, p=0.01) and CPR (aOR 0.88, 95% CI 0.79 to 0.98, p=0.02) (table 2).

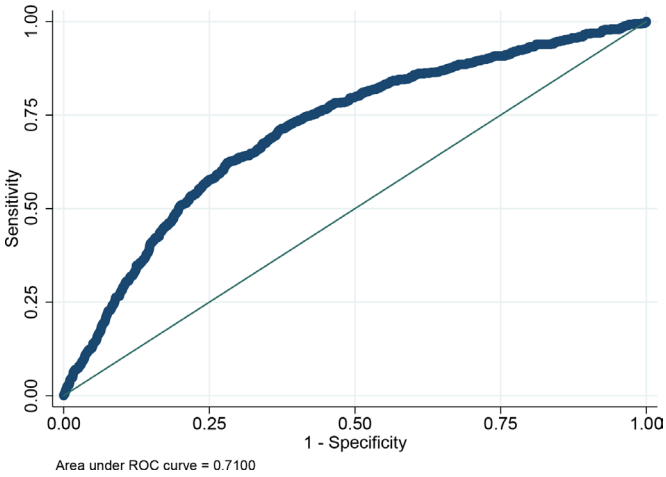

Figure 1 Receiver operating characteristics (ROC) for prediction of serious adverse neonatal outcome.

The final model had an area under the receiver operating characteristic (AUROC) curve of 0.71 (95% CI 0.69 to 0.73) (figure 1). Using a fixed false positive cut-off of 10%, the model demonstrated a sensitivity of 28.2% (95% CI 24.7 to 31.8), a PLR of 2.8 (95% CI 2.4 to 3.3) and NLR of 0.80 (95% CI 0.76 to 0.84). The PPV was 27.3% (95% CI 23.9 to 30.8), NPV of 90.4% (95% CI 89.5 to 91.2).

We also assessed the performance of the model in high-risk cohorts (EFW <10th centile and CPR <10th centile). Overall, there was negligible improvement in performance in any of the AUROC curves, but there was substantial improvement in the PPV for a cohort with an EFW <10th centile as well as those with both an EFW <10th centile and CPR <10th centile. There was also improvement in the PLR observed in the EFW <10th centile cohort (table 3).

Cross-validation of the model showed accurate and robust performance of the model with little difference between the final model (AUROC curve 0.71, 95% CI 0.69 to 0.73) compared with the cross-validation model (AUROC curve 0.70, 95% CI 0.68 to 0.72) (figure 2). Confusion matrices of the comparisons of predicted and true outcome of the SANO for the final and cross-validation model can be found in table 4, with diagnostic accuracies presented in table 5.

DISCUSSION

The results of this study demonstrate it is possible to predict with moderate accuracy, infants that are at risk of SANO at term using a combination of maternal, intrapartum and ultrasound variables. Cross-validation analysis suggests a high calibration of the model (table 4, table 5, figure 2).

Table 3 Diagnostic evaluation

	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Correctly classified	PLR (95% CI)	NLR (95% CI)	PPV (95% CI)	NPV (95% CI)
Final model	0.71 (0.69 to 0.73)	66.0% (62.2 to 69.7)	66.4% (65.1 to 67.8)	66.2%	1.97 (1.84 to 2.11)	0.51 (0.46 to 0.57)	20.7% (19.0 to 22.6)	93.6% (92.8 to 94.4)
Cohort CPR<10th centile	0.70 (0.65 to 0.75)	65.9% (56.9 to 74.1)	65.9% (62.1 to 69.5)	65.4%	1.93 (1.64 to 2.27)	0.52 (0.40 to 0.66)	26.9% (22.0 to 32.2)	91.0% (88.1 to 93.4)
Cohort EFW<10th centile	0.73 (0.67 to 0.78)	67.6% (57.9 to 76.3)	68.5% (63.9 to 72.9)	68.5%	2.15 (1.77 to 2.60)	0.47 (0.36 to 0.63)	35.3% (28.8 to 42.2)	89.3% (85.4 to 92.4)
Cohort with CPR<10th and EFW<10th centiles	0.74 (0.65 to 0.83)	64.4% (48.8 to 78.1)	65.5% (56.0 to 74.2)	65.4%	1.87 (1.34 to 2.61)	0.54 (0.36 to 0.82)	42.6% (30.7 to 55.2)	82.2% (72.7 to 89.5)
Cohort with CPR<10th or EFW<10th centiles	0.69 (0.65 to 0.73)	63.5% (56.2 to 70.4)	63.8% (60.7 to 66.8)	63.4%	1.75 (1.53 to 2.01)	0.57 (0.47 to 0.70)	25.4% (21.5 to 29.5)	90.0% (87.5 to 92.1)

AUC, area under the curve; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value.

There is increasing demand for a test to predict adverse late pregnancy outcome and the EFW and CPR are often being used to guide clinical management.[9 11 13 17 27] We have previously shown that both these variables identify separate cohorts of infants at risk of SANO and emphasise the need to incorporate both in risk stratification models.[15] In this paper, we extend our previous findings and use a variety of maternal, intrapartum and ultrasound-derived variables to develop a model for the prediction of a composite of adverse outcomes.

More complex predictive models have recently been developed to identify fetuses at risk of neonatal care unit admission and operative delivery for intrapartum fetal compromise, although in SGA cohorts.[16 28] Evaluation of our model within high-risk cohorts (SGA or low CPR) saw an improvement in the PPV as well as the PLR but only a small increase in the AUROC. Our results demonstrate that the relationship between EFW as well as CPR and SANO is linear (illustrated in figure 3) and suggests that using a threshold to categorise a higher risk cohort (eg, EFW <10th centile) based on fetal weight will affect the accuracy of a model and fail to identify fetuses that have an increased risk when their weights are close to but do not exceed the threshold.[29] Indeed, there is good evidence that the incidence of adverse outcomes including perinatal death rises when birth weight is <20th centile for gestation.[30–32] Using a predictive model that incorporates risk factors as continuous variables is more reflective of the true 'real-life' relationship with adverse outcomes. While creating predictive models in high-risk cohorts using predetermined cut-offs may provide superficially more impressive model diagnostics, they are arguably misleading and may provide false reassurance for individuals that fall outside, but are very close to the cut-off threshold.[29]

The strengths of this study lie in the large study cohort and development of a regression model which was not subjected to overfitting. We also chose components of

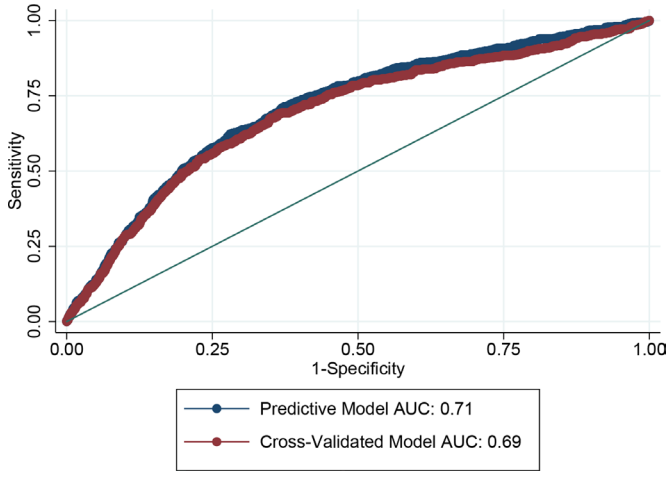

Figure 2 Comparison of predictive model and cross-validated model. AUC, area under the curve.

Table 4 Cross-validation—confusion matrix

	Predictive outcome		
True outcome	SANO	No SANO	Total
Predicted model			
SANO	422	217	639
No SANO	1611	3188	4799
Total	2033	3405	5438
Cross-validation model			
SANO	414	225	639
No SANO	1603	3196	4799
Total	2017	3421	5438

SANO, severe adverse neonatal outcome.

Table 5 Cross-validation—diagnostic evaluation

	Predictive model	Cross-validated model
Sensitivity	66.0% (62.2 to 69.7)	64.8% (60.9 to 68.5)
Specificity	66.4% (65.1 to 67.8)	66.6% (65.2 to 67.9)
PPV	20.8% (19.0 to 22.6)	20.5% (18.8 to 22.4)
NPV	93.6% (92.8 to 94.4)	93.4% (92.5 to 94.2)
PLR	2.0 (1.8 to 2.1)	1.9 (1.8 to 2.1)
NLR	0.51 (0.46 to 0.57)	0.53 (0.48 to 0.59)
Correctly classified	66.2%	66.4%
AUROC curve	0.71 (0.69 to 0.73)	0.70 (0.68 to 0.72)

NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value.

the composite outcome to reflect poor condition at birth and the association with hypoxic birth injury which are important clinically relevant outcomes. These outcomes are also correlated with both short-term morbidity such as hypoxic ischaemic encephalopathy as well as long-term complications including cerebral palsy. We also used a reasonably contemporary cohort of women so that perinatal outcomes should not have been significantly influenced by evolution in obstetric or neonatal practices. Nevertheless, there are several limitations that must be acknowledged. Although the CPR was not reported, the EFW and UA PI were, which sometimes may have influenced management decisions. Furthermore, as routine late third trimester scans are not normally performed at our institution, by definition our study cohort cannot be truly considered an unselected or low-risk population. Although the AUROC curve for our model was good, the PLR was modest suggesting only a small increase in the likelihood of the outcome. When combined with a low pretest probability of adverse outcomes at term, the veracity and clinical utility of any model needs to be interpreted with caution.[33]

Clearly, any screening test has potential for harm from false positive or false negative results. During pregnancy, a positive screen result is often followed considerable maternal anxiety, increased obstetric intervention and early term birth. Indeed, there is evidence that children born at early term gestations have higher rates of neonatal complications[34 35] and are at risk for long-term adverse neurodevelopmental sequelae.[36–38] The low rates of serious outcomes for term births constrains the development of any screening test for use in the general obstetric population and clinicians need to be cognizant of the limitations of these tests. It is possible however, that the addition of placental biomarkers may improve the performance of such models.[39 40] Despite the above-mentioned caveats, our model could be used to guide prenatal decision-making and may help guide clinical practice.

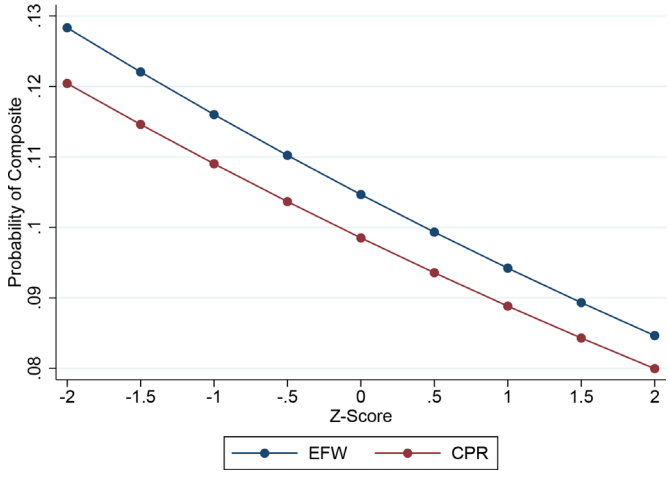

Figure 3 Adjusted probabilities of serious adverse neonatal outcome for estimated fetal weight and cerebroplacental ratio z-scores. CPR, cerebroplacental ratio; EFW, estimated fetal weight; SANO, serious adverse neonatal outcome.

Patient and public involvement

There was no patient or public involvement in the research process.

Acknowledgements The authors acknowledge research support by the Mater Foundation.

Contributors CF and SK conceived the study. CF, KG and CH designed the statistical analysis plan. CF prepared the data and carried out the analysis. CF and SK interpreted the results. VF assisted in the concept and interpretation. All authors drafted and finalised the paper.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study meets national guidelines set out by the National Health and Medical Research Council of Australia. Ethics approval was granted by The Mater Misericordiae Ltd Human Research Ethics Committee (reference number HREC/14/MHS/37).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data used throughout this study are available from Mater Research but restrictions apply to the availability. Data were used under an agreement of confidentiality and privacy and therefore not available publicly. Data are available from the authors with privacy agreements and with permission from Mater Research. Statistical code is available from authors on request.

REFERENCES

1. Low JA, Pickersgill H, Killen H, *et al*. The prediction and prevention of intrapartum fetal asphyxia in term pregnancies. *Am J Obstet Gynecol* 2001;184:724–30.
2. RCOG. *Each baby counts: 2015 full report*. London: Royal College of Obstetricians and Gynaecologists, 2017.
3. Moons KG, Royston P, Vergouwe Y, *et al*. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.
4. Flenady V, Wojcieszek AM, Middleton P, *et al*. Stillbirths: recall to action in high-income countries. *Lancet* 2016;387:691–702.
5. Baschat AA, Gembruch U. The cerebroplacental Doppler ratio revisited. *Ultrasound Obstet Gynecol* 2003;21:124–7.

6. Morales-Roselló J, Khalil A, Morlando M, *et al*. Changes in fetal Doppler indices as a marker of failure to reach growth potential at term. *Ultrasound Obstet Gynecol* 2014;43:303–10.
7. Prior T, Paramasivam G, Bennett P, *et al*. Are fetuses that fail to achieve their growth potential at increased risk of intrapartum compromise? *Ultrasound Obstet Gynecol* 2015;46:460–4.
8. Dall'Asta A, Ghi T, Rizzo G, *et al*. Early labor cerebroplacental ratio assessment in uncomplicated term pregnancies and prediction of adverse perinatal outcomes: a prospective, multicentre study. *Ultrasound Obstet Gynecol* 2018 [Epub ahead of print 13 Jun 2018].
9. Prior T, Mullins E, Bennett P, *et al*. Prediction of intrapartum fetal compromise using the cerebroumbilical ratio: a prospective observational study. *Am J Obstet Gynecol* 2013;208:124.e1–6.
10. Conde-Agudelo A, Villar J, Kennedy SH, *et al*. Predictive accuracy of cerebroplacental ratio for adverse perinatal and neurodevelopmental outcomes in suspected fetal growth restriction: systematic review and meta-analysis. *Ultrasound Obstet Gynecol* 2018;52:430–41.
11. Dunn L, Sherrell H, Kumar S. Review: Systematic review of the utility of the fetal cerebroplacental ratio measured at term for the prediction of adverse perinatal outcome. *Placenta* 2017;54:68–75.
12. Akolekar R, Syngelaki A, Gallo DM, *et al*. Umbilical and fetal middle cerebral artery Doppler at 35-37 weeks' gestation in the prediction of adverse perinatal outcome. *Ultrasound Obstet Gynecol* 2015;46:82–92.
13. Flatley C, Greer RM, Kumar S. Magnitude of change in fetal cerebroplacental ratio in third trimester and risk of adverse pregnancy outcome. *Ultrasound Obstet Gynecol* 2017;50:514–9.
14. Gibbons A, Flatley C, Kumar S. Cerebroplacental ratio in pregnancies complicated by gestational diabetes mellitus. *Ultrasound Obstet Gynecol* 2017;50:200–6.
15. Flatley C, Kumar S. Is the fetal cerebroplacental ratio better that the estimated fetal weight in predicting adverse perinatal outcomes in a low risk cohort? *The Journal of Maternal-Fetal & Neonatal Medicine* 2018;84:1–7.
16. Kalafat E, Morales-Rosello J, Thilaganathan B, *et al*. Risk of neonatal care unit admission in small for gestational age fetuses at term: a prediction model and internal validation. *J Matern Fetal Neonatal Med* 2018;15:1–8.
17. Morales-Roselló J, Khalil A, Morlando M, *et al*. Changes in fetal Doppler indices as a marker of failure to reach growth potential at term. *Ultrasound Obstet Gynecol* 2014;43:1–11.
18. Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology* 1991;181:129–33.
19. ABS. *Building on SEIFA: finer levels of socio-economic summary measures*. 1st edn. Canberra: Australian Bureau of Statistics, 2013.
20. Flatley C, Kumar S, Greer RM. Reference centiles for the middle cerebral artery and umbilical artery pulsatility index and cerebro-placental ratio from a low-risk population - a Generalised Additive Model for Location, Shape and Scale (GAMLSS) approach. *J Matern Fetal Neonatal Med* 2018:1–8.
21. Stirnemann J, Villar J, Salomon LJ, *et al*. International estimated fetal weight standards of the INTERGROWTH-21st Project. *Ultrasound Obstet Gynecol* 2017;49:478–86.
22. Peduzzi P, Concato J, Kemper E, *et al*. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373–9.
23. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* 2007;26:5512–28.
24. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974;19:716–23.
25. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Stat Surv* 2010;4:40–79.
26. Steyerberg EW, Harrell FE, Borsboom GJ, *et al*. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.
27. Cruz-Martínez R, Figueras F, Hernandez-Andrade E, *et al*. Fetal brain Doppler to predict cesarean delivery for nonreassuring fetal status in term small-for-gestational-age fetuses. *Obstet Gynecol* 2011;117:618–26.
28. Kalafat E, Morales-Rosello J, Thilaganathan B, *et al*. Risk of operative delivery for intrapartum fetal compromise in small-for-gestational-age fetuses at term: an internally validated prediction model. *Am J Obstet Gynecol* 2018;218:134.e1–34 e8.
29. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;332:1080.1.
30. Yu J, Flatley C, Greer RM, *et al*. Birth-weight centiles and the risk of serious adverse neonatal outcomes at term. *J Perinat Med* 2018;46:1048–56.
31. Dowdall D, Flatley C, Kumar S. Birth weight centiles, risk of intrapartum compromise, and adverse perinatal outcomes in term infants. *J Matern Fetal Neonatal Med* 2017;30:2126–32.
32. Moraitis AA, Wood AM, Fleming M, *et al*. Birth weight percentile and the risk of term perinatal death. *Obstet Gynecol* 2014;124(2 Pt 1):274–83.
33. Kumar S, Figueras F, Ganzevoort W, *et al*. Using cerebroplacental ratio in non-SGA fetuses to predict adverse perinatal outcome: caution is required. *Ultrasound Obstet Gynecol* 2018;52:427–9.
34. Madden JV, Flatley CJ, Kumar S. Term small-for-gestational-age infants from low-risk women are at significantly greater risk of adverse neonatal outcomes. *Am J Obstet Gynecol* 2018;218:525.e1–9.
35. Seikku L, Gissler M, Andersson S, *et al*. Asphyxia, neurologic morbidity, and perinatal mortality in early-term and postterm birth. *Pediatrics* 2016;137:e20153334.
36. Dueker G, Chen J, Cowling C, *et al*. Early developmental outcomes predicted by gestational age from 35 to 41weeks. *Early Hum Dev* 2016;103:85–90.
37. Shapiro-Mendoza C, Kotelchuck M, Barfield W, *et al*. Enrollment in early intervention programs among infants born late preterm, early term, and term. *Pediatrics* 2013;132:e61–9.
38. Spong CY, Mercer BM, D'alton M, *et al*. Timing of indicated late-preterm and early-term birth. *Obstet Gynecol* 2011;118(2 Pt 1):323–33.
39. Bligh LN, Alsolai AA, Greer RM, *et al*. Prelabor screening for intrapartum fetal compromise in low-risk pregnancies at term: cerebroplacental ratio and placental growth factor. *Ultrasound Obstet Gynecol* 2018;52:750–6.
40. Gaccioli F, Aye I, Sovio U, *et al*. Screening for fetal growth restriction using fetal biometry combined with maternal biomarkers. *Am J Obstet Gynecol* 2018;218:S725–37.