Predicting Perinatal Morbidity in Fetal Growth Restriction: Evidence, Challenges, and Opportunities

Sara E. Post, MD,† and Nathan R. Blue, MD, MS**

Abstract: Risk stratification is a core challenge in fetal growth restriction (FGR) care, in part because FGR does not represent a single diagnosis but instead is a finding that is associated with morbidity. Considerable effort has been invested in the development and study of methods to identify fetuses at risk of morbidity and who warrant intervention across multiple domains: Doppler ultrasound, maternal biomarkers, multivariable modeling, and artificial intelligence. It is likely that the most promising advances will integrate findings from across these domains, but further investigation remains necessary.

Key Words: fetal growth restriction, stillbirth, perinatal morbidity, prediction, ultrasound artificial intelligence

(*Clin Obstet Gynecol* 2025;68:571–578)

Fetal growth restriction (FGR) is a leading risk factor for stillbirth and is associated with long-term cardiovascular morbidity and early mortality among survivors, leading to a considerable global burden of disease across the lifespan.[1–3] Because FGR affects as many as 1 in 10 pregnancies globally, prenatal fetal growth assessment and screening for FGR are a core element of prenatal care.[4,5] The majority of fetuses measuring below the 10th percentile are apparently healthy and do not experience morbidity, making risk stratification and morbidity prediction a critical aspect of care in FGR.[6–8] However, accurately distinguishing between pathologic and constitutional smallness has proved to be a persistent problem with no clear solution, as neither of the leading diagnostic strategies has good performance for the prediction of perinatal morbidity in FGR.[9–11] The ongoing controversy and lack of progress are due, in no small part, to 2 facts. First, FGR is not a condition per se, but a finding that can reflect any one of multiple underlying pathologies. Second, there is no clinical outcome that is unique to FGR, such that the most practical and relevant clinical endpoints used in studies have multiple other possible causes, making studies of morbidity prediction in FGR difficult to interpret.[12] Despite these challenges, there remains considerable opportunity to improve prognostication and enable better obstetric decision making. The objectives of this review are to define the current challenges, synthesize available evidence, and highlight future directions for the prediction of perinatal morbidity in FGR.

From the *Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, University of Utah Health; and †Intermountain Health, Maternal Fetal Medicine, Salt Lake City, Utah.

The authors declare that they have nothing to disclose.

Correspondence: Nathan R. Blue, MD, MS, Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, University of Utah Health, Salt Lake City, UT. E-mail: nblue1297@gmail.com

DOI: 10.1097/GRF.0000000000000974

COMPETING DEFINITIONS OF FGR

FGR is clinically defined in 1 of 2 ways. The Society for Maternal Fetal Medicine (SMFM) defines FGR based on fetal size alone: an ultrasonographically estimated fetal weight (EFW) or abdominal circumference (AC) less than the 10th percentile, while severe FGR is defined as EFW less than the third percentile. Small for gestational age (SGA) is defined as birth weight < 10th percentile.[5] This framework considers FGR and SGA to be conceptually similar, differing only in terms of prenatal or postnatal application and not distinguishing pathologic from constitutional smallness. Alternatively, the International Society for Ultrasound in Obstetrics and Gynecology (ISUOG) attempts to distinguish pathologic from constitutional smallness by defining FGR as small fetal size at high risk of morbidity, as indicated by the presence of a finding suggestive of fetal compromise.[4] Conversely, SGA is defined as fetal or newborn size < 10th percentile in the absence of such high-risk findings, irrespective of pre- or postnatal application. These approaches are philosophical and not empirical; however, both approaches endorse frequent antenatal surveillance for all fetuses with EFW or AC below the 10th percentile and have comparably poor performance to predict composite perinatal morbidity (Table 1).[9–11]

FGR IS A FINDING, NOT A DIAGNOSIS

FGR is not a diagnosis per se, as it does not represent a single unified pathophysiology. Instead, it is a manifestation of several possible underlying pathologies, including genetic conditions, transplacental infection, or intrinsic or secondary placental dysfunction. Each of these potential etiologies can lead to perinatal morbidity in the absence of FGR. The nature of FGR as a finding rather than a distinct pathophysiology is reflected by 2 important concepts. The first is that there is no EFW percentile threshold below which the risk of perinatal morbidity changes from low to high. Instead, the risk of morbidity gradually decreases between the 10th and 80th percentiles. Therefore, the threshold by which to define FGR (or SGA, depending on one's preferred professional society) is an empirical decision based on the optimal balance of sensitivity and specificity. While this may be true in concept, it does not appear to be true in practice, as our group found that the standard threshold of EFW < 10th percentile does not perform as well as higher percentile thresholds to predict both BW < 10th percentile and perinatal morbidity.[6,13]

The second important concept underlying the nature of FGR as a finding rather than a diagnosis is reflected in the findings of the COSGROVE study, which used the DELPHI method to achieve expert consensus on the core outcomes recommended for reporting in studies of FGR.[12] Because

TABLE 1. Studies Comparing ISUOG and SMFM FGR Diagnostic Approaches for Prediction of Composite Perinatal Morbidity in FGR

Study	N	Outcome	ISUOG AUC (95% CI)	SMFM AUC (95% CI)
Roeckner et al[9]	1054	Composite perinatal morbidity	0.53 (0.50-0.55)	0.51 (0.48-0.54)
Schreiber et al[10]	19,436	Composite perinatal morbidity	0.61 (0.54-0.69)	0.63 (0.55-0.71)
Rodriguez-Sibaja et al[11]	1748	Composite perinatal morbidity	0.59 (0.56-0.61)	0.60 (0.57-0.62)

An AUC of 0.5 or a 95% confidence interval that crosses 0.5 is consistent with a test that is no better than chance at predicting the outcome of interest. AUC ≥0.8 is consistent with good prediction.

AUC indicates area under the receiver operating characteristic curve; CI, confidence interval; ISUOG, International Society for Ultrasound in Obstetrics and Gynecology; SMFM, Society for Maternal Fetal Medicine.

there is no postnatal outcome specific to FGR, the COSGROVE core outcome set largely consisted of outcomes that are associated with multiple obstetric complications (Table 2).

The lack of a clear gold standard for the diagnosis of FGR makes it difficult to interpret the reporting test characteristics for diagnostic or risk stratification approaches in FGR. This is illustrated by the following thought experiment: suppose we are designing a study assessing the diagnostic accuracy of a hypothetical test for FGR. In designing such a study, what endpoint for prediction should be used? Birth weight below a certain threshold, such as <10th or less than the third percentile, will always misclassify normal infants as abnormal and abnormal infants as normal to varying degrees, with stricter thresholds favoring specificity but sacrificing sensitivity. Predicting birth weight <10th percentile with morbidity excludes small infants with growth pathology who were not affected badly enough to experience postnatal morbidity, and those affected who experienced morbidity from growth pathology but whose birth weight was >10th percentile. Each possible endpoint contains limiting assumptions that predispose to a form of misclassification, with drawbacks for either relevance or validity (Table 3). The result of this is that all FGR studies of prediction or diagnostic accuracy suffer limitations that can never be fully mitigated due to the lack of a gold standard outcome for FGR.

DOPPLER VELOCIMETRY FOR RISK STRATIFICATION IN FGR

No matter which diagnostic strategy for FGR one chooses, the ultimate goal is the same: early identification of fetal compromise to inform carefully timed delivery and reduce perinatal morbidity and mortality. Pulse wave Doppler velocimetry allows for noninvasive interrogation of maternal and fetal vessels to detect the hemodynamic adaptations to or consequences of FGR. The most widely accepted use of Doppler for FGR surveillance is the interrogation of umbilical arterial (UA) waveforms, which has been shown to reduce perinatal mortality while being associated with reductions in both the use of labor induction and cesarean.[14] Despite the widespread support for UA Doppler surveillance in FGR, none of the 19 trials included in the Cochrane review were of high quality, and many included pregnancies affected by conditions other than FGR, such as maternal hypertension.[14] Even so, the clear signal for benefit was adequate to support widespread adoption in FGR management guidelines.

Recognition of the brain-sparing phenomenon in FGR led to the study of fetal middle cerebral arterial (MCA) waveforms and other derivatives, such as the cerebroplacental ratio (CPR), or the ratio of the MCA pulsatility index (PI) to the UA PI, to improve the identification of fetal compromise not otherwise detected by UA Doppler indices. Such a capability would be especially relevant in late-onset FGR, in which UA Doppler indices are known to be less informative in early-onset FGR.[15] Reduced pulsatility in the MCA reflects increased diastolic flow suggestive of cerebral redistribution, since the normal hemodynamic pattern in the MCA is of high pulsatility. Alternatively, a CPR <1.0 reflects lower pulsatility in the MCA than in the UA, a reversal of the typical relationship. Studies assessing low MCA PI and low CPR consistently demonstrate an association with adverse perinatal outcomes, but direct comparisons of each with UA PI do not consistently show improved prediction of adverse perinatal outcomes.[16] In an individual patient meta-analysis of 18,731 participants assessing the added value of CPR in addition to UA PI for prediction of adverse perinatal outcomes, Vollgraff Heidweiller-Schreurs et al[17] found that MCA PI, CPR, and UA+CPR predicted perinatal morbidity no better than UA PI alone, though they were not able to sub-analyze early and late onset FGR groups (Table 4). The same research group

TABLE 2. Core Outcomes for FGR Reporting From the COSGROVE Study

Domain	Outcome
Maternal	Preeclampsia
	Eclampsia
	Death
	Mode of birth
Fetal	Stillbirth/live birth
Newborn	GA at birth
	PTB <37 wk
	PTB <28 wk
	BW
	BW <10th percentile
	BW <3rd percentile
	Mechanical ventilation
	BPD/chronic lung disease
	Necrotizing enterocolitis
	Neonatal seizures
	HIE
	Neonatal death
Childhood	Cognitive impairment
	Cerebral palsy
	Hearing impairment
	Visual impairment

Adapted from Healy et al, *AJOG* 2019.[12]Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.

BPD indicates bronchopulmonary dysplasia; BW, birth weight; GA, gestational age; HIE, hypoxic ischemic encephalopathy; PTB, preterm birth.

TABLE 3. All Possible Prediction Endpoints in FGR Include Flawed Assumptions With Major Limitations

Prediction endpoint	Underlying assumptions	Limitations
SGA (birth weight < 10th percentile)	1. Birth weight percentile is clinically important, independent of perinatal morbidity	1. Includes constitutionally small, healthy newborns 2. Excludes newborns with growth pathology weighing > 10th percentile 3. BW centile is a surrogate outcome with limited clinical utility
Perinatal morbidity	1. All perinatal morbidity is due to FGR, or perinatal morbidity not related to FGR is important in FGR risk stratification 2. True growth pathology is only present if it leads to perinatal morbidity	1. Includes cases of morbidity not caused by FGR 2. Excludes cases of "true" FGR* not severe enough to cause perinatal morbidity
SGA+perinatal morbidity	1. Morbidity occurring in the setting of fetal smallness is due to FGR 2. Morbidity occurring in newborns > 10th percentile is not due to FGR	1. Includes constitutionally small newborns experiencing morbidity not caused by FGR 2. Excludes newborns > 10th percentile affected by FGR-related morbidity

*Here, "true" FGR refers to fetal smallness caused by an underlying abnormality or pathology, as opposed to constitutional smallness, in which a fetus is small but unaffected by pathology.

found possible evidence of publication bias in studies of MCA Doppler, favoring studies with higher negative predictive values.[18] However, there were no differences in positive predictive values, sensitivity, specificity, or "conclusion positivity" between published and unpublished studies. Available data on the use of uterine artery (UtA) Doppler in FGR is more limited, with studies showing an association with adverse outcomes, but no available studies assessing the added value of UtA Doppler in addition to UA Doppler alone.[19]

BIOMARKERS FOR RISK STRATIFICATION IN FGR

One potential means of risk stratification to address the prognostic challenges in FGR care is the use of biomarkers that are measurable in maternal blood, which can be used to identify placental dysfunction. One such biomarker is pregnancy-associated plasma protein A (PAPP-A), a placenta-derived protein known to reflect placental growth and function and in longstanding use for aneuploidy screening. In addition to its association with aneuploidy, first-trimester serum PAPP-A below 0.4 multiples of the median (MoM) has been associated with SGA at birth.[20] In a systematic review and meta-analysis of 32 studies and 175,240 pregnancies, first-trimester PAPP-A less than the fifth percentile conferred an OR of 2.08 for SGA at birth. While the association of PAPP-A with growth abnormalities and other placenta-mediated adverse outcomes is consistent, its predictive ability for relevant adverse outcomes is modest at best (Table 5), and when added to multivariable models with other recognized risk factors, it increases the screen-positive rate without increasing sensitivity.[21,35]

Angiogenesis-related biomarkers have also been a focus of study for FGR. Placental growth factor (PlGF) is often significantly lower in the first trimester in women who ultimately deliver an SGA infant than in those who do not, with low sensitivity (27%, 95% CI: 20%-36%) and good specificity (90%, 95% CI: 83%-94%).[22] When measured later in pregnancy, PlGF measurements have considerably better performance to predict placenta-mediated disease. In a small case-control study of N = 114 participants, PlGF had a sensitivity of 84.2% at a preset specificity of 90% to detect FGR and/or preeclampsia at < 32 weeks. However, prediction of late onset disease was poor, with a sensitivity of 0% at the preset specificity of 90%.[23] In a small secondary analysis of banked blood samples and placental tissue with suspected prenatal FGR, maternal serum PlGF had excellent performance to predict abnormal placental histopathology (AUC: 0.96, 95% CI: 0.93-0.98; Table 5), and low PlGF was associated with considerably shorter interval to delivery (13.0 vs. 29.5 d, $P < 0.001$).[24] A Cochrane review from 2019 reported that PlGF was associated with SGA at birth, but there was too much variability in the thresholds used across studies for meaningful meta-analysis.[36] Before clinical use, larger and more rigorous studies of PlGF are needed to demonstrate clinical benefit.

Soluble fms-like tyrosine kinase 1 (sFlt1) is a placentally secreted soluble vascular endothelial growth factor (VEGF) receptor that improves identification of preeclampsia when used in a ratio with PlGF (sFlt1:PlGF). In a systematic review and meta-analysis of 8 studies with 238 FGR cases (variably defined) and 5111 controls, the authors found that an sflt1/PlGF ratio > 33 had good overall prediction of FGR (Table 5).[25] In a single-center case-control study of N = 125 participants with early onset FGR defined by ISUOG criteria, sFlt1/PlGF > 85 was associated with shorter interval to delivery or stillbirth (13.3 vs. 39.9 d, $P < 0.001$) and the ratio had good overall prediction of delivery for fetal indications within 1 week (AUC: 0.84, 95% CI: 0.73-0.96).[26]

TABLE 4. Added Value of MCA PI and CPR in Addition to UA PI for Prediction of Adverse Perinatal Outcome According to EFW Percentile

	UA PI AUC (95% CI)	MCA PI AUC (95% CI)	CPR AUC (95% CI)	UA PI+CPR AUC (95% CI)
EFW < 10th percentile (n = 589)	0.78 (0.74-0.81)	0.68 (0.64-0.72)	0.77 (0.77-0.82)	0.78 (0.74-0.82)
EFW > 10th percentile (n = 3995)	0.60 (0.53-0.62)	0.57 (0.55-0.60)	0.59 (0.56-0.61)	0.60 (0.57-0.62)

Compiled from Vollgraff Heidweiller-Schreurs et al, *BJOG* 2021.[17]

AUC indicates area under the receiver operating characteristic curve; CI, confidence interval; CPR, cerebroplacental ratio; EFW, estimated fetal weight; MCA, middle cerebral artery; PI, pulsatility index.

TABLE 5. Summary of Biomarker Associations With SGA or FGR

Biomarker	Physiology	Association with FGR or SGA	Additional findings
PAPP-A[21]	Placenta-derived, involved in placental growth and function	PAPP-A < 5 centile for SGA at birth: OR: 2.08, +LR: 1.96, −LR: 0.93	PAPP-A < 5 centile for stillbirth > 24 wk: OR: 2.4, +LR: 1.58, −LR: 0.92
PIGF	Placenta-derived, important for angiogenesis	Low PIGF in first trimester for SGA: 27% sensitivity at 90% specificity;[22] Low PIGF at 24 wk for early-onset FGR/ preeclampsia: 84% sensitivity, 90% specificity[23]	Low PIGF prediction of placental lesions in suspected FGR: AUC 0.96 (0.93-0.98), sensitivity 98.2% (90.5-99.0), specificity 75.1% (67.6-81.7)[24]
SFlt1/PIGF	SFlt1: placentally secreted protein that binds VEGF	SFlt1/PIGF > 33 for FGR: 63% sensitivity, 84% specificity, AUC 0.834[25]	sFlt1/PIGF > 85 predicted shorter latency to delivery for fetal indications: +HR: 9.869, AUC: 0.847 (0.73-0.96);[26] Elevated sFlt1/PIGF (> 5.78 at < 28 wk, > 38 at > 28 wk): RR: 10.75 for medically indicated vs. spontaneous PTB, RR: 2.71 for delivery within 3 wk[27]
Dlk1	Fetal protein that preserves progenitor cells until differentiation, drives lipid metabolism and adipogenesis	Pham et al (N = 563):[28] DLK1 in the bottom quartile associated with SGA (OR: 1.98, 1.15-3.37), AUC: 0.58; not associated with SGA with placental vascular dysfunction ($P = 0.5$). Macdonald et al (N = 999):[29] DLK1 lower in SGA vs. AGA, AUC 0.64, $P < 0.001$. Cleaton et al (N = 127):[30] DLK1 lower in SGA vs. AGA (AUC: 0.65, $P = 0.01$), lower in SGA with abnormal Doppler findings (AUC: 0.71, $P = 0.001$).	Page et al (N = 468):[31] DLK1 levels were no different between stillbirths and live births; DLK1 was lower in live births with SGA < 5th centile of histologic placental insufficiency vs. live births with BW > 5th centile and no placental lesions
SPINT1	Placenta membrane-bound protease involved in placental integrity	Kaitu'u-Lino et al (N = 2003):[32] SPINT1 was lower in SGA vs. AGA newborns: AUC: 0.66 (0.61-0.71), especially among those requiring delivery < 34 wk vs. controls (AUC: 0.95, 0.85-1.0).	Kaitu'u-Lino et al (N = 2003):[32] low SPINT1 confers RR of 14.2 (3.4-58.5) for BW < 5th centile with NICU admission (sensitivity: 62.5%, specificity: 89.9%). Murphy et al:[33] SPINT1 is lower in FGR/ SGA in cases of preeclampsia $P < 0.001$)
NrCAM	Neural cell adhesion molecule, highly expressed in the placenta	Bartho et al (N = 2841):[34] NrCAM was lower in SGA < 3rd centile than controls (AUC range: 0.70-0.76 across 3 cohorts).	Bartho et al (N = 2841):[34] hypoxia reduced NrCAM expression in ex vivo trophoblasts and in a mouse model of FGR

The sFlt1:PlGF may prove clinically useful for risk stratification in FGR, but small sample sizes, variable definitions of abnormality, and a lack of prospective trials assessing the impact on outcomes make it too early to recommend for general clinical implementation.

More recently identified potential biomarkers include DLK1, SPINT1, and NrCAM. DLK1, or delta-like noncanonical notch ligand 1, is a transmembrane protein of fetal origin that is cleaved by extracellular proteases into fetal and maternal circulation, with important functions for progenitor cell preservation and fetal lipid metabolism and adipogenesis.[37] Low DLK1 levels at various time points in pregnancy have been shown to be associated with the delivery of an SGA newborn.[28,29] One small study (N = 127)[30] found that DLK1 was helpful to distinguish "pathologic" FGR from "non-pathologic" SGA, though 2 larger studies (N = 563,[28] N = 999[29]) found no difference between SGA newborns with and without evidence of compromise (Table 5). Finally, a secondary analysis of the landmark Stillbirth Collaborative Research Network study compared maternal DLK1 levels between cases of stillbirth and live birth (N = 468).[31] The authors found no difference in DLK1 levels between stillbirths and live births, though in exploratory analyses, DLK1 was lower in live births with birth weight < 5th centile or histologic evidence of placental insufficiency, compared with newborns without (Table 5). While DLK1 has considerable biological plausibility and early demonstration of an association with SGA at birth, further study is necessary to clarify its role in clinical risk stratification.

SPINT1 (serine peptidase inhibitor, Kunitz type-1) is a membrane-bound cell surface protease inhibitor that is highly expressed on the placental surface, which is essential for placental development.[38] Kaitu'u-Lino and colleagues screened 22 proteins for differences in maternal blood between SGA and AGA newborns. In both discovery (n = 1001) and validation cohorts (n = 1002), they found that maternal plasma SPINT1 at 36 weeks was significantly lower in women who ultimately delivered an SGA infant (validation cohort AUC: 0.66, 95% CI: 0.61-0.71), which performed better than PlGF concentrations ($P = 0.03$).[32] A noteworthy finding of the same study arose in a subgroup of FGR requiring delivery at < 34 weeks, in which SPINT1 concentrations were consistently lower than gestational age-matched controls (AUC: 0.95, 95% CI: 0.85-1.0). An important point that the authors make is that SPINT1 should be measured in plasma rather than serum, as associations with serum values were much less consistent. The association of SPINT1 with SGA was later confirmed by the same group in participants diagnosed with pre-eclampsia (N = 232).[33] As with the other biomarkers associated with small size at birth, the key gap remains whether it can be used for risk stratification for more concrete perinatal morbidities beyond birth weight percentile.

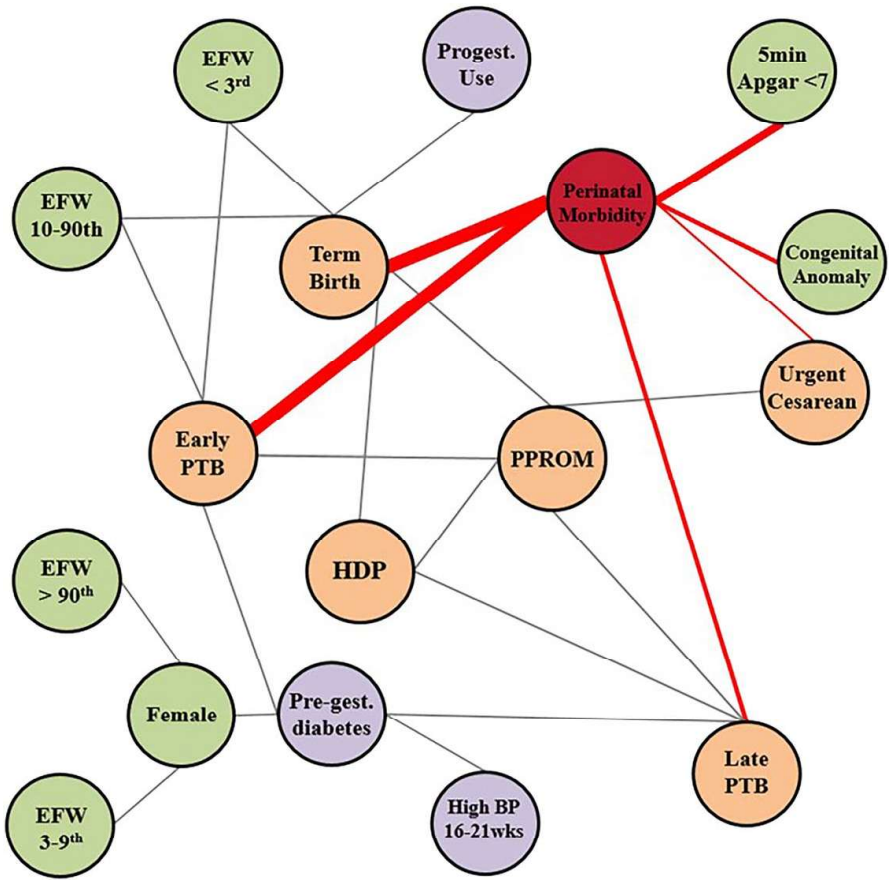

FIGURE 1. Graphical structure of a probabilistic graphical model for perinatal morbidity and mortality prediction in fetal growth restriction. Nodes in the model represent variables, with lines representing conditional dependencies between variables. Among variables with direct relationships to perinatal morbidity, red lines signify association with higher risk, while the green line signifies association with lower risk. The thickness of the colored lines reflects the strength of association with perinatal morbidity. BP indicates blood pressure; EFW, estimated fetal weight; HDP, hypertensive disorder of pregnancy (any of: gestational hypertension, preeclampsia, superimposed preeclampsia, eclampsia); PPROM, preterm premature rupture of membranes; PTB, preterm birth. Figure from Zimmerman et al, *BMC Pregnancy and Childbirth* 2025, printed with permission. full color online

NrCAM, or neuronal cell adhesion molecule, is a cell adhesion molecule known to have important functions for neuronal development and cell signaling. NrCAM is highly expressed in the human placenta and has been explored as a biomarker for abnormal fetal growth.[39] The same laboratory that published extensively on SPINT1, led by Dr Kaitu'u-Lina, also assessed NrCAM levels in multiple independent obstetric cohorts and found that maternal circulating NrCAM was significantly lower in all cohorts, with an AUC range of 0.70 to 0.76 to predict birth weight less than the third percentile.[34] They also demonstrated that hypoxia reduced NrCAM expression in human trophoblast cells and in a mouse model of FGR. NrCAM has not been further evaluated and requires further study.

In summary, multiple biomarkers that are measurable in maternal blood have been associated with fetal or newborn smallness, but none have been comprehensively evaluated for clinical application or risk stratification to guide management once FGR is suspected. This represents an area of considerable investigational opportunity to address the risk stratification gaps and prognostic uncertainty that persist in FGR care.

MULTIVARIABLE MODELING FOR RISK STRATIFICATION IN FGR

While many studies have been published deriving or validating models to predict FGR or SGA itself, available data targeted for risk stratification and prediction of concrete perinatal morbidity is sparse.[40,41] The lack of an outcome unique to FGR means that any risk stratification tool designed for application in the setting of FGR must account for other potential contributors to morbidity to achieve good performance. Our group performed a study in which we identified multiple variables that differentiated SGA newborns who experienced morbidity from SGA newborns who did not and used those variables to develop a multivariable model for perinatal morbidity prediction.[42] We intended the model to only include factors known prenatally to maximize clinical relevance, but we found that excluding gestational age at birth as a variable resulted in a model that was essentially predicting prematurity. As a result, we concluded that it is critical to include other key drivers of morbidity in multivariable models, even if they are not explicitly related to FGR, to optimize both the clinical relevance and predictive performance of risk stratification tools.

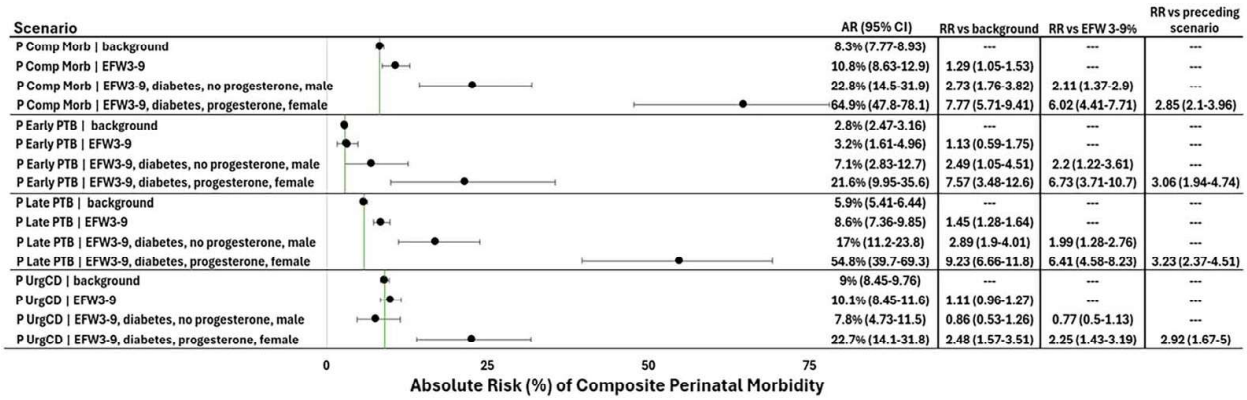

FIGURE 2. Probabilistic graphical models generate scenario-specific probabilities for any variable in the model. The scenarios are framed as probability expressions where "P Comp Morb I EFW 3-9" would be written as "the probability of composite morbidity given the presence of EFW third to ninth percentile." The green lines represent the cohort's background absolute risk of the associated outcome (composite perinatal morbidity, early PTB, late PTB, or urgent cesarean), allowing for visual interpretation based on confidence intervals that overlap with the background risk estimate. Absolute risks are expressed as percentages. The number of derivation cohort participants meeting criteria for each scenario in order of presentation is: 7645, 334, 11, and 0, respectively. "RR versus background" expresses the relative risk of a given factor or scenario over the cohort's background risk of the same outcome (green line). "RR versus EFW 3% to 9%" expresses the relative risk of each scenario over the risk conferred by EFW third to ninth percentile alone. "RR versus preceding scenario" expresses the relative risk of the final scenario over the preceding scenario, in which the only differing factors are progesterone use and fetal sex. All risks (AR, RR) are followed by 95% confidence intervals. Diabetes refers to pregestational diabetes. AR indicates absolute risk; CI, confidence interval; RR, relative risk; EFW, estimated fetal weight; *P*, probability; PTB, preterm birth; UrgCD, urgent cesarean delivery. Figure from Zimmerman et al, *BMC Pregnancy and Childbirth* 2025, printed with permission. full color online

An excellent example of this framework is the body of work recently published by Papastefanou and colleagues, who developed a Fetal Medicine Foundation (FMF) "competing risks model" wherein SGA and gestational age at birth are both considered as closely interrelated variables that are both endpoints for prediction.[35,43–46] While this framework does not target other measures of perinatal morbidity, it uses SGA with preterm delivery at varying time points as relevant targets for morbidity prediction. The model incorporates a variety of maternal, obstetric, and ultrasound factors. In an external validation study that analyzed the model's performance in both the original FMF derivation cohort (n = 96,678) and an external nulliparous cohort (n = 8974), the overall performance in the nulliparous validation cohort was similar to that of the FMF derivation cohort's nulliparous participant subset, though performance in nulliparous participants appears to be lower than in participants of mixed parity.[47] In another external validation study by Chaveeva et al[48] (N = 35,170), application of the FMF model in the first trimester had moderate-good prediction of preterm SGA at BW centiles < 3rd and < 10th at < 32 and < 37 weeks of gestation, with AUCs as high as 0.86. However, prediction of SGA in the absence of preeclampsia, which is arguably of more interest for FGR-specific risk stratification than SGA and preterm birth in the setting of preeclampsia, was generally not as high (maximum AUC: 0.81). The most recent publication from this proliferative group assesses the model's performance using maternal data, UtA PI, and EFW percentile at midgestation and achieves high detection rates for delivery of an SGA newborn at < 28, < 32, and < 36 weeks. For example, the model achieved a detection rate of 80% for BW less than the third percentile of 80% with a screen positive rate (SPR) of just 6.5%.[45] This approach is promising as these data can be routinely ascertained to guide the timing of surveillance based on the likelihood of preterm SGA delivery.

A group led by Powel et al[49] used a cohort of N = 875 consecutive cases of FGR (ISUOG criteria) to derive an index to predict adverse perinatal outcome in FGR. Their index included variables that were weighted according to influence on adverse outcome: UA absent/reversed diastolic flow (+8 points), hypertensive disorder of pregnancy at diagnosis (+8 points), GA < 32 weeks at diagnosis (+5 points), chronic hypertension (+4 points), non-Hispanic Black race (−2 points), and isolated AC less than third centile (−4 points). This model had good performance with an AUC of 0.86 (95% CI: 0.83-0.88), though its performance in the setting of FGR as defined by SMFM criteria is uncertain. Another group[50] derived a model for adverse perinatal outcome prediction from 1453 pregnancies in the OPTICORE study, to be applied at the time of admission for monitoring of early onset FGR (ISUOG criteria). Their model contained 14 variables, including maternal and pregnancy characteristics, UA and MCI findings, concomitant hypertensive disorders, and fetal sex, among others. The model achieved an AUC of 0.83 (0.79-0.87). These approaches offer the potential for promising advances in morbidity risk stratification, but their findings can only currently be generalized to FGR as defined by ISUOG criteria.

MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Given the complexity of and large number of morbidity-associated findings, machine learning (ML) and artificial intelligence (AI)-based methods are attractive to optimize the prediction of adverse outcomes. A 2023 systematic review and meta-analysis identified 20 studies of AI/ML approaches to predict FGR itself rather than for risk stratification once FGR is suspected.[51,52] Only 2 of the included studies had sample sizes > N = 1000, and many used experimental data points and biomarkers, illustrating that this line of inquiry is yet in early stages.

So far, a key limitation of many AI-based approaches is a limited ability to apply intellectual oversight, as most

approaches do not allow for direct interrogation of how the outputs are arrived at.[53] Explainable AI describes AI-driven approaches that are transparent in how they arrive at a given output, allowing for intellectual oversight to mitigate bias or other downsides of "black box" approaches. Our group used probabilistic graphical models (PGMs), a form of explainable AI, to reduce prognostic uncertainty and improve risk stratification in FGR in a prospectively observed nulliparous cohort (N = 9558).[54] We analyzed > 4000 unique variables to identify the 16 most informative features to predict composite perinatal morbidity, defined as perinatal mortality or NICU stay > 7 days or other severe complications such as intraventricular hemorrhage. We derived a PGM in a derivation subgroup (n = 7645) before testing it in a separate validation subgroup (n = 1912). This approach achieved good overall performance for composite morbidity prediction (AUC: 0.83, 95% CI: 0.79-0.87), including in "N of 1" scenarios, or combinations of variables that only occurred a single time in the validation cohort (AUC: 0.81, 0.72-0.90). The PGM graphical structure is shown in Figure 1. Our PGM identified an unexpected interaction between fetal sex, EFW percentile, and maternal diabetes, whereby the expected protective effect of female sex was reversed in the setting of maternal pregestational diabetes, in which female sex conferred an increased risk compared with male sex. This interaction illustrates the potential of AI-based methods to identify previously unrecognized but potentially relevant risk relationships. Two noteworthy findings were that (1) in our feature selection process, EFW percentile at 22 to 29 weeks was more informative than birth weight percentile, and (2) other variables that we suspected would be highly informative, including UtA PI and maternal socioeconomic factors, were ultimately not informative enough for inclusion in the model. One helpful feature of PGMs is that they can be used to estimate the probability of any variable without having to compute a new model. This capability is demonstrated in Figure 2. Beyond this, studies on AI/ML for risk stratification in FGR are limited, and future efforts should continue to develop transparent approaches where intellectual oversight can be maintained.

SUMMARY AND CONCLUSIONS

Risk stratification is a core challenge in the clinical care of FGR, with many sources of opportunity and discovery with the potential to improve outcomes for those affected. A key obstacle to the development of accurate prediction tools is the fact that FGR is a finding that can reflect underlying fetal compromise rather than a single unified pathology with a gold standard postnatal definition. Because the adverse perinatal outcomes associated with FGR can also be caused by unrelated conditions, any risk stratification effort for FGR will have a reduced potential for prediction performance unless such efforts account for the other sources of morbidity. As a result, approaches that integrate variables from multiple domains, including ultrasound, biomarkers in maternal blood, and maternal and obstetric characteristics, are most likely to yield meaningful advances and enable newer computational AI/ML technologies to further improve risk assessment.

REFERENCES

1. Flenady V, Koopmans L, Middleton P, et al. Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. *Lancet*. 2011;377:1331–1340.
2. Crispi F, Miranda J, Gratacós E. Long-term cardiovascular consequences of fetal growth restriction: biology, clinical implications, and opportunities for prevention of adult disease. *Am J Obstet Gynecol*. 2018;218:S869–S879.
3. Barker DJP, Godfrey KM, Gluckman PD, et al. Fetal nutrition and cardiovascular disease in adult life. *Lancet*. 1993;341: 938–941.
4. Lees CC, Stampalija T, Baschat AA, et al. ISUOG Practice Guidelines: diagnosis and management of small-for-gestational-age fetus and fetal growth restriction. *Ultrasound Obstet Gynecol*. 2020;56:298–312.
5. Society for Maternal-Fetal Medicine. Electronic address pso, Martins JG, Biggio JR, Abuhamad A. Society for Maternal-Fetal Medicine Consult Series #52: diagnosis and management of fetal growth restriction: (Replaces Clinical Guideline Number 3, April 2012). *Am J Obstet Gynecol*. 2020;223: B2–B17.
6. Blue NR, Beddow ME, Savabi M, et al. Comparing the Hadlock fetal growth standard to the Eunice Kennedy Shriver National Institute of Child Health and Human Development racial/ethnic standard for the prediction of neonatal morbidity and small for gestational age. *Am J Obstet Gynecol*. 2018;219: 474.e1–474.e12.
7. Blue NR, Grobman WA, Larkin JC, et al. Customized versus population growth standards for morbidity and mortality risk stratification using ultrasonographic fetal growth assessment at 22 to 29 weeks' gestation. *Am J Perinatol*. 2021;38(S 01): e46–e56.
8. Blue NR, Mele L, Grobman WA, et al. Predictive performance of newborn small for gestational age by a United States intrauterine vs birthweight-derived standard for short-term neonatal morbidity and mortality. *Am J Obstet Gynecol MFM*. 2022;4:100599.
9. Roeckner JT, Pressman K, Odibo L, et al. Outcome-based comparison of SMFM and ISUOG definitions of fetal growth restriction. *Ultrasound Obstet Gynecol*. 2021;57:925–930.
10. Schreiber V, Hurst C, Costa FDS, et al. Definitions matter: detection rates and perinatal outcomes for infants prenatally identified with late fetal growth restriction using biometric vs consensus Delphi criteria. *Ultrasound Obstet Gynecol*. 2023;61: 377–385.
11. Rodriguez-Sibaja MJ, Lopez-Diaz AJ, Valdespino-Vazquez MY, et al. Placental pathology lesions: International Society for Ultrasound in Obstetrics and Gynecology vs Society for Maternal-Fetal Medicine fetal growth restriction definitions. *Am J Obstet Gynecol MFM*. 2024;6:101422.
12. Healy P, Gordijn SJ, Ganzevoort W, et al. A core outcome set for the prevention and treatment of fetal GROwth restriction: deVeloping Endpoints: the COSGROVE study. *Am J Obstet Gynecol*. 2019;221:339.e1–339.e10.
13. Blue NR, Savabi M, Beddow ME, et al. The Hadlock method is superior to newer methods for the prediction of the birth weight percentile. *J Ultrasound Med*. 2019;38:587–596.
14. Alfirevic Z, Stampalija T, Dowswell T. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. *Cochrane Database Syst Rev*. 2017;6:CD007529.
15. Figueras F, Caradeux J, Crispi F, et al. Diagnosis and surveillance of late-onset fetal growth restriction. *Am J Obstet Gynecol*. 2018;218(2s):S790–S802.e1.
16. Vollgraff Heidweiller-Schreurs CA, De Boer MA, Heymans MW, et al. Prognostic accuracy of cerebroplacental ratio and middle cerebral artery Doppler for adverse perinatal outcome: systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2018;51:313–322.
17. Vollgraff Heidweiller-Schreurs C, Van Osch I, Heymans M, et al. Cerebroplacental ratio in predicting adverse perinatal outcome: a meta-analysis of individual participant data. *BJOG*. 2021;128:226–235.
18. Vollgraff Heidweiller-Schreurs CA, Korevaar DA, Mol BWJ, et al. Publication bias may exist among prognostic accuracy studies of middle cerebral artery Doppler ultrasound. *J Clin Epidemiol*. 2019;116:1–8.

19. Martinez-Portilla RJ, Caradeux J, Meler E, et al. Third-trimester uterine artery Doppler for prediction of adverse outcome in late small-for-gestational-age fetuses: systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2020;55:575–585.
20. Tong S, Joy Kaitu'U-Lino TU, Walker SP, et al. Blood-based biomarkers in the maternal circulation associated with fetal growth restriction. *Prenatal Diagn*. 2019;39:947–957.
21. Morris RK, Bilagi A, Devani P, et al. Association of serum <scp>PAPP-A</scp> levels in first trimester with small for gestational age and adverse pregnancy outcomes: systematic review and meta-analysis. *Prenatal Diagn*. 2017;37:253–265.
22. Zhong Y, Zhu F, Ding Y. Serum screening in first trimester to predict pre-eclampsia, small for gestational age and preterm delivery: systematic review and meta-analysis. *BMC Pregn Childbirth*. 2015;15:191.
23. Crispi F, Llurba E, Domínguez C, et al. Predictive value of angiogenic factors and uterine artery Doppler for early- versus late-onset pre-eclampsia and intrauterine growth restriction. *Ultrasound Obstet Gynecol*. 2008;31:303–309.
24. Benton SJ, McCowan LM, Heazell AE, et al. Placental growth factor as a marker of fetal growth restriction caused by placental dysfunction. *Placenta*. 2016;42:1–8.
25. Chen W, Wei Q, Liang Q, et al. Diagnostic capacity of sFlt-1/PlGF ratio in fetal growth restriction: A systematic review and meta-analysis. *Placenta*. 2022;127:37–42.
26. Palma dos Reis C, Brás S, Meneses T, et al. The sFlt1/PlGF ratio predicts faster fetal deterioration in early fetal growth restriction: a historical cohort study. *Acta Obstet Gynecol Scand*. 2023;102:635–643.
27. Hong J, Crawford K, Cavanagh E, et al. Prediction of preterm birth in growth-restricted and appropriate-for-gestational-age infants using maternal <scp>PlGF</scp> and the <scp>sFlt</scp>-1/<scp>PlGF</scp> ratio—A prospective study. *BJOG*. 2024;131:1089–1101.
28. Pham A, Mitanchez D, Forhan A, et al. Low maternal DLK1 levels at 26 weeks is associated with small for gestational age at birth. *Front Endocrinol*. 2022;13:836731.
29. MacDonald TM, Walker SP, Hiscock R, et al. Circulating Delta-like homolog 1 (DLK1) at 36 weeks is correlated with birthweight and is of placental origin. *Placenta*. 2020;91:24–30.
30. Cleaton MAM, Dent CL, Howard M, et al. Fetus-derived DLK1 is required for maternal metabolic adaptations to pregnancy and is associated with fetal growth restriction. *Nat Genet*. 2016;48:1473–1480.
31. Page JM, Allshouse AA, Gaffney JE, et al. DLK1: a novel biomarker of placental insufficiency in stillbirth and live birth. *Am J Perinatol*. 2024;41(S 01):e221–e229.
32. Kaitu'U-Lino TUJ, Macdonald TM, Cannon P, et al. Circulating SPINT1 is a biomarker of pregnancies with poor placental function and fetal growth restriction. *Nat Commun*. 2020;11:2411.
33. Murphy CN, Cluver CA, Walker SP, et al. Circulating SPINT1 is reduced in a preeclamptic cohort with co-existing fetal growth restriction. *J Clin Med*. 2022;11:901.
34. Bartho LA, Walker SP, Idzes D, et al. Reduced circulating NrCAM as a biomarker for fetal growth restriction. *EBioMedicine*. 2025;118:105854.
35. Papastefanou I, Wright D, Syngelaki A, et al. Competing-risks model for prediction of small-for-gestational-age neonate from maternal characteristics and serum pregnancy-associated plasma <scp>protein-A</scp> at 11–13 weeks' gestation. *Ultrasound Obstet Gynecol*. 2020;56:541–548.
36. Heazell AE, Hayes DJ, Whitworth M, et al. Biochemical tests of placental function versus ultrasound assessment of fetal size for stillbirth and small-for-gestational-age infants. *Cochrane Database Syst Rev*. 2019;14:CD012245.
37. Traustadóttir GÁ, Lagoni LV, Ankerstjerne LBS, et al. The imprinted gene Delta like non-canonical Notch ligand 1 (Dlk1) is conserved in mammals, and serves a growth modulatory role during tissue development and regeneration through Notch dependent and independent mechanisms. *Cytok Growth Factor Rev*. 2019;46:17–27.
38. Szabo R, Molinolo A, List K, et al. Matriptase inhibition by hepatocyte growth factor activator inhibitor-1 is essential for placental development. *Oncogene*. 2007;26:1546–1556.
39. Özkan S, Vural B, Dalçık C, et al. Placental expression of insulin-like growth factor-I, fibroblast growth factor-basic and neural cell adhesion molecule in pregnancies with small for gestational age fetuses. *J Perinatol*. 2008;28:468–474.
40. Tarca AL, Hernandez-Andrade E, Ahn H, et al. Single and serial fetal biometry to detect preterm and term small- and large-for-gestational-age neonates: a longitudinal cohort study. *PLOS One*. 2016;11:e0164161.
41. Crovetto F, Triunfo S, Crispi F, et al. Differential performance of first-trimester screening in predicting small-for-gestational-age neonate or fetal growth restriction. *Ultrasound Obstet Gynecol*. 2017;49:349–356.
42. Blue NR, Allshouse AA, Grobman WA, et al. Developing a predictive model for perinatal morbidity among small for gestational age infants. *J Matern Fetal Neonatal Med*. 2022;35:8462–8471.
43. Papastefanou I, Nowacka U, Buerger O, et al. Evaluation of the RCOG guideline for the prediction of neonates that are small for gestational age and comparison with the competing risks model. *BJOG*. 2021;128:2110–2115.
44. Papastefanou I, Nowacka U, Syngelaki A, et al. Competing risks model for prediction of small-for-gestational-age neonates from biophysical markers at 19 to 24 weeks' gestation. *Am J Obstet Gynecol*. 2021;225:530.e1–530.e19.
45. Papastefanou I, Syngelaki A, Logdanidis V, et al. Assessment of risk for small for gestational age at midgestation to define subsequent care. *Am J Obstet Gynecol* 2025. doi:10.1016/j.ajog.2025.06.016
46. Papastefanou I, Wright D, Syngelaki A, et al. Personalized stratification of pregnancy care for small for gestational age neonates from biophysical markers at midgestation. *Am J Obstet Gynecol*. 2023;229:57.e1–57.e14.
47. Bligard KH, Papastefanou I, Odibo AO. Fetal Medicine Foundation competing-risks model for small-for-gestational-age neonate in midgestation: external validation in nulliparous patients. *Ultrasound Obstet Gynecol*. 2025;66:324–330.
48. Chaveeva P, Papastefanou I, Dagklis T, et al. External validation and comparison of <scp>Fetal Medicine Foundation</scp> competing-risks model for small-for-gestational-age neonate in the first trimester: multicenter cohort study. *Ultrasound Obstet Gynecol*. 2025;65:729–737.
49. Powel JE, Zantow EW, Bialko MF, et al. Predictive index for adverse perinatal outcome in pregnancies complicated by fetal growth restriction. *Ultrasound Obstet Gynecol*. 2023;61:367–376.
50. Van De Meent M, Bel EW, Ganzevoort W, et al. Perinatal risk assessment in pregnancies complicated by early-onset fetal growth restriction: development and internal validation of a prediction model for composite adverse perinatal outcome. *Ultrasound Obstet Gynecol*. 2025;66:175–185.
51. Rescinito R, Ratti M, Payedimarri AB, et al. Prediction models for intrauterine growth restriction using artificial intelligence and machine learning: a systematic review and meta-analysis. *Healthcare*. 2023;11:1617.
52. Vasilache I-A, Scripcariu I-S, Doroftei B, et al. Prediction of intrauterine growth restriction and preeclampsia using machine learning-based algorithms: a prospective study. *Diagnostics*. 2024;14:453.
53. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. *J Glob Health*. 2019;9:010318.
54. Zimmerman RM, Hernandez EJ, Yandell M, et al. AI-based analysis of fetal growth restriction in a prospective obstetric cohort quantifies compound risks for perinatal morbidity and mortality and identifies previously unrecognized high risk clinical scenarios. *BMC Pregn Childbirth*. 2025;25. doi:10.1186/s12884-024-07095-6