BMC Pregnancy and Childbirth

RESEARCH

Open Access

A machine learning model based on placental magnetic resonance imaging and clinical factors to predict fetal growth restriction

Jida Wang[1†], Zhuying Chen[1†], Hongxi Zhang[2], Weikang Li[2], Kui Li[1], Meixiang Deng[1] and Yu Zou[1*]

Abstract

Objectives To create a placental radiomics-clinical machine learning model to predict FGR.

Materials and methods Retrospectively analyzed placental MRI and clinical data of 110 FGR cases and 158 healthy controls at 28–37 weeks of gestation from two campuses of ZWH. 227 cases from Hubin campus were randomly divided into training ($n=182$) and internal testing set ($n=45$). 41 cases from Xiaoshan campus were included in external testing set. Placental MRI features were extracted from sagittal T2WI. Mann–Whitney U test, redundancy analysis, and LASSO were used to identify the radiomics signature, and the best-performing radiomics model was constructed by comparing eight machine learning algorithms. Clinical factors determined by univariate and multivariate analyses. A united model and nomogram combining the radiomics Rad-score and clinical factors were established. The performance of the models was assessed by DeLong test, calibration curve and decision curve analysis.

Results Of 1561 radiomics features, 10 strongly correlated with FGR were selected. The radiomics model using logistic regression performed best compared eight algorithms. 5 important clinical features identified by analysis. The united model demonstrated a good predictive performance in the training, internal testing and external testing sets, with AUC 0.941 (95% CI, 0.0.904–0.977), 0.899 (95% CI, 0.789–1) and 0.861 (95% CI 0.725–0.998), prediction accuracies 0.885, 0.844 and 0.805, precisions 0.871, 0.789 and 0.867, recalls 0.836, 0.833 and 0.684, and F1 scores 0.853, 0.811 and 0.765, respectively. The calibration and decision curves of the united model also showed good performance. Nomogram confirmed clinical applicability of the model.

Conclusions The proposed placental radiomics-clinical machine learning model is simple yet effective to predict FGR.

Keywords Fetal growth restriction, Radiomics, Placenta, Magnetic resonance imaging, Machine learning

†Jida Wang and Zhuying Chen contributed equally to this work and should be considered co-first authors.

*Correspondence:
Yu Zou
zouyuzju@zju.edu.cn
Full list of author information is available at the end of the article

Introduction

Fetal growth restriction (FGR) is defined as the failure of the fetus to meet growth potential due to a pathological factor, most commonly placental dysfunction [1]. FGR increases the risks of perinatal morbidity and mortality [2, 3], as well as adverse long-term health outcomes, especially impaired neurological and cognitive development [4], in addition to cardiovascular and endocrine diseases in adulthood [5]. Early and accurate diagnosis of FGR is essential for obstetric care. However, it is difficult to determine fetal growth potential in clinical practice, which is usually defined by the statistical deviation of fetal size from population-based references, such as the 10th, 5th, or 3rd percentile, to better identify "small for gestational age" (SGA) fetuses [6]. However, as compared to FGR, SGA also includes most small but otherwise healthy fetuses, which have a lower risk of abnormal perinatal outcomes [7]. To overcome these limitations and avoid false-positive and false-negative diagnoses of FGR, a consensus definition of placental-mediated FGR was proposed based on estimated fetal weight (EFW), abdominal circumference (AC), and abnormal Doppler findings of the umbilical, uterine, and middle cerebral arteries [1, 8]. In practice, the Hadlock formula, which is based on prenatal fetal ultrasound parameters, is applied to evaluate EFW. However, ultrasound-based EFW and birth weight (BW) may differ by more than 15% [9]. Therefore, there is an urgent need to seek more accurate techniques or indicators to predict FGR.

FGR is associated with pathophysiological alterations [10, 11], such as reduced perfusion of the maternal placenta as a result of transformation of the spiral artery into a low-resistance vessel during placental development [12], which is insufficient to support fetal growth. Functional MRI, including T2* and intravoxel incoherent movement, can be used to evaluate FGR-related placental dysfunction [13–15]. Several studies have shown significant differences in shape and textural characteristics between healthy and growth-restricted placentas [16, 17]. By mining a large number of quantitative features from medical imaging data, radiomics has the potential to reflect the heterogeneity of disease and can be used to improve the predictive accuracy of diagnosis and prognosis [18, 19]. Recent studies have focused on radiomics and machine learning techniques to evaluate the placenta. For instance, Dahdouh et al. [16] proposed an FGR prediction model based on 23 placental shape and textural features. However, these features were manually selected, which may introduce redundancy and lead to model overfitting. Another study by Song et al. [20] introduced a radiomics-based FGR prediction model that combined placental radiomic features with fetal ultrasound measurements. This model achieved high diagnostic accuracy for FGR by single-center validation. However, these studies focused only on changes in placental morphology, structure and texture, and did not take into account the effect of maternal physiological factors on fetal development and their potential evaluation value for FGR [21, 22].

Therefore, the aim of the present study was to develop a radiomics-clinical model that integrated the most representative placental radiomics features with the Rad-score and the most relevant clinical factors to predict FGR. The accuracy of the model was verified using internal testing and external testing sets.

Materials and methods

The study protocol was approved by the Institutional Review Board of Women's Hospital School of Medicine Zhejiang University (also known as Zhejiang Women's Hospital, ZWH) and conducted in accordance with the tenets of the Declaration of Helsinki. Due to the retrospective design of this study, the requirement for written informed consent was waived.

Study population

The study cohort included pregnant women at 28–37 weeks of gestation who underwent placental MRI from January 2017 to March 2024 in Hubin campus ($n=227$), and from August 2023 to May 2024 in Xiaoshan campus ($n=41$). MR images were retrieved from a picture archiving and communication system, and ultrasound and clinical data from an electronic medical record system. The inclusion criteria for FGR were: 1) singleton pregnancy; 2) BW below the 3rd percentile of gestational age; 3) AC from prenatal ultrasound below the 3rd percentile of gestational age; 4) AC from prenatal ultrasound or BW below the 10th percentile of gestational age with an umbilical artery pulse index higher than the 95th percentile; and 5) no congenital malformations. The inclusion criteria for the control group were: 1) singleton pregnancy; 2) delivery of healthy newborns; 3) no serious maternal complications during pregnancy; and 4) availability of placental MRI data. The exclusion criteria were: 1) missing delivery records; 2) insufficient clinical and prenatal examination data due to incomplete medical records; 3) treatment to improve placental microcirculation during MRI and delivery; 4) poor MR image quality due to artifacts or motion; 5) twin or multiple pregnancies; 6) placental lesions, such as placental abruption, placenta accreta spectrum disorders and evident placental hemorrhage; and 7) fetal malformations. The detailed data flow is shown in Fig. 1. Finally, 268 participants were selected. The 227 pregnant women from Hubin campus were randomly divided into the training set ($n=182$) and internal testing set ($n=45$). The 41 pregnant women

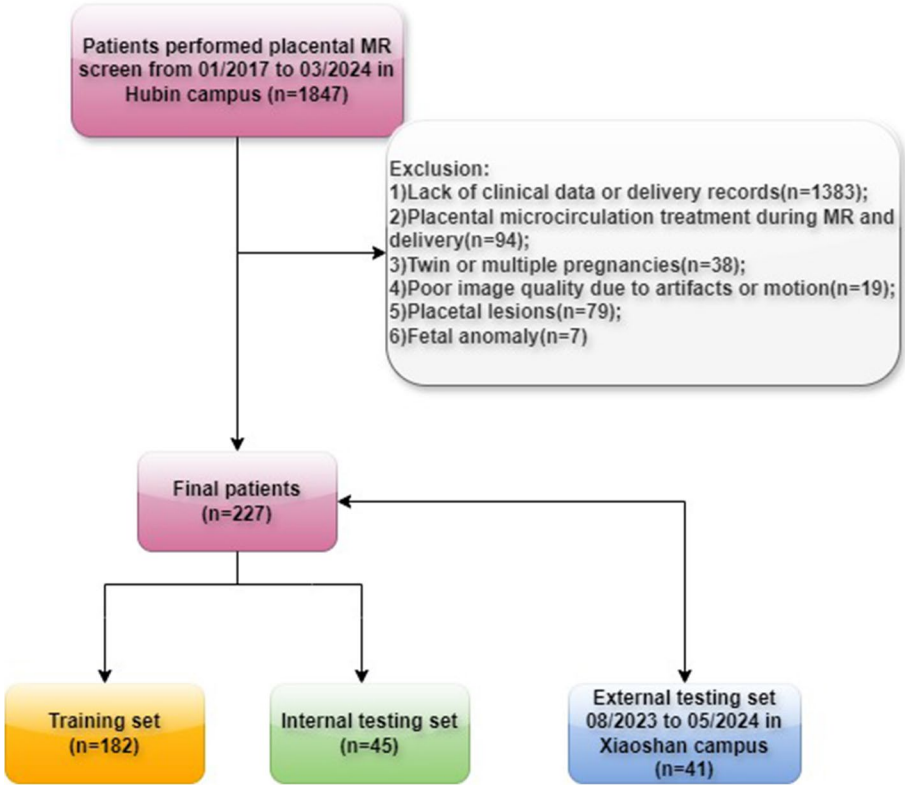

Fig. 1 Flowchart of study population selection

from Xiaoshan campus were included in external testing set. The clinical characteristics of the study population are presented in Table 1.

MRI protocol

For suspected FGR based on ultrasonography or further confirmation of other structural abnormalities, MRI was performed. Placental MRI data were obtained in the oblique sagittal view with a field of view (FOV) covering the entire organ as much as possible. Participants were placed in the supine or left lateral decubitus position depending on comfort level. T2-weighted single-shot fast self-selective echo sequences were used for radiomics analysis. The patients from Hubin campus used a 1.5-T SIGNA HDxt scanner (GE Healthcare, Milwaukee, WI, USA) with an 8-channel cardiac coil, and the specific parameters were as follows: repetition time, 2400 ms; echo time, 128 ms; flip angle, 90; matrix size, 320×224; FOV, 40 cm^2; slice thickness, 6–7 mm; slice gap, 0.5–1 mm; number of slices, 10–28; and acquisition time, 56 s. The patients from Xiaoshan campus used a 3.0-T Premier scanner (GE Healthcare, Milwaukee, WI, USA) with 30-channel body anterior Air coil and 60-channel spine posterior Air coil, and the specific parameters were as follows: repetition time, 1500 ms; echo time, 100 ms; matrix size, 384×320; FOV, 36–38 cm^2; slice thickness, 6–8 mm; slice gap, 1–1.5 mm; number of slices, 15–32; and acquisition time, 50 s.

Image segmentation and extraction of radiomics features

All placental tissue images were manually segmented by an experienced radiologist with more than 10 years of experience in fetal and placental MRI using the open source ITK-SNAP software (version 4.0.1; http://www.itksnap.org/pmwiki/pmwiki.php) [23]. The feature extraction process was performed using the open-source Python package, pyradiomics (version 3.1.0, https://github.com/AIM-Harvard/pyradiomics) [24]. To minimize the impact of heterogeneous datasets due to different scanners and MRI protocols, images were preprocessed prior to feature extraction, including normalization, discretization, and resampling to a voxel size of 3×3×3 mm, which is recommended to improve the reliability and robustness of radiomics analysis. In this study, 1561 features of each patient were extracted, which included shape and the first-order and texture of the original images and the derived images generated using the wavelet and Laplacian of Gaussian filters.

Table 1 Baseline characteristics of the study population in the training, internal testing, and external testing sets

cohort (N)		training (N=182)			Internal testing (N=45)			external testing (N=41)			p
name	levels	non-FGR (N=109)	FGR (N=73)	p	non-FGR (N=27)	FGR (N=18)	p	non-FGR (N=22)	FGR (N=19)	p	
maternal.age(year)	Mean ± SD	31.2 ± 3.7	31.1 ± 4.3	.920	31.0 ± 4.0	31.0 ± 5.1	.978	29.7 ± 3.7	32.1 ± 4.0	.056	.899
height(mm)	Mean ± SD	161.1 ± 5.0	158.3 ± 4.5	<.001*	160.9 ± 5.0	159.0 ± 4.0	.202	159.6 ± 4.8	158.7 ± 6.3	.625	.613
weight(kg)	Mean ± SD	67.9 ± 8.7	63.7 ± 9.3	.002*	68.7 ± 6.1	63.9 ± 9.3	.040*	65.5 ± 11.3	66.1 ± 12.7	.875	.884
WTgrow	Mean ± SD	13.1 ± 3.9	11.6 ± 4.2	.013*	14.6 ± 6.8	11.2 ± 3.4	.029*	12.3 ± 5.1	10.6 ± 3.4	.215	.197
fetal.sex	female	62 (56.9%)	40 (54.8%)	.900	13 (48.1%)	12 (66.7%)	.358	9 (40.9%)	12 (63.2%)	.268	.853
	male	47 (43.1%)	33 (45.2%)		14 (51.9%)	6 (33.3%)		13 (59.1%)	7 (36.8%)		
ODS	No	102 (93.6%)	51 (69.9%)	<.001*	26 (96.3%)	14 (77.8%)	.146	21 (95.5%)	9 (47.4%)	.002*	.129
	Yes	7 (6.4%)	22 (30.1%)		1 (3.7%)	4 (22.2%)		1 (4.5%)	10 (52.6%)		
PREE	No	108 (99.1%)	58 (79.5%)	<.001*	27 (100%)	15 (83.3%)	.113	22 (100%)	16 (84.2%)	.182	.875
	Yes	1 (0.9%)	15 (20.5%)		0 (0%)	3 (16.7%)		0 (0%)	3 (15.8%)		
GH	No	105 (96.3%)	53 (72.6%)	<.001*	26 (96.3%)	13 (72.2%)	.060	19 (86.4%)	15 (78.9%)	.831	.805
	Yes	4 (3.7%)	20 (27.4%)		1 (3.7%)	5 (27.8%)		3 (13.6%)	4 (21.1%)		
AIP	No	96 (88.1%)	66 (90.4%)	.801	21 (77.8%)	16 (88.9%)	.577	20 (90.9%)	15 (78.9%)	.524	.435
	Yes	13 (11.9%)	7 (9.6%)		6 (22.2%)	2 (11.1%)		2 (9.1%)	4 (21.1%)		
GD	No	90 (82.6%)	62 (84.9%)	.828	22 (81.5%)	13 (72.2%)	.714	17 (77.3%)	16 (84.2%)	.870	.640
	Yes	19 (17.4%)	11 (15.1%)		5 (18.5%)	5 (27.8%)		5 (22.7%)	3 (15.8%)		
ABPLH	No	87 (79.8%)	58 (79.5%)	1.000	18 (66.7%)	14 (77.8%)	.638	20 (90.9%)	11 (57.9%)	.037*	.442
	Yes	22 (20.2%)	15 (20.5%)		9 (33.3%)	4 (22.2%)		2 (9.1%)	8 (42.1%)		
ABPU	No	99 (90.8%)	50 (68.5%)	<.001*	25 (92.6%)	10 (55.6%)	.010*	19 (86.4%)	15 (78.9%)	.831	.788
	Yes	10 (9.2%)	23 (31.5%)		2 (7.4%)	8 (44.4%)		3 (13.6%)	4 (21.1%)		
gravidity	1	47 (49.5%)	41 (60.3%)	.307	11 (50%)	9 (56.2%)	.560	13 (59.1%)	8 (42.1%)	.439	.705
	2	29 (30.5%)	14 (20.6%)		7 (31.8%)	6 (37.5%)		7 (31.8%)	7 (36.8%)		
	>=3	19 (20%)	13 (19.1%)		4 (18.2%)	1 (6.2%)		2 (9.1%)	4 (21.1%)		
parity	1	59 (54.1%)	53 (72.6%)	.026*	17 (63%)	14 (77.8%)	1.000	13 (59.1%)	18 (94.7%)	.005*	.297
	2	43 (39.4%)	19 (26%)		10 (37%)	4 (22.2%)		9 (40.9%)	0 (0%)		
	3	7 (6.4%)	1 (1.4%)		0 (0%)	0 (0%)		0 (0%)	1 (5.3%)		

ABPLH abnormal pregnancy-labor history, *ABPU* abnormal placenta or umbilical cord, *AIP* anemia in pregnancy, *GD* gestational diabetes, *GH* gestational hypertension, *ODS* oligohydramnios, *PREE* preeclampsia, *SD* standard deviation, *WTgrow* maternal weight grows in pregnancy

Selection of radiomics features

First, a radiologist with more than 10 years of experience in fetal and placental MRI completed all image segmentation and feature extraction, randomly selected 30 cases, defined the first segmentation as region of interest (ROI)1, and the second segmentation at one month later is defined as ROI2. The third segmentation is completed by another radiologist with the same qualification, defined as ROI3. Each segmentation ROI covers the entire placenta of each layer image. The radiomics features of all ROIs were extracted to calculate the intra-group consistency between ROI1 and ROI2, and the inter-group consistency between ROI1 and ROI3 [25, 26]. Features with the intraclass correlation coefficient (ICC) value > 0.75 indicated satisfactory repeatability and were selected for further analysis.

Second, the radiomic features were preprocessed, including centering and scaling. Radiomics features may have different dimensions, thus, the differences between feature values may vary widely [27]. Two hundred twenty-seven MR images from Hubin campus were assigned to training set or internal testing set at a ratio of 8:2 by stratified random sampling. The radiomic features in the FGR group and the non-FGR group from training set were compared by the Mann–Whitney U test. After that, redundancy analysis of the training set was conducted. First step, the normality of the features was analyzed. When the feature satisfied the normal distribution, Pearson correlation was be used for analysis. Otherwise, used Spearman's correlation analysis. When the correlation coefficient > 0.8, the system would randomly delete one feature and retain the other. Subsequently, least absolute shrinkage and selection operator (LASSO) regression was performed, using a recursive feature elimination method with tenfold cross-validation to filter out features with non-zero coefficients as radiomics signature when $\lambda = 1$ times the standard error (λ1SE).The Rad-score was calculated for each patient using a linear combination of the radiomics signature weighted by the corresponding coefficients.

Construction and evaluation of prediction models

Based on the workflow of the tidymodels framework (version 1.2.0; https://tidymodels.tidymodels.org/), eight machine learning algorithms (naive_Bayes, svm_rbf, rand_forest, logistic_reg, nearest_neighbor, boost_tree, mlp,

and decision_tree) were used to construct the prediction models with the selected radiomics signature. The algorithm with the best AUC, brier score, and accuracy value was selected to construct the final radiomics model. The DALEXtra package (version 2.3.0; https://modeloriented.github.io/DALEXtra/) was used to explain the rank of the importance of the variables in the final model. Univariate and multivariate analyses with stepwise backward regression were performed using the autoReg package (version 0.3.4; https://cran.r-project.org/web/packages/autoReg/) to select the clinical factors, the factors with probability (p) values < 0.1 were selected for the clinical model. Finally, the Rad-score and the screened clinical factors were integrated into a united model. Receiver operating characteristic curves (ROC) were constructed for each model in the training set, internal testing set and external testing set, and the accuracy, precision, recall and F1 score of the models were calculated. Subsequently, calibration and decision curves were constructed to further evaluate the performance of the united model. Also, a nomogram of the united model was established based on the training set. An overall flowchart of the study is presented in Fig. 2.

Statistical analysis

Python software (version 3.9.10; http://www.python.org) was used for image normalization and radiomics feature extraction, partitioning of training and internal testing sets, and subsequent features Mann–Whitney U test. The follow-up statistical analysis and image visualization were performed using R software (version 4.3.3; http://www.R-project.org). The following packages in R were used: caret, glmnet, tidymodels, bonsai, parsinip, discrim, baguette, corrplot, DALEXtra, pROC, iBreakDown, vivo, ggsignif, ggpubr, RColorbrewer, autoReg, dplyr, rms, rmda. Continuous variables were expressed as mean ± standard deviation (SD) and categorical variables were summarised using counts and proportions. The statistical analysis included the t-Test, Chi-square,

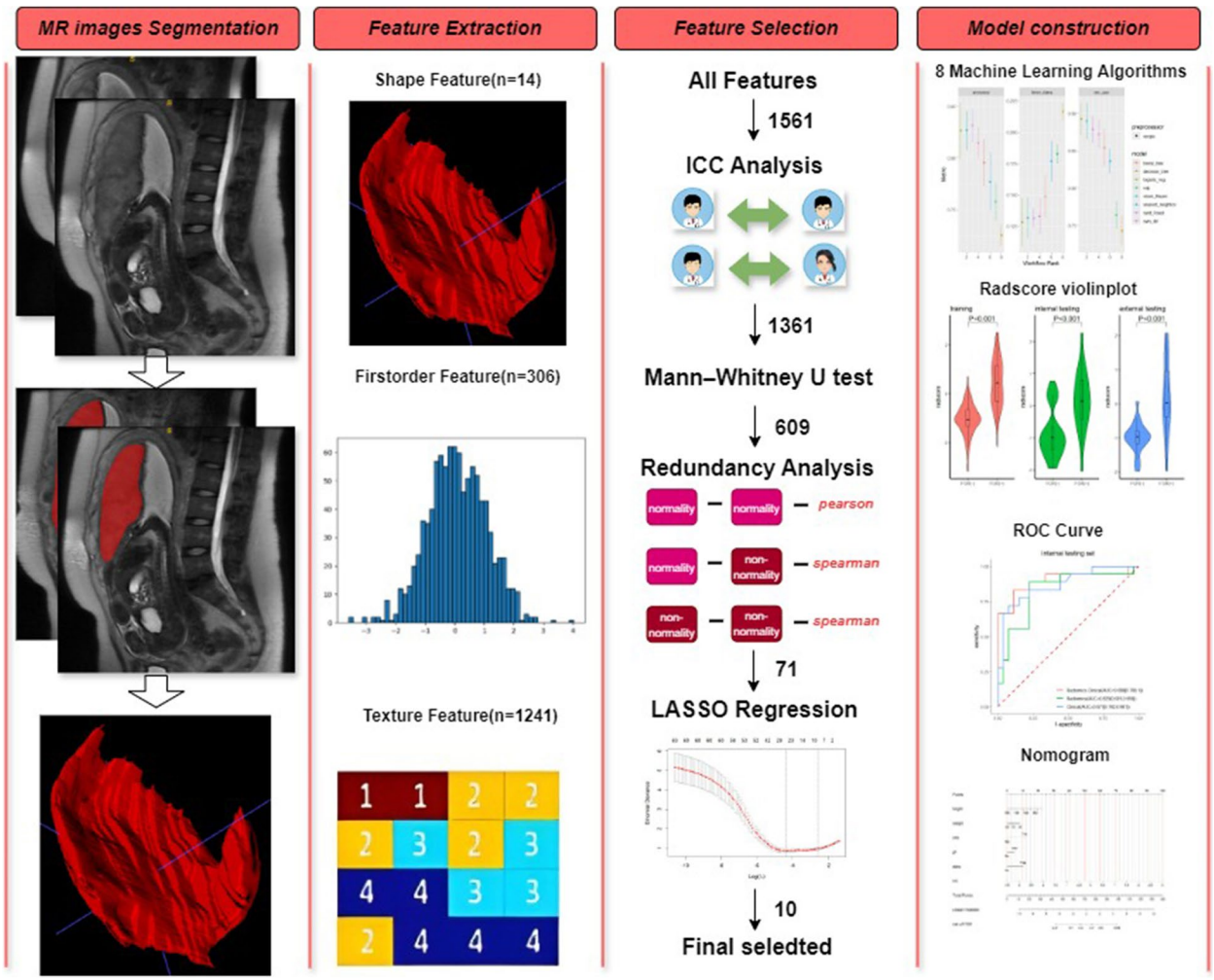

Fig. 2 The overall workflow of radiomics processing and model construction

and receiver operating characteristic (ROC) curves. The ROC curve was analyzed to evaluate the area under the curve (AUC). A two-sided p value < 0.05 was considered statistically significant. Pearson and Spearman's correlation analysis was used in redundancy analysis. DeLong test was used to evaluate the performance of models.

Results

Population and cohort characteristics

Table 1 shows that there are no significant differences in all clinical characteristics between the training set, internal test set and external test set ($p = 0.129–0.899$). In total, five important clinical factors related to FGR were selected by univariable and multivariable analysis (Table 2) to construct a clinical logical regression model.

Selection of radiomic features

ICC analysis of 1561 radiomics features was performed. The reliability of radiomics features with an ICC value < 0.75 was considered moderate to poor [26]. Based on this criterion, 179 and 134 features identified by inter- and intra-group comparisons, respectively. For a total of 200 features after duplication was combined, were excluded. while the remaining 1361 features were centralized and scaled. After the Mann–Whitney U test, 609 radiomic features remained in the training set. Following redundancy analysis, the features with a correlation coefficient > 0.8 were eliminated, resulting in 71 non-redundant features. Then, LASSO regression using the tenfold cross validation method and λ1SE as the criterion identified 10 features as the radiomics signature (Fig. 3). Finally, the Rad-score was calculated for the

Table 2 Clinical factors selected by univariable and multivariable analysis

Dependent: label		Non-FGR (N=109)	FGR (N=73)	OR (univariable)	OR (multivariable)	OR (final)
age	Mean ± SD	31.2 ± 3.7	31.1 ± 4.3	1.00 (0.92-1.07, p=.919)		
height	Mean ± SD	161.1 ± 5.0	158.3 ± 4.5	0.89 (0.83-0.95, p<.001)	0.92 (0.84-1.00, p=.053)	0.93 (0.85-1.01, p=.083)
weight	Mean ± SD	67.9 ± 8.7	63.7 ± 9.3	0.95 (0.91-0.98, p=.003)	0.96 (0.91-1.01, p=.115)	0.95 (0.90-0.99, p=.025)
WTgrow	Mean ± SD	13.1 ± 3.9	11.6 ± 4.2	0.91 (0.84-0.98, p=.015)	0.96 (0.87-1.07, p=.502)	
fetal.sex	male	47 (43.1%)	33 (45.2%)	1.09 (0.60-1.98, p=.781)		
ODS	Yes	7 (6.4%)	22 (30.1%)	6.29 (2.52-15.69, p<.001)	5.63 (1.75-18.13, p=.004)	7.15 (2.26-22.67, p<.001)
PREE	Yes	1 (0.9%)	15 (20.5%)	27.93 (3.60-216.80, p=.001)	6.92 (0.44-108.93, p=.169)	
GH	Yes	4 (3.7%)	20 (27.4%)	9.91 (3.22-30.46, p<.001)	6.72 (1.14-39.50, p=.035)	17.36 (4.75-63.44, p<.001)
AIP	Yes	13 (11.9%)	7 (9.6%)	0.78 (0.30-2.07, p=.622)		
GD	Yes	19 (17.4%)	11 (15.1%)	0.84 (0.37-1.89, p=.674)		
ABPLH	Yes	22 (20.2%)	15 (20.5%)	1.02 (0.49-2.13, p=.952)		
ABPU	Yes	10 (9.2%)	23 (31.5%)	4.55 (2.01-10.30, p<.001)	5.36 (1.91-15.09, p=.002)	4.48 (1.72-11.65, p=.002)
gravidity	2	29 (30.5%)	14 (20.6%)	0.55 (0.26-1.19, p=.129)		
	>=3	19 (20%)	13 (19.1%)	0.78 (0.35-1.78, p=.562)		
parity	2	43 (39.4%)	19 (26%)	0.49 (0.26-0.95, p=.034)	0.60 (0.26-1.39, p=.232)	
	3	7 (6.4%)	1 (1.4%)	0.16 (0.02-1.34, p=.090)	0.03 (0.00-1.41, p=.075)	

ABPLH abnormal pregnancy-labor history, *ABPU* abnormal placenta or umbilical cord, *AIP* anemia in pregnancy, *GD* gestational diabetes, *GH* gestational hypertension, *ODS* oligohydramnios, *PREE* preeclampsia, *SD* standard deviation, *WTgrow* maternal weight grows in pregnancy

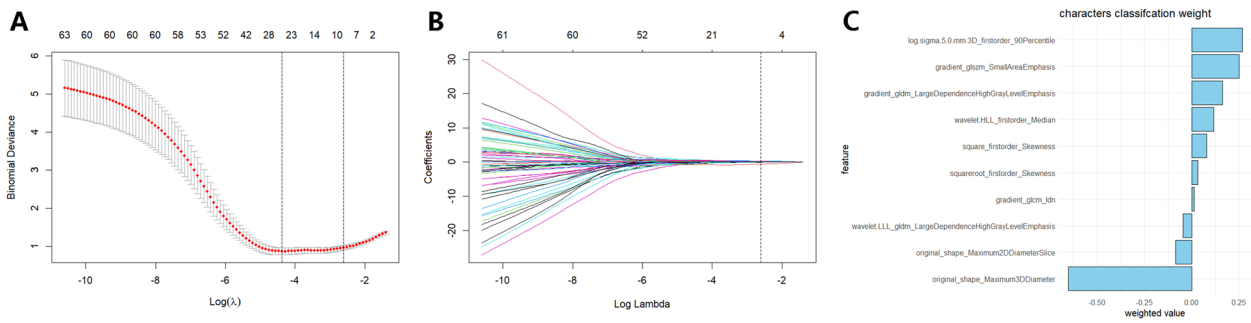

Fig. 3 Best match features screening by LASSO regression. **A** LASSO regression cross-validation curve. **B** LASSO regression path diagram. **C** The selected radiomics features and coefficient weighted values

radiomic signature and corresponding coefficients of the training, internal testing, and external testing sets. A correlation heat map of the radiomic signature is presented in Figure S1, and the Rad-score calculation formula is showed in Table S1.

Comparative study results

Waterfall plots of the Rad-score of the training, internal testing, and external testing sets are shown in Figure S2. Overall, FGR cases (purple) had higher Rad-scores than the non-FGR cases (blue). Violin-and-box (Figure S3) plots show significant differences ($p < 0.001$) in Rad-scores between the FGR and non-FGR case. The radiomics model was constructed with 10 selected features using the workflow approach of the tidymodels package with eight machine learning algorithms and tenfold cross validation. Finally, compared with the ROC-AUC, Brier-class and accuracy (Fig. 4, Table S2), the logistic regression algorithm was used in the radiomics model, which had the highest AUC (0.893), the lowest Brier score (0.128), and the higher accuracy value (0.826). By explained the final radiomics model using DALEXtra package, the importance ranking of variables was presented in Figure S4.

Next, the indices of all prediction models were compared (Table S3). The united model demonstrated the best performance across the training, internal testing and external testing sets, with AUC 0.941 (95% CI, 0.0.904–0.977), 0.899 (95% CI, 0.789–1) and 0.861 (95% CI 0.725–0.998), prediction accuracies 0.885, 0.844 and 0.805, precisions 0.871, 0.789 and 0.867, recalls 0.836, 0.833 and 0.684, and F1 scores 0.853, 0.811 and 0.765, respectively. Figure 5 shows the ROC curves of each prediction model with the training, internal testing, and external testing sets. the calibration and decision curves of the united model are shown in Figure S5. The nomogram of the united model developed with the training set is presented in Fig. 6.

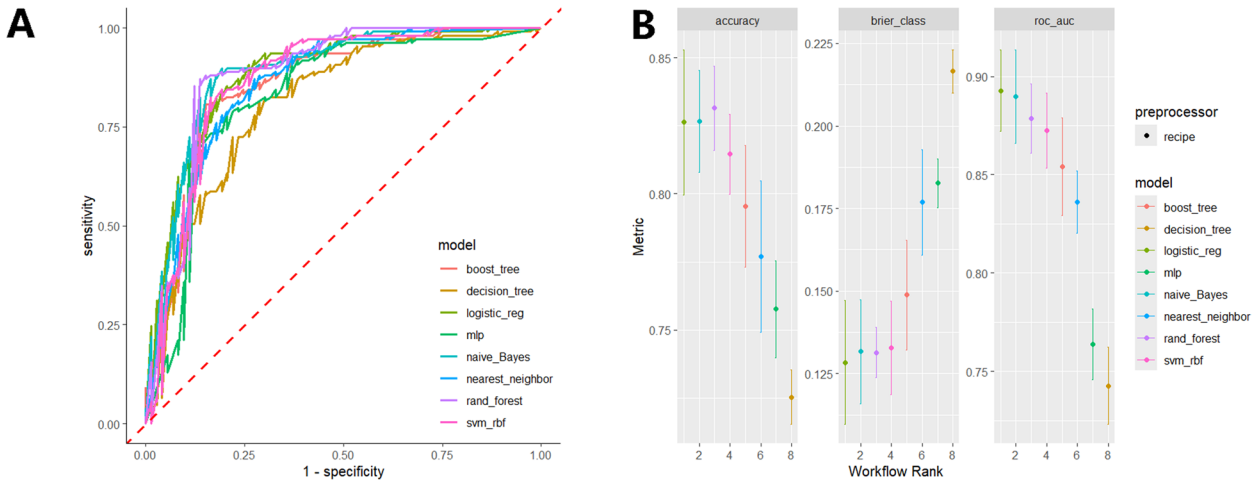

Fig. 4 Comparisons of the (**A**) ROC curves, (**B**) accuracies, Brier scores, and AUC values of the eight machine learning algorithms constructed radiomics model

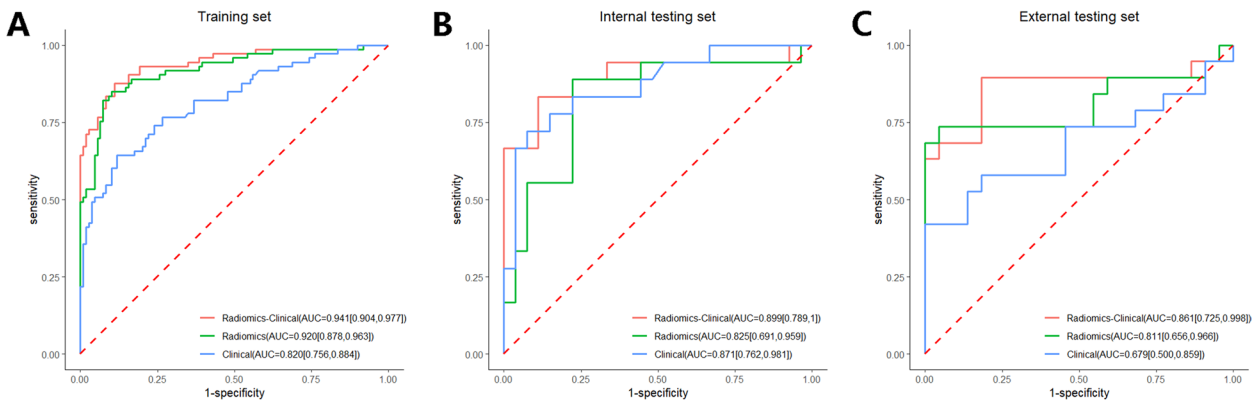

Fig. 5 ROC curves of the clinical, radiomics, and united models with the (**A**) training, (**B**) internal testing, and (**C**) external testing sets

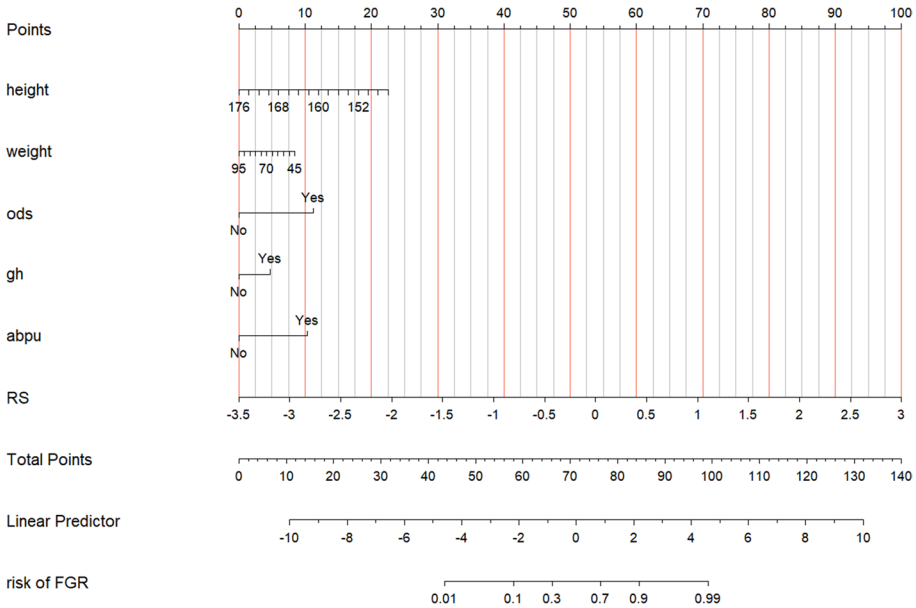

Fig. 6 Nomogram of the united model developed with the training set

Discussion

In this study, a placental MRI-based radiomics signature model and a united model based on the Rad-score combined with the clinical factors were developed to predict FGR prenatally. After model validation and visualization, the radiomics model using the logistic regression obtained the highest evaluation indices. The waterfall plot and violin box plot show that Rad-score has a high ability to identify FGR. Compared with the radiomics signature model, the performance of the united model in predicting FGR was further improved. The calibration curve showed that the model fitting line was basically consistent with the diagonal line, and the decision curve showed that the net gains of the united model with the training, internal testing, and external testing sets were much higher than the treat-all and treat-none lines. The developed nomograms facilitated clinical application and decision-making of the united model.

Generalization of machine learning-based models is key. In clinical practice, FGR assessment models based on retrospective analysis are needed to predict future outcomes. However, it is difficult to fully validate the generalization ability of the model only with training and internal testing sets that are derived from the same MR images. Therefore, an external testing set with different scanners and parameters is used to more comprehensively validate the effectiveness and generalization of each model.

In the placenta, the T2 value reflects the volume fraction of the maternal and fetal blood and villous tissue, as well as the magnetic field spatial changes of each voxel, and Oxyhemoglobin have slightly diamagnetic (negative), magnetic susceptibility, whereas deoxyhemoglobin has paramagnetic (positive) magnetic susceptibility [28]. Therefore, any spatial variations in the concentration of deoxyhemoglobin due to changes in placental blood flow perfusion could lead to the changes of T2 value in the placental magnetic resonance images.

Radiomics analysis of the T2-weighted sequence was conducted to identify placenta-based features, while ICC analysis was performed to eliminate unstable features caused by differences in the ROI by artificial segmentation, thereby enhancing robustness. The diagnostic criteria for FGR was adopted from the consensus definition of fetal growth restriction [8], the International Federation of Gynecology and Obstetrics initiative on fetal growth [1], and more accurate BW reference data based on a Chinese population [29] rather than the unreliable EFW. The reference data for fetal abdominal circumference corresponding to gestational age were also based on an Asian population [30]. The umbilical artery pulse index refers to the standards established by the Fetal Medicine Foundation [31]. The above reference criteria ensured the accuracy of the dependent variables of the model.

Ultrasound is a commonly used method for monitoring of fetal development monitoring and screening of diseases in clinical practice [32]. However, fetal-based

measurements obtained by ultrasound do not guarantee accurate assessment of FGR [33]. Therefore, more effective detection tools and prediction methods are needed. Various placental-based radiomics features have been proposed to complement fetal-based measurements to predict FGR [16, 17, 20]. Clinically, the physical foundation of the pregnant woman, in addition to the fetus and placenta, also affects fetal development [21, 22]. The nomogram constructed in this study showed that the risk of FGR increased with decreasing maternal height and weight. Therefore, the radiomics signature combined with easily accessible maternal clinical characteristics can improve the prediction capabilities of the model.

In this study, the FGR prediction united model based on placental MRI radiomics features and clinical factors obtained high ROC-AUC, accuracy, precision and F1 scores, reflecting the good efficacy of the model.

Limitations

This study included only T2-weighted sequences in the MRI-based radiomics analysis, thereby overlooking the potential contributions of multimodal or multi-sequence imaging in assessing FGR. Additionally, the study did not differentiate between early and late-onset FGR, which could have resulted in variations in placental texture. Furthermore, the study was constrained by the limited variety of MRI equipment manufacturers. The study cohort was also relatively small. Consequently, a larger testing set is required to confirm the generalizability of the proposed model.

Conclusion

The machine learning model based on placental MRI features performed well. The inclusion of the clinical characteristics of pregnant women further improved the ability of the model to predict FGR. The proposed model is suitable for use in clinical practice.

Abbreviations

AC	Abdominal circumference
AUC	Area under the curve
BW	Birth weight
CI	Confidence interval
EFW	Estimated fetal weight
FGR	Fetal growth restriction
FOV	Field of view
GA	Gestational age
ICC	Intraclass correlation coefficient
LASSO	Least absolute shrinkage and selection operators
MRI	Magnetic resonance imaging
ROC	Receiver operating characteristic curve
ROI	Region of interest
SE	Standard error
SGA	Small for gestational age

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-025-07450-1.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.
Supplementary Material 7.

Acknowledgements

Not applicable.

Authors' contributions

Jida Wang and Zhuying Chen designed the study and drafted/revised the manuscript. Kui Li and Meixiang Deng made contributions to the acquisition of clinical study data. Hongxi Zhang and Weikang Li analysed the imaging data. Yu Zou contributed to reviewing and revising the manuscript.All authors reviewed the manuscript.

Funding

The authors state that this work has not received any funding.

Data availability

Data will be provided by the correspondence author (zouyuzju@zju.edu.cn) upon request.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Institutional Review Board of Women's Hospital School of Medicine Zhejiang University (approval no. IRB-20240075-R) and conducted in accordance with the tenets of the Declaration of Helsinki. Due to the retrospective design of this study, the requirement for written informed consent was waived.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Radiology, Women'S Hospital, Zhejiang University School of Medicine, Zhejiang Provincial Clinical Research Center for Obstetrics and Gynecology, Hangzhou, Zhejiang 310006, China. [2]Department of Radiology, Children'S Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, Zhejiang, China.

Received: 20 July 2024 Accepted: 10 March 2025
Published online: 20 March 2025

References

1. Melamed N, Baschat A, Yinon Y, et al. FIGO (International Federation of Gynecology and Obstetrics) initiative on fetal growth: Best practice advice for screening, diagnosis, and management of fetal growth restriction. Int J Gynecol Obstet. 2021;152(S1):3–57. https://doi.org/10.1002/ijgo.13522.
2. Lees C, Marlow N, Arabin B, et al. Perinatal morbidity and mortality in early-onset fetal growth restriction: cohort outcomes of the trial of

randomized umbilical and fetal flow in Europe (TRUFFLE). Ultrasound Obst Gyn. 2013;42(4):400–8. https://doi.org/10.1002/uog.13190.
3. Pérez-Cruz M, Cruz-Lemini M, Fernández MT, et al. Fetal cardiac function in late-onset intrauterine growth restriction vs small-for-gestational age, as defined by estimated fetal weight, cerebroplacental ratio and uterine artery Doppler. Ultrasound Obst Gyn. 2015;46(4):465–71. https://doi.org/10.1002/uog.14930.
4. Meher S, Hernandez-Andrade E, Basheer SN, Lees C. Impact of cerebral redistribution on neurodevelopmental outcome in small-for-gestational-age or growth-restricted babies: a systematic review. Ultrasound Obst Gyn. 2015;46(4):398–404. https://doi.org/10.1002/uog.14818.
5. Jaddoe VWV, de Jonge LL, Hofman A, Franco OH, Steegers EAP, Gaillard R. First trimester fetal growth restriction and cardiovascular risk factors in school age children: population based cohort study. BMJ. 2014;348: g14. https://doi.org/10.1136/bmj.g14.
6. Unterscheider J, Daly S, Geary MP, et al. Optimizing the definition of intrauterine growth restriction: the multicenter prospective PORTO Study. Am J Obstet Gynecol. 2013;208(4):290.e1-290.e6. https://doi.org/10.1016/j.ajog.2013.02.007.
7. 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 in Obstet & Gyne. 2020;56(2):298–312. https://doi.org/10.1002/uog.22134.
8. Gordijn SJ, Beune IM, Thilaganathan B, et al. Consensus definition of fetal growth restriction: a Delphi procedure. Ultrasound Obst Gyn. 2016;48(3):333–9. https://doi.org/10.1002/uog.15884.
9. Melamed N, Ryan G, Windrim R, Toi A, Kingdom J. Choice of Formula and Accuracy of Fetal Weight Estimation in Small-for-Gestational-Age Fetuses. J Ultrasound Med. 2016;35(1):71–82. https://doi.org/10.7863/ultra.15.02058.
10. Gj B, E J. Pathophysiology of placental-derived fetal growth restriction. Am J Obstet Gynecol. 2018;218(2S). https://doi.org/10.1016/j.ajog.2017.11.577.
11. Sun C, Groom KM, Oyston C, Chamley LW, Clark AR, James JL. The placenta in fetal growth restriction: What is going wrong? Placenta. 2020;96:10–8. https://doi.org/10.1016/j.placenta.2020.05.003.
12. Parks WT. Placental hypoxia: the lesions of maternal malperfusion. Semin Perinatol. 2015;39(1):9–19. https://doi.org/10.1053/j.semperi.2014.10.003.
13. Hansen DN, Sinding M, Petersen A, et al. T2*-weighted placental magnetic resonance imaging: a biomarker of placental dysfunction in small-for-gestational-age pregnancies. American journal of obstetrics & gynecology MFM. 2022;4(3): 100578. https://doi.org/10.1016/j.ajogmf.2022.100578.
14. Baadsgaard K, Hansen DN, Peters DA, Frøkjær JB, Sinding M, Sørensen A. T2* weighted fetal MRI and the correlation with placental dysfunction. Placenta. 2023;131:90–7. https://doi.org/10.1016/j.placenta.2022.12.002.
15. Liu XL, Feng J, Huang CT, Mei YJ, Xu YK. Use of intravoxel incoherent motion MRI to assess placental perfusion in normal and Fetal Growth Restricted pregnancies on their third trimester. Placenta. 2022;118:10–5. https://doi.org/10.1016/j.placenta.2021.12.019.
16. Dahdouh S, Andescavage N, Yewale S, et al. In vivo placental MRI shape and textural features predict fetal growth restriction and postnatal outcome. J Magn Reson Imaging. 2018;47(2):449–58. https://doi.org/10.1002/jmri.25806.
17. Andescavage N, Dahdouh S, Jacobs M, et al. In vivo textural and morphometric analysis of placental development in healthy & growth-restricted pregnancies using magnetic resonance imaging. Pediatr Res. 2019;85(7):974–81. https://doi.org/10.1038/s41390-019-0311-1.
18. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–62. https://doi.org/10.1038/nrclinonc.2017.141.
19. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. Radiology. 2016;278(2):563–77. https://doi.org/10.1148/radiol.2015151169.
20. Song F, Li R, Lin J, et al. Predicting the risk of fetal growth restriction by radiomics analysis of the placenta on T2WI: A retrospective case-control study. Placenta. 2023;134:15–22. https://doi.org/10.1016/j.placenta.2023.02.007.
21. Fetal Growth Restriction. ACOG Practice Bulletin, Number 227. Obstet Gynecol. 2021;137(2):e16–28. https://doi.org/10.1097/AOG.0000000000004251.
22. Vayssière C, Sentilhes L, Ego A, et al. Fetal growth restriction and intrauterine growth restriction: guidelines for clinical practice from the French College of Gynaecologists and Obstetricians. Eur J Obstet Gynecol Reprod Biol. 2015;193:10–8. https://doi.org/10.1016/j.ejogrb.2015.06.021.
23. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage. 2006;31(3):1116–28. https://doi.org/10.1016/j.neuroimage.2006.01.015.
24. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res. 2017;77(21):e104–7. https://doi.org/10.1158/0008-5472.CAN-17-0339.
25. Liljequist D, Elfving B, Skavberg RK. Intraclass correlation - A discussion and demonstration of basic features. PLoS ONE. 2019;14(7):e0219854. https://doi.org/10.1371/journal.pone.0219854.
26. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
27. Koçak B, Durmaz EŞ, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. Diagnostic and Interventional Radiology (Ankara, Turkey). 2019;25(6):485–95. https://doi.org/10.5152/dir.2019.19321.
28. Sørensen A, Hutter J, Seed M, Grant PE, Gowland P. T2*-weighted placental MRI: basic research tool or emerging clinical test for placental dysfunction? Ultrasound in Obstet & Gyne. 2020;55(3):293–302. https://doi.org/10.1002/uog.20855.
29. L D, C D, Y L, et al. Birth weight reference percentiles for Chinese. PloS one. 2014;9(8). https://doi.org/10.1371/journal.pone.0104779.
30. Buck Louis GM, Grewal J, Albert PS, et al. Racial/ethnic standards for fetal growth: the NICHD Fetal Growth Studies. Am J Obstet Gynecol. 2015;213(4):449.e1-449.e41. https://doi.org/10.1016/j.ajog.2015.08.032.
31. Ciobanu A, Wright A, Syngelaki A, Wright D, Akolekar R, Nicolaides KH. Fetal Medicine Foundation reference ranges for umbilical artery and middle cerebral artery pulsatility index and cerebroplacental ratio. Ultrasound Obst Gyn. 2019;53(4):465–72. https://doi.org/10.1002/uog.20157.
32. Salomon LJ, Alfirevic Z, Da Silva CF, et al. ISUOG Practice Guidelines: ultrasound assessment of fetal biometry and growth. Ultrasound Obstet Gynecol. 2019;53(6):715–23. https://doi.org/10.1002/uog.20272.
33. Whitham MD, Reynolds DM, Urban AR, Ennen CS, Dudley DJ. Comparative Diagnostic Performance of Estimated Fetal Weight and Isolated Abdominal Circumference for the Detection of Fetal Growth Restriction. J Ultrasound Med. 2023;42(2):477–85. https://doi.org/10.1002/jum.16001.

Publisher's Note