International Journal of Women's Health

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Comparison of Interpretable Machine Learning Models Using Systemic Inflammation Index to Predict Preterm Birth in Gestational Diabetes Mellitus

Qinxia Pang*, Lei Peng*, Jianfa Wu, Ying Wang, Rong Zhang, Zhou Liu, Lingli Jiang

Department of Obstetrics and Gynecology, Shanghai University of Medicine & Health Sciences Affiliated Zhoupu Hospital, Shanghai, People's Republic of China

*These authors contributed equally to this work

Correspondence: Lingli Jiang; Zhou Liu, Department of Obstetrics and Gynecology, Shanghai University of Medicine & Health Sciences affiliated Zhoupu Hospital, No. 1500, Zhouyuan Road, Pudong New Area, Shanghai, 201318, People's Republic of China, Tel +86-13817459299; +86-18930837705, Email jiangll@sumhs.edu.cn; zp_liuz@sumhs.edu.cn

Background: Gestational diabetes mellitus (GDM) elevates preterm birth risk, highlighting the need for improved prediction methods to enhance outcomes. Current models show limited accuracy by ignoring some inflammatory biomarkers (eg, PLR, LMR, SII). Machine learning (ML) can better analyze complex patterns but remains underused for GDM preterm birth prediction.

Objective: This study develops an interpretable ML model combining systemic inflammatory indices and traditional clinical markers to predict preterm birth in GDM. Enabling early risk stratification at diagnosis, it facilitates timely interventions for this high-risk population.

Methods: This retrospective study analyzed 389 GDM patients, stratified into training (n=272) and temporal external validation (n=117) cohorts, and further classified by birth outcome (term/preterm). Using the training cohort, we developed and internally validated multiple ML models incorporating: (1) systemic inflammation indices, (2) traditional clinical indicators, and (3) their combination. The optimal model underwent temporal external validation and subsequent Shapley Additive Explanations (SHAP) analysis for feature interpretation. To assess the robustness of our findings, sensitivity analyses were conducted.

Results: Our cohort of 389 GDM patients included 53 preterm births (13.6%). Analysis revealed seven significant predictors combining systemic inflammatory markers and traditional clinical parameters. The extreme gradient boosting (XGBoost) model outperformed comparative algorithms (AUC-ROC: 0.932 vs Logit: 0.871, SVM: 0.847, RF: 0.917; AUC-PRC: 0.754 vs Logit: 0.686, SVM: 0.582, RF: 0.670). SHAP analysis identified five key determinants (two clinical and three inflammatory markers) as most influential for preterm birth prediction. Sensitivity analyses were conducted to assess the robustness of the results.

Conclusion: The XGBoost model outperforms in predicting GDM-related preterm birth by integrating traditional clinical and systemic inflammatory markers, enabling precise risk assessment to guide clinical management.

Keywords: machine learning, preterm birth, gestational diabetes mellitus, systemic inflammation index

Introduction

Gestational diabetes mellitus (GDM) is a common metabolic disorder characterized by carbohydrate intolerance and represents one of the fastest-growing pregnancy complications. It is defined as the first occurrence or detection of abnormal glucose metabolism during pregnancy.[1,2] Beyond its metabolic implications, this condition significantly elevates the risk of adverse perinatal outcomes, such as preterm birth (<37 weeks' gestation).[3,4] Although current management combining lifestyle modifications and pharmacological therapy can improve glycemic control and reduce preterm birth risk, a clinically significant proportion of GDM patients still experience preterm delivery.[5–7] Preterm birth imposes a dual burden, constituting both a major financial strain on families and the leading cause of under-5 mortality

https://doi.org/10.2147/IJWH.S541610
Received: 20 May 2025
Accepted: 23 January 2026
Published: 10 February 2026

worldwide.[8] Early prediction and prevention of preterm delivery in GDM patients could significantly improve perinatal outcomes while alleviating healthcare system pressures. This underscores the critical need for reliable preterm birth prediction methods in this high-risk population.

GDM pathogenesis demonstrates significant associations with both established clinical predictors (including maternal age ≥35 years, smoking status, and insulin resistance)[9,10] and inflammatory alterations.[11,12] Although current prediction models based on clinical parameters achieve only modest sensitivity (51.5%) for preterm birth,[13] the incorporation of routinely measured inflammatory markers (such as lymphocyte and neutrophil counts) may enhance predictive accuracy. In GDM, circulating inflammatory cells (such as monocytes and neutrophils) and proinflammatory cytokines (eg, IL-1β, IL-6, TNF-α) are elevated. As pregnancy progresses, increasing placental and adipose tissue further boosts cytokine secretion, while elevated glucose levels activate inflammatory pathways, leading to the release of additional cytokines and chemokines (eg, CXCL1, CXCL5, CXCL8, CCL2).[14] Parturition involves uterine contractions, cervical dilation, and membrane rupture, all driven by inflammation.[15–20] GDM may prematurely trigger this inflammatory process, increasing the risk of preterm birth. Recent studies highlight the prognostic value of hematologic inflammatory indices - including neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), lymphocyte-to-monocyte ratio (LMR), and systemic immune-inflammation index (SII) - in assessing systemic inflammation and disease severity.[21] Higher levels of NLR, PLR, LMR, and SII are associated with GDM.[22] Derived from routine complete blood count parameters (neutrophils, lymphocytes, and monocytes), these indices demonstrate significant clinical advantages: they are cost-efficient, easily obtainable, and provide comprehensive profiling. Importantly, their potential association with preterm birth pathogenesis suggests these biomarkers may improve risk prediction in GDM patients. Nevertheless, the complex, nonlinear relationships between multidimensional clinical data and patient outcomes present significant challenges for conventional linear models (eg, logistic regression [Logit]), limiting their predictive accuracy. In this context, machine learning (ML) methods offer distinct advantages due to their capacity to identify sophisticated patterns in high-dimensional biomedical data.[23] Although ML has demonstrated success in various medical domains, including oncology[24] and cardiology,[25] its application for preterm birth prediction in GDM using systemic inflammation biomarkers and traditional clinical predictors remains underexplored.

In this context, this study aims to develop and validate an interpretable ML model that integrates systemic inflammation indices with traditional clinical predictors to identify novel and practical obstetric biomarkers for preterm birth risk assessment in GDM patients. By enabling early risk stratification at the time of GDM diagnosis, this approach will facilitate timely clinical interventions to prevent preterm delivery in this high-risk population.

Methods

Ethics Statement

This study received approval from the medical ethics committee of Shanghai University of Medicine & Health Sciences affiliated Zhoupu Hospital (Approval No. 2024-C-160-E01) and complied with the Declaration of Helsinki. Given the retrospective, observational design, the ethics committee granted a waiver of individual informed consent. Patient confidentiality was maintained through comprehensive deidentification procedures, with systematic removal of all personal identifiers from electronic health records prior to analysis in accordance with institutional privacy standards.

Subjects and Study Design

This retrospective cohort study was conducted from August 2019 to August 2024, during which 568 patients with GDM were consecutively screened from Shanghai University of Medicine & Health Sciences affiliated Zhoupu Hospital. Participants met the following criteria: (1) age ≥18 years; (2) GDM diagnosis confirmed by 75g oral glucose tolerance test (OGTT) at 24–28 weeks' gestation (fasting glucose ≥5.1 mmol/L, 1-hour ≥10.0 mmol/L, or 2-hour ≥8.5 mmol/L); and (3) singleton pregnancy. Exclusion criteria included pre-existing diabetes, multifetal gestation, prior preterm birth, incomplete medical records, and significant comorbidities (renal, hepatic, or cardiovascular diseases). Following strict inclusion/exclusion criteria, 389 GDM patients were enrolled and divided into two cohorts: (1) a training cohort (n=272; August 2019 to March 2023), which included subgroups of term birth and preterm birth; and (2) a temporal external validation cohort (n=117; April 2023 to August 2024) (Figure 1A–C).

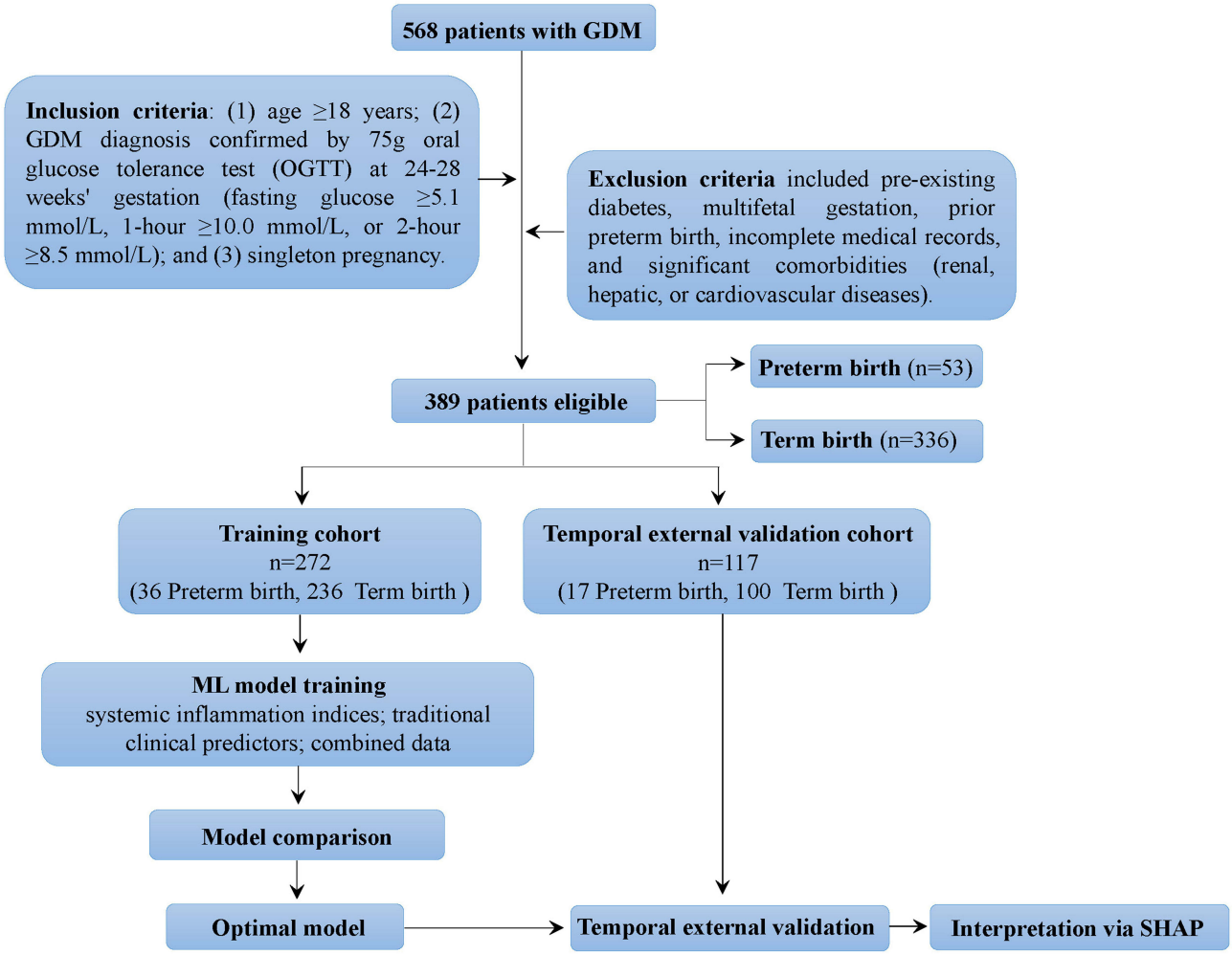

Figure 1 Flowchart of patient selection and cohort distribution for developing and validating predictive models to assess preterm birth risk in patients with GDM.
Abbreviations: GDM, gestational diabetes mellitus; ML, machine learning; SHAP, Shapley Additive Explanations.

Data Collection

This study collected comprehensive maternal data including demographic characteristics (age, pre-pregnancy BMI), medical history (diabetes family history, smoking/alcohol use, hypertension, parity, uterine curettage, IVF-ET), OGTT results, and hematological indices at GDM diagnosis (complete blood counts and derived inflammatory ratios: NLR, PLR, LMR, SII).

Feature Preprocessing and Selection in the Training Cohort

Before model development, data preprocessing was performed to ensure fairness. Continuous variables were standardized to Z-scores (mean = 0, standard deviation = 1), and categorical variables were binarized to "0" or "1," ensuring fair comparison across scales.

To simplify the model, we first identified statistically significant variables distinguishing between term and preterm birth groups using univariate logistic regression, followed by appropriate statistical tests (Student's *t*-test, Mann–Whitney *U*-test, or chi-square test based on data distribution) to validate the results of the univariate logistic regression. Variables with $P < 0.05$ in the univariate analysis were then included in a multivariable logistic regression to identify independent risk factors and construct the predictive model.

https://doi.org/10.2147/IJWH.S541610

Development and Internal Validation of Prediction Models

In our study, a marked class imbalance was observed, with far fewer preterm births than term births. Such disparities are common in medical datasets, where the prevalence of non-cases typically outweighs cases, often impairing predictive accuracy.[26] To address this issue, we applied the Synthetic Minority Oversampling Technique (SMOTE), which generates synthetic samples based on k-nearest neighbors to balance minority and majority classes.[27,28] This method has been shown to enhance disease prediction and reduce model overfitting. Importantly, SMOTE was implemented only in the training cohort, while the validation cohort remained untouched to preserve its natural outcome distribution.

We developed four ML models (Logit, random forest [RF], support vector machines [SVM], and extreme gradient boosting [XGBoost]) to predict preterm birth risk in GDM patients using: (1) traditional clinical parameters, (2) systemic inflammatory markers, and (3) their combination. A triple five-fold cross-validation strategy prevented model overfitting. A nested cross-validation framework was employed during model development, using stratified 5-fold cross-validation to preserve outcome distribution and evaluate generalization performance. The outer loop assessed performance, while the inner loop, optimized through grid search, focused on hyperparameter tuning to prevent data leakage. This process was repeated across 1000 bootstrap iterations. Implemented with the StratifiedKFold function in scikit-learn, it separates model selection from evaluation, enhancing generalization and reducing overfitting risk.

Hyperparameter optimization for the ML models was performed using randomized grid search with 5-fold cross-validation to avoid information leakage. The LR grid included regularization strength and penalty type; RF focused on tree depth, estimators, and minimum samples per leaf; XGBoost tuned learning rate, max depth, and estimators; and SVM adjusted kernel, regularization, and kernel coefficient. Each model underwent 100 iterations, with area under the receiver operating characteristic curve (AUC-ROC) as the optimization metric. After identifying optimal parameters, final models were trained on the training cohort to maximize training data use.

Temporal External Validation and Interpretability of ML Models

The model's performance was evaluated using discrimination metrics (AUC-ROC, area under the precision-recall curve [AUC-PRC], sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV], F1 score), calibration measures (brier score, calibration curves), and clinical utility assessment via decision curve analysis (DCA). The Brier score quantifies prediction accuracy (lower values indicating better calibration), while DCA estimates net clinical benefit.

Based on cooperative game theory,[29] SHAP (SHapley Additive exPlanations) analysis quantifies variable importance through Shapley values, representing each feature's predictive contribution. This method provides both: (i) quantitative assessment of directional feature effects (protective/risk factors), and (ii) visual interpretation via summary plots (population-level importance) and force plots (individual-case predictions).[30–32]

Statistical Analysis

We conducted additional sensitivity analyses to strengthen the reliability of our findings. First, we excluded participants with a history of hypertension, smoking, or alcohol use at baseline to minimize potential confounding. Second, to address missing data, we used multiple imputation by chained equations (MICE), which preserved statistical power and minimized bias associated with missing values.

Data were analyzed using IBM SPSS (v26.0), R (v4.2.3), and Python (v3.10.0). Continuous variables were assessed for normality via the Shapiro–Wilk test and summarized as mean ± SD (normal distribution) or median (interquartile range [IQR]) (non-normal distribution), with between-group comparisons conducted using Student's *t*-test or Mann–Whitney *U*-test, respectively. Categorical variables were presented as n (%) and compared using χ^2-tests. Statistical significance was set at $p < 0.05$ (two-tailed).

Results

Patient Characteristics

Figure 1A–C outlines the participant selection process, with 389 GDM patients (of 568 screened) meeting eligibility criteria. The preterm birth prevalence was similar between cohorts: 13.2% (36/272) in the training set versus 14.5% (17/117) in the temporal external validation set (χ^2=0.077, p=0.782). Table 1 demonstrates comparable baseline characteristics between cohorts, including traditional clinical parameters and systemic inflammation indices (all p>0.05).

Table 1 Baseline Characteristics of Patients in the Training and Validation Cohorts

Variable	Training Cohort (n=272)	Validation Cohort (n=117)	*P* value
Traditional clinical parameters			
Maternal age, years, median (IQR)	29.00 (26.00, 36.00)	30.00 (26.00, 35.00)	0.872#
BMI before pregnancy, kg/m^2, median (IQR)	23.60 (19.88, 25.72)	23.90 (21.00, 26.60)	0.294#
Family history of diabetes mellitus, n (%)	31 (11.4)	12 (10.3)	0.879*
Smoking history, n (%)	3 (1.1)	2 (1.7)	1.000*
Alcohol drinking history, n (%)	26 (9.6)	14 (12.0)	0.593*
History of hypertension, n (%)	14 (5.1)	6 (5.1)	1.000*
Parity ≥1, n (%)	83 (30.5)	40 (34.2)	0.551*
Uterine curettage, n (%)			0.990*
0	187 (68.8)	79 (67.5)	
1	73 (26.8)	32 (27.4)	
2	8 (2.9)	4 (3.4)	
≥3	4 (1.5)	2 (1.7)	
IVF-ET, n (%)	41 (15.1)	18 (15.4)	1.000*
FBG, mmol/L, median (IQR)	5.30 (4.60, 6.00)	5.20 (4.60, 5.90)	0.721#
1H-OGTT, mmol/L, median (IQR)	11.10 (10.10, 12.00)	11.20 (10.10, 12.30)	0.309#
2H-OGTT, mmol/L, median (IQR)	10.10 (9.40, 10.80)	10.40 (9.60, 11.00)	0.056#
Systemic inflammation indices at GDM diagnosis			
WBC, 10^9/L, median (IQR)	10.30 (6.00, 13.30)	9.30 (6.30, 13.00)	0.580#
Lymphocyte count, ×10^9/L, median (IQR)	2.04 (0.95, 2.95)	1.90 (0.98, 2.86)	0.580#
Neutrophil count, ×10^9/L, median (IQR)	6.77 (4.07, 9.13)	6.53 (4.24, 8.83)	0.637#
Platelet, 10^9/L, median (IQR)	241.64 (130.12, 366.11)	250.23 (124.30, 360.45)	0.781#
Monocyte count, ×10^9/L, median (IQR)	0.65 (0.53, 0.78)	0.65 (0.53, 0.78)	0.841#
Eosinophil, 10^9/L, median (IQR)	0.40 (0.23, 0.53)	0.34 (0.20, 0.53)	0.230#
Basophil, 10^9/L, median (IQR)	0.08 (0.04, 0.12)	0.08 (0.04, 0.11)	0.645#
NLR, median (IQR)	3.30 (2.80, 4.10)	3.40 (2.80, 4.20)	0.563#

(*Continued*)

Table 1 (Continued).

Variable	Training Cohort (n=272)	Validation Cohort (n=117)	*P* value
PLR, median (IQR)	124.00 (93.65, 151.98)	122.40 (98.20, 154.10)	0.562[#]
LMR, median (IQR)	2.80 (1.50, 4.60)	2.60 (1.50, 4.10)	0.603[#]
SII, median (IQR)	784.85 (435.75, 1206.26)	816.93 (522.77, 1165.31)	0.919[#]

Notes: *For chi-square test; [#]For Mann–Whitney *U*-test.
Abbreviations: IQR, inter-quartile range; BMI, body mass index; IVF-ET, in vitro fertilization and embryo transfer; FBG, fasting blood glucose; OGTT, oral glucose tolerance test; GDM, gestational diabetes mellitus; WBC, white blood cell; NLR, neutrophil-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; LMR, lymphocyte-to-monocyte ratio; SII, systemic immune-inflammation index.

Feature Selection in the Training Cohort

Table 2 presents the logistic regression analyses identifying preterm birth risk factors in the training cohort. Univariate analysis revealed significant associations (P<0.05) for eight variables: maternal age, uterine curettage, and six inflammatory markers (neutrophil count, monocyte count, NLR, PLR, LMR, SII), with consistent non-parametric test results (Table S1). Multivariate analysis identified seven independent predictors: maternal age, uterine curettage, and five

Table 2 Univariate and Multivariate Logistic Analyses to Determine the Independent Predictors Associated with Preterm Birth in the Training Cohort

Variables	Univariate Logistic Regression			Multivariate Logistic Regression		
	P value	OR	95% CI	P value	OR	95% CI
Traditional clinical parameters						
Maternal age, years	<0.001	1.081	1.037–1.129	0.005	1.150	1.051–1.282
BMI before pregnancy, kg/m^2	0.409	1.037	0.951–1.134			
Family history of diabetes mellitus	0.109	2.132	0.791–5.183			
Smoking history	0.330	3.343	0.153–35.789			
Alcohol drinking history	0.789	0.842	0.192–2.596			
History of hypertension	0.905	1.098	0.166–4.259			
Parity ≥1	0.250	0.613	0.251–1.353			
Uterine curettage						
0	Reference			Reference		
1	0.039	2.228	1.029–4.748	0.017	7.020	1.576–40.592
2	0.025	5.633	1.085–24.966	0.112	52.500	0.899–113.643
≥3	0.334	3.130	0.151–25.934	0.002	218.000	5.858–788.046
IVF-ET	0.774	1.149	0.408–2.797			
FBG, mmol/L	0.511	1.165	0.738–1.840			
1H-OGTT, mmol/L	0.476	1.121	0.820–1.542			
2H-OGTT, mmol/L	0.533	1.149	0.743–1.785			

(*Continued*)

Table 2 (Continued).

Variables	Univariate Logistic Regression			Multivariate Logistic Regression		
	P value	OR	95% CI	P value	OR	95% CI
Systemic inflammation indices at GDM diagnosis						
WBC, 10^9/L	0.100	1.072	0.988–1.167			
Lymphocyte count, $\times10^9$/L	0.109	0.762	0.540–1.055			
Neutrophil count, $\times10^9$/L	0.037	1.124	1.008–1.256	0.369	3.870	0.130–16.428
Platelet, 10^9/L	0.100	1.002	0.999–1.005			
Monocyte count, $\times10^9$/L	<0.001	76.037	21.109–310.209	0.015	335.000	3.854–470.814
Eosinophil, 10^9/L	0.924	0.913	0.141–5.910			
Basophil, 10^9/L	0.745	0.267	0.009–787.836			
NLR	<0.001	2.480	1.842–3.481	0.013	2.250	1.835–6.561
PLR	<0.001	1.031	1.021–1.042	0.013	1.080	1.026–1.160
LMR	<0.001	0.583	0.433–0.748	0.037	0.470	0.066–0.658
SII	<0.001	1.002	1.001–1.003	0.012	1.990	1.155–10.008

Abbreviations: BMI, body mass index; IVF-ET, in vitro fertilization and embryo transfer; FBG, fasting blood glucose; OGTT, oral glucose tolerance test; GDM, gestational diabetes mellitus; WBC, white blood cell; NLR, neutrophil-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; LMR, lymphocyte-to-monocyte ratio; SII, systemic immune-inflammation index; OR, odds ratio; CI, confidence interval.

inflammatory indices (monocyte count, NLR, PLR, LMR, and SII) (all P<0.05; Table 2). Multicollinearity was assessed using the variance inflation factor (VIF), with all independent predictors showing VIF values below 2, indicating no multicollinearity. Therefore, LASSO regression is likely unnecessary. These seven key variables were subsequently incorporated into the final ML model. These inflammatory indices can be derived from routine blood tests, making them practical, rapid, and cost-effective tools. This is particularly advantageous in healthcare settings with limited resources, where access to advanced imaging or molecular diagnostics may be restricted.

Comparing Models for Predicting Preterm Birth Risk

We evaluated four ML algorithms (Logit, SVM, RF, and XGBoost) for preterm birth prediction in GDM patients using three predictor sets: traditional clinical parameters, systemic inflammation indices, and their combined features. The combined-feature models demonstrated superior performance (AUC-ROC 0.847–0.932; AUC-PRC 0.582–0.754) versus clinical-only (AUC-ROC 0.761–0.820; AUC-PRC 0.394–0.486) or inflammation-only (AUC-ROC 0.786–0.872; AUC-PRC 0.475–0.708) approaches (DeLong's test, p<0.05), as shown in Table 3 (performance metrics) and Figures 2–4 (ROC, PRC, calibration, and decision curves).

XGBoost demonstrated superior predictive performance among combined clinical-inflammatory models, achieving the highest discriminative accuracy (AUC-ROC = 0.932; AUC-PRC=0.754) with excellent calibration, particularly in the clinically relevant risk range below 30%. DCA confirmed robust clinical utility across all models, with XGBoost consistently outperforming in sensitivity, specificity, PPV, NPV, F1-score, and Brier score. These findings support XGBoost as the preferred model for preterm birth risk prediction in GDM patients. Furthermore, our findings align with a previous study indicating that ML algorithms often outperform traditional linear approaches in terms of accuracy.[33]

Temporal External Validation of ML Model Performance

The XGBoost model's predictive performance was externally validated using ROC analysis (AUC-ROC=0.893, AUC-PRC=0.667; Figures 5A and B), calibration curves (Figure 5C), and DCA (Figure 5D). While showing marginally

Table 3 Performance of ML Classifiers in Predicting Preterm Birth Risk in GDM Using Traditional Clinical Data, Systemic Inflammation Markers, and Combined Datasets

Data Type	ML Classifier	AUC-ROC	AUC-PRC	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	F1 Score	Brier Score
Traditional clinical parameters	Logit	0.761	0.473	0.650 (0.188, 0.752)	0.890 (0.823, 0.997)	0.667 (0.301, 0.954)	0.880 (0.795, 0.938)	0.658	0.023
	SVM	0.792	0.486	0.699 (0.412, 0.792)	0.811 (0.614, 0.991)	0.701 (0.376, 0.964)	0.900 (0.808, 0.946)	0.7	0.039
	RF	0.820	0.394	0.575 (0.291, 0.768)	0.882 (0.699, 0.961)	0.529 (0.291, 0.768)	0.930 (0.841, 0.969)	0.551	0.019
	XGBoost	0.811	0.416	0.625 (0.254, 0.796)	0.893 (0.712, 0.952)	0.556 (0.254, 0.746)	0.910 (0.825, 0.960)	0.588	0.020
Systemic inflammation indices	Logit	0.828	0.564	0.675 (0.344, 0.824)	0.800 (0.677, 0.924)	0.77 (0.510, 0.882)	0.900 (0.810, 0.947)	0.719	0.016
	SVM	0.786	0.475	0.651 (0.331, 0.821)	0.960 (0.621, 0.991)	0.801 (0.521, 0.941)	0.850 (0.761, 0.914)	0.718	0.049
	RF	0.872	0.594	0.750 (0.551, 0.892)	0.882 (0.669, 0.992)	0.667 (0.497, 0.918)	0.910 (0.900, 0.992)	0.706	0.007
	XGBoost	0.864	0.708	0.750 (0.552, 0.953)	0.840 (0.669, 0.902)	0.6 (0.454, 0.883)	0.970 (0.899, 0.992)	0.667	0.015
Combined traditional clinical and systemic inflammation data	Logit	0.871	0.686	0.750 (0.553, 0.899)	0.970 (0.902, 0.992)	0.7 (0.409, 0.929)	0.920 (0.834, 0.962)	0.724	0.009
	SVM	0.847	0.582	0.790 (0.451, 0.901)	0.970 (0.902, 0.992)	0.8 (0.453, 0.937)	0.930 (0.849, 0.970)	0.795	0.046
	RF	0.917	0.670	0.759 (0.451, 0.882)	0.940 (0.862, 0.978)	0.860 (0.453, 0.902)	0.970 (0.899, 0.992)	0.806	0.004
	XGBoost	0.932	0.754	0.859 (0.523, 0.907)	0.910 (0.825, 0.960)	0.867 (0.386, 0.915)	0.970 (0.896, 0.992)	0.863	0.004

Abbreviations: ML, machine learning; GDM, gestational diabetes mellitus; AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; PPV, positive predictive value; NPV, negative predictive value; CI confidence interval; Logit, logistic regression; SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

https://doi.org/10.2147/IJWH.S541610

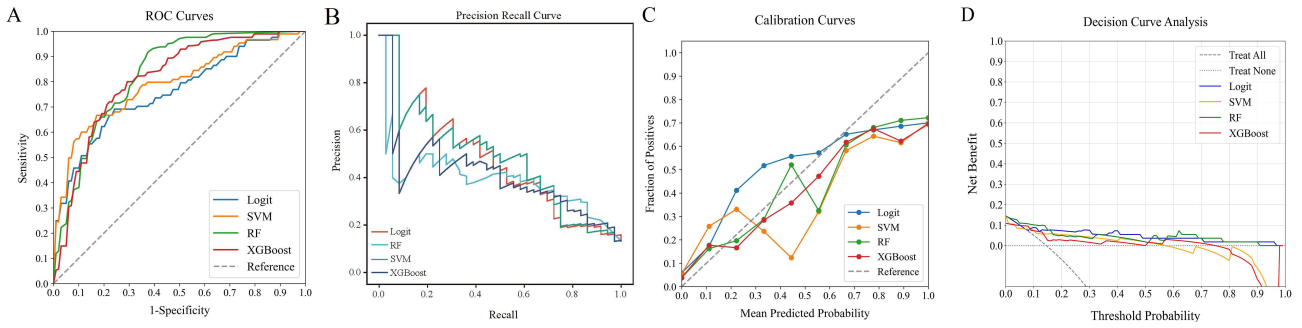

Figure 2 Performance comparison of ML classifiers (Logit, SVM, RF, XGBoost) on clinical data: (**A**) ROC curves (AUC-ROC: 0.761–0.820), (**B**) precision-recall curve (AUC-PRC: 0.394–0.486), (**C**) calibration plots, and (**D**) DCA.
Abbreviations: ML, machine learning; AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; DCA, decision curve analysis; Logit, logistic regression; SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

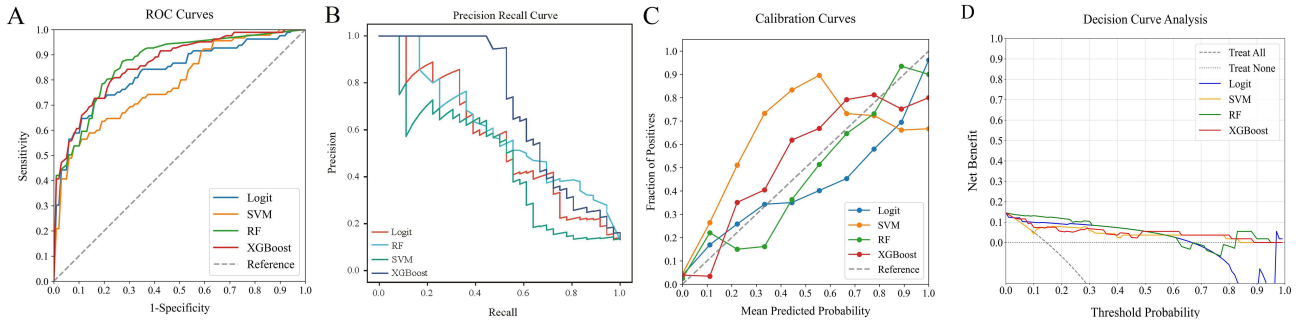

Figure 3 Performance comparison of ML classifiers (Logit, SVM, RF, XGBoost) on systemic inflammation index: (**A**) ROC curves (AUC-ROC: 0.786–0.872), (**B**) precision-recall curve (AUC-PRC: 0.475–0.708), (**C**) calibration plots, and (**D**) DCA.
Abbreviations: ML, machine learning; AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; DCA, decision curve analysis; Logit, logistic regression; SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

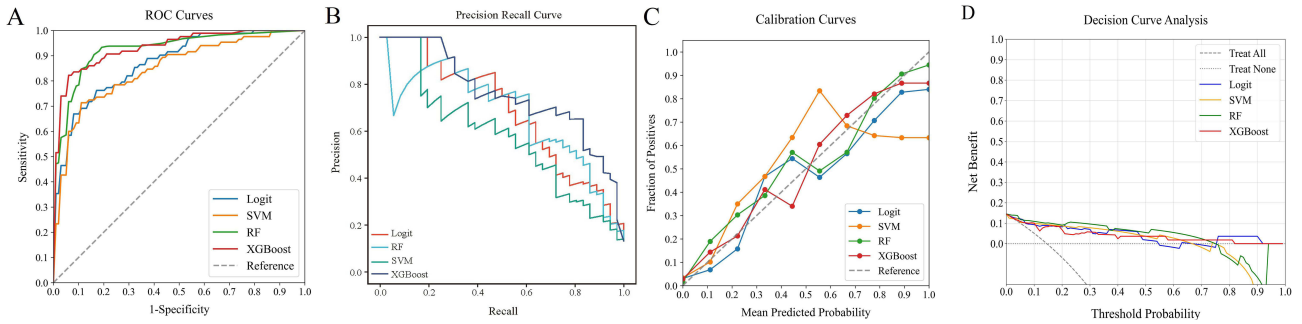

Figure 4 ML classifiers (Logit, SVM, RF, XGBoost) on combined traditional clinical and systemic inflammation data: (**A**) ROC curves (AUC-ROC: 0.847–0.932), (**B**) precision-recall curve (AUC-PRC: 0.582–0.754), (**C**) calibration, (**D**) DCA.
Abbreviations: ML, machine learning; AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; DCA, decision curve analysis; Logit, logistic regression; SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

reduced accuracy compared to the training set, the model demonstrated: (1) maintained discriminative capacity, (2) excellent prediction-observation concordance, and (3) clinically meaningful net benefits. These results confirm the model's robustness for preterm birth risk stratification in GDM patients.

https://doi.org/10.2147/IJWH.S541610

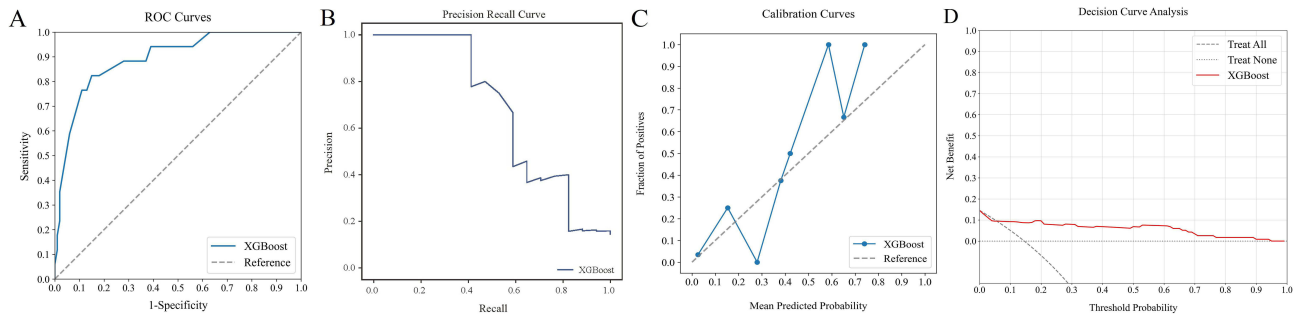

Figure 5 Validation of the optimal ML model in a temporal external cohort: (**A**) ROC (AUC-ROC=0.893), (**B**) precision-recall curve (AUC-PRC=0.667), (**C**) calibration, (**D**) DCA. **Abbreviations**: ML, machine learning; AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; DCA, decision curve analysis.

Interpretation of the Model

To evaluate feature importance in the XGBoost model, we applied SHAP analysis, which quantifies each predictor's contribution using absolute mean SHAP values. This revealed the top five influential factors: two clinical parameters (maternal age and uterine curettage history) and three systemic inflammation markers (SII, PLR, and NLR) (Figure 6). The SHAP summary plot (Figure 6A) displays individual feature impacts across patients, with point colors indicating feature values (yellow = high, blue = low). The horizontal axis represents SHAP values, with feature importance visually emphasized by larger data point clusters - the wider the spread, the stronger the predictive influence on preterm birth risk in GDM patients. Complementing this, the importance bar chart (Figure 6B) ranks variables by their overall predictive power. Key features, in descending order of significance, were: SII, PLR, maternal age, NLR, uterine curettage history, monocyte count, and LMR. Notably, inflammatory indices dominated the top predictors, underscoring their clinical relevance.

The SHAP force plot (Figure 7) visualizes feature contributions to individual patient predictions, where yellow and red regions respectively indicate risk-increasing and protective factors for preterm birth in GDM patients, with region width reflecting effect magnitude. The output value f(x) aggregates all feature contributions (SHAP values) for a given patient, while the base value represents the population average prediction. The SHAP force plot demonstrates XGBoost's predictive accuracy through two representative cases: the upper panel (Figure 7A) correctly predicts preterm birth based on elevated SII and PLR values, advanced maternal age, and other contributing factors, while the lower panel (Figure 7B) accurately identifies a term birth case characterized by lower SII, younger maternal age, and additional protective indicators. This method enables precise differentiation between the risks of preterm birth and term birth, offering personalized risk assessments for GDM patients and facilitating early clinical intervention.

Sensitivity Analyses

Sensitivity analyses confirmed the robustness of our findings. After excluding participants with a history of hypertension, smoking, or alcohol use, the XGBoost model demonstrated performance consistent with the primary analysis in the combined clinical-inflammatory models (AUC-ROC=0.891, AUC-PRC=0.670, Figures 8A–D). Furthermore, the use of MICE for handling missing data, compared with complete-case analysis, did not materially affect model performance (AUC-ROC=0.893, AUC-PRC=0.620, Figures 9A–D).

Discussion

GDM, a prevalent metabolic disorder of pregnancy, significantly increases preterm birth risk and associated adverse outcomes despite current management strategies.[7,34,35] Developing reliable prediction methods for preterm birth in GDM patients is crucial to improve perinatal outcomes and reduce healthcare burdens. GDM pathogenesis is significantly associated with both established clinical predictors (such as maternal age ≥35 years and smoking status)[9,36] and inflammatory biomarkers,[11,12,22,37] and may enhance the modest predictive accuracy of current preterm birth models, which currently neglect some inflammatory biomarkers like PLR, LMR, SII.[13] While conventional linear models (eg,

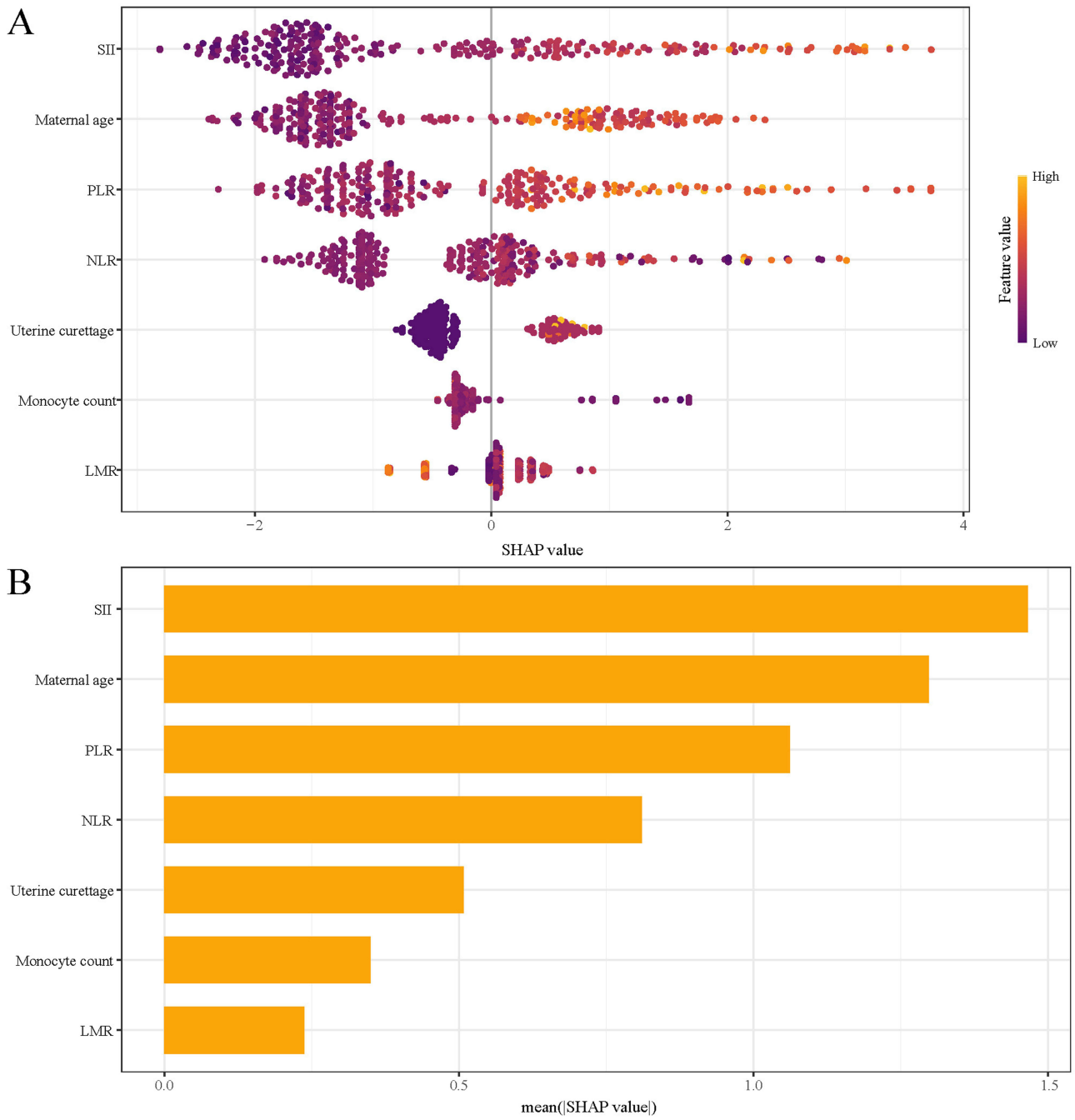

Figure 6 SHAP analysis of XGBoost for preterm birth prediction: (**A**) summary plot, (**B**) feature importance.
Abbreviations: SHAP, Shapley Additive Explanations; XGBoost, extreme gradient boosting; SII, systemic immune-inflammation index; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; LMR, lymphocyte-to-monocyte ratio.

Logit) struggle with the complex nonlinear relationships between multidimensional clinical data and patient outcomes, ML methods remain underexplored for preterm birth prediction in GDM despite their superior ability to decipher sophisticated patterns in combined clinical and systemic inflammation biomarker data. Our study addressed this critical gap by developing and comparing four ML algorithms that synergistically combined conventional clinical factors with systemic inflammatory markers. The XGBoost model demonstrated superior performance in predicting preterm birth among GDM patients by effectively integrating both data types. Through SHAP value analysis, we enhanced model interpretability and gained novel insights into the contribution of inflammatory pathways to preterm birth risk. Sensitivity

https://doi.org/10.2147/IJWH.S541610

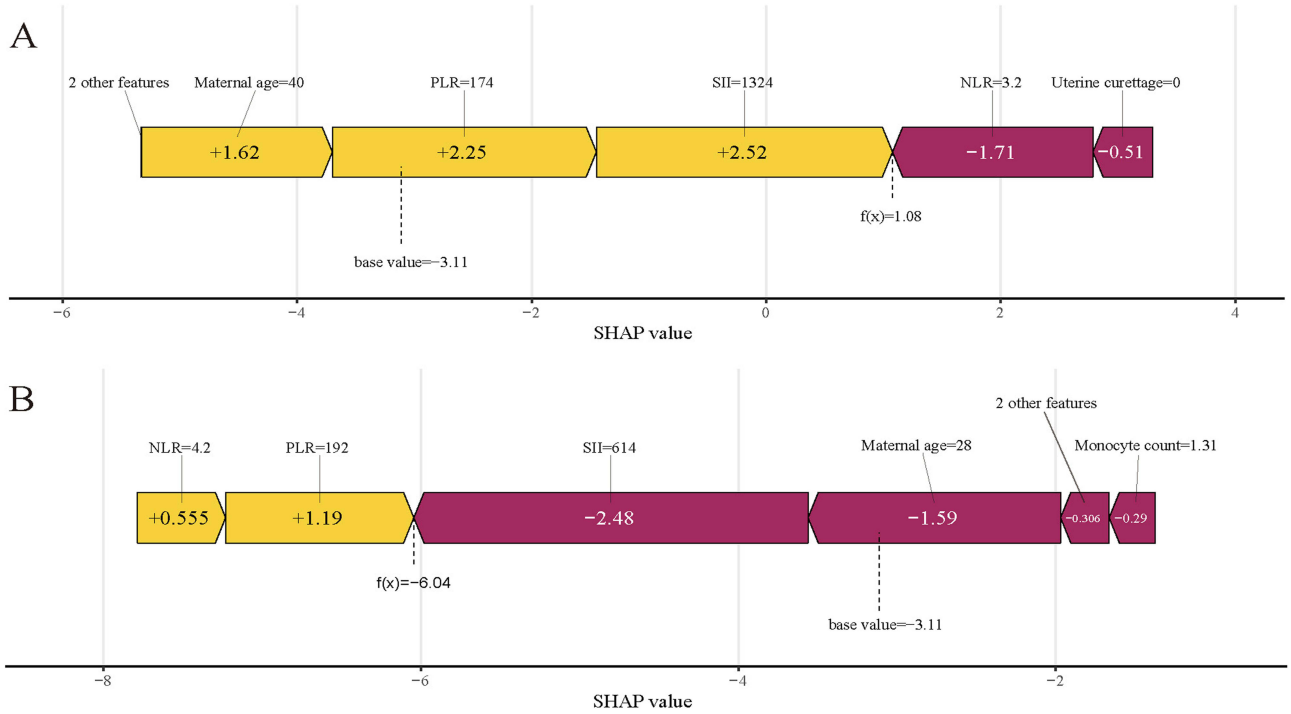

Figure 7 SHAP force plots for individual predictions: (**A**) preterm birth case (**B**) term birth case.
Abbreviations: SHAP, Shapley Additive Explanations; SII, systemic immune-inflammation index; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio.

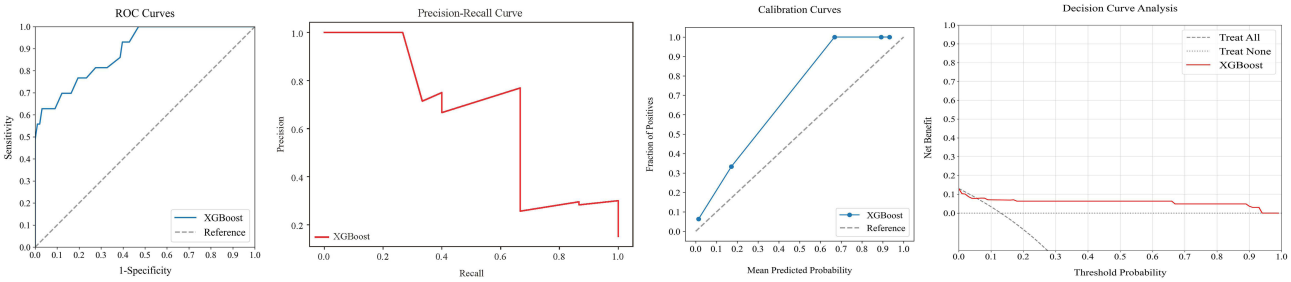

Figure 8 Sensitivity analyses were performed in participants without a history of hypertension, smoking, or alcohol use at baseline (N=330). The XGBoost model demonstrated performance consistent with the primary combined clinical-inflammatory analysis, as assessed by (**A**) ROC curve (AUC-ROC=0.891), (**B**) precision-recall curve (AUC-PRC=0.670), (**C**) calibration plots, and (**D**) DCA.
Abbreviations: AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; DCA, decision curve analysis; XGBoost, extreme gradient boosting.

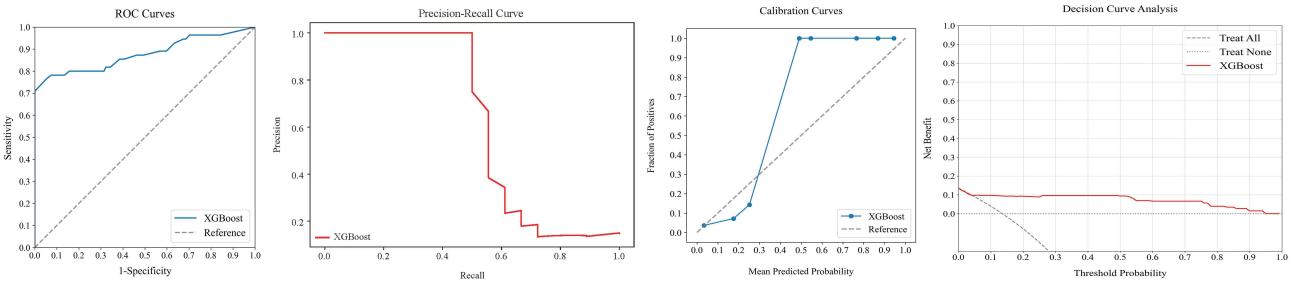

Figure 9 Sensitivity analyses were performed after imputation of missing data (N=405). Compared with the complete-case analysis, imputation did not materially affect the performance of the XGBoost model, as shown by (**A**) ROC curve (AUC-ROC=0.893), (**B**) precision-recall curve (AUC-PRC=0.620), (**C**) calibration plots, and (**D**) DCA.
Abbreviations: AUC-ROC, area under the receiver operating characteristic curve; AUC-PRC, area under the precision-recall curve; DCA, decision curve analysis; XGBoost, extreme gradient boosting.

https://doi.org/10.2147/IJWH.S541610

analyses further confirmed the reliability of these results. These findings advance the field by establishing an interpretable ML framework that combines traditional clinical and systemic inflammatory markers for early risk stratification, enabling more targeted interventions in GDM management.

Our study employed ML algorithms to overcome the limitations of conventional linear models in analyzing complex non-linear relationships.[38] We evaluated four ML models comparing three approaches: clinical indicators alone, systemic inflammation indices alone, and their combined integration. The integrated models achieved superior predictive performance for preterm birth, as their multidimensional analysis captures synergistic interactions between clinical and inflammatory factors that singular approaches miss. This comprehensive integration strategy enhanced prediction accuracy.

XGBoost emerged as the optimal model in our evaluation, demonstrating consistent high accuracy in temporal external validation through its integration of traditional clinical indicators with systemic inflammation indices. Li et al developed a nomogram for predicting preterm birth in pregnant women with GDM using clinical risk factors, but its performance (AUC = 0.722) was notably lower than that of our model (AUC = 0.932).[3] While prior research has primarily employed linear models using traditional clinical and inflammatory markers, our study advances the field by applying ML to analyze a broader spectrum of inflammatory biomarkers, such as PLR, LMR, and SII, significantly enhancing prediction accuracy for preterm birth in GDM patients (AUC=0.932 vs 0.885).[39] To improve interpretability of the complex ML model, we used SHAP analysis. The SHAP feature importance map visually shows each feature's impact on model outputs through SHAP values, representing their influence range and direction.[40] Each plotted point corresponds to a sample, with colored bars indicating feature values and their distribution. Left-positioned bars denote negative impacts, while right-positioned bars show positive effects.[41,42] This approach helps identify key features for model optimization and selection. The SHAP analysis identified five key preterm birth predictors: three systemic inflammation indices and two traditional clinical indicators. Our analysis identified elevated NLR, PLR, and SII levels as significant predictors of preterm birth in GDM patients. These composite inflammatory indices demonstrate greater clinical reliability than isolated hematological parameters (including neutrophil, lymphocyte, and monocyte counts) due to their stability against physiological fluctuations, pathological interferences, and technical measurement variations.[43,44] The SII demonstrates particular clinical value by quantitatively integrating the synergistic interactions among platelets, neutrophils, and lymphocytes.[45] This composite index may provide a more comprehensive assessment of inflammatory-immune crosstalk compared to isolated ratios like NLR or PLR. In addition to systemic inflammation indices, our study identified maternal age and uterine curettage as additional preterm birth risk factors. Global trends show increasing maternal age, a recognized independent risk factor for GDM. Meta-analyses indicate 2-fold and 4-fold higher GDM risks for women aged 35–39 and >40 years, respectively, compared to those <35 years.[46–50] Additionally, women aged 35–39 face 1.4% preterm birth risk.[51] Building on prior research, our study confirmed maternal age as an independent preterm birth risk factor in GDM patients, enabling better pregnancy outcome assessment. Uterine curettage, particularly for miscarriage or pregnancy termination, is an independent preterm birth risk factor due to potential cervical trauma.[52] This underscores the need to document termination procedures and evaluate their preterm birth risks. The SHAP-interpreted XGBoost model effectively identified and quantified key predictive factors for preterm birth in GDM patients, providing clinically valuable risk stratification. This ML approach facilitates timely preventive interventions, such as intensive monitoring, corticosteroid administration, or transfer to a higher-level care facility. While these measures improve outcomes, they also carry risks, including unnecessary treatments or hospitalizations for patients who do not deliver preterm.

While demonstrating the predictive potential of ML models combining traditional clinical indicators and systemic inflammation indices for preterm birth in GDM, this study has two main limitations. First, the single-center design with 389 participants may limit generalizability due to potential regional biases, the relatively small sample size, spectrum bias, and limited transportability. Second, the retrospective nature and exclusion of cases with incomplete records could introduce selection bias. These constraints notwithstanding, our findings underscore the clinical value of this integrated approach. Future multi-center prospective studies with larger cohorts are warranted to validate and optimize the model's performance, ideally including external or geographic validation from additional centers.

In summary, the XGBoost model demonstrates superior predictive performance for preterm birth in GDM patients by synergistically combining traditional clinical parameters with systemic inflammation indices. This integrated ML approach enables precise differentiation between the risks of preterm birth and term birth, offering personalized risk assessments for GDM patients. It supports clinical decision-making, facilitates early intervention, and has the potential to enhance perinatal outcomes.

Data Sharing Statement

All data supporting this study are included in the article, and additional inquiries can be addressed to either of the two corresponding authors.

Ethical Approval

This study received approval from the medical ethics committee of Shanghai University of Medicine & Health Sciences affiliated Zhoupu Hospital (Approval No. 2024-C-160-E01) and complied with the Declaration of Helsinki.

Informed Consent

The requirement for informed consent was waived by the Ethics Committee of Shanghai University of Medicine & Health Sciences affiliated Zhoupu Hospital due to the retrospective nature of the study.

Funding

Key Discipline Group Construction Project of Pudong New Area Health Commission (Grant No.: PWZxq2022-15).

Disclosure

The authors have no conflicts of interest to declare.

References

1. Sweeting A, Hannah W, Backman H, et al. Epidemiology and management of gestational diabetes. *Lancet*. 2024;404(10448):175–192. doi:10.1016/s0140-6736(24)00825-0
2. Durnwald C, Beck RW, Li Z, et al. Continuous glucose monitoring profiles in pregnancies with and without gestational diabetes mellitus. *Diabetes Care*. 2024;47(8):1333–1341. doi:10.2337/dc23-2149
3. Li H, Gao L, Yang X, Chen L. Development and validation of a risk prediction model for preterm birth in women with gestational diabetes mellitus. *Clin Endocrinol*. 2024;101(3):206–215. doi:10.1111/cen.15044
4. Gou W, Xiao C, Liang X, et al. Physical activity during pregnancy and preterm birth among women with gestational diabetes. *JAMA Network Open*. 2024;7(12):e2451799. doi:10.1001/jamanetworkopen.2024.51799
5. Caughey AB, Turrentine M. ACOG practice bulletin no. 190: gestational diabetes mellitus. *Obstet Gynecol*. 2018;131(2):e49–e64. doi:10.1097/aog.0000000000002501
6. American Diabetes Association. 14. Management of diabetes in pregnancy: standards of medical care in diabetes-2019. *Diabetes Care*. 2019;42(Suppl 1):S165–s172. doi:10.2337/dc19-S014
7. Teede HJ, Bailey C, Moran LJ, et al. Association of antenatal diet and physical activity-based interventions with gestational weight gain and pregnancy outcomes: a systematic review and meta-analysis. *JAMA Intern Med*. 2022;182(2):106–114. doi:10.1001/jamainternmed.2021.6373
8. Huang S, Tian J, Liu C, et al. Elevated C-reactive protein and complement C3 levels are associated with preterm birth: a nested case-control study in Chinese women. *BMC Pregnancy Childbirth*. 2020;20(1):131. doi:10.1186/s12884-020-2802-9
9. Ye W, Luo C, Huang J, et al. Gestational diabetes mellitus and adverse pregnancy outcomes: systematic review and meta-analysis. *Bmj*. 2022;377:e067946 doi:10.1136/bmj-2021-067946
10. Tagami K, Iwama N. Advanced maternal age is a risk factor for both early and late gestational diabetes mellitus: the Japan Environment and Children's Study. *J Diab Investigat*. 2025;16(4):735–743. doi:10.1111/jdi.14400
11. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes mellitus. *Int J Mol Sci*. 2018;19(11). doi:10.3390/ijms19113342
12. Cinkajzlová A, Anderlová K, Šimják P, et al. Subclinical inflammation and adipose tissue lymphocytes in pregnant females with gestational diabetes mellitus. *J Clin Endocrinol Metab*. 2020;105(11). doi:10.1210/clinem/dgaa528
13. Mehta-Lee SS, Palma A, Bernstein PS, Lounsbury D, Schlecht NF. A preconception nomogram to predict preterm delivery. *Matern Child Health J*. 2017;21(1):118–127. doi:10.1007/s10995-016-2100-3
14. Saucedo R, Ortega-Camarillo C, Ferreira-Hermosillo A, et al. Role of oxidative stress and inflammation in gestational diabetes mellitus. *Antioxidants*. 2023;12(10). doi:10.3390/antiox12101812
15. Pavlidis I, Stock SJ. Preterm birth therapies to target inflammation. *J Clin Pharmacol*. 2022;62(Suppl 1):S79–s93. doi:10.1002/jcph.2107
16. Parturition SR. Parturition. *N Engl J Med*. 2007;356(3):271–283. doi:10.1056/NEJMra061360

17. Romero R, Manogue KR, Mitchell MD, et al. Infection and labor. IV. Cachectin-tumor necrosis factor in the amniotic fluid of women with intraamniotic infection and preterm labor. *Am J Obstet Gynecol*. 1989;161(2):336–341. doi:10.1016/0002-9378(89)90515-2
18. Romero R, Brody DT, Oyarzun E, et al. Infection and labor. III. Interleukin-1: a signal for the onset of parturition. *Am J Obstet Gynecol*. 1989;160 (5):t1):1117–1123. doi:10.1016/0002-9378(89)90172-5
19. Romero R, Avila C, Santhanam U, Sehgal PB. Amniotic fluid interleukin 6 in preterm labor. Association with infection. *J Clin Invest*. 1990;85 (5):1392–1400. doi:10.1172/jci114583
20. Ghezzi F, Gomez R, Romero R, et al. Elevated interleukin-8 concentrations in amniotic fluid of mothers whose neonates subsequently develop bronchopulmonary dysplasia. *Eur J Obstet Gynecol Reprod Biol*. 1998;78(1):5–10. doi:10.1016/s0301-2115(97)00236-4
21. Ulugün F, Özdemir N. The change of systemic inflammation response index in the treatment of patients with myasthenia gravis undergoing thymectomy: a retrospective, follow-up study. *Turkish J Thorac Cardiovasc Surg*. 2023;31(4):547–555. doi:10.5606/tgkdc.dergisi.2023.24588
22. Bozbay N, Medinaeva A. The role of first-trimester systemic immune-inflammation index for the prediction of gestational diabetes mellitus. *Revista da Associação Médica Brasileira*. 2024;70(10):e20240532. doi:10.1590/1806-9282.20240532
23. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. *Behav Ther*. 2020;51(5):675–687. doi:10.1016/j.beth.2020.05.002
24. Radak M, Lafta HY, Fallahi H. Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies. *J Cancer Res Clin Oncol*. 2023;149(12):10473–10491. doi:10.1007/s00432-023-04956-z
25. Chen J, Yang L, Han J, et al. Interpretable machine learning models using peripheral immune cells to predict 90-day readmission or mortality in acute heart failure patients. *Clin Appl Thromb Hemost*. 2024;30:10760296241259784. doi:10.1177/10760296241259784
26. Yu M, Yuan Z, Li R, et al. Interpretable machine learning model to predict surgical difficulty in laparoscopic resection for rectal cancer. *Front Oncol*. 2024;14:1337219. doi:10.3389/fonc.2024.1337219
27. Chen PN, Lee CC, Liang CM, et al. General deep learning model for detecting diabetic retinopathy. *BMC Bioinf*. 2021;22(Suppl 5):84. doi:10.1186/s12859-021-04005-x
28. Wang K, Tian J. Improving risk identification of adverse outcomes in chronic heart failure using SMOTE+ENN and machine learning. *Risk Manage Healthcare Policy*. 2021;14:2453–2463. doi:10.2147/rmhp.s310295
29. Roth AE. Lloyd Shapley (1923-2016). *Nature*. 2016;532(7598):178. doi:10.1038/532178a
30. Li W, Song Y. Predictive model and risk analysis for diabetic retinopathy using machine learning: a retrospective cohort study in China. *BMJ Open*. 2021;11(11):e050989. doi:10.1136/bmjopen-2021-050989
31. Ogami C, Tsuji Y, Seki H, et al. An artificial neural network-pharmacokinetic model and its interpretation using Shapley additive explanations. *CPT Pharmacometrics Syst Pharmacol*. 2021;10(7):760–768. doi:10.1002/psp4.12643
32. Zheng P, Yu Z, Li L, et al. Predicting blood concentration of tacrolimus in patients with autoimmune diseases using machine learning techniques based on real-world evidence. *Front Pharmacol*. 2021;12:727245. doi:10.3389/fphar.2021.727245
33. Liu H, Li J, Leng J. Machine learning risk score for prediction of gestational diabetes in early pregnancy in Tianjin, China. *Diabet Metabol Res Rev*. 2021;37(5):e3397. doi:10.1002/dmrr.3397
34. Xu H, Liu R. Comprehensive management of gestational diabetes mellitus: practical efficacy of exercise therapy and sustained intervention strategies. *Front Endocrinol*. 2024;15:1347754. doi:10.3389/fendo.2024.1347754
35. Huang S, Guo Y, Xu X, Jiang L, Yan J. Gestational diabetes complicated with preterm birth: a retrospective cohort study. *BMC Pregnancy Childbirth*. 2024;24(1):631. doi:10.1186/s12884-024-06810-7
36. Schummers L, Hutcheon JA, Hacker MR, et al. Absolute risks of obstetric outcomes by maternal age at first birth: a population-based cohort. *Epidemiology*. 2018;29(3):379–387. doi:10.1097/ede.0000000000000818
37. Chen B, Chen X, Hu R, et al. Alternative polyadenylation regulates the translation of metabolic and inflammation-related proteins in adipose tissue of gestational diabetes mellitus. *Comput Struct Biotechnol J*. 2024;23:1298–1310. doi:10.1016/j.csbj.2024.03.013
38. Ramaswamy SM, Kuizenga MH, Weerink MAS, et al. Frontal electroencephalogram based drug, sex, and age independent sedation level prediction using non-linear machine learning algorithms. *J Clin Monit Comput*. 2022;36(1):121–130. doi:10.1007/s10877-020-00627-3
39. Huang Y, Cai F, Zhang W, Shen R, Jin L. Development and validation of nomogram for the prediction of preterm delivery based on patient characteristics and circulating inflammatory cells in patients with gestational diabetes mellitus. *Ann Transl Med*. 2023;11(2):70. doi:10.21037/atm-22-6223
40. Zhang J, Niu W, Yang Y, Hou D, Dong B. Machine learning prediction models for compressive strength of calcined sludge-cement composites. *Constr Build Mater*. 2022;346:128442.
41. Nohara Y, Matsumoto K, Soejima H, Nakashima N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput Methods Programs Biomed*. 2022;214:106584. doi:10.1016/j.cmpb.2021.106584
42. Xue B, Li D, Lu C, et al. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Network Open*. 2021;4(3):e212240. doi:10.1001/jamanetworkopen.2021.2240
43. Schiefer S, Wirsik NM, Kalkum E, Seide SE, Nienhüser H. Systematic review of prognostic role of blood cell ratios in patients with gastric cancer undergoing surgery. *Diagnostics*. 2022;12(3). doi:10.3390/diagnostics12030593
44. Almășan O, Leucuța DC. Blood cell count inflammatory markers as prognostic indicators of periodontitis: a systematic review and meta-analysis. *J Personal Med*. 2022;12(6). doi:10.3390/jpm12060992
45. Wu X, Wang H, Xie G, Lin S, Ji C. Increased systemic immune-inflammation index can predict respiratory failure in patients with Guillain-Barré syndrome. *Neurological Sci*. 2022;43(2):1223–1231. doi:10.1007/s10072-021-05420-x
46. Lean SC, Derricott H, Jones RL, Heazell AEP. Advanced maternal age and adverse pregnancy outcomes: a systematic review and meta-analysis. *PLoS One*. 2017;12(10):e0186287. doi:10.1371/journal.pone.0186287
47. Khalil A, Syngelaki A, Maiz N, Zinevich Y, Nicolaides KH. Maternal age and adverse pregnancy outcome: a cohort study. *Ultrasound Obstet Gynecol*. 2013;42(6):634–643. doi:10.1002/uog.12494
48. Cleary-Goldman J, Malone FD, Vidaver J, et al. Impact of maternal age on obstetric outcome. *Obstet Gynecol*. 2005;105(5). doi:10.1097/01.aog.0000158118.75532.51
49. Jacobsson B, Ladfors L, Milsom I. Advanced maternal age and adverse perinatal outcome. *Obstet Gynecol*. 2004;104(4):727–733. doi:10.1097/01.AOG.0000140682.63746.be

50. Waldenström U, Aasheim V, Nilsen ABV, et al. Adverse pregnancy outcomes related to advanced maternal age compared with smoking and being overweight. *Obstet Gynecol*. 2014;123(1):104–112. doi:10.1097/aog.0000000000000062
51. Frederiksen LE, Ernst A, Brix N, et al. Risk of adverse pregnancy outcomes at advanced maternal age. *Obstet Gynecol*. 2018;131(3):457–463. doi:10.1097/aog.0000000000002504
52. Jiang L, Peng L, Rong M, et al. Nomogram incorporating multimodal transvaginal ultrasound assessment at 20 to 24 weeks' gestation for predicting spontaneous preterm delivery in low-risk women. *Int J Women's Health*. 2022;14:323–331. doi:10.2147/ijwh.s356167