



OPEN

Early prediction of postpartum dyslipidemia in gestational diabetes using machine learning models

Zhifa Jiang^{1,3}, Xiekun Chen^{2,3}, Yuhang Lai², Jingwen Liu¹, Xiangyun Ye¹, Ping Chen¹ & Zhen Zhang²✉

This study addresses a gap in research on predictive models for postpartum dyslipidemia in women with gestational diabetes mellitus (GDM). The goal was to develop a machine learning-based model to predict postpartum dyslipidemia using early pregnancy clinical data, and the model's robustness was evaluated through both internal and temporal validation. Clinical data from 15,946 pregnant women were utilized. After cleaning, the data were divided into two sets: Dataset A (n = 1,116), used for training and evaluating the model, and Dataset B (n = 707), used for temporal validation. Several machine learning algorithms were applied, and the performance of the model was assessed with Dataset A, while Dataset B was used to validate the model across a different time period. Feature significance was evaluated through Information Value (IV), model importance analysis, and SHAP (SHapley Additive exPlanations) analysis. The results showed that among the five machine learning algorithms tested, tree-based ensemble models, such as XGBoost, LightGBM, and Random Forest, outperformed others in predicting postpartum dyslipidemia. In Dataset A, these models achieved accuracies of 70.54%, 70.54%, and 69.64%, respectively, with AUC-ROC values of 73.10%, 71.94%, and 76.14%. Temporal validation with Dataset B indicated that XGBoost performed best, achieving an accuracy of 81.05% and an AUC-ROC of 87.92%. The predictive power of the model was strengthened by key variables such as total cholesterol, fasting glucose, triglycerides, and BMI, with total cholesterol being identified as the most important feature. Further IV and SHAP analyses confirmed the pivotal role of these variables in predicting dyslipidemia. The study concluded that the XGBoost-based predictive model for postpartum dyslipidemia in GDM showed strong and consistent performance in both internal and temporal validations. By introducing new variables, the model can identify high-risk groups during early pregnancy, supporting early intervention and potentially improving pregnancy outcomes and reducing complications.

Keywords Machine learning, Ensemble learning algorithms, Diabetes, Gestational, Postpartum period, Dyslipidemias

Gestational diabetes mellitus (GDM) is the most prevalent metabolic disorder during pregnancy, significantly increasing the incidence of preeclampsia, cesarean section, preterm birth, and macrosomia. Furthermore, it elevates the long-term risk of cardiovascular diseases for both the mother and her offspring, severely impacting pregnancy outcomes and maternal and child health^{1,2}. Offspring of mothers with GDM are at a higher risk of impaired glucose tolerance, obesity, and diabetes in both childhood and adulthood³. The pathogenesis of GDM is heterogeneous, typically involving insulin resistance and insufficient compensatory insulin secretion from pancreatic β -cells⁴. GDM affects both glucose and lipid metabolism⁵.

Dyslipidemia refers to elevated serum levels of cholesterol and/or triglycerides and includes hypercholesterolemia, hypertriglyceridemia, and mixed hyperlipidemia⁶. Type 2 diabetes mellitus (T2DM), cardiovascular disease (CVD), and non-alcoholic fatty liver disease (NAFLD) are significant long-term complications of GDM. Atherosclerotic cardiovascular disease (ASCVD) is the leading cause of mortality among urban and rural residents in China, and the burden of ASCVD continues to rise⁷. Epidemiological, genetic, and

¹Obstetrics and Gynaecology, Huizhou First Maternal and Child Health Care Hospital, Huizhou 516000, Guangdong, China. ²School of Computer Science and Engineering, Huizhou University, Huizhou 516000, Guangdong, China.

³These authors contributed equally: Zhifa Jiang and Xiekun Chen. ✉email: zzsjbme@sjtu.edu.cn

clinical intervention studies have consistently demonstrated that dyslipidemia is a pathogenic risk factor for ASCVD, and diabetes is an important independent risk factor for ASCVD⁸. Therefore, delaying the progression of GDM-related complications to T2DM, CVD, and even ASCVD is of critical importance.

Insulin resistance is closely related to various diseases, including cardiovascular diseases, diabetes, dyslipidemia, and non-alcoholic fatty liver disease^{9–11}. During dyslipidemia, excessive fat accumulation leads to abnormal expansion of white adipose tissue, with hypertrophy and hyperplasia of adipocytes¹². Activation of the pathway results in impaired insulin signaling leading to insulin resistance, which in turn leads to impaired insulin-induced inhibition of white adipose tissue lipolysis^{13,14}. Recent HAPO-FUS studies have shown that maternal hyperglycemia has long-term metabolic impacts on mothers and their offspring^{15,16}. Depending on the severity of maternal hyperglycemia and related vascular risk factors, GDM complications may be more severe. Pregnancy has been demonstrated to result in physiologic lipid elevations, with severe hypertriglyceridemia (HTG) having the potential to induce acute pancreatitis and a maternal mortality rate of up to 20%¹⁷.

Postpartum dyslipidemia is a significant health issue for mothers with GDM¹⁸. Recent studies have confirmed the metabolic trajectory of dyslipidemia following glucose metabolism disorders during pregnancy. Lowe et al. tracked mothers with GDM and found that women with a history of GDM had significantly higher levels of total cholesterol and LDL-C 10–14 years postpartum compared to controls, even after adjusting for traditional cardiovascular and diabetes risk factors¹⁹. Wang et al. studied the lipid profiles of GDM mothers 6–9 weeks and 6 months postpartum, showing that serum triglycerides and LDL-C levels were significantly elevated at both time points, unrelated to obesity, and associated with higher postpartum glucose levels²⁰. Insulin resistance and hyperglycemia during pregnancy promote dyslipidemia, which often persists postpartum in women with prior GDM. These longitudinal and cross-sectional studies demonstrate that postpartum dyslipidemia in mothers with GDM persists, significantly impacting future health, thereby highlighting the importance of early identification and intervention through predictive modeling.

Researchers have recently sought to develop predictive models to identify which mothers with GDM are most likely to experience postpartum dyslipidemia. These models typically include lipid metabolites and clinical indicators during and after pregnancy, such as glucose control levels, gestational weight gain, insulin resistance index, and lipid levels. Pei et al. retrospectively examined the incidence and potential predictors of early postpartum dyslipidemia in mothers with GDM, finding that maternal age, systolic blood pressure (SBP), glycated hemoglobin (HbA1c), and LDL-C were independent risk factors, with an AUC of 0.733 when LDL-C was combined with age, SBP, and HbA1c¹⁸. Mi et al. employed lipidomic to analyze the serum of GDM cohorts during the postpartum period. They followed up for 8 years, identifying 107 lipids significantly associated with future diabetes and establishing a panel of 11 lipids with a predictive AUC of 0.739²¹. Although various predictive models have been established using different parameters, methodological flaws remain, with most models lacking temporal validation.

With the rapid development of machine learning technology, more artificial intelligence methods can be used to construct predictive models and validate them from multiple perspectives. This study aims to identify and analyze high-risk factors for postpartum dyslipidemia in GDM using machine learning methods and to construct and validate a predictive model.

Methods and materials

Dataset description

The study cohort comprised singleton pregnant women who received antenatal care and delivered at Huizhou First Maternal and Child Health Hospital in Guangdong Province, China. The participants' ages ranged from 18 to 45 years. These women underwent standard prenatal examinations and had blood tests for glucose and lipids at various stages of pregnancy, with delivery dates spanning from October 2021 to December 2023. Exclusion criteria included a history of diabetes, impaired fasting glucose, prior drug treatment for dyslipidemia, or a miscarriage or elective termination of pregnancy before 28 weeks of gestation.

The study variables primarily encompassed demographic characteristics such as maternal age, gravidity, parity, pre-pregnancy body mass index (BMI), and gestational weight gain. Blood test results during pregnancy and postpartum were collected, including complete blood count, coagulation function, lipid profile, HbA1c, and oral glucose tolerance tests (OGTT) at 24–28 weeks of gestation and postpartum.

Following the Standards of Care in Diabetes—2024, all participants underwent a 75g OGTT at 24–28 weeks of gestation to diagnose GDM, utilizing the International Association of Diabetes and Pregnancy Study Groups Consensus Panel diagnostic criteria. The criteria included fasting glucose ≥ 5.1 mmol/L, 1-h post-glucose ≥ 10.0 mmol/L, or 2-h post-glucose ≥ 8.5 mmol/L; meeting any one of these criteria confirmed a diagnosis of GDM²².

Postpartum dyslipidemia was defined according to the National Cholesterol Education Program Adult Treatment Panel III guidelines²³: total cholesterol (TC) ≥ 6.22 mmol/L, triglycerides (TG) ≥ 2.26 mmol/L, low-density lipoprotein cholesterol (LDL-C) ≥ 4.14 mmol/L, and high-density lipoprotein cholesterol (HDL-C) ≤ 1.04 mmol/L.

This study was approved by the Medical Ethics Committee of Huizhou First Maternal and Child Health Hospital (Ethics Approval No. 20240328A14), and all participants provided written informed consent.

Data preprocessing

This retrospective study gathered two batches of raw data: 10,283 samples collected from October 2021 to December 2022 for model development and training, and 5,663 samples collected from January 2023 to December 2023 for testing model reliability and generalizability. No data overlap between the two batches of raw data. Duplicate samples were removed, and missing data in the remaining samples were addressed. Samples with more than 30% missing data were deemed invalid and excluded. For the retained samples, missing values

were imputed using the mean value. The processed data were then filtered to identify GDM cases based on mid-pregnancy oral glucose tolerance test results.

To further investigate the relationship between study variables and postpartum dyslipidemia, new variables were calculated using early pregnancy clinical indicators: Platelet/Lymphocytes, Neutrophils/Platelet, Urea/Creatinine, Alanine Aminotransferase/Aspartate Aminotransferase (ALT/AST), and the Triglycerides Glucose Index (TyG Index). The variables were derived through ratio calculations. Additionally, based on clinical expertise, triglycerides, total cholesterol, fasting glucose, and pre-pregnancy BMI were utilized to calculate three new study variables:

$$\text{TyG Index} = \ln[\text{Fasting triglyceride (mg/dL)} \times \text{Fasting glucose (mg/dL)} / 2], \quad (1)$$

$$\text{Feature1} = \text{Exp}[(\text{Triglycerides} + \text{Total Cholesterol} + \text{Fasting Glucose}) / \text{BMI}], \quad (2)$$

$$\text{Feature2} = \ln(\text{Triglycerides} + \text{Total Cholesterol} + \text{Fasting Glucose}), \quad (3)$$

$$\text{Feature3} = \ln(\text{Total Cholesterol}) \times \text{TyG Index}. \quad (4)$$

Figure 1 presents the data processing workflow employed in this study.

Prediction model

Five machine learning algorithms were employed to develop a predictive model for postpartum dyslipidemia in GDM: Support Vector Machine (SVM), random forest, decision tree, Light Gradient Boosting Machine (LightGBM), and extreme gradient boosting (XGBoost). Random forest mitigates overfitting by incorporating randomness during training, while both LightGBM and XGBoost utilize gradient boosting techniques²⁴. Prior research indicates that LightGBM and XGBoost significantly outperform others in predictive accuracy, stability, and generalization ability²⁵. The models were implemented in Python using the Scikit-Learn, XGBoost, and LGBM libraries.

Model evaluation

Dataset A was divided into training and test sets in a 9:1 ratio, with dataset B utilized for temporal validation. The models were rigorously evaluated on the test set of Dataset A, using AUC-PR (area under the Precision-Recall curve) and AUC-ROC (area under the Receiver Operating Characteristic curve) to comprehensively assess overall performance. Additionally, metrics such as Accuracy, Precision, Recall, F1_score, Specificity, Sensitivity, Negative Predictive Value (NPV), and Positive Predictive Value (PPV) were calculated to evaluate predictive performance. PPV denotes the proportion of true positive samples among all predicted positive samples, reflecting the accuracy of positive predictions, while NPV indicates the accuracy of negative predictions.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}), \quad (5)$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}), \quad (6)$$

$$\text{Recall} = (\text{Sensitivity} + \text{Specificity}) / 2, \quad (7)$$

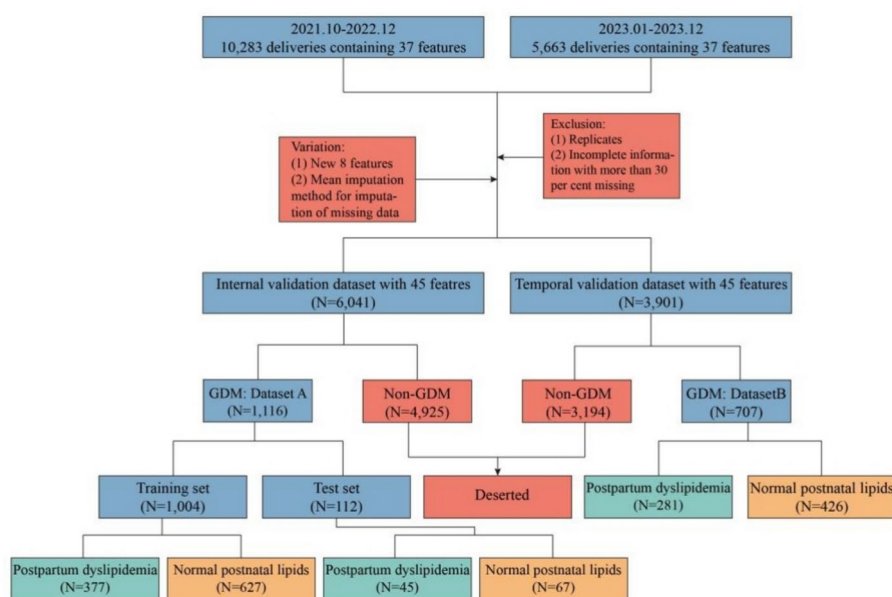


Fig. 1. Process of data processing in this study.

$$F1_score = 2 \times Precision \times Recall / (Precision + Recall), \tag{8}$$

$$Specificity = TN / (TN + FP), \tag{9}$$

$$Sensitivity = TP / (TP + FN), \tag{10}$$

$$Negative\ Predictive\ Value = TN / (TN + FN). \tag{11}$$

$$Positive\ Predictive\ Value = TP / (TP + FP), \tag{12}$$

Furthermore, we will use the model after training and testing on Dataset A to assess stability of the training results by performing a ten-fold cross-validation on the entire Dataset A. In ten-fold cross-validation, the dataset is partitioned into ten subsets, with each subset used once as the validation set and the remaining nine subsets serving as the training set. This process is repeated ten times with different validation sets, ensuring every sample is validated once. This method provides a more reliable and robust performance evaluation, minimizing the impact of a single data split.

Feature analysis

Correlation plots visually reveal potential relationships and trends among variables within a dataset. Before analyzing individual variables, it is essential to explore and understand their interrelationships through correlation plots, thereby forming reasonable hypotheses about these connections²⁶. This preliminary exploration enhances our understanding of the data and provides a solid, reliable foundation for subsequent data analysis and decision-making. Nonetheless, the utilization of correlation plots in this study is intended to elucidate the correlation between the objectives and the study characteristics, thereby facilitating an analysis of the mathematical and statistical validity of our proposed hypothesis.

IV (Information Value) is an important metric used to measure the ability of a feature to predict a target variable²⁷. IV helps identify which features are important for predicting the target variable²⁸. The calculation of IV values is not contingent upon the ordering of the variables and can be utilized to assess the predictive power of different variables. By calculating the IV value, we can assess the extent to which each feature explains the target variable. The higher the IV value, the closer the relationship between the feature and the target variable. An IV value of less than 0.02 is considered to be of limited utility for prediction; 0.02 to 0.1 is weak predictive power; 0.1 to 0.3 is moderate predictive power; 0.3 to 0.5 is strong predictive power; and greater than 0.5 is questionable predictive power and may require further examination²⁹.

SHAP (SHapley Additive exPlanations) is a game-theoretic approach to explaining the output of any machine learning model by assigning each feature an importance value for a given prediction, ensuring interpretability while maintaining accuracy³⁰. SHAP helps in understanding how the model makes predictions by identifying which features contribute most to the prediction and how they influence the prediction results²⁴. Additionally, it identifies the most important features for the model's predictions. By understanding the features that contribute most to the model's predictions, we can better grasp the essence of the problem and extract key features for further analysis³¹.

Results

Participant characteristics

Dataset A, utilized for model development, comprised 1116 samples, with 422 cases (37.8%) of postpartum dyslipidemia. This prevalence rate aligns closely with findings from Pei et al.¹⁸. Dataset B, employed for temporal validation, included 707 samples, with 281 cases (39.7%) of postpartum dyslipidemia.

Model development and evaluation

Table 1 presents the performance of the models evaluated on the independent test set of Dataset A. Tree-based ensemble learning algorithms, specifically XGBoost, LightGBM, and Random Forest, exhibited superior performance in predicting postpartum dyslipidemia in GDM patients, with accuracies of 70.54%, 70.54%, and 69.64%, and AUC-ROC values of 0.7310, 0.7194, and 0.7614, respectively. This demonstrates that ensemble

	SVM	Decision tree	Random forest	LightGBM	XGBoost
Accuracy	65.18%	66.96%	69.64%	70.54%	70.54%
Precision	65.78%	66.41%	68.78%	70.83%	71.42%
Recall	58.86%	62.17%	66.24%	66.25%	65.89%
F1_score	56.94%	61.85%	66.56%	66.47%	65.98%
Specificity	91.04%	86.57%	83.58%	88.06%	89.55%
Sensitivity	26.67%	37.78%	48.89%	44.44%	42.22%
NPV	64.89%	67.44%	70.89%	70.24%	69.77%
PPV	66.67%	65.38%	66.67%	71.43%	73.08%
AUC-ROC	0.5886	0.6285	0.7614	0.7194	0.7310
AUC-PR	0.4724	0.5833	0.6839	0.6324	0.6772

Table 1. Results of evaluating model performance using an independent test set of Dataset A.

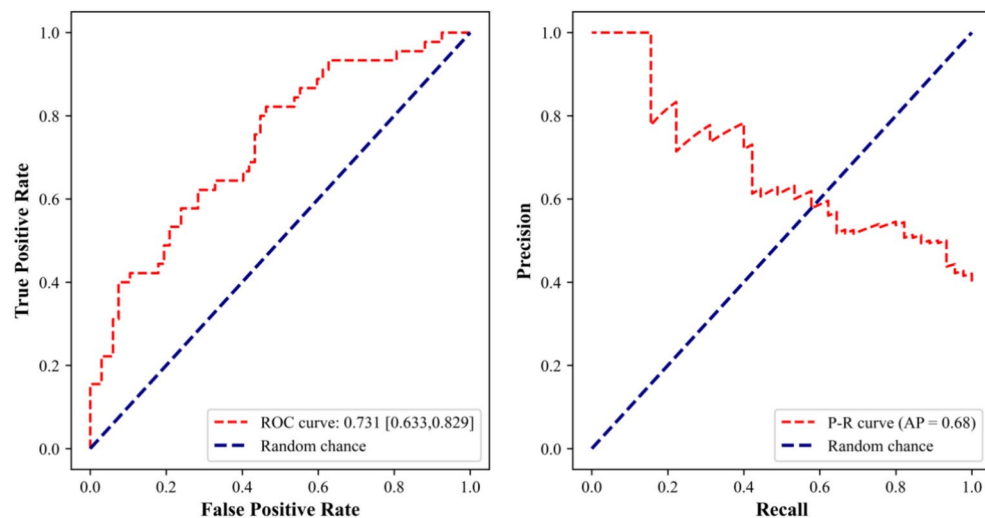


Fig. 2. ROC curve and P-R curve for XGBoost model.

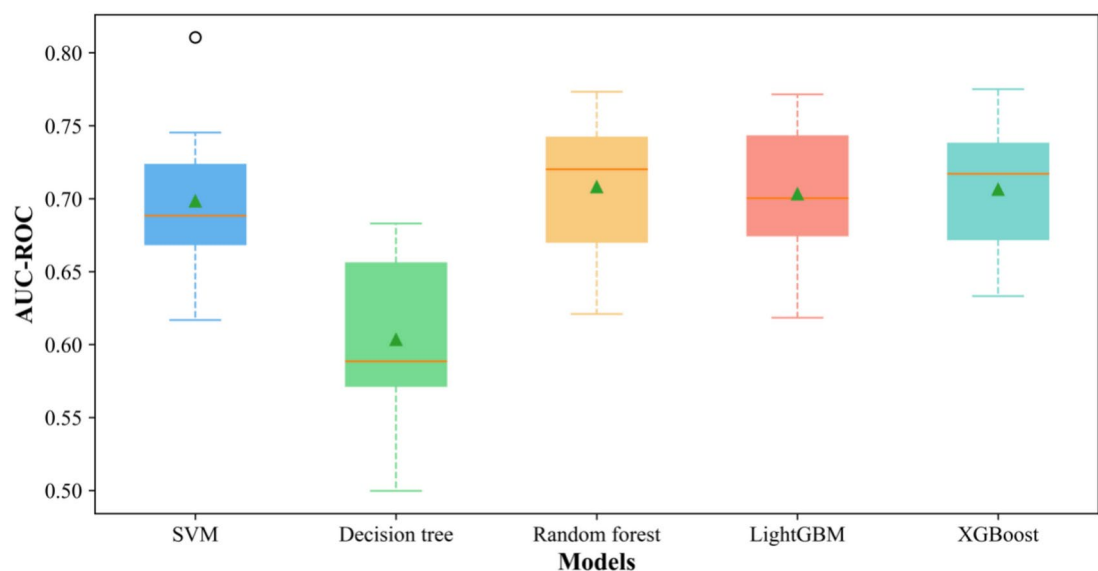


Fig. 3. AUC-ROC distribution of the model under tenfold cross-validation.

learning models outperform single decision trees and SVM models. The results preliminarily indicate that the XGBoost model achieves robust overall performance. Figure 2 depicts the ROC curve and P-R curve of the XGBoost model.

Figure 3 illustrates the AUC-ROC distribution of the models across tenfold cross-validation on Dataset A, demonstrating that XGBoost, LightGBM, and Random Forest consistently outperformed other algorithms. These results, derived from the entire Dataset A, underscore the superior performance of ensemble methods regarding AUC-ROC.

The models were then applied to the entire Dataset B for temporal validation to further evaluate their predictive performance. Table 2 showcases that the XGBoost model exhibited the best performance metrics, achieving an accuracy of 81.05% and an AUC-ROC of 0.8792, both the highest among all models evaluated. Moreover, we observed that the results of temporal validation were generally superior to those from the evaluation on the independent test set of Dataset A. This finding suggests that the evaluation on the test set of Dataset A may not fully demonstrate the model's performance due to the limited number of test samples.

In Dataset B, we employed a bootstrapping method without replacement, ensuring each sample accounted for 40% of the total dataset size in every iteration. This approach aimed to provide an objective and fair evaluation of the models. The entire evaluation process was repeated 1000 times, each time using a different random seed to ensure the stability and reliability of the evaluation results. Figure 4 presents a boxplot visualization of the evaluation results from the 1000 bootstrap samples. It demonstrates that the performance metrics of

	SVM	Decision tree	Random forest	LightGBM	XGBoost
Accuracy	65.21%	70.16%	79.77%	76.66%	81.05%
Precision	65.81%	68.79%	81.45%	76.94%	81.83%
Recall	58.35%	68.57%	76.31%	73.49%	78.28%
F1_score	56.17%	68.67%	77.38%	74.26%	79.23%
Specificity	91.78%	76.29%	93.19%	88.97%	91.78%
Sensitivity	24.91%	60.85%	59.43%	58.01%	64.77%
NPV	64.95%	74.71%	77.69%	76.26%	79.80%
PPV	66.67%	62.87%	85.20%	77.62%	83.87%
AUC-ROC	0.5835	0.7153	0.8656	0.8249	0.8792
AUC-PR	0.4645	0.6152	0.8277	0.7762	0.8453

Table 2. Results of evaluating model performance using dataset B.

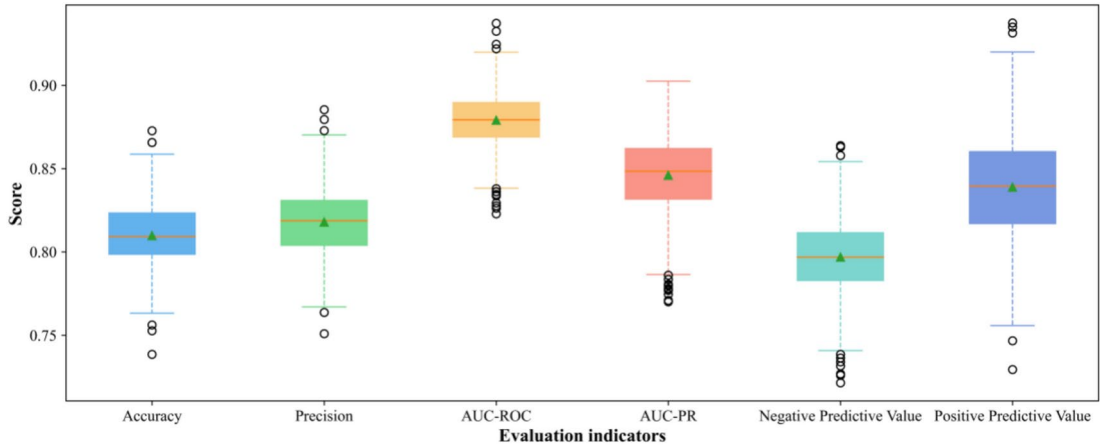


Fig. 4. Boxplot of model evaluation metrics validated by 1000 random samples.

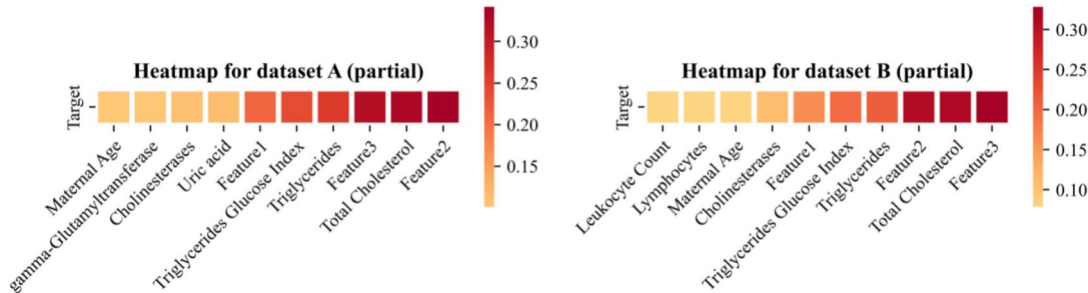


Fig. 5. Correlation Heatmap of Features with Target Variable.

the XGBoost model consistently fell within the 80%-90% range, indicating robust overall performance across multiple evaluations.

Feature importance analysis

Top 10 features

Figure 5 illustrates the top 10 variables most strongly correlated with postpartum dyslipidemia in both Dataset A and Dataset B. Notably, Feature2, total cholesterol, and Feature3 exhibit the highest correlations with the target variable, indicating their potential value in predicting postpartum dyslipidemia.

IV analysis

To further validate the importance of Feature2, total cholesterol, and Feature3 in predicting postpartum dyslipidemia, we calculated the IV values for all features. Table 3 lists the IV values for features in Dataset A,

Features	IV Score	Features	IV Score
Feature2	0.825	Thrombin time	0.130
Feature1	0.775	Gamma-glutamyltransferase	0.125
Feature3	0.752	Neutrophils	0.121
Total cholesterol	0.625	Total bile acid	0.121
Triglycerides glucose index	0.502	Total bilirubin	0.120
Triglycerides	0.464	Alanine aminotransferase	0.118
Free thyroxine	0.244	Urea/creatinine	0.117
Free triiodothyronine	0.235	ALT/AST	0.116
Partial thromboplastin time	0.221	Urea	0.115
Uric acid	0.211	Fasting blood glucose	0.114
Pre-pregnancy BMI	0.209	Thyrotropin	0.113
Total protein	0.201	Maternal age	0.113
Platelet count	0.194	Erythrocyte count	0.109
Cholinesterases	0.191	Placental weight	0.107
Birth weight	0.165	Creatinine	0.093
Lymphocytes	0.165	Aspartate Aminotransferase	0.091
Neutrophils/platelet	0.156	Albumins	0.091
Alkaline phosphatase	0.146	Fibrinogen	0.082
Hematocrit	0.144	Hemoglobins	0.063
Leukocyte count	0.144	Monocytes	0.060
Platelet/lymphocytes	0.141	Gestation	0.049
Gestational weight gain	0.138	Parity	0.045
Direct bilirubin	0.131		

Table 3. IV scores between Features and target in Dataset A.

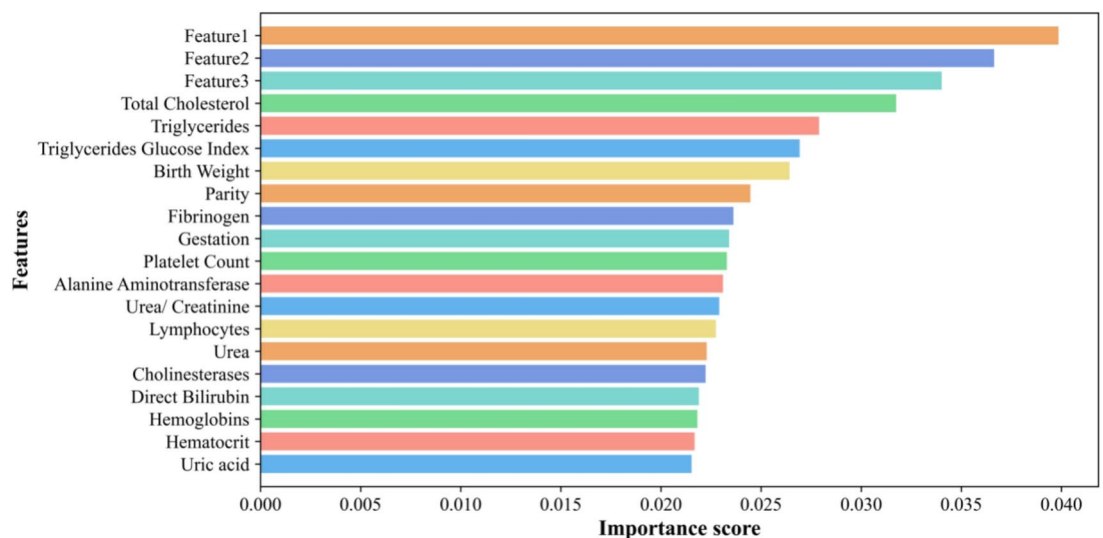


Fig. 6. Feature Importance Analysis of XGBoost.

sorted in descending order. The results indicate that the IV values for these features exceed 0.5, highlighting their superior predictive capability compared to other variables.

XGBoost feature importance

Using the XGBoost model, which demonstrated optimal predictive performance in our evaluations, we analyzed the influence of each feature on the prediction outcomes. Figure 6 displays a bar chart of feature importance from the XGBoost model. The findings corroborate that Feature1, Feature2, Feature3, and total cholesterol are critical in predicting postpartum dyslipidemia and should be prioritized.

SHAP analysis

SHAP values were computed to understand the marginal contribution of each feature in the trained XGBoost and LightGBM models on Dataset A. Figure 7 shows the global importance plot with scatter points for the top 20 features. Total cholesterol emerges as the most globally important, followed by Feature1, Feature2, and Feature3 in the SHAP values ranking. Moreover, it was notable to find that the feature values of the aforementioned features exhibit a positive correlation with their influence on the model's predicted value. Specifically, higher feature values in total cholesterol, Feature1, Feature2, and Feature3 are correlated with positive predicted values and positive SHAP values, and this positive impact is particularly pronounced.

Figure 8 shows the global importance plot of the positive class scatter for the top 20 features. Similar to the results shown in Fig. 7, we find that Feature1, Total Cholesterol, Feature3 and Feature2 were ranked in the top four SHAP values. The difference between Figs. 7 and 8 is that Feature1 is considered the most important feature in the LightGBM model, while total cholesterol is ranked second. However, the feature values corresponding to the four features still demonstrate a positive correlation with the model's predictive outputs. This result further supports our hypothesis.

In Fig. 9, the construction processes of prediction results between the LightGBM and XGBoost models are compared using waterfall plots generated from SHAP model outputs. It was observed that in the construction process of the prediction results for both models, Feature1 and Feature2 positively contributed to the prediction outcomes. In the XGBoost model, total cholesterol made the largest positive contribution, followed by Feature1, Feature2, and Feature3. This indicates that total cholesterol is the main driving factor for the prediction results in the XGBoost model.

Discussion

Clinical implications

Genetic factors, environmental changes, and lifestyle choices all play roles in the development of GDM^{32,33}. Observational studies have shown that in GDM, dysregulation of lipid levels, including elevated plasma triglyceride and fatty acid concentrations, decreased HDL cholesterol, and smaller LDL particles, leads to insulin resistance, which is of course one of the pathogenic mechanisms of GDM³⁴. This is characterized by impaired insulin-mediated glucose processing and inhibition of nitric oxide production, with insulin signaling playing a key role in activating nitric oxide. Impaired insulin signaling not only inhibits glucose metabolism but also stimulates the MAPK pathway, directly promoting atherosclerosis³⁵. Epigenetic studies showed that MTNR1B, TCF7L2, and IRS1 were significantly associated with GDM, supporting that insulin resistance and impaired insulin secretion play an important role in the pathogenesis of GDM and type 2 diabetes, and further subgroup analyses showed a significant role for two genes, TCF7L2 and PPARG³⁶. Currently, postpartum follow-up models for GDM in China include 42-day postpartum outpatient visits, community postpartum visits, and postpartum self-management. While these approaches have contributed to improving maternal and infant health, certain challenges remain, such as insufficient patient awareness, underestimation of the condition, inconvenient access to medical care, and the lack of timely reminders from healthcare workers, all contributing to a high rate of loss to follow-up. Therefore, early prediction and stratified management are particularly critical. The predictive model developed in this study is undoubtedly timely and highly relevant.

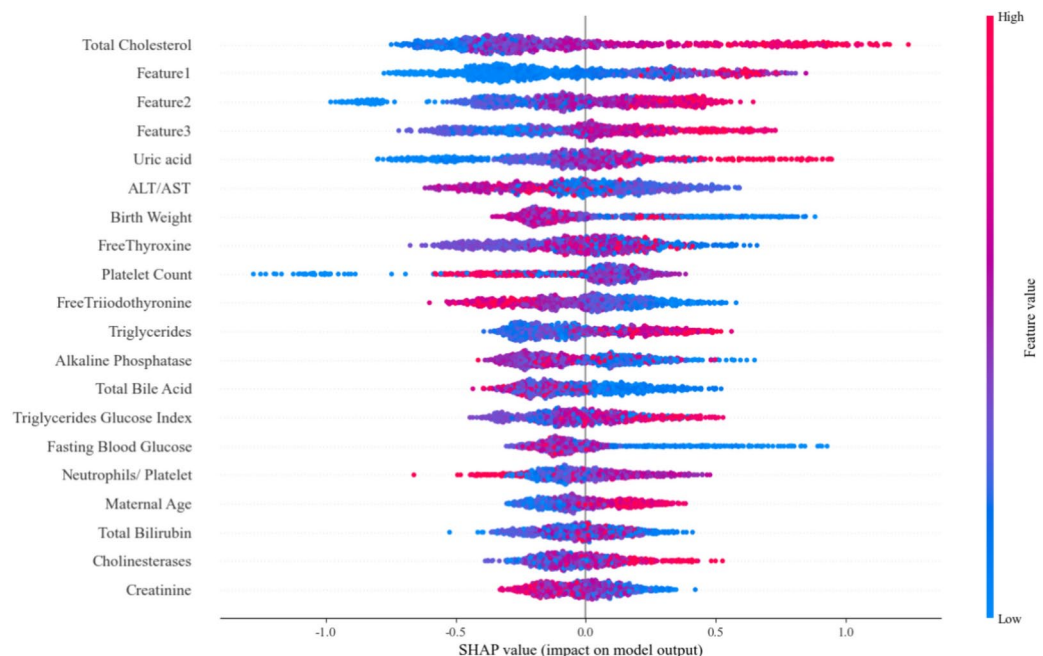


Fig. 7. Global diagram of the XGBoost model.

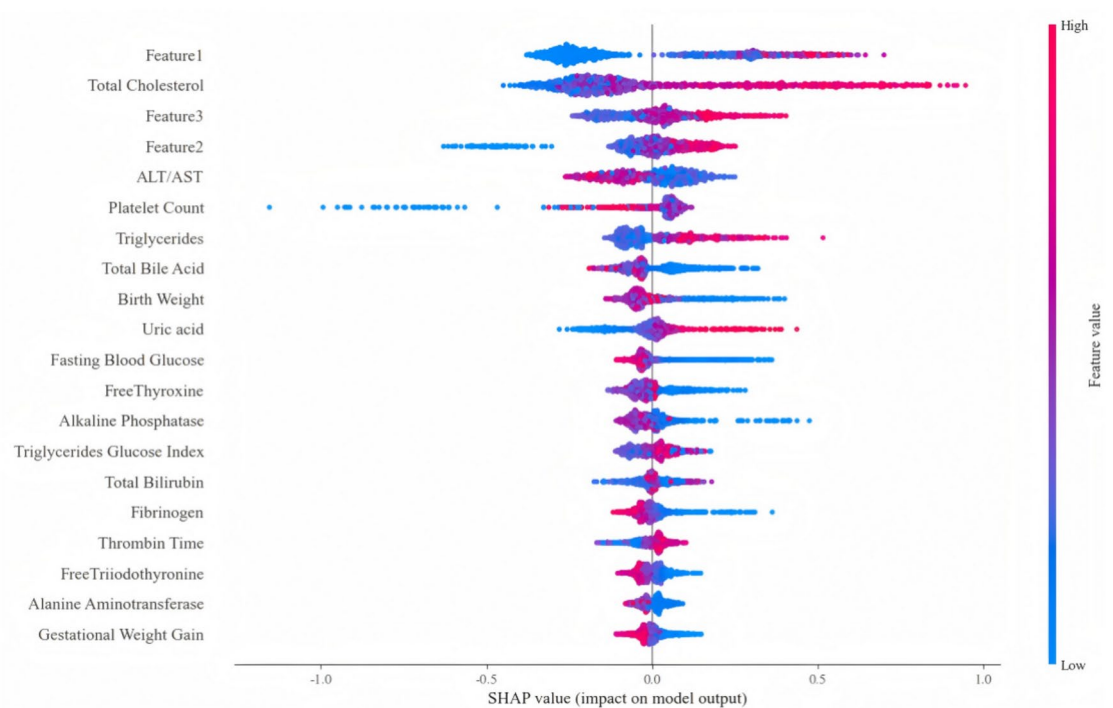


Fig. 8. Global diagram of the LightGBM model in the positive class.

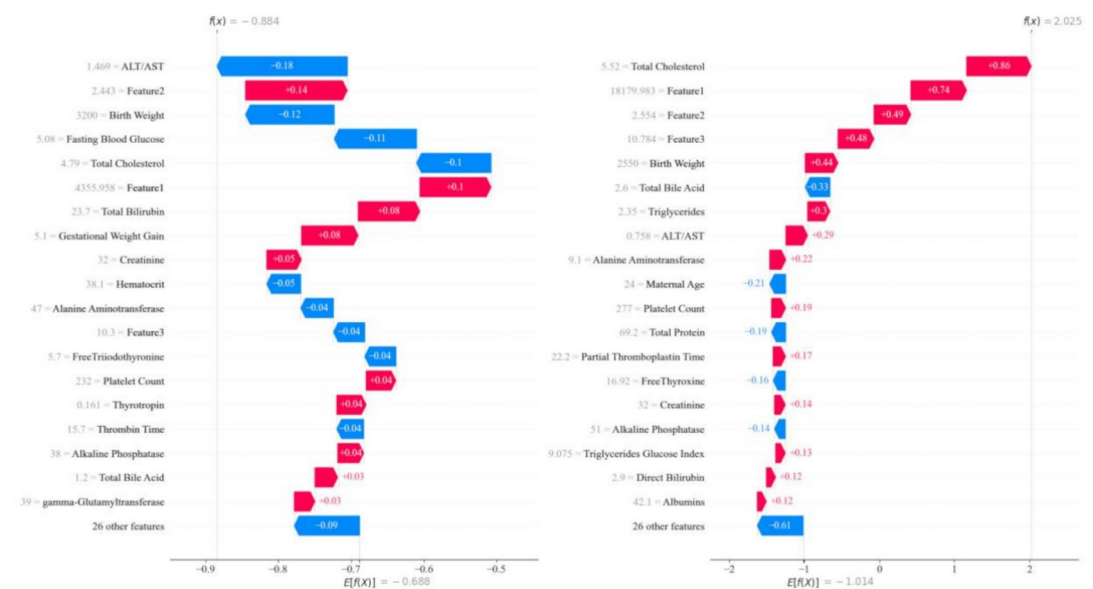


Fig. 9. Waterfall diagram of LightGBM (left) and XGBoost (right).

Insulin resistance and hyperglycemia during pregnancy in GDM patients can lead to persistent lipid metabolism disorders postpartum. Elevated triglycerides can coexist with other lipid or lipoprotein disorders³⁷. Patients with elevated plasma cholesterol and triglycerides can be categorized into three groups, with the third group including those with severe and very severe hypertriglyceridemia, where increased plasma cholesterol results from elevated VLDL and chylomicron cholesterol³⁸. In the real clinical world, severe hypertriglyceridemia (TG > 10 mmol/L or > 885 mg/dL) has a clear causal relationship with acute pancreatitis. In the results of this study, Feature1, Feature2, and Feature3, which were calculated using triglycerides, total cholesterol, fasting glucose, and BMI as blueprints, all played important roles in the model, which may be attributed to the fact that the early dyslipidemia induced insulin resistance, which subsequently caused a series of imbalances in the regulation between the glucose and lipid risk factors and the protective factors. Given that statins can cause fetal developmental abnormalities, dietary interventions have become the primary means of managing dyslipidemia

during pregnancy. Time-restricted feeding represents a new approach to evaluating the impact of eating times on metabolism, with animal models and human trials showing significant improvements in obesity and metabolic diseases³⁹.

For model selection, Pei and Lai et al. used a logistic regression model, whereas in this study, the XGBoost model of the integrated learning was used^{18,21}. The choice of this model allows us to optimize based on the inherent characteristics of the dataset, which in turn improves the generalizability and robustness of the model. In addition, the XGBoost model used in this study is able to handle datasets with category imbalance by adjusting the parameters, which is particularly important in clinical data processing, as certain diseases or conditions may be less frequent in the aggregate. In terms of model performance evaluation, compared to the models of^{18,21}, the XGBoost model of this study achieved 81.05% accuracy on dataset B, with an AUC-ROC score of 0.8792 and an NPV of 79.80%. In comparison, the best performing model in²¹ has an AUC score of only 0.842. In terms of predictive power, the present model has a significant advantage.

We strongly believe that our model is characterized by high accuracy and recall, which means that the model can help physicians to identify patients with dyslipidemia early and intervene in a timely manner. In addition, the high specificity and high NPV will reduce unnecessary treatment and further diagnostic testing, thereby reducing healthcare costs.

Limitations

Machine learning plays a pivotal role in advancing artificial intelligence, particularly in the field of medical AI where it enhances disease diagnosis and prediction accuracy through data learning and pattern recognition⁴⁰. Ye et al. demonstrated the superiority of ensemble learning algorithms, such as decision trees and random forests, in developing GDM prediction models⁴¹. Consistent with these findings, our study affirms that tree-based ensemble learning models outperform single decision trees and SVM models in terms of predictive accuracy, stability, and generalization ability.

Most medical data contain nonlinearities, non-normality, and inherent correlated structures⁴². SVMs are linear classifiers that can handle nonlinearly separable problems using kernel functions. Although kernel functions are an effective solution, SVMs have limitations when dealing with highly complex datasets; their performance heavily depends on the choice of kernel function and its parameters⁴³. Common kernel functions may not capture all fundamental features of the data.

XGBoost is a powerful and efficient ensemble learning technique known for its scalability and effectiveness in solving complex machine-learning problems. This method has proven reliable and effective in various applications, especially in handling large datasets and providing high predictive performance. Candice Bentéjac et al. compared the performance of XGBoost, LightGBM, CatBoost, random forests, and gradient boosting regarding training speed and accuracy, concluding that XGBoost ranked second in both accuracy and training speed⁴⁴. Moreover, XGBoost optimizes the learning process of decision trees by reducing unnecessary computations and using a histogram-based algorithm, thus improving training speed. Our experimental results also indicate that models constructed using tree-based ensemble learning algorithms outperform single decision tree and SVM models in terms of prediction accuracy, stability, and generalization ability.

XGBoost and LightGBM are two efficient machine learning algorithms based on gradient boosting frameworks that perform well on classification problems. XGBoost significantly improves the efficiency of processing large datasets by optimising the gradient boosting algorithm, which is well suited to our data characteristics. LightGBM is optimized for speed and memory usage, which we expect to significantly speed up the training of our models. In the top features comparison, XGBoost better controls the structure of the tree and provides high accuracy and stability in handling the dataset we adopted. This is the main reason why we believe that XGBoost is superior to LightGBM.

It must be acknowledged that the present study is subject to certain limitations. It was conducted in a single center and comprised a time-cohort study. Patients from a single healthcare organization may exhibit similar cohorts, despite being from different time periods. Consequently, further validation is necessary to assess the stability and predictive power of the model when applied to patients from other healthcare organizations.

Future directions

The standard approach to the management of women with the GDM has historically centered on the reduction of glycemic excursions during pregnancy, with the primary objective being the mitigation of pregnancy complications and the prevention of fetal macrosomia. However, extant evidence suggests that the lifelong metabolic sequelae of GDM have the potential to adversely impact a woman's overall health⁴⁵. Contemporary care for women with GDM during the postpartum period is not optimal. A seamless transition from obstetric care to primary care focused on the metabolic and cardiovascular health of women with GDM is not currently in place. It is well known that GDM significantly increases the metabolic health burden in women of reproductive age, including pre-diabetes and obesity. The clinical focus of GDM prevention and treatment must shift to early postpartum prevention to slow the progression from GDM to type 2 diabetes and reduce the risk of heart and metabolic diseases in mothers and their offspring through life-cycle management strategies⁴⁶. Stratified management using time-restricted feeding for high-risk individuals with early pregnancy dyslipidemia is undoubtedly of great value in the real clinical world.

It must be acknowledged, however, that while the capacity to identify at-risk populations may be realized, challenges emerge in clinical practice due to the absence of consensus on guidelines for postnatal follow-up, the variability in adherence to these guidelines by GDM mothers, and the ambiguity surrounding the allocation of responsibilities between primary and secondary care for postnatal screening.

Conclusion

We utilized five machine learning algorithms to establish a predictive model for postpartum dyslipidemia in GDM. Through rigorous internal and temporal validation, we found that the model established using the XGBoost algorithm demonstrated excellent and stable predictive performance. Additionally, we designed three new research variables and verified their strong positive predictive contributions in identifying high-risk groups for dyslipidemia in early pregnancy through methods such as XGBoost's feature importance analysis, correlation analysis, IV analysis, and SHAP analysis. By using the XGBoost model and the new research variables, we can identify and stratify high-risk groups for early intervention with time-restricted feeding methods. This approach aims to improve pregnancy outcomes and reduce both short-term and long-term complications for mothers and their offspring.

In the future, we plan to conduct multi-center temporal validation to confirm the strength and reliability of the model. Simultaneously, our focus will be on implementing deep learning techniques to combine data with images and create multimodal models, thus improving our ability to predict postpartum dyslipidemia in GDM.

Data availability

Data are accessible through corresponding authors upon reasonable request.

Received: 7 August 2024; Accepted: 26 February 2025

Published online: 07 March 2025

References

1. Ye, W. et al. Gestational diabetes mellitus and adverse pregnancy outcomes: Systematic review and meta-analysis. *BMJ* **377**, 1 (2022).
2. Xie, W. et al. Association of gestational diabetes mellitus with overall and type specific cardiovascular and cerebrovascular diseases: systematic review and meta-analysis. *BMJ* **378**, 1 (2022).
3. Xiang, A. H. Diabetes in pregnancy for mothers and offspring: Reflection on 30 years of clinical and translational research: The 2022 norbert freinkel award lecture. *Diabetes Care* **46**(3), 482–489 (2023).
4. Liu, Y. et al. Heterogeneity of insulin resistance and beta cell dysfunction in gestational diabetes mellitus: a prospective cohort study of perinatal outcomes. *J. Transl. Med.* **16**, 1–9 (2018).
5. Luo, M. et al. The mediating role of maternal metabolites between lipids and adverse pregnancy outcomes of gestational diabetes mellitus. *Front. Med.* **9**, 925602 (2022).
6. Jacobson, T. A., Ito, M. K., Maki, K. C. et al. National lipid association recommendations for patient-centered management of dyslipidemia: Part 1—Full Report[J]. *J. Clin. Lipidol.* (2015).
7. Zhao, D., Liu, J., Wang, M., Zhang, X. & Zhou, M. Epidemiology of cardiovascular disease in China: Current features and implications. *Nat. Rev. Cardiol.* **16**(4), 203–212 (2019).
8. Ference, B. A., Ginsberg, H. N., & Graham, I. et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel[J]. *Eur. Heart J.* (2017).
9. Ormazabal, V. et al. Association between insulin resistance and the development of cardiovascular disease. *Cardiovasc. Diabetol.* **17**, 1–14 (2018).
10. Zhao, X. et al. The crucial role and mechanism of insulin resistance in metabolic disease. *Front. Endocrinol.* **14**, 1149239 (2023).
11. Szablewski, L. Changes in cells associated with insulin resistance. *Int. J. Mol. Sci.* **25**(4), 2397 (2024).
12. Vekic, J., Stefanovic, A. & Zeljkovic, A. Obesity and dyslipidemia: A review of current evidence. *Curr. Obes. Rep.* **12**(3), 207–222 (2023).
13. Lyu, K. et al. Short-term overnutrition induces white adipose tissue insulin resistance through sn-1, 2-diacylglycerol/pkce/insulin receptor thr160 phosphorylation. *JCI Insight* **6**(4), 1 (2021).
14. Roden, M. & Shulman, G. I. The integrative biology of type 2 diabetes. *Nature* **576**(7785), 51–60 (2019).
15. Lowe, W. L. Jr. et al. Hyperglycemia and adverse pregnancy outcome follow-up study (hapo fus): Maternal gestational diabetes mellitus and childhood glucose metabolism. *Diabetes care* **42**(3), 372–380 (2019).
16. Lowe, W. L. et al. Maternal glucose levels during pregnancy and childhood adiposity in the hyperglycemia and adverse pregnancy outcome follow-up study. *Diabetologia* **62**, 598–610 (2019).
17. Jacobson, T. A. et al. National Lipid Association Recommendations for Patient-Centered Management of Dyslipidemia: Part 2[J]. *J. Clin. Lipidol.* **8**(5), 473–488 (2014).
18. Pei, L. et al. Early postpartum dyslipidemia and its potential predictors during pregnancy in women with a history of gestational diabetes mellitus. *Lipids Health Dis.* **19**, 1–8 (2020).
19. Lowe, L. P. et al. Associations of glycemia and lipid levels in pregnancy with dyslipidemia 10–14 years later: The hapo follow-up study. *Diabetes Res. Clin. Pract.* **185**, 109790 (2022).
20. Wang, C. et al. Recommended reference values for serum lipids during early and middle pregnancy: A retrospective study from china. *Lipids Health Dis.* **17**, 1–16 (2018).
21. Lai, M. et al. Underlying dyslipidemia postpartum in women with a recent gdm pregnancy who develop type 2 diabetes. *Elife* **9**, e59153 (2020).
22. Metzger, B. E. et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy: response to weinert. *Diabetes Care* **33**(7), e98–e98 (2010).
23. on Detection Evaluation EP, of High Blood Cholesterol in Adults T. Executive summary of the third report of the national cholesterol education program (ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii). *JAMA* **285**(19), 2486–2497 (2001).
24. Kang, B. S. et al. Prediction of gestational diabetes mellitus in asian women using machine learning algorithms. *Sci. Rep.* **13**(1), 13356 (2023).
25. Rathakrishnan, V. Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms. *Sci. Rep.* **12**(1), 9539 (2022).
26. Du, Y. et al. An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus. *Sci. Rep.* **12**(1), 1170 (2022).
27. Lund, B., Solutions, M. A., & Marketings, A. D. O. et al. Information Value Statistic[J]. (2013).
28. Zhang, J. Prediction of gestational diabetes mellitus under cascade and ensemble learning algorithm. *Comput. Intell. Neurosci.* **1**, 3212738 (2022).
29. Rojas, H., Alvarez, C., & Rojas, N. Statistical Hypothesis Testing for Information Value (IV)[J]. arXiv preprint [arXiv:2309.13183](https://arxiv.org/abs/2309.13183) (2023).

30. Lundberg, S. A unified approach to interpreting model predictions[J]. arXiv preprint [arXiv:1705.07874](https://arxiv.org/abs/1705.07874) (2017).
31. Lundberg, S. M. & Lee, S. I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30**, 1 (2017).
32. Plows, J. F. et al. The pathophysiology of gestational diabetes mellitus. *Int. J. Mol. Sci.* **19**(11), 3342 (2018).
33. Hu, G. et al. Effects of a lifestyle intervention in young women with gdm and subsequent diabetes. *Nutrients* **14**(24), 5232 (2022).
34. Layton, J. et al. Maternal lipid profile differs by gestational diabetes physiologic subtype. *Metabolism* **91**, 39–42 (2019).
35. Di Pino, A. & DeFronzo, R. A. Insulin resistance and atherosclerosis: Implications for insulin-sensitizing agents. *Endocrine Rev.* **40**(6), 1447–1467 (2019).
36. Wu, L. et al. Genetic variants associated with gestational diabetes mellitus: A meta-analysis and subgroup analysis. *Sci. Rep.* **6**(1), 30539 (2016).
37. Yang, T. et al. Correlation between the triglyceride-to-high-density lipoprotein cholesterol ratio and other unconventional lipid parameters with the risk of prediabetes and type 2 diabetes in patients with coronary heart disease: A rcsd-tcm study in china. *Cardiovasc. Diabetol.* **21**(1), 93 (2022).
38. Berglund, L. et al. Evaluation and treatment of hypertriglyceridemia: An endocrine society clinical practice guideline. *J. Clin. Endocrinol. Metab.* **97**(9), 2969–2989 (2012).
39. Manoogian, E. N. et al. Time-restricted eating for the prevention and management of metabolic diseases. *Endocrine Rev.* **43**(2), 405–436 (2022).
40. Zhu, T. et al. Deep learning for diabetes: A systematic review. *IEEE J. Biomed. Health Inf.* **25**(7), 2744–2757 (2020).
41. Ye, Y. et al. (2020) Comparison of machine learning methods and conventional logistic regressions for predicting gestational diabetes using routine clinical data: A retrospective cohort study. *J. Diabetes Res.* **1**, 4168340 (2020).
42. Maniruzzaman, M. et al. Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm. *Comput. Methods Prog. Biomed.* **152**, 23–34 (2017).
43. Roman, I. et al. In-depth analysis of svm kernel learning and its components. *Neural Comput. Appl.* **33**(12), 6575–6594 (2021).
44. Bentejac, C., Csorgo, A., Martinez-Munoz, G. A comparative analysis of gradient boosting algorithms. *Artif. Intell. Rev.* **54**, 1937–1967 (2021).
45. Kramer, C. K., Campbell, S. & Retnakaran, R. Gestational diabetes and the risk of cardiovascular disease in women: A systematic review and meta-analysis[J]. *Diabetologia* **1**, 1 (2019).
46. Sweeting, A. et al. A clinical update on gestational diabetes mellitus. *Endocrine Rev.* **43**(5), 763–793 (2022).

Acknowledgements

We thank the doctors from the Department of Obstetrics and Gynecology and the Department of Pathology at the Huizhou First Maternal and Child Health Care Hospital for their assistance in clinical information collection and blood testing. We would also like to thank all the staff of the Medical Ethics Committee of the Huizhou First Maternal and Child Health Care Hospital for their supervision and support of this study.

Author contributions

ZF Jiang and XK Chen contributed equally and were jointly responsible for the original draft, Resources, Methodology, and Conceptualization. YH Lai: Writing – review & editing, Visualization. JW Liu: Resources, Methodology, Investigation. XY Ye: Writing – original draft, Formal analysis. Ping Chen: Resources, Investigation, Data curation. Zhen Zhang: Supervision, Methodology, Conceptualization. All authors read and approved the final manuscript.

Funding

This work was supported by [the Guangdong Basic and Applied Research Foundation] grant number [2023A1515140146; 2022A1515110138].

Declarations

Competing interests

The authors declare no competing interests.

Ethics declarations

This study was approved by the Medical Ethics Committee of Huizhou First Maternal and Child Health Hospital (Ethics Approval No. 20240328A14), and all participants provided written informed consent. All methods were performed in accordance with the relevant guidelines and regulations.

Additional information

Correspondence and requests for materials should be addressed to Z.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025