BMC Pregnancy and Childbirth

RESEARCH

Open Access

Building a machine learning-based risk prediction model for second-trimester miscarriage

Sangsang Qi[1], Shi Zheng[1], Mengdan Lu[1], Aner Chen[1], Yanbo Chen[1] and Xianhu Fu[1*]

Abstract

Background Second-trimester miscarriage is a common adverse pregnancy outcome that imposes substantial economic and psychological pressures on both the physical and mental well-being of patients and their families. Currently, there is a scarcity of research on predictive models for the risk of second-trimester miscarriage.

Methods Clinical data were retrospectively collected from patients who were in the second trimester of pregnancy (between 14+0 and 27+6 weeks gestation), whose main diagnosis was "threatened abortion" and who were hospitalized at the Women and Children's Hospital of Ningbo University from January 2020 to October 2023. Following preliminary data processing, the patient cohort was randomly stratified into a training cohort and a validation cohort at proportions of 70% and 30%, respectively. The Boruta algorithm and multifactor analysis were used to refine feature factors and determine the optimal features linked to second-trimester miscarriages. The imbalanced dataset from the training cohort was rectified by applying the SMOTE oversampling approach. Seven machine-learning models were built and subjected to a comprehensive analysis to validate and evaluate their predictive capabilities. Through this rigorous assessment, the optimal model was selected. Shapley additive explanations (SHAP) were generated to provide insights into the model's predictions, and a visual representation of the predictive model was built.

Results A total of 2006 patients were included in the study; 395 (19.69%) of them had second-trimester miscarriages. XGBoost was shown to be the optimal model after a comparison of seven different models utilizing metrics such as accuracy, precision, recall, the F1 score, precision-recall average precision, the receiver operating characteristic-area under the curve, decision curve analysis, and the calibration curve. The most significant feature was cervical length, and the top ten features of second-trimester miscarriage were found using the SHAP technique based on relevance rankings.

Conclusion The risk of a second-trimester miscarriage can be accurately predicted by the visual risk prediction model, which is based on the machine learning mentioned above.

Keywords Second-trimester miscarriage, Machine learning, Prediction models

Introduction

Second-trimester miscarriage (STM) is a common adverse pregnancy outcome. Miscarriages occur in approximately 11-20% of clinical pregnancies, with STM accounting for 2-3% of all pregnancy losses [1, 2]. Miscarriage is undervalued in clinical practice because of its low prevalence. Nevertheless, with the prolongation of pregnancy, the effects of STM on a patient's physical

*Correspondence:
Xianhu Fu
fuxianhu2005@sina.com
[1] Department of Obstetrics and Gynecology, Women and Children's Hospital of Ningbo University, No. 339 Liuting Street, Haishu District, Ningbo 315012, Zhejiang, China

and mental health become more pronounced. Notably, certain early clinical symptoms and pregnancy complications may predict adverse clinical pregnancy outcomes [3, 4]. Current medical technology and research have made some progress in predicting miscarriage using machine learning (ML) models. Most related studies have focused on early miscarriage [5] and premature delivery [6], with little research on intermediate miscarriage. The risk factors for STMs are not fully known, and new predictive models and approaches must be developed to detect STMs.

Clinical risk prediction models are ubiquitous in many medical domains, and ML algorithms are gaining popularity as alternative approaches for prediction and classification problems [7, 8]. Since there are few studies on ML prediction models for STM, this research aims to establish a visual ML prediction model that can be used to predict the risk of STM.

Methods

The overall workflow is illustrated in Fig. 1.

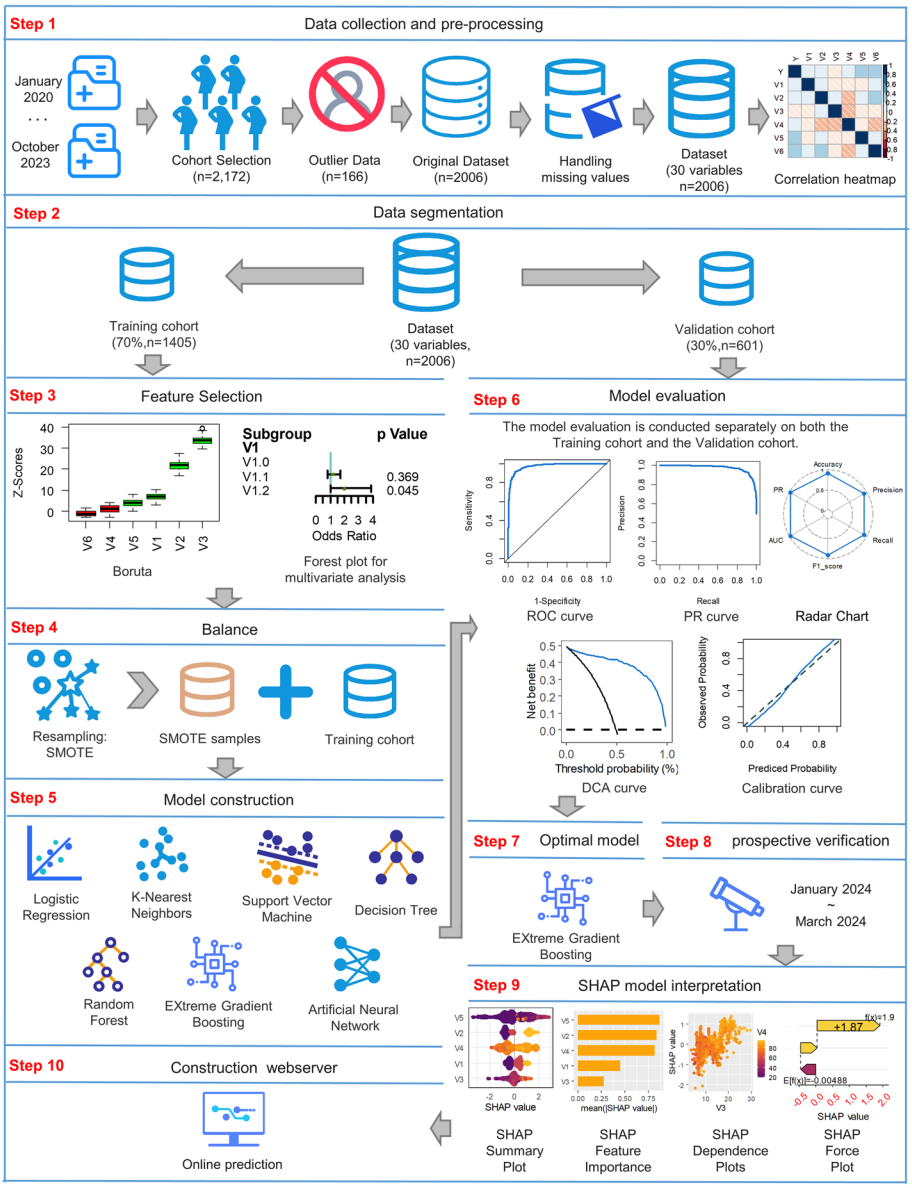

Fig. 1 Flowchart of the predictive model

Participants

Patients admitted to Ningbo University's Women and Children's Hospital with a primary diagnosis of "threatened miscarriage" were enrolled between January 2020 and October 2023.

Threatened miscarriage is defined as follows: There is initial vaginal bleeding or spotting indicators, followed by abdominal pain or lower back pain. The uterus size matches the gestational age, the cervix stays closed, the membranes are unbroken, and no foetal tissue has passed through. If symptoms persist or worsen, threatened miscarriage may escalate to a complete abortion.

Inevitable abortion is defined as follows: If the cervical opening is dilated, there is tissue blockage or water outflow, or an amniotic sac is visible, then miscarriage is unavoidable [9].

Inclusion criteria

Patients who were between 14+0 and 27+6 weeks gestation and whose main diagnosis was consistent with "threatened miscarriage".

Exclusion criteria

(1) Patients who were diagnosed with "inevitable abortion" at the time of admission. (2) Patients with systemic diseases, including multiple malignant tumours, mental illness, clotting issues, severe cardiovascular and liver conditions, haematologic conditions, and severe surgical diseases. (3) Patients with definite embryonic or foetal chromosomal abnormalities. (4) Patients who were voluntarily discharged from the hospital or transferred to a higher-level hospital for treatment with unclear pregnancy outcomes.

In addition, to avoid bias from multiple hospitalizations, only the most recent hospitalization information was recorded for the same patient. This research was approved by the Ethics Review Board of Women and Children's Hospital of Ningbo University. Given the retrospective nature of the study, informed consent from the participants was not required.

Predictive variables and outcome variables

We identified the predictive variables for this investigation by collecting and summarizing the current literature, as well as clinical experience. The data included maternal age, gestational week, parity, maternal history, prepregnancy body mass index (BMI), assisted pregnancy, and multiple pregnancies. The patient's chief complaint was abdominal pain, vaginal bleeding, and vaginal discharge. The following laboratory tests were performed on the day of admission: white blood cell (WBC) count, neutrophil percentage, C-reactive protein (CRP) level, and presence of vaginitis during pregnancy. B-scan ultrasonography revealed the presence of a subchorionic haematoma and cervical length. The complications of pregnancy included the presence or absence of uterine abnormalities, abnormal amniotic fluid volume, diabetes, hypertension, anaemia, thrombophilia, thyroid disease, scarred uterus, and uterine fibroids. The anamnesis included the history of preterm birth, history of second trimester miscarriage, number of intrauterine procedures, and presence or absence of cervical neoplasms after cervical surgery.

The outcome variable was whether the pregnancy outcome progressed to spontaneous abortion.

Sample size calculation

According to the literature [10], the prevalence of outpatient STM was $p \approx 25\%$, the tolerated error was $\delta \approx 0.01$ P, and the sample size calculation indicated that 1153 patients were needed. Given that approximately 10% of patients drop out and are lost to follow-up, at least 1281 patients had to be recruited in the training cohort consecutively [11].

Statistical analysis

The R (version 4.3.1) "VIM" package was used to count the missing values for each factor in the cases examined. By using the "zoo" package in R to fill in missing data, the createDataPartition function in the R "caret" package randomly divided all the data into two groups: a training cohort and a validation cohort (7:3). SPSS 26.0 software was used to compare the training and validation cohorts; the chi-square test was used for categorical variables, and the results are reported as quartiles; the Mann-Whitney U test was used to describe nonnormal counting data between groups, and the results are reported as frequencies (percentages).

Feature selection

To avoid factor collinearity, a collinearity test was performed on the features included in the study using the "corrplot" package. The "Boruta algorithm" was applied to the modelling group through the "Boruta" package to obtain the primary screening variables. Multivariate analysis was then performed on the filtered variables using SPSS 26.0, and the variables with a statistically significant difference of $P<0.05$ were selected as the final optimal variables.

The SmoteClassif function in the "UBL" package was utilized to create fresh minority samples for the modelling group to address the issue of class imbalance. The newly generated data were then combined with the modelling group's original data to regenerate the modelling group data [12]. All analyses were performed using R software and SPSS.

Model construction and model evaluation

The model for this study was built using seven distinct ML techniques: Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Deci-sion Tree (DT), Random Forest (RF), EXtreme Gradient Boosting (XGBoost), and Artificial Neural Network (ANN). The model hyperparameters were tuned for optimization, and model performance was evaluated based on accuracy, precision, recall, the F1 score, the receiver operating characteristic (ROC)-AUC, the precision-recall area under the curve (PR-AP), decision curve analysis (DCA), and the calibration curve.

Prospective verification

From January to March 2024, a pregnancy outcome prediction study was conducted on patients who were diagnosed with threatened abortion in the second trimester of pregnancy and who met the inclusion criteria at the Women and Children's Hospital of Ningbo University. The predicted outcomes are expressed as the accuracy, precision, recall, F1 score, ROC-AUC, and PR-AP.

SHAP model explained

Using the "shapviz" package in R software, we employed XGBoost to interpret the Shapley additive explanations (SHAP) of the importance and contribution of the features to the overall model. In addition, we performed SHAP demonstrations on individual samples and created simple application software.

Results

Patient characteristics

Between January 2020 and October 2023, a total of 2,172 patients diagnosed with "threatened abortion" between 14+0 and 27+6 weeks gestation were admitted to the Women and Children's Hospital of Ningbo University. After applying the exclusion criteria, we excluded 166 patients: two who had mental disorders, 153 who were hospitalized repeatedly, and eleven who were automatically discharged from the hospital for specific reasons. Ultimately, our study included a total of 2006 patients; among them, spontaneous abortion occurred in 395 (19.69%) patients. We considered thirty potential predictive variables for analysis while statistically accounting for missing values within each factor across all included cases. The missing data for cervical length (4.99%), CRP (3.54%), abnormal amniotic fluid volume (1.94%), prepregnancy BMI (0.25%), WBC (0.15%), and neutrophil percentage (0.15%) accounted for less than five percent of the individual data, as shown in Fig. 2. We filled missing values using the mean for numerical variables and the mode for categorical variables [13].

This study examined 30 risk variables that may contribute to STM. To avoid collinearity among the numerous features, a correlation heatmap was created to predict the risk features for STM. A correlation heatmap is a visualization tool that displays the correlation coefficient between features in the form of a heatmap and represents the degree of correlation between features using colour shading [14]. Figure 3 depicts the correlation analysis of

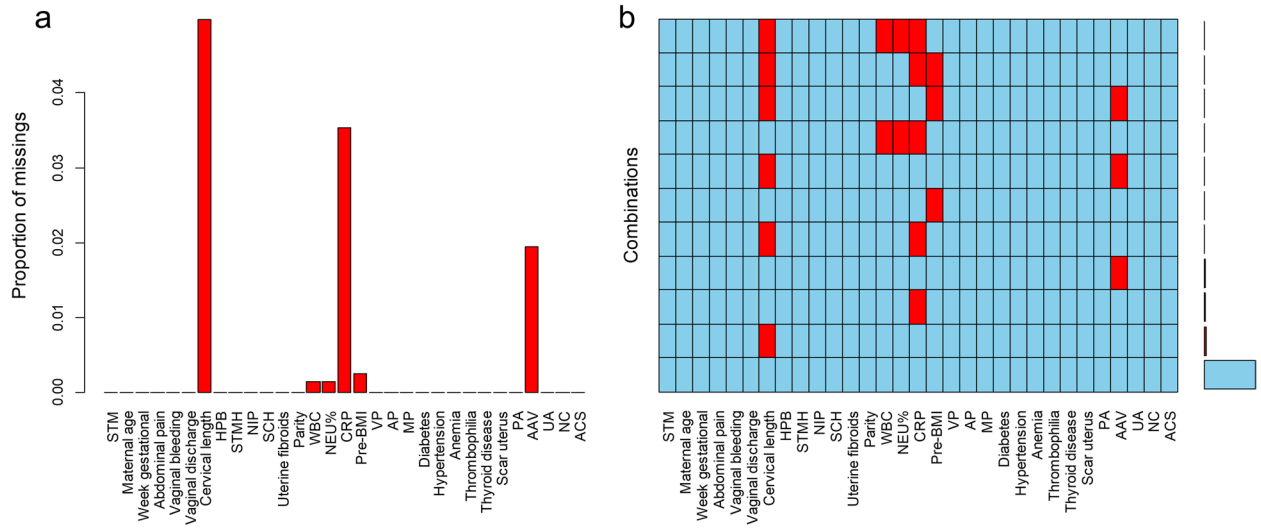

Fig. 2 **a** The percentage of missing datas in each variable, **b** With the red squares representing the missing values STM: Second-trimester miscarriage HPB: History of preterm birth STMH: Second-trimester miscarriage history NIP: Number of intrauterine procedures SCH: Subchorionic haematoma WBC: White blood cell count NEU%: Neutrophil percentage CRP: C-reactive protein BMI: Body mass index VP: Vaginitis during pregnancy AP: Assisted pregnancy MP: Multiple pregnancies PA: Placental abnormalities AAV: Abnormal amniotic fluid volume UA: Uterine abnormalities NC: Neoplasms of cervix ACS: After cervical surgery

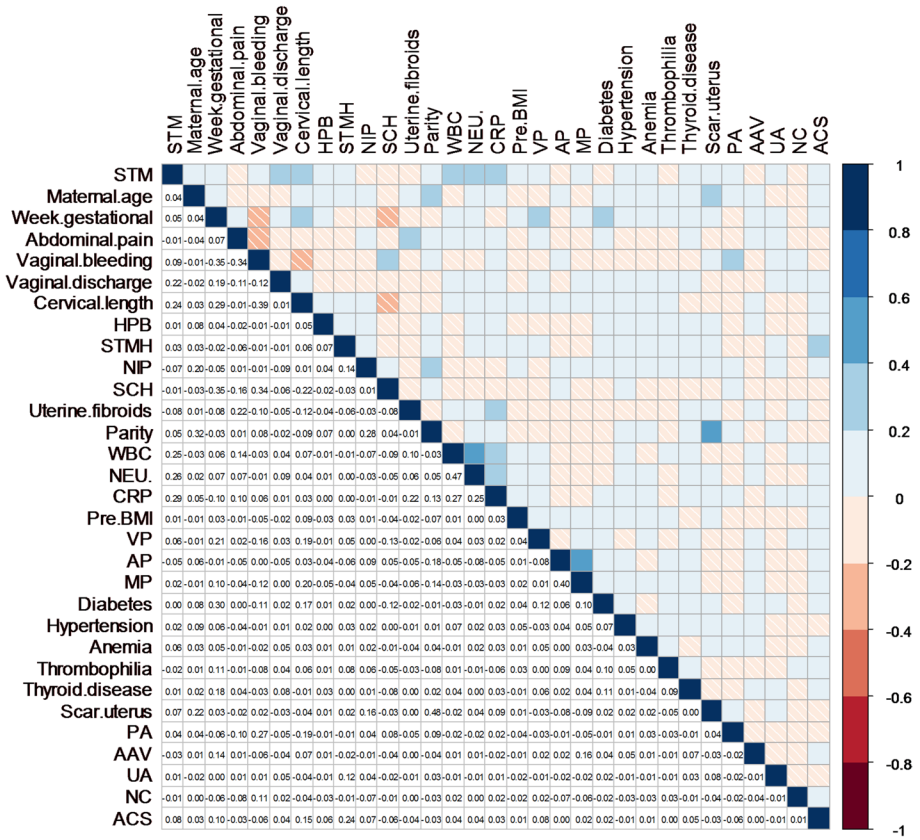

Fig. 3 STM: Second-trimester miscarriage HPB: History of preterm birth STMH: Second-trimester miscarriage history NIP: Number of intrauterine procedures SCH: Subchorionic haematoma WBC: White blood cell count NEU%: Neutrophil percentage CRP: C-reactive protein BMI: Body mass index VP: Vaginitis during pregnancy AP: Assisted pregnancy MP: Multiple pregnancies PA: Placental abnormalities AAV: Abnormal amniotic fluid volume UA: Uterine abnormalities NC: Neoplasms of cervix ACS: After cervical surgery

risk features in the included studies, with the Spearman correlation coefficient being less than 0.6, showing a low correlation between the included features.

Following data preprocessing, all the data were randomly divided into a training cohort (1405 patients) and a validation cohort (601 patients) at a 7:3 ratio, and there was no significant difference between the training and validation cohorts ($P>0.05$). (Additional file 1: Table S1)

Feature selection

Boruta is a feature selection algorithm based on the random forest classifier. This method iteratively compares the importance of each variable with randomly shaded attributes to identify significant relevant variables [15]. In our study, Boruta performed 500 iterations, and the selection results were summarized in Fig. 4. Variables having box plot in green shows all predictors are important. If boxplots are in red, it shows they are rejected. And yellow color of box plot indicates they are tentative. The following features were associated with the risk of STM: maternal age, abdominal pain, vaginal bleeding, vaginal discharge, cervical length, number of intrauterine procedures, subchorionic haematoma, uterine fibroids, WBC, neutrophil percentage, CRP, anaemia, and placental abnormalities. A multivariate analysis of the 13 features is presented in Fig. 5. The features with statistically significant differences of $P<0.05$ were selected as the final features. The resulting 10 ideal features were as follows: abdominal pain, vaginal bleeding, vaginal discharge, cervical length, subchorionic haematoma, uterine fibroids, WBC, neutrophil percentage, CRP, and placental abnormalities. According to the statistical findings, the risk of STM was 19.69%, which was significantly lower than the probability of no miscarriage occurring-a sign of an unbalanced data sample. The classifier is biased towards a high number of samples after training, which damages the classification result if the issue of sample imbalance is not addressed. From the standpoint of data, the random sampling technique is a way to address the issue of data imbalance and can be broadly classified into two types: undersampling and oversampling. Oversampling

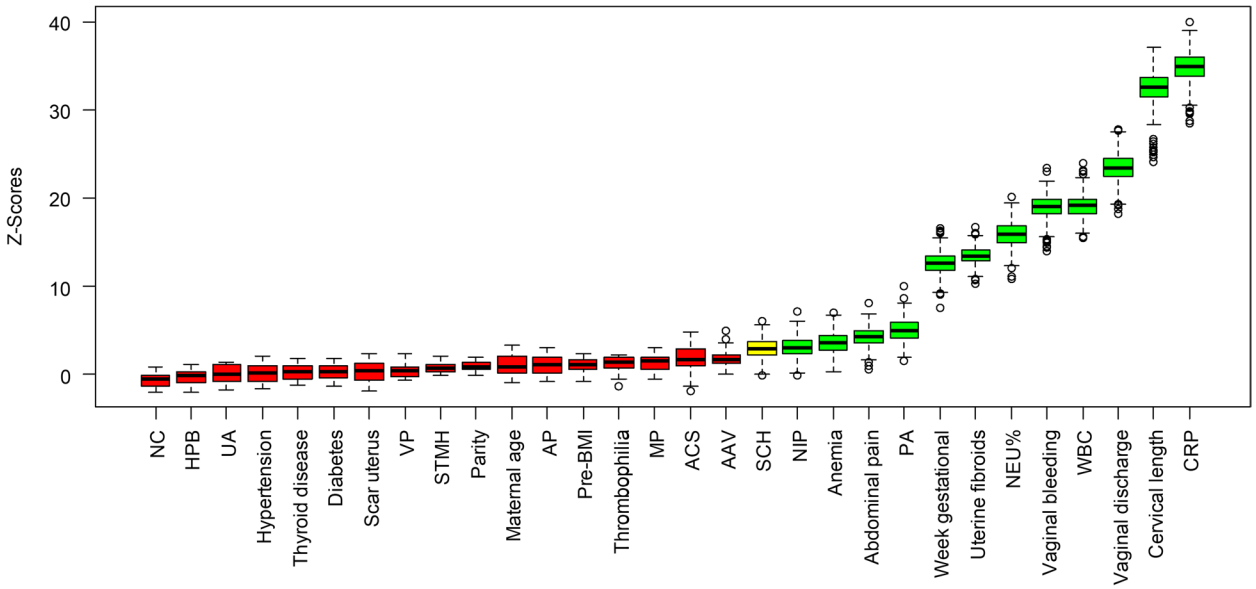

Fig. 4 Variables having box plot in green shows all predictors are important. If boxplots are in red, it shows they are rejected. And yellow color of box plot indicates they are tentative. STM: Second-trimester miscarriage HPB: History of preterm birth STMH: Second-trimester miscarriage history NIP: Number of intrauterine procedures SCH: Subchorionic haematoma WBC: White blood cell count NEU%: Neutrophil percentage CRP: C-reaction protein BMI: Body mass index VP: Vaginitis during pregnancy AP: Assisted pregnancy MP: Multiple pregnancies PA: Placental abnormalities AAV: Abnormal amniotic fluid volume UA: Uterine abnormalities NC: Neoplasms of cervix ACS: After cervical surgery

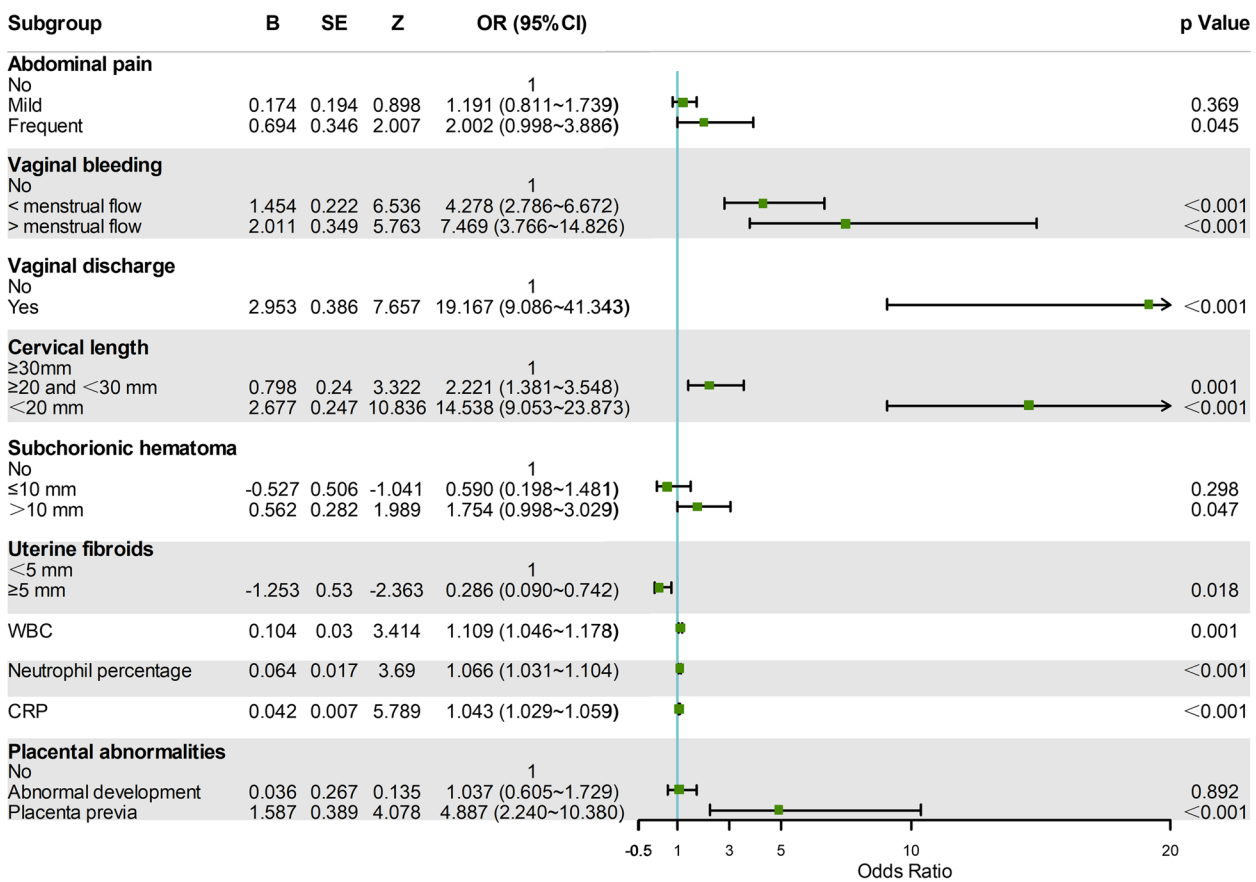

Fig. 5 White blood cell; CRP: C-reactive protein <menstrual flow: less than menstrual flow >menstrual flow: more than menstrual flow

is a frequently used technique that can effectively prevent overfitting. In this study, SMOTE oversampling technology was utilized to generate minority samples, ensuring that the ratio of positive to negative samples in the second trimester of pregnancy was 1:1 and that the data structure was balanced [16], generating a post-SMOTE dataset (Additional file 2: Table S2).

Model construction and evaluation

In this study, seven different ML techniques were used to establish the model: LR, KNN, SVM, DT, RF, XGBoost, and ANN. The performance of the seven models in terms of the ROC-AUC and PR-AP for both the training cohort and validation cohort is displayed in Fig. 6. The accuracy, precision, recall, and F1 score are compared in Table 1,

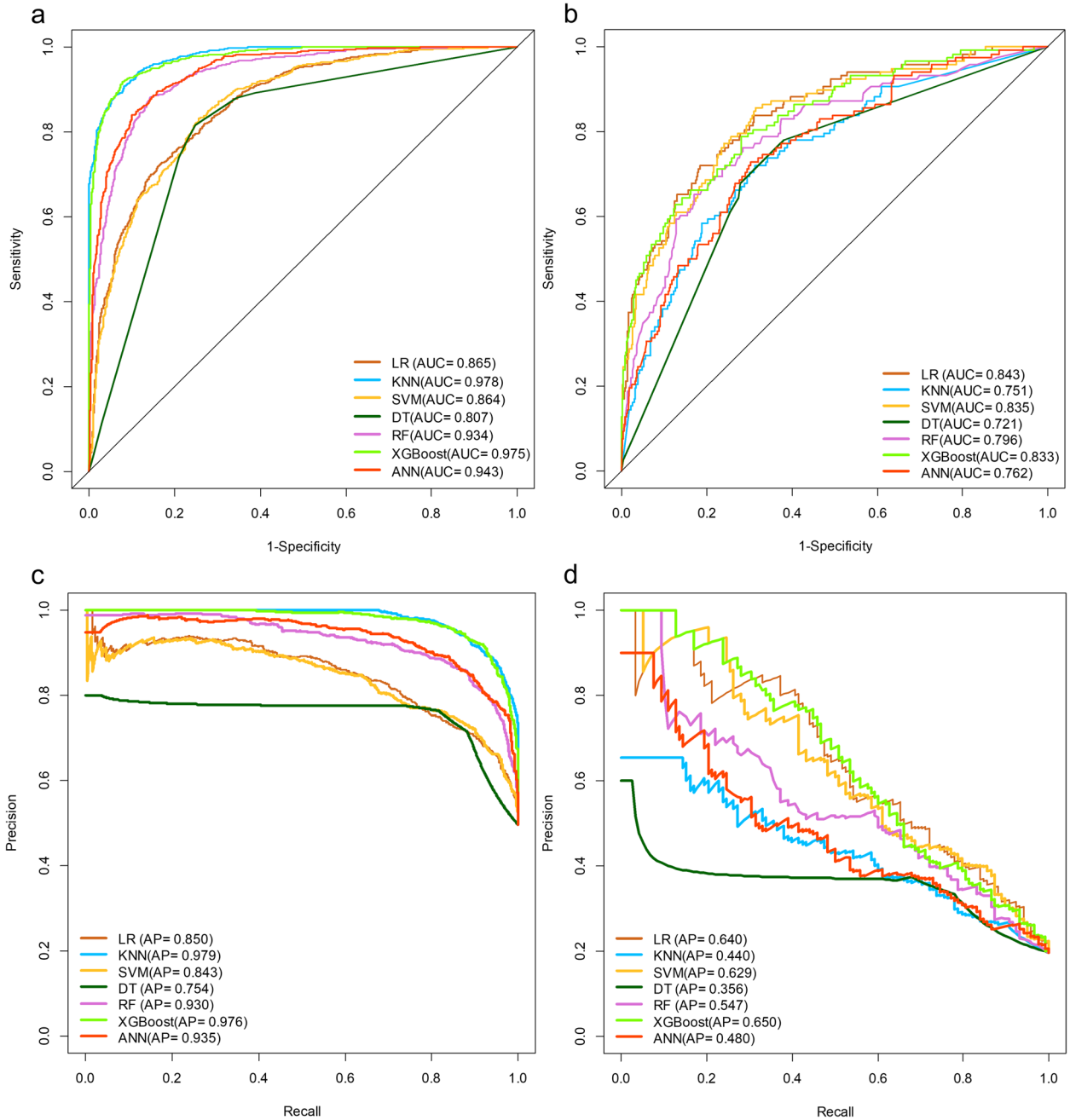

Fig. 6 Receiver operating characteristic curves and precision-recall curves. **a**-**b** Receiver operating characteristic curves in the training cohort and validation cohort. **c**-**d** Precision-recall curves in the training cohort and validation cohort. LR: Logistic Regression, KNN: K-Nearest Neighbors, SVM: Support Vector Machine DT: Decision Tree RF: Random Forest, XGBoost: EXtreme Gradient Boosting ANN: Artificial Neural Network

Table 1 Performance metrics for prediction models in the training and validation cohort

	Accuracy	Precision	Recall	F1-score	ROC-AUC	PR-AP
Training cohort						
LR	0.780	0.810	0.727	0.766	0.865	0.850
KNN	0.916	0.885	0.955	0.918	0.978	0.979
SVM	0.768	0.770	0.760	0.765	0.864	0.843
DT	0.784	0.764	0.816	0.789	0.807	0.754
RF	0.869	0.857	0.882	0.869	0.934	0.930
XGBoost	0.918	0.923	0.911	0.917	0.975	0.976
ANN	0.870	0.890	0.841	0.865	0.943	0.935
Validation cohort						
LR	0.795	0.486	0.720	0.580	0.843	0.640
KNN	0.729	0.380	0.602	0.466	0.751	0.440
SVM	0.707	0.389	0.864	0.537	0.835	0.629
DT	0.714	0.374	0.678	0.482	0.721	0.356
RF	0.484	0.653	0.653	0.556	0.796	0.547
XGBoost	0.817	0.529	0.627	0.574	0.833	0.650
ANN	0.651	0.332	0.771	0.464	0.762	0.480

LR: Logistic Regression KNN: K-Nearest Neighbors SVM: Support Vector Machine DT: Decision Tree RF: Random Forest XGBoost: EXtreme Gradient Boosting ANN: Artificial Neural Network PR-AP: Precision-Recall Average Precision ROC-AUC: Area under the receiver operating characteristic curve

and the performance data of each model are compared in several dimensions using the radar map depicted in Fig. 7 to enhance the clarity and visual appeal of the results. Out of the seven models, three models all showed high predictive performance: the LR model (AUC = 0.843), SVM (AUC = 0.835), and XGBoost (AUC = 0.833). An AUC greater than 0.8 indicates good predictive performance. The prediction capability of the KNN (AUC = 0.751), DT (AUC = 0.721), RF (AUC = 0.796), and ANN (AUC = 0.771) models is moderate. The RF model has a highest precision, with a value of 0.653, followed by the XGBoost (precision of 0.529). However, XGBoost has the

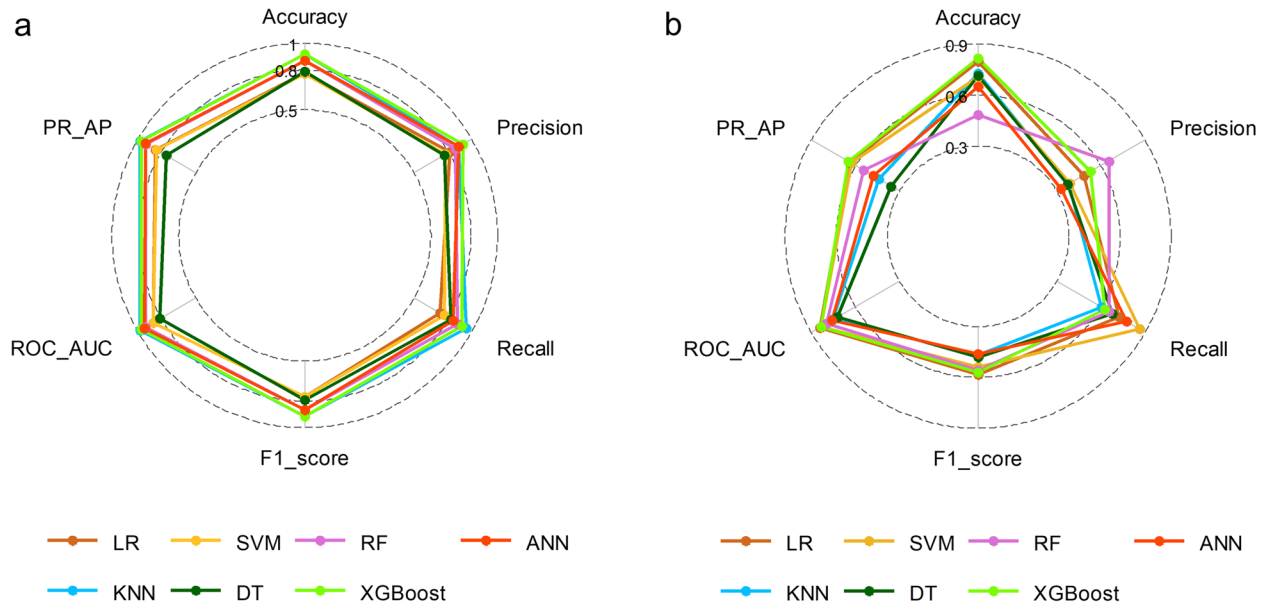

Fig. 7 Radar map for comparative performance analysis of machine learning models. LR: Logistic Regression, KNN: K-Nearest Neighbors, SVM: Support Vector Machine DT: Decision Tree RF: Random Forest, XGBoost: EXtreme Gradient Boosting ANN: Artificial Neural Network ROC-AUC: Area under the precision-recall curve PR-AP: Precision-Recall Average Precision

highest PR and accuracy, with values of 0.650 and 0.817, respectively. When the clinical effectiveness of the prediction model was assessed using clinical decision curve analysis (DCA), the XGBoost model achieved a greater net benefit within a specific range than another models (Fig. 8), indicating that the XGBoost model has good clinical utility. Brier Score is a metric used to evaluate the accuracy of probability predictions, the lower the Brier-score, the more accurate the model's predictions are. The calibration curve analysis reveals that the XGBoost model

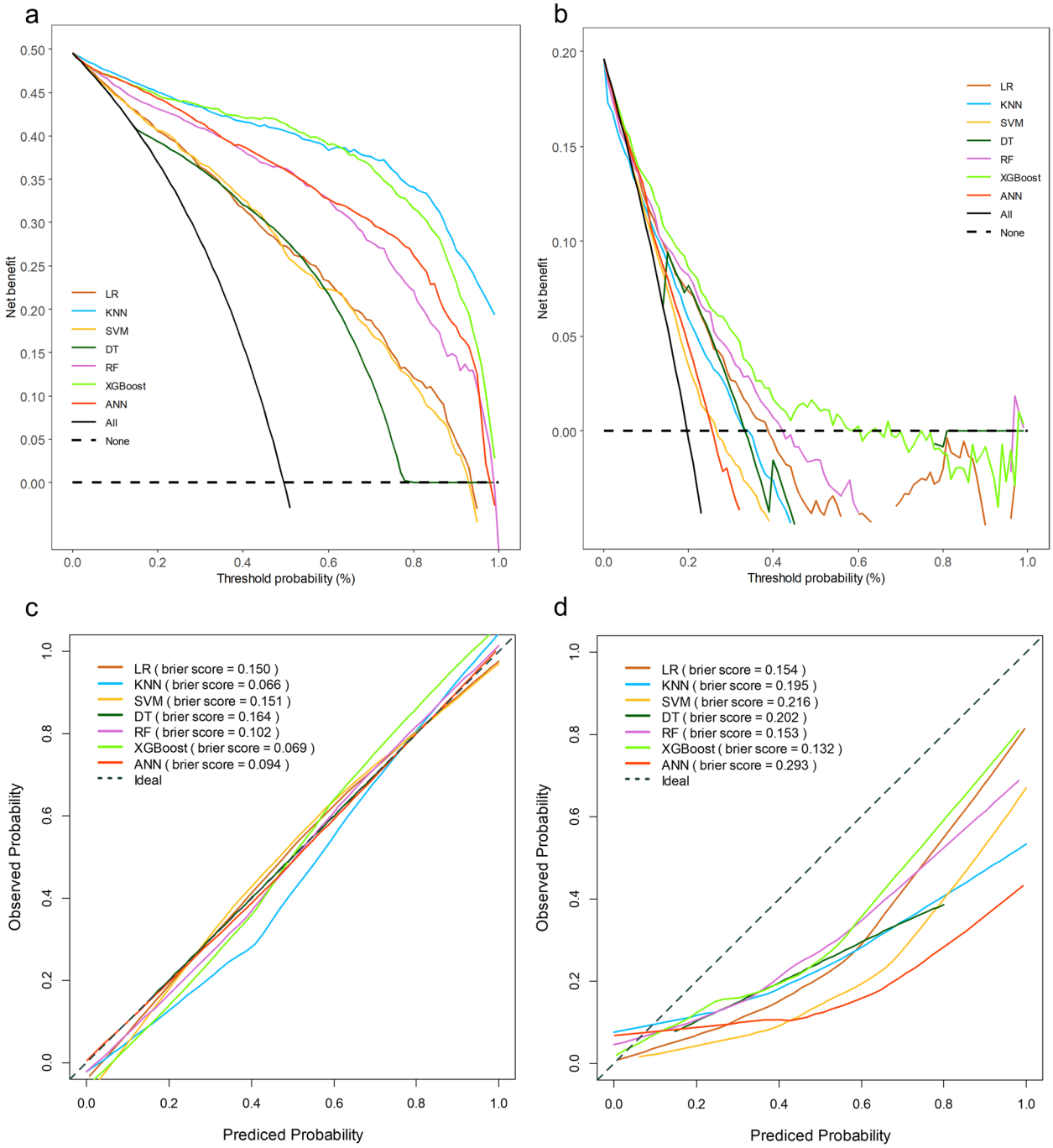

Fig. 8 Decision curve analysis and calibration plots. **a**-**b** Decision curve analysis in the training cohort and validation cohort **c**-**d** Calibration plots in the training cohort and validation cohort LR: Logistic Regression, KNN: K-Nearest Neighbors, SVM: Support Vector Machine DT: Decision Tree RF: Random Forest, XGBoost: EXtreme Gradient Boosting ANN: Artificial Neural Network

outperforms the others, boasting the lowest Brier Score (0.132 in the validation group). The aforementioned findings unequivocally demonstrate that the XGBoost model is theoptimal model.

Prospective verification

The prediction model was clinically evaluated, and data from 120 patients who satisfied the inclusion and exclusion criteria of the Women's and Children's Hospital of Ningbo University from January to March 2024 were prospectively collected. Of these, 15.0% (18/120) had STM. The model had an accuracy of 0.858, precision of 0.519, recall of 0.778, F1 score of 0.622, ROC-AUC of 0.883, and PR-AP of 0.702, as shown in Table 2.

SHAP model interpretation

Interpretation of SHAP scores is as follows. For the SHAP value (Fig. 9a), each row presents the distribution of SHAP values assigned to a feature across all cases. The x-axis denotes the SHAP value, which indicates the magnitude and direction of a feature's contribution to the model's prediction. The absolute value of SHAP values reflects the strength of a feature's impact on the model's prediction. A large positive value suggests the feature significantly boosts the prediction, whereas a large negative value indicates a substantial decrease in the prediction due to the feature. For example, for cervical length, the patient's SHAP value is positive when the cervical length is shorter, and negative when the cervical length is lower. In the SHAP interpretation for our prediction model, the shorter the cervical length is, the greater the risk of STM. The weight of each feature is determined by the average absolute value of the SHAP value for each feature, multiplied by the length of the band. The importance matrix plot (Fig. 9b) ranked the variables contributing to STM risk prediction from most to least important as cervical length, CRP, neutrophil percentage, vaginal bleeding, and vaginal discharge emerged as the five most influential variables in predictive power. We also use SHAP dependency graphs to evaluate the nonlinear impacts of features (Fig. 9c). The SHAP analysis revealed how each input feature influenced the model's output, providing potential explanations for the predictions. Yellow-colored features that increase the likelihood of an STM are shown on the left side, indicating a positive value. On the other hand, qualities that reduce the likelihood of STM are highlighted in purple and on the right, leading to a negative score. The band length of each trait indicates the value of the input feature for each patient, the longer the arrow, the bigger the impact of the feature on the output. Furthermore, the SHAP values for two typical samples were determined using the XGBoost model. Based on Fig. 9d, the pregnant woman did not report any vaginal bleeding, yet displayed a cervical length of less than 20 mm, a high neutrophil percentage of 82.8%, a WBC count of $13.8 \times 10^9/L$, and a CRP level of 2.27 mg/L. Among these factors, cervical canal length, WBC count, and neutrophil percentage contributed positively to the prediction of STM, with contribution values of 1.54, 0.312, and 1.62 respectively. In contrast, the absence of vaginal bleeding, CRP, and the remaining five factors exerted a negative influence, characterized by contribution values of -0.504, -0.204, and -0.164 respectively. Since the total positive contribution (Yellow-colored stripes) is larger than the negative contribution (purple stripe), the final value is greater than the base value. The point in the figure represents that the baseline value (expected value) of the model is -0.0146, and the total aggregate output value of the model is $f(x) = 2.58$, indicating that the patient was at a high risk of STM, that's how SHAP works. In another pregnant woman (as illustrated in Fig. 9e), the cervical length was measured at over 30 mm, the CRP level was 12 mg/L, the WBC count was $8 \times 10^9/L$, the neutrophil percentage was 66%, and the SCH was larger than 10 mm, the total value of the other five factors has a lesser impact. Specifically, the cervical canal length, neutrophil percentage, and WBC count had a negative contribution of -0.91, -0.953, and -1.34 respectively, while SCH and CRP had positive contributions of 0.438 and 0.661. The point depicted in the figure signifies the model's baseline value at -0.0146. The cumulative output of the model is represented by $f(x) = -2.07$, suggesting that the patient was at a lower risk for STM.

In addition, a web-based tool was built for clinicians to use the proposed model (available at https://qisangsang.shinyapps.io/STMRISK/) (Fig. 10).

Discussion

Different countries have different definitions of the second trimester according to the level of medical care and the ability to rescue newborns [17, 18]. At present, termination of pregnancy at 14+0 to 27+6 weeks in China is

Table 2 Prospective verification of XGBoost

	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score	ROC-AUC	PR-AP
XGBoost	14	89	4	13	0.858	0.519	0.778	0.622	0.883	0.702

TP: True Positives TN: True Negatives FP: False Positives FN: False Negatives ROC-AUC: Area under the receiver operating characteristic curve PR-AP: Precision-Recall Average Precision XGBoost: EXtreme Gradient Boosting

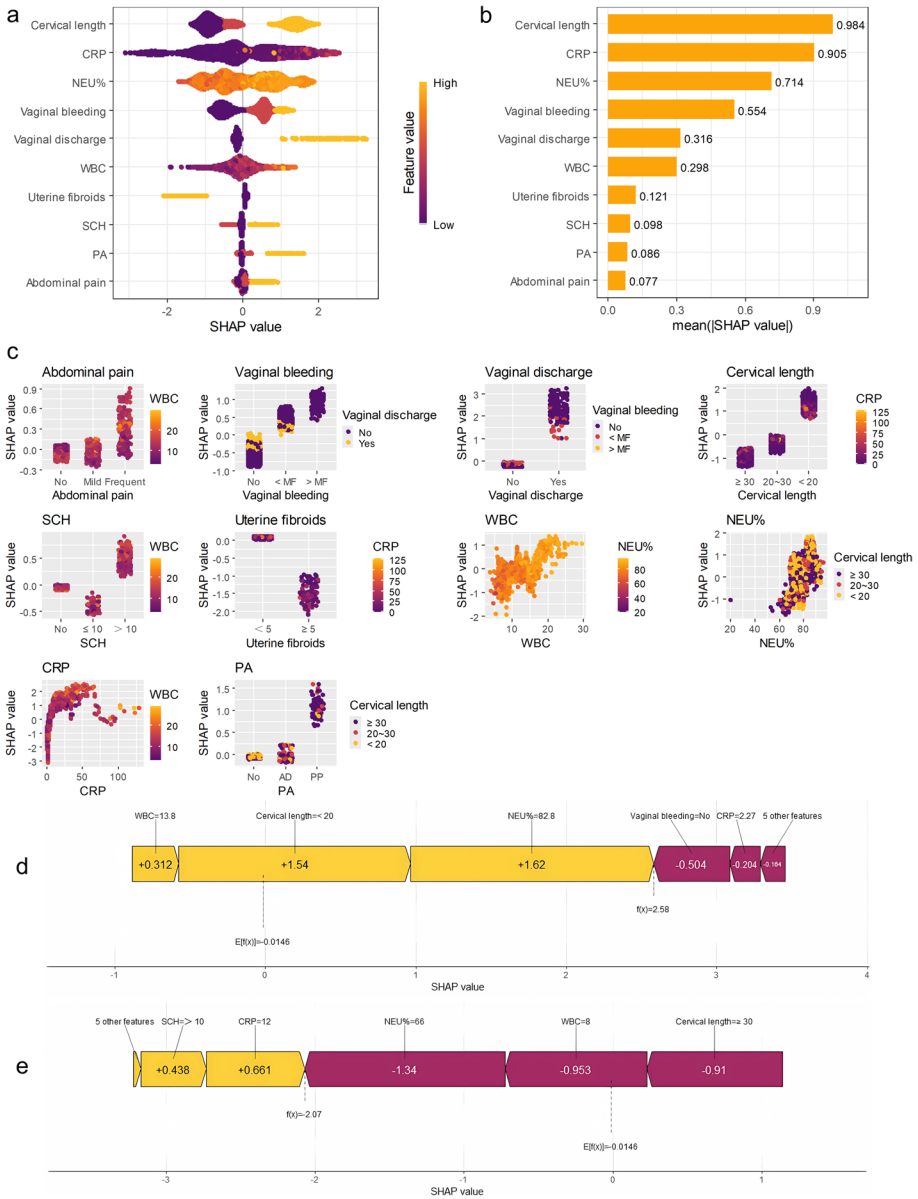

Fig. 9 **a** Shapley's additive interpretation. The positive and negative SHAP values of a feature indicate the degree to which the feature increases or decreases the predicted value, respectively. **b** SHAP feature importance matrix. Each bar represents the contribution of a feature to a particular prediction. **c** SHAP dependence plots showing the predicted risk versus the feature value. It can reveal the relationship between features and predictions, as well as the impact of different intervals of eigenvalues on predictions. **d**, **e** SHAP model for two typical predictions: SCH: subchorionic haematoma; WBC: white blood cell count; NEU%: neutrophil percentage; CRP: C-reactive protein <MF: less than menstrual flow >MF: more than menstrual flow; PA: placental abnormalities; AD: abnormal placental development; PP: placenta previa

defined as miscarriage in the second trimester [19]. Several studies have explored the risk factors for STM and constructed pregnancy risk indicators. Isenlik et al. [20]. conducted a prospective cohort study and discovered that obstetric characteristics (number of pregnancies, living children, miscarriages, dilatation and curettage, gestational age at admission, and foetal crown-rump length) and laboratory data (complete blood count; haematocrit; leukocyte, neutrophil, lymphocyte, and platelet counts; C-reactive protein; neutrophil-lymphocyte and platelet-lymphocyte ratios; and serum AA values) can be used as biomarkers. Cheung, K W, et al. [21] conducted a retrospective analysis of clinical data from 2012 to 2021 and found that foetal abnormalities, suspected cervical insufficiency, diabetes mellitus, and unexplained miscarriage were all risk factors for STM. Kelly M. McNamee

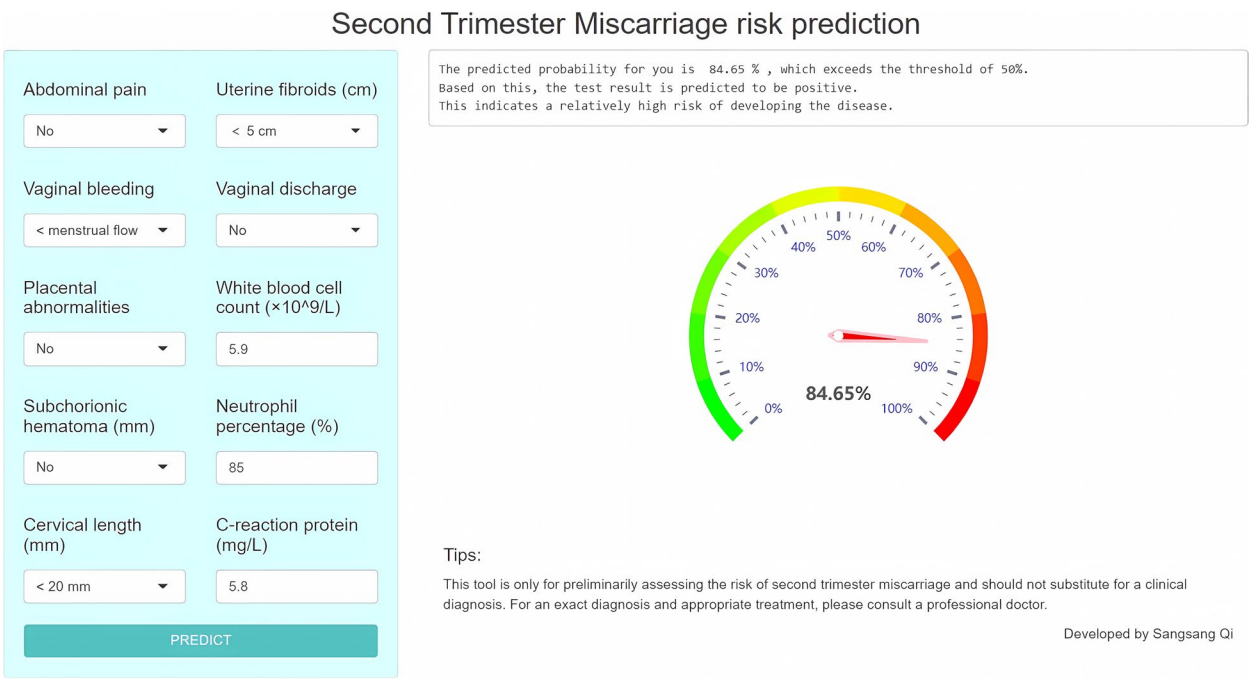

Fig. 10 A web-based risk assessment tool for second-trimester miscarriage

et al. [22] proposed that women experiencing STM are a diverse group with various causes of STM and that the presence of dual or even triple pathologies considerably enhances the probability of subsequent STM. In this retrospective study, we developed and validated a machine learning model for predicting the risk of second-trimester miscarriage across 30 variables. The XGBoost model outperformed the other models tested in terms of predictive performance and showed that the model is highly accurate and stable in predicting poor pregnancy outcomes, which may assist physicians in identifying the risk of miscarriage in a timely manner for further prevention and treatment.

Similar to the results of earlier research [23, 24], the length of the cervical canal was found to be a risk factor for STM since a short cervical length can cause early opening of the cervix and hence an incapacity to contain the foetus. The study revealed that pregnant women with cervical canal lengths between 20 mm and 30 mm had a 2.221-fold greater risk of developing STMs than did those with canal lengths $\geq$ 30 mm. The risk of pregnancy with a length of less than 20 mm was 14.538 times greater than that of pregnant women with a length of $\geq$ 30 mm, proving the feasibility of using ultrasound technology to measure the length of the cervical canal in the first and second trimesters to predict miscarriage. There is a close relationship between the length of the closed cervical canal and pregnancy outcomes; the shorter the length of the cervical canal is, the greater the risk of STM. Routine blood tests are frequent clinical tests that can assist in monitoring changes in a pregnant woman's body and may indirectly indicate the risk of miscarriage [25]. Markers of inflammation in whole blood, such as hypersensitive C-reactive protein and absolute neutrophils, have also been shown to have important effects on second-trimester miscarriage. Higher clinical markers of inflammation indicate the presence of infection, which can lead to miscarriage as a factor in STM [26]. In a retrospective study of placental histology from more than 7000 spontaneous abortions, chorioamnionitis was found to be present in 77% of patients with STM; that study also revealed that most infections, particularly early occult infections, frequently result in STMs [22, 27]. Infection and inflammation in patients make it easy for pathogens to enter the area between the decidua and the chorion, where they can release toxins, produce different cytokines, generate and release prostaglandins, induce contractions in the uterus, and ultimately result in miscarriage [28]. According to a statistical study of 4510 pregnant patients who experienced vaginal bleeding or spotting during pregnancy, there is a threefold increased risk of miscarriage when the amount of bleeding resembles or exceeds heavy menstrual bleeding [29]. The remodelling of the placental spiral artery caused by placental blood vessel rupture increases blood flow to the expanding placenta and significantly boosts oxygen tone. The placenta and foetus may be exposed to dangerously high amounts of oxidative stress, which could increase the risk of miscarriage.

Attention is warranted when a patient reports vaginal discharge during pregnancy, as pathological vaginal discharge during pregnancy is more commo n and is linked to poor outcomes for both the mother and the foetus [30]. Examples of these diseases include bacterial vaginosis, candidiasis, and trichomoniasis. Premature rupture of the membranes and miscarriage can occur due to pathological vaginal discharge during pregnancy and neglect of vaginal hygiene, which can easily lead to the development of secondary vaginal inflammation, ascending infection secondary to intrauterine infection, and decreased resistance of the foetal membranes [31]. Placental malfunction is a common cause of STM. The placenta is a crucial organ for the transmission of oxygen and nutrients between the embryo and the mother. A healthy placenta is also essential for maintaining pregnancy. Inadequate placental development can make it more difficult for the foetus to receive enough oxygen and nutrients to maintain its growth, which increases the risk of miscarriage [32]. Patients suffering from placenta previa, particularly those with the placenta covering the intrauterine orifice, have poor contractility. The placenta in this location is prone to peeling off during a contraction, which can result in severe prenatal haemorrhage and foetal distress, prolonged bleeding, an easily visible genital tract infection that causes inflammation, and a secondary intrauterine infection that can cause miscarriage. To lower the chance of miscarriage and safeguard the health of the mother and unborn child, pregnant women who are at risk for placental anomalies should be closely monitored and treated with care.

This research not only advances our understanding of STM but also creates a visual prediction model and a web server that can anticipate STM. Provided as a web-based calculator for ease of use, the tool estimates the probability of an STM by analyzing inputted data on a pregnant woman's general condition. Accessible via any web browser, the calculator can be bookmarked on computers or mobile phones for quick reference. However, our model which we have built to supplement rather than replace the clinical decision-making process. There are certain obstacles to overcome, such as the need for healthcare professionals to receive the necessary training, and we are aware of potential challenges like technology acceptance, data privacy, and applicability of the model for different populations.

This study has several limitations. First, the data was retrospectively collected, some data were incomplete or missing (e.g., cervical length, CRP, neutrophil percentage) in our data represented a potential bias. Second, it was conducted at a single institution, external validation using multicenter studies is required to establish stability in the performance of our prediction model. Third, the current study's results indicate that the causes of STM are complex, the significance of each variable may fluctuate throughout modeling populations, and various researchers may find that different factors have varying effects on STM, which affects the prediction performance of the model. In future research, we need to gather more detailed and varied data, such as information on various demographic, geographical, and environmental factors, as well as lifestyle choices, to estimate the risk of STM more precisely. Second, it is possible to tailor models for certain populations. For example, distinct risk prediction models can be constructed based on various ethnic groups and gestational weeks. Furthermore, ML can be applied to the long-term dynamic monitoring of pregnant women's real-time status. Future research can investigate how to integrate real-time monitoring technology with predictive models to deliver personalized health interventions in real-time. We anticipate that increased research and innovation will lead to the creation of more precise and useful risk prediction models, which will benefit public health.

Conclusion

This work not only increased our knowledge of the risk of miscarriage in the second trimester but also allowed us to create a visual prediction model with good performance and interpretability for predicting the risk of STM.

Abbreviations

STM	Second-trimester miscarriage
ML	Machine learning
LR	Logistic regression
BMI	Body mass index
WBC	White blood cell count
CRP	C-reactive protein
KNN	K-nearest neighbors
SVM	Support vector machine
DT	Decision tree
RF	Random forest
XGBoost	EXtreme gradient boosting
ANN	Artificial neural network
ROC-AUC	Area under the precision-recall curve
PR-AP	Precision-recall average precision
DCA	Decision curve analysis
SHAP	Shapley additive explanations

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-024-06942-w.

Additional file 1: Table S1.

Additional file 2: Table S2.

Acknowledgements

Not applicable.

Authors' contributions

SQ: collected the data, conceived the study, performed the analysis, and drafted the manuscript. SZ and ML: collected the data XF: helped revise the

manuscript (including the research concepts and designs), obtained funding and ethical approval, and oversaw data collection and implementation AC and YC: oversaw data collection and implementation. All authors have read and approved the manuscript.

Funding
This study was supported by Ningbo Key Medical Discipline (No. 2022-B16), Zhejiang Health Science and Technology Program (2022KY1155) and Ningbo key research and development program (2023Z183).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Women and Children's Hospital of Ningbo University approved this study. The Institutional Review Board of Women and Children's Hospital of Ningbo University waived the need for informed consent because of the retrospective nature of this study. This study was conducted according to the ethical standards of the Helsinki Declaration.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 10 May 2024 Accepted: 29 October 2024
Published online: 09 November 2024

References

1. Niinimäki M, Mentula M, Jahangiri R, Männistö J, Haverinen A, Heikinheimo O. Medical treatment of second-trimester fetal miscarriage; A retrospective analysis. PLoS ONE. 2017;12(7):e0182198.
2. Bottomley C, Bourne T. Diagnosing miscarriage. Best Pract Res Clin Obstet Gynaecol. 2009;23(4):463–77.
3. Saraswat L, Bhattacharya S, Maheshwari A, Bhattacharya S. Maternal and perinatal outcome in women with threatened miscarriage in the first trimester: a systematic review. BJOG Int J Obstet Gynaecol. 2010;117(3):245–57.
4. Weiss JL, Malone FD, Vidaver J, Ball RH, Nyberg DA, Comstock CH, et al. Threatened abortion: a risk factor for poor pregnancy outcome, a population-based screening study. Am J Obstet Gynecol. 2004;190(3):745–50.
5. Aljameel SS, Aljabri M, Aslam N, Alomari DM, Alyahya A, Alfaris S, et al. An Automated System for Early Prediction of Miscarriage in the First Trimester Using Machine Learning. CMC-Comput Mater Continua. 2023;75(1):1291–304.
6. Zhang Y, Du S, Hu T, Xu S, Lu H, Xu C, et al. Establishment of a model for predicting preterm birth based on the machine learning algorithm. BMC Pregnancy Childbirth. 2023;23(1):779.
7. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019;110:12–22.
8. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015;349(6245):255–60.
9. Rao VA. Early Pregnancy: Miscarriage. Gynecol Emergencies. 2020;23.
10. Drakeley A, Quenby S, Farquharson R. Mid-trimester loss-appraisal of a screening protocol. Hum Reprod (Oxf). 1998;13(7):1975–80.
11. Kotrlik J, Higgins C. Organizational research: Determining appropriate sample size in survey research appropriate sample size in survey research. Inf Technol Learn Perform J. 2001;19(1):43.
12. LemaÃŽtre G, Nogueira F, Aridas CK. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. J Mach Learn Res. 2017;18(17):1–5.
13. Emmanuel T, Maupong T, Mpoeleng D, Semong T, Mphago B, Tabona O. A survey on missing data in machine learning. J Big Data. 2021;8:1–37.
14. Wu HM, Tien YJ, Ho MR, Hwu HG, Lin WC, Tao MH, et al. Covariate-adjusted heatmaps for visualizing biological data via correlation decomposition. Bioinformatics. 2018;34(20):3529–38.
15. Kursa MB, Jankowski A, Rudnicki WR. Boruta-a system for feature selection. Fundam Informaticae. 2010;101(4):271–85.
16. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57.
17. Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller AB, et al. Born too soon: the global epidemiology of 15 million preterm births. Reprod Health. 2013;10:1–14.
18. Cullen S, Sobczyk K, Elebert R, Tarleton D, Casey B, Doyle S, et al. Second-trimester miscarriage: a review of postnatal investigations and subsequent pregnancy outcomes. Ir J Med Sci (1971-). 2023;192(4):1757–60.
19. Gu X. Chinese expert consensus on standardized diagnosis and treatment of missed miscarriage in the second trimester. Chin J Pract Gynecol Obstet. 2021;37(9):928.
20. Isenlik BS, Sarica MC, Kaygun BC, Inal HA. An evaluation of serum blood parameters and amyloid A levels in pregnant women with threatened miscarriage. Am J Reprod Immunol. 2024;91(3):e13829.
21. Cheung KW, Seto MTY, Wang W, Mok YK, Cheung VY. Clinical presentation, investigation, underlying causes, and subsequent pregnancy outcomes among different phenotypes of second trimester miscarriage. J Obstet Gynaecol Res. 2023;49(2):539–47.
22. McNamee KM, Dawood F, Farquharson RG. Mid-trimester pregnancy loss. Obstet Gynecol Clin. 2014;41(1):87–102.
23. Wikström T, Hagberg H, Jacobsson B, Kuusela P, Wesström J, Lindgren P, et al. Effect of second-trimester sonographic cervical length on the risk of spontaneous preterm delivery in different risk groups: A prospective observational multicenter study. Acta Obstet Gynecol Scand. 2021;100(9):1644–55.
24. Kuusela P, Jacobsson B, Hagberg H, Fadl H, Lindgren P, Wesström J, et al. Second-trimester transvaginal ultrasound measurement of cervical length for prediction of preterm birth: a blinded prospective multicentre diagnostic accuracy study. BJOG Int J Obstet Gynaecol. 2021;2:195–206.
25. Bas FY, Tola EN, Sak S, Cankaya BA. The role of complete blood inflammation markers in the prediction of spontaneous abortion. Pak J Med Sci. 2018;34(6):1381.
26. Yazdizadeh M, Hivehchi N, Ghaemi M, Azizi S, Saeedzarandi M, Afrooz N, et al. Platelet to lymphocyte and neutrophil to lymphocyte ratio in the first trimester of pregnancy, are they useful for predicting spontaneous miscarriage? A case-control study. Int J Reprod BioMed. 2023;21(6):463.
27. Quinn P, Butany J, Taylor J, Hannah W. Chorioamnionitis: its association with pregnancy outcome and microbial infection. Am J Obstet Gynecol. 1987;156(2):379–87.
28. Ugwumadu A. Chorioamnionitis and mid-trimester pregnancy loss. Gynecol Obstet Investig. 2010;70(4):281–5.
29. Hasan R, Baird DD, Herring AH, Olshan AF, Funk MLJ, Hartmann KE. Association between first-trimester vaginal bleeding and miscarriage. Obstet Gynecol. 2009;114(4):860–7.
30. Khaskheli M, Baloch S, Baloch AS, Shah SGS. Vaginal discharge during pregnancy and associated adverse maternal and perinatal outcomes. Pak J Med Sci. 2021;37(5):1302.
31. Wikström T, Abrahamsson S, Bengtsson-Palme J, Ek J, Kuusela P, Rekabdar E, et al. Microbial and human transcriptome in vaginal fluid at midgestation: Association with spontaneous preterm delivery. Clin Transl Med. 2022;12(9):e1023.
32. Odendaal H, Wright C, Brink L, Schubert P, Geldenhuys E, Groenewald C. Association of late second trimester miscarriages with placental histology and autopsy findings. Eur J Obstet Gynecol Reprod Biol. 2019;243:32–5.

Publisher's Note