RESEARCH



Integrating SHAP analysis with machine learning to predict postpartum hemorrhage in vaginal births

Zixuan Song¹, Hong Lin², Mengyuan Shao³, Xiaoxue Wang⁴, Xueting Chen⁴, Yangzi Zhou¹ and Dandan Zhang^{1*}

Abstract

Objective This study aimed to develop a machine learning (ML) model integrated with SHapley Additive exPlanations (SHAP) analysis to predict postpartum hemorrhage (PPH) following vaginal deliveries, offering a potential tool for personalized risk assessment and prevention in clinical settings.

Methods We conducted a retrospective multicenter cohort study in Northeast China, including women who had vaginal deliveries at three tertiary hospitals from September 2018 to December 2023. Data were extracted from electronic medical records. The dataset was split into a training set (70%) and an internal validation set (30%) to prevent overfitting. External validation was performed on a separate dataset. Several evaluation metrics, including the area under the receiver operating characteristic curve (AUC), were used to compare prediction performance. Features were ranked using SHAP, and the final model was explained.

Results The XGBoost model demonstrated superior predictive accuracy for PPH, with an AUC of 0.997 in the training set. SHAP value-based feature selection identified 15 key features contributing to the model's predictive power. SHAP dependence and summary plots provided intuitive insights into each feature's contribution, enabling the identification of anomalies. The final model maintained high predictive power, with an AUC of 0.894 in internal validation and 0.880 in external validation.

Conclusion This study successfully developed an interpretable ML model that predicts PPH with high accuracy. Future studies with larger and more diverse datasets are necessary to further validate and refine the model, particularly to assess its generalizability across different populations and healthcare settings.

Keywords Postpartum hemorrhage, Vaginal births, Interpretable model, XGBoost, SHAP

*Correspondence:

Dandan Zhang zhangdd@sj-hospital.org

¹Department of Obstetrics and Gynecology, Shengjing Hospital of China

Medical University, Shenyang, China

²Department of Obstetrics and Gynecology, Liaoning Maternal and Child

Health Hospital, Shenyang, China ³Department of Obstetrics and Gynecology, Shenyang Women's and

Children's Hospital, Shenyang, China

⁴Department of Health Management, Shengjing Hospital of China

Medical University, Shenyang, China



© The Author(s) 2025. 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://creati vecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Postpartum hemorrhage (PPH) is a significant global health concern that can lead to severe and potentially fatal complications for women, particularly in low-resource settings. It has been extensively studied due to its status as one of the leading causes of maternal mortality, particularly in developing countries [1]. The majority of these deaths are preventable through the establishment of clinical guidelines and policies [2], as well as the promotion of relevant research and training.

In routine clinical practice, physicians typically estimate the probability of PPH by assessing clinical history, conducting physical examinations, and performing laboratory tests. However, the limited sensitivity and specificity of these assessments, combined with the low incidence of PPH, mean that traditional bleeding assessment tools [3, 4], such as structured history taking and systematic evaluation scales, have shown low efficacy in assessing the incidence of PPH. With the advancement of medical big data and artificial intelligence, predictive models for PPH have begun to emerge, offering new opportunities for early risk assessment and intervention. However, most current studies do not differentiate between PPH following cesarean section and vaginal delivery [4]. Moreover, the selection of predictive factors in model construction is often constrained by data collection limitations and sample size [5], resulting in a lack of comprehensive assessment in most clinical predictive models.

In recent years, the rise of smart medicine and artificial intelligence has highlighted the unparalleled advantages of machine learning (ML) techniques over traditional statistics [6]. ML involves fitting predictive models to data or identifying informative patterns within datasets [7], leveraging data features to establish automated data analysis processes that enhance predictive capabilities for new data. Scholars have applied ML algorithms to various fields to construct predictive models, such as disease prediction and diagnosis [8], prognosis or mortality prediction [9], drug interaction prediction [10], rehospitalization prediction [11], and patient care needs prediction [12], all of which have shown good predictive performance [13].

Despite its advantages, ML research in the medical field faces several challenges, including handling missing data, avoiding model overfitting, and accounting for interrelationships among dataset attributes [14]. Additionally, the "black box" issue, where model inputs and operations are not visible to users or stakeholders, complicates interpretability [15]. Due to the complexity and multi-dimensionality of its algorithmic structure, understanding ML models can be difficult for clinicians. SHapley Additive exPlanations (SHAP), a method inspired by game theory and proposed by Lundberg et al. [16], addresses this issue by assigning a value to each input feature, indicating how the feature contributes to the prediction for a specific data point. Some factors positively impact the prediction probability, while others have a negative effect [17]. This can help clinicians quantify risk factors, improving their ability to focus on and prevent them in clinical practice.

This study aims to use various ML algorithms to construct an optimal PPH prediction model for vaginal delivery. Additionally, it seeks to evaluate and quantify risk factors, providing a highly reliable reference for personalized assessment and prevention of PPH in high-risk pregnant women.

Methods

Study population

This retrospective multicenter cohort study was conducted in Northeast China, focusing on women who underwent vaginal deliveries to develop and validate predictive models for PPH. The derivation cohort included women who delivered vaginally at three independent tertiary hospitals (Shengjing Hospital of China Medical University, Liaoning Maternal and Child Health Hospital, and Shenyang Women's and Children's Hospital) from September 2018 to December 2023. At the time of admission, all women were informed that their clinical data, excluding personally identifiable information, might be used for research purposes. Those who consented after being fully informed were included in the study. Exclusion criteria were: (1) age less than 18 years or more than 50 years; (2) gestational age at delivery less than 37 weeks or more than 42 weeks; (3) multiple births; and (4) stillbirth, neonatal death, or any induced labor performed with the intention of terminating the fetus's life.

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Boards of Shengjing Hospital (No. 2016PS344K, Date: 17/12/2016).

Data collection and processing

Using the electronic medical systems of the hospitals, data were collected on basic characteristics, obstetric history, pregnancy complications, delivery processes, and neonatal conditions to identify features for constructing predictive models. The data categories included: **1.Basic Characteristics**: Age, ethnicity, education level, occupation (classified into three categories based on physical labor intensity: light physical labor (LPL), moderate physical labor (MPL), and heavy physical labor (HPL)), family per capita monthly income, pre-pregnancy Body Mass Index (BMI), smoking status, and alcohol consumption status. **2.Obstetric History and Pregnancy Complications**: Gravidity, parity, history of miscarriage, spontaneous abortion history, induced and medical abortion history, history of labor induction for fetal demise

(induced with the intention of terminating a nonviable or deceased fetus), use of assisted reproductive technology, gestational age at delivery, gestational diabetes, pregnancy-induced hypertension (PIH, including gestational hypertension, preeclampsia, and eclampsia), anemia during pregnancy, coagulation dysfunction, uterine fibroids or adenomyosis, polyhydramnios, umbilical cord entanglement, premature rupture of membranes, placental abruption, vaginal bleeding during pregnancy, and presence of a scarred uterus. 3.Delivery Process and Neonatal Conditions: Delivery time, total duration of labor, first stage of labor time (including latent and active phases), second stage duration, third stage duration, placental retention/adhesion/implantation, instrumental assistance in delivery, cervical, vaginal, and perineal lacerations, newborn weight, and newborn length.

PPH was defined as vaginal bleeding exceeding 500 ml within 24 h after vaginal delivery, corresponding to the clinical concept of early postpartum hemorrhage. PPH was primarily measured using the weighing method, which calculates the difference in weight of the absorbent materials before and after blood collection. In cases of heavier bleeding, blood was collected in a container and measured using a graduated cup. Due to the potential impact of multicollinearity on predictive accuracy, features with a high correlation (correlation coefficient > 0.6) in Spearman's correlation analysis were handled by removing one of the two correlated features based on its lower correlation with the outcome. This result is illustrated in Supplementary Figure S1.

Model development and comparison

Data from the derivation cohort, collected from September 2018 to December 2022 at the three independent tertiary hospitals, was divided into a training set (70%) and a validation set (30%) to prevent over-fitting. An additional test data-set from admissions between January 2023 and December 2023, with the same inclusion and exclusion criteria as the derivation cohort, was used for external validation.

A total of 34 features were used to develop predictive models. Missing data were handled using median imputation, a common approach for dealing with missing values in clinical datasets. Six ML models were employed to predict PPH in critically ill pregnant women: eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LGBM), Gradient Boosting Decision Tree (GBDT), Gradient Boosting Machine (GBM), Adaptive Boosting (AdaBoost), and Bernoulli Naive Bayes (BNB). Common evaluation metrics, including the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, precision, recall, and F1 score, were employed to assess the reliability and performance of the models. Additionally, Decision Curve Analysis (DCA) and calibration curves were applied to validate the predictive models on both the internal validation dataset and the external validation dataset.

Feature selection and model explanation

To ensure clinicians can accept and understand the predictive models, the SHapley Additive exPlanations (SHAP) methodology was used to calculate the contribution of each variable to the prediction, thereby explaining the output of the final model. This interpretability approach provides two types of explanations: a global explanation that describes the overall functionality of the model at the feature level, and a local explanation that shows how individual features impact the model's output through a dependence plot.

SHAP values were utilized to assist in feature selection, ranking the features of the predictive model by importance and selecting those with the strongest predictive power for further analysis. The non-parametric method of Delong et al. was used to compare differences in AUC using MedCalc version 19.6 (https://www.medcalc.or g). Features of the selected ML models were gradually reduced until a significant decrease in AUC occurred.

Webpage deployment tool

To enhance the clinical utility of the model, the final predictive model was implemented into a web application using the Streamlit Python framework. This application allows users to input values for the corresponding features of the final model, returning the probability of PPH and a force plot for individual sub-items.

Statistical analysis

Data analysis was conducted using Python version 3.6.5 (https://www.python.org) and SPSS statistical software version 23.0 (https://www.ibm.com/spss). Continuous variables with a skewed distribution are presented as medians with interquartile ranges and were compared using the Mann-Whitney U test or Kruskal-Wallis H test. Categorical variables are presented as numbers with percentages and were compared using the chi-square test or Fisher's exact test. Analysis of covariance (ANCOVA) was used to adjust for confounding factors. AUCs were used to evaluate predictive efficacy. DCA was performed using R version 4.1.0 (https://www.r-project.org). A two-tailed p-value < 0.05 was considered statistically significant.

Results

Patient characteristics

This retrospective study included a total of 30,745 parturients for the identification of the predictive model cohort. During the study period from September 2018 to December 2022, 27,389 parturients were admitted to the obstetrics departments of the three hospitals, with 2,556 excluded based on the study's exclusion criteria. The remaining 24,833 parturients who met the inclusion criteria were randomly assigned to separate training and internal validation groups (see Table 1). Additionally, in the external validation cohort admitted from January 2023 to December 2023, 257 parturients were excluded, resulting in 3,099 parturients included (Supplementary Table 1). Study design details are shown in Fig. 1.

Model development and performance comparison

Data collected during pregnancy and within 24 h after delivery were used to generate six ML models to predict the likelihood of PPH in parturients during the perinatal period. Among the six models, the XGBoost model (AUC = 0.997, CI: 0.997–0.998) demonstrated the best predictive performance for PPH, followed by the LGBM model (AUC = 0.980, CI: 0.977–0.984) and the GBDT model (AUC = 0.966, CI: 0.960–0.972). The ROC curves for all six ML models with all features included are shown in Fig. 2, with predictive values detailed in Table 2.

An initial predictive model incorporating all 34 identified risk factors was constructed, and SHAP value analysis was applied for feature selection. By plotting the SHAP values for each feature across all samples, we gained an intuitive understanding of the overall patterns in the data, which also facilitated the detection of outlier predictions. In these visualizations (Fig. 3 and Supplementary Fig. 2), each row corresponds to a different feature, with the horizontal axis representing the SHAP values. Individual data points represent samples, colorcoded to indicate the magnitude of feature values—red for high and blue for low. This approach allowed us to discern the contribution of each feature to the model's predictions and to identify any anomalies that might suggest a need for further investigation.

Identification of the final model

During the feature reduction process, based on feature importance ranking, the XGBoost model's AUC and F1 score demonstrated that the model maintained good predictive power, with no significant change in predictive ability when the number of features was reduced to 15 (Fig. 4A-B, Supplementary Table 2, Supplementary Fig. 3). Thus, the final model was selected when the feature set was narrowed down to 15 features.

Multicollinearity among the 15 features was assessed to determine its potential impact on predictive accuracy. A correlation coefficient close to 0 indicates low correlation, with values less than 0.8 generally considered not correlated. Figure 4C shows that each feature exhibits independence, suggesting that multicollinearity is not a significant issue in this model.

Model explanation

As illustrated in the SHAP summary plot (Fig. 5), the contribution of the 15 selected features to the model was evaluated using average SHAP values, displayed in descending order. Figure 6 depicts the relationship between the actual values and SHAP values of these 15 features. SHAP values above zero correspond to a higher risk of PPH in the model's positive class prediction. For instance, parturients with a newborn weight \geq 3500 g or a second stage of labor \geq 100 min have SHAP values above zero, pushing the decision towards PPH.

Local explanations analyze how specific predictions for individual patients are made by combining personalized input data. Figure 7A and C, and 7E show parturients who did not experience PPH within 24 h postpartum, illustrating the impact of the selected features on the model's output. According to the predictive model, the x-axis in Fig. 8A represents the probability of the sample being predicted as non-PPH, and the y-axis represents the selected features and their corresponding values. The waterfall plot starts with the expected model output on the x-axis (E[f(X)] = -3.051). This "baseline" value of -3.051 is the average predicted probability of the test set. The combination of positive contributions (red) and negative contributions (blue) shifts the expected value output to the final model output (f(x) = -6.327). Positive SHAP values increase the probability of the sample being classified as PPH, while negative SHAP values decrease it. The force plot provides further insights through an additive force layout (Fig. 7B).

Similarly, Fig. 7B and D, and 7F show parturients who experienced PPH within 24 h postpartum. Figure 7B highlights the features that push or pull the decision towards the PPH category and their actual measured values, indicating that the decision for this case inclines towards PPH, with a probability of 32.3%.

Validation of the final model

To verify the robustness of the model and ensure an adequate sample size, we applied internal and external validation datasets. The AUC for the internal validation dataset was 0.894 (95% CI: 0.875–0.912) and for the external validation dataset was 0.880 (95% CI: 0.855–0.905), as Fig. 8. Although these AUC values are slightly lower than those observed in the training set, they still indicate strong predictive performance in both internal and external validations. The calibration curves and DCA also showed improvement in the internal and external validation datasets, addressing the imbalance in positive data seen in the test set.

Convenient application for clinical utility

The final predictive model has been implemented into a web application to facilitate practical use in clinical Table 1 Comparison of clinical characteristics and outcomes between non-PPH and PPH in the training and internal validation cohort

Age (pars)200[270, 32.0]200[20, 32.0]<000**		Non-PPH(<i>n</i> = 23,210)	PPH(n=1,623)	Р
Enhistic	Age (years)	29.0[27.0, 32.0]	30.0[28.0, 32.0]	< 0.001*
Han20,948 (89.99)1,54 (97.99)1,54 (97.99)Manchu1,630 (7.39)7.9 (1.096)Coure ethnic groups660 (2.99)7.9 (1.096)Educational AttainmentEducational Attainment12.713 (64.396)9.26 (57.196)Bachelors/Segnee2.55 (1.01.90)2.60 (7.196)Detragmatine or higher2.55 (1.01.90)2.20 (2.596)CourseUnemployed10.44 (45.098)2.20 (5.596)-Moderate physical isbor0.50 (7.11.90)2.50 (5.96)-Heave physical isbor0.50 (7.11.90)0.50 (7.11.90)-Course1.23 (5.690)2.10 (7.96)-Course1.23 (5.690)2.20 (7.96)-Course2.32 (5.690)2.20 (7.96)-Course1.23 (5.690)2.00 (7.90)-Course2.128 (0.969)2.00 (7.90)-Course2.128 (0.969)2.00 (7.90)-Course2.128 (0.969)1.00 (7.90)-Course2.128 (0.969) <t< td=""><td>Ethnicity</td><td></td><td></td><td>< 0.001*</td></t<>	Ethnicity			< 0.001*
Manchu1,693 (2.3%)12 /01 (1.3%)Other achinic groups669 (2.9%)70 (4.3%)Other achinic groups669 (2.9%)93 (3.0.4%)Bachalaris digme12,73 (4.8.3%)93 (3.0.4%)Bachalaris digme2,415 (4.8.3%)20 (2.5.%)Potgraduate or higher2,415 (4.8.3%)20 (2.5.%)Compation2,57 (1.1.%)20 (2.5.%)Unamployscal labor2,57 (1.1.%)20 (2.5.%)Heavy physical labor2,55 (1.1.%)20 (2.5.%)Heavy physical labor2,55 (1.1.%)20 (2.5.%)Heavy physical labor2,55 (1.5.%)3.65 (1.2.%)Heavy physical labor2,55 (1.5.%)3.65 (1.2.%)Heavy physical labor2,55 (1.5.%)3.65 (1.2.%)Heavy physical labor2,55 (1.5.%)3.65 (1.2.%)Social Construction2,21 (1.5.%)3.65 (1.2.%)Social Construction2,29 (1.5.%)3.65 (1.2.%)Social Construction2,21 (1.5.%)3.65 (1.2.%)Social Construction2,21 (1.5.%)3.65 (1.2.%)No2,12 (1.9.%)11 (2.7.%)11 (2.7.%)No2,12 (1.5.%)1,12 (1.9.%)1.1.No2,13 (1.9.%)1,12 (Han	20,848 (89.8%)	1,374 (84.7%)	
Other ethnic groups669 (2.9%)70 (4.3%)Educational AttainmentEducational AttainmentItiph school or tobow8.142 (35.1%)9.05 (57.1%)Bachelors/sdegnee12.713 (51.8%)9.05 (57.1%)Ortganduite or higher2.353 (0.1%)201 (12.5%)Occupation10.441 (45.0%)9.32 (57.4%)Unemployed10.441 (45.0%)9.32 (57.4%)Unemployed2.573 (11.1%)2.22 (15.5%)Moderate physical labor8.25 (26.0%)10.10%)Howay physical labor8.25 (26.0%)10.10%)Family Per Capita Monthly Income (10.000 yuan)6.68 (41.1%)10.25 (20.0%)C33.68 (16.3%)6.63 (31.7%)10.00%)S-203.362 (16.5%)22 (13.7%)10.00%)S-201.205 (56.0%)70 (20%)10.00%)S-501.205 (56.0%)70 (20%)10.00%)Soking1.20 (99.3%)10.00%)10.00%)No2.128 (19.2%)10.10%)10.00%)No2.128 (19.2%)10.20 (19.3%)10.00%)No2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)Yes2.128 (19.2%)10.00%)10.00%)<	Manchu	1,693 (7.3%)	179 (11.0%)	
Educational Attainment <td>Other ethnic groups</td> <td>669 (2.9%)</td> <td>70 (4.3%)</td> <td></td>	Other ethnic groups	669 (2.9%)	70 (4.3%)	
Bit bachol or below B,142 (35.1%) 493 (30.4%) 1 Bacholor S degree 12,713 (64.4%) 926 (57.1%) 1 Occupation	Educational Attainment			< 0.001*
Bachelor's degree 12/13 (54.8%) 926 (57.1%) Postgraduate or higher 2,355 (10.1%) 244 (12.5%) Cocupation - - Unemployed 10,441 (45.0%) 922 (57.4%) - Light physical labor 9,597 (41.5%) 420 (23.5%) - Moderate physical labor 9,597 (41.5%) 420 (23.5%) - Family Per Capita Monthly Income (10.000 yuan) 665 (24%) 668 (41.1%) - Solid	High school or below	8.142 (35.1%)	493 (30,4%)	
bottom besideLast of the part of the pa	Bachelor's degree	12 713 (54 8%)	926 (57.1%)	
Companie Control (Companie) Control (Companie) Control (Companie) Occupation 10.441 (45.0%) 322 (5.5%) Companie) Unemployed 10.441 (45.0%) 322 (5.5%) Companie) Moderate physical labor 9.591 (41.3%) 420 (25.9%) Companie) Family Per Cata Monthly Income (10.000 yuan) 0.887 Gold (41.1%) Companie) <0.5-20	Postoraduate or higher	2 355 (10 1%)	204 (12 5%)	
Unemplayed 10,441 (45,0%) 932 (57,4%) Unemplayed Light physical labor 2,573 (11.9%) 252 (15.5%) Image: Complayed in the complayed in t	Occupation	2,555 (10.176)	201 (12.570)	< 0.001*
basis basis basis basis basis 2,573 (11.186) 252 (15.5%) 11 basis 2,552 (11.186) 252 (15.5%) 11 basis 2,552 (2.6%) 19 (2.9%) 11 Family Per Capita Monthly Income (10.000 yuan) 6.65 (30.2%) 5.65 5	Unemployed	10 441 (45 0%)	932 (57.4%)	(0.001
lagna para basis and bas	Light physical labor	2 573 (11 1%)	252 (37.176)	
Induction of the state of the stat	Moderate physical labor	9 501 (/1 3%)	420 (25.9%)	
Instry production of 10,000 yuan) 0.657 (41.66) 684.11%) <0.5		605 (2 6%)	10 (1 2%)	
c5 6,57 (41,6%) 668 (41,1%) 0.5-20 8,896 (83,3%) 668 (41,1%) 0.5-20 3,362 (14,5%) 222 (13,7%) 2.0-5.0 1,393 (5,4%) 222 (13,7%) 5.50 20,213,7%) 20,00% 5.50 222 (13,7%) 20,00% Smoking 20,218,97%) 20,128,93%) 20,00% Smoking 23,128 (99,6%) 1,612 (99,3%) 20,00% Yes 20,04%) 110,7%) 20,00% Yes 20,04%) 10,02% 20,00% Yes 57 (0,2%) 3(0,2%) 20,00% Yes 0,0,1 1,020 (99,8%) 20,000% Yes 0,0,1 1,020 (99,8%) 20,000% Yes 10,0,1 0,0,01% 20,000% Yes 0,0,1 1,12 0,000% Yes 0,0,1 1,200 (99,8%) 20,001% Yes 0,0,1 1,200 (90,30) 30,201 (90,30) Yes 1,200 (90,30) 3,201 (90,30) 20,001%	Family Ber Capita Monthly Income (10,000 yuan)	005 (2.070)	19 (1.270)	0.887
Sub Sub Sub Sub Sub Sub Size Sub Sub Sub Sub Sub 20-50 3,82 (14,5%) 22 (13,7%) Sub Sub >5.0 1,23 (5%) 97 (6.0%) Cob Sub Su		0.657 (41.6%)	669 (41 104)	0.007
0.9-2.0 0,96 (93.9%) 0.95 (93.9%) 20-5.0 3,362 (14.9%) 222 (13.7%) >5.0 1,293 (5.6%) 97 (6.0%) Smoking 0.001* 0.001* Smoking 22 (13.7%) 0.001* No 23.128 (99.6%) 1.612 (99.3%) 0.001* Yes 0.204%) 1.602 (99.3%) 0.25%) Alcohol Consumption 23.153 (99.8%) 1.602 (99.2%) 0.001* Yes 57 (0.02%) 3 (0.2%) 0.001* Pregnancy History 10.12 0.001* 0.001* Asisted reproductive technology 0.011 0.013 0.001* No 22.595 (97.4%) 3.589 (97.9%) 0.001* Parity (number of deliveries) 0.01,1 0.001 0.001* No 22.059 (97.4%) 3.589 (97.9%) 0.001* No 20.007 (87.1%) 3.00 (80.1%) 0.001* No 20.207 (87.1%) 3.283 (85.3%) 0.001* No 3.003 (12.9%) 3.283 (85.3%) 0.001*	< 0.5 0.5 20	9,037 (41.0%)	626 (20 20%)	
20-30 3,302 (14.3%) 24/137.%) >50 1,293 (56%) 97 (60%) BMI(kg/m^2) 1,938 (18.4, 22.4) 20.4 (18.5, 23.3) <0.001*	20.50	0,090 (30.3%)	050 (59.2%)	
>3.0 1,29 (3.0%) 9 (0.0%) BM(lkg/m^2) 198[18.4, 22.4] 20.4[18.5, 23.3] <0.001*	2.0-5.0	3,302 (14.5%)	222 (13.7%)	
Binking/mc/1 (0,01/18), 2,33) (0,00/11 Smoking (0,00/11) (0,00/11) (0,00/11) No 23,128 (99,6%) 1,612 (99,3%) (0,00/11) Yes 20,04%) 1,620 (99,8%) (99,8%) Alcohol Consumption 23,153 (99,8%) 1,620 (99,8%) (99,8%) Yes 57 (0,2%) 3 (0,2%) (0,01)* Pregnancy History [11, 2] (11, 2) (0,01)* Parity (number of deliveries) (0,01) (0,00)* (0,00)* Assisted reproductive technology 22595 (97,4%) 1,589 (97,9%) (0,00)* Polivery (weeks) 395(30,04,02) 40,0(3),4,0,3) (0,00)* GDM 22,595 (97,4%) 1,580 (97,9%) (0,00)* Yes 30,00,12,9%) 323 (19,9%) (0,00)* GDM 20,007 (87,1%) 1,300 (0,01%) (0,00)* No 21,649 (93,3%) 1,385 (85,3%) (0,00)* Yes 3,020 (67,1%) 1,385 (85,3%) (0,00)* No 21,649 (93,3%) 1,385 (85,3%)	> 5.0	1,293 (5.6%)	97 (6.0%)	-0.001*
Smoxing 0.038 No 23,128 (99.6%) 16.12 (99.3%) Yes 82 (0.4%) 11 (0.7%) Acconstant 0.927 No 23,133 (99.8%) 16.20 (99.8%) Yes 57 (0.2%) 3 (0.2%) Pregnancy History 1[1, 2] 1[1, 2] <0.001*	BMI(kg/m^2)	19.8[18.4, 22.4]	20.4[18.5, 23.3]	< 0.001*
No 25,128 (99.0%) 16.12 (99.3%) 16.12 (99.3%) Yes 32 (0.4%) 11.0.7%) 0 Acohol Consumption 23,153 (99.8%) 16.20 (99.8%) 0 Yes 57 (0.2%) 30.02%) - Pregnancy History 11.2 0.011 0.011 <0.001*	Smoking	22,122,022,024	1 (12 (02 20))	0.036
Yes 0.04%) 1.04% Alcohol Consumption 0.927 Alcohol Consumption 23,153 (99.8%) 1.620 (99.8%) Yes 57 (0.2%) 3 (0.2%) Yes 57 (0.2%) 3 (0.2%) Pregnancy History 1[1, 2] (11, 2] <0.01*	No	23,128 (99.6%)	1,612 (99.3%)	
Alconstrumption 0927 No 23,153 (99,8%) 1,620 (99,3%) Yes 57 (0.2%) 3 (0.2%) Pregnancy History [1, 2] [1, 2] <0.001*	Yes	82 (0.4%)	11 (0.7%)	
No 23,153 (99.88) 1,620 (99.88) Yes 57 (0.298) 3 (0.296) Pregnancy History [1, 2] [1, 1, 2] <0.01*	Alcohol Consumption			0.927
Yes 57 (0.2%) 3 (0.2%) Pregnancy History [1, 2] [1, 2] <0.01	No	23,153 (99.8%)	1,620 (99.8%)	
Pregnacy History [1, 2] [1, 2] <.001* Parity (number of deliveries) 0[0, 1] 0[0, 1] <.0001*	Yes	57 (0.2%)	3 (0.2%)	
Parity (number of deliveries) 0[0, 1] 0[0, 1] <0001*	Pregnancy History	1[1, 2]	1[1, 2]	< 0.001*
Assisted reproductive technology 0,299 No 22,595 (07.4%) 1,589 (07.9%) Yes 615 (2.6%) 34 (2.1%) Delivery (weeks) 395[39,04.0.2] 40.0[39.1,40.3] 0.001* GDM 20,207 (87.1%) 1,300 (80.1%) 0.001* Yes 3,003 (12.9%) 323 (19.9%) 0.001* No 21,649 (93.3%) 1,385 (85.3%) 0.001* Yes 1,561 (67.9%) 28 (14.7%) 0.001* No 1,561 (67.9%) 28 (14.7%) 0.001* Yes 1,561 (67.9%) 28 (14.7%) 0.001* No 1,581 (95.9%) 1,889 (73.9%) 0.001* Yes 1,9381 (83.5%) 1,189 (73.3%) 0.001* No 1,320 (16.9%) 1,588 (97.8%) 0.001* Yes 1,260 (5%) 35 (2.2%) 0.001* No 1,260 (5%) 1,581 (62.9%) 0.001* No 2,595 (74.4%) 1,581 (62.9%) 0.001* No 2,595 (74.9%) 1,581 (62.9%) 0.01* No 2,595 (74.9%) 1,581 (62.9%) 0.01*	Parity (number of deliveries)	0[0, 1]	0[0, 1]	< 0.001*
No 22,595 (97,4%) 1,589 (97,9%) Yes 615 (2,6%) 34 (2,1%) Delivery (weeks) 35 (30,0,40.2) 40 (03),1 (40.3) <0001*	Assisted reproductive technology			0.299
Yes 615 (2.6%) 34 (2.1%) Delivery (weeks) 39,5[39,0,40,2] 40.0[39.1,40,3] <.0001*	No	22,595 (97.4%)	1,589 (97.9%)	
Delivery (weeks) 39:5[39.0, 40.2] 40.0[39.1, 40.3] <0.001*	Yes	615 (2.6%)	34 (2.1%)	
GDM	Delivery (weeks)	39.5[39.0, 40.2]	40.0[39.1, 40.3]	< 0.001*
No 20,207 (87.1%) 1,300 (80.1%) Yes 3,003 (12.9%) 323 (19.9%) PIH <.0001*	GDM			< 0.001*
Yes 3,003 (12.9%) 323 (19.9%) PH No 1,364 (93.3%) 1,385 (85.3%) Yes 1,561 (6.7%) 238 (14.7%) Anemia 1,561 (6.7%) 238 (14.7%) No 1,381 (83.5%) 1,189 (73.3%) Yes 3,829 (16.5%) 434 (26.7%) Yes 3,829 (16.5%) 434 (26.7%) Coagulation disorder No 1,381 (83.5%) 1,189 (73.3%) Yes 3,829 (16.5%) 434 (26.7%) Yes 1,260,5%) 434 (26.7%) Yes 1,260,5%) 1,588 (97.8%) Yes 1,260,5%) 1,588 (97.8%) Yes 1,260,5%) 1,501 (92.5%) Yes 1,260,5%) 1,201 (92.5%) Yes 1,501 (92.5%) 1,501 (92.5%) Yes 1,501 (92.5%) 1,501 (92.5%) Yes 2,0967 (90.3%) 1,399 (86.2%) Yes 2,439 (7%) 2,41 (13.8%) Yes 2,439 (7%) 24 (13.8%) Yes 2,430 (7%) 24 (13.8%) <td>No</td> <td>20,207 (87.1%)</td> <td>1,300 (80.1%)</td> <td></td>	No	20,207 (87.1%)	1,300 (80.1%)	
PIH >	Yes	3,003 (12.9%)	323 (19.9%)	
No 21,649 (93,3%) 1,385 (85,3%) Yes 1,561 (6.7%) 238 (14.7%) Anemia No 19,381 (83,5%) 1,189 (73,3%) Yes 3,829 (16.5%) 434 (26.7%) Cogulation disorder No 23,084 (99.5%) 1,588 (97.8%) Yes 126 (0.5%) 35 (2.2%) Uterine fibroids/adenomyosis 22,595 (97.4%) 1,501 (92.5%) Yes 1501 (92.5%) 430 (0.01) Polyhydramnios 22,595 (97.4%) 1,20 (7.5%) No 20,697 (90.3%) 1,399 (86.2%) Yes 05 (2.6%) 122 (7.5%) No 20,967 (90.3%) 1,399 (86.2%) Yes 20,967 (90.3%) 1,399 (86.2%) Yes 20,976 (0.02) 24 (13.8%) Umblical cord entanglement	PIH			< 0.001*
Yes 1,561 (6.7%) 238 (14.7%) Anemia No 1,9381 (83.5%) 1,189 (73.3%) Yes 3,829 (16.5%) 434 (26.7%) Coagulation disorder No 23,084 (99.5%) 1,588 (97.8%) Yes 126 (0.5%) 35 (2.2%) Uterine fibroids/adenomyosis No 22,595 (97.4%) 1,501 (92.5%) Yes 1,501 (92.5%) Polydramnios No 20,967 (90.3%) 1,399 (86.2%) Yes 20,967 (90.3%) 1,399 (86.2%) Yes 20,967 (90.3%) 1,399 (86.2%) No 20,967 (90.3%) 1,399 (86.2%) Yes 1,43 (70.4%) 1,973	No	21,649 (93.3%)	1,385 (85.3%)	
Anemia	Yes	1,561 (6.7%)	238 (14.7%)	
No 19,381 (83.5%) 1,189 (73.3%) Yes 3,829 (16.5%) 434 (26.7%) Coagulation disorder < 0.001*	Anemia			< 0.001*
Yes 3829 (16.5%) 434 (26.7%) Coagulation disorder < 0.001*	No	19,381 (83.5%)	1,189 (73.3%)	
Coagulation disorder < 0,001*	Yes	3,829 (16.5%)	434 (26.7%)	
No 33,084 (99.5%) 1,588 (97.8%) Yes 35 (2.2%) Uterine fibroids/adenomyosis < 0.001*	Coagulation disorder			< 0.001*
Yes 35 (2.2%) Uterine fibroids/adenomyosis < 0.001*	No	23,084 (99.5%)	1,588 (97.8%)	
Uterine fibroids/adenomyosis < 0.001* No 22,595 (97.4%) 1,501 (92.5%) Yes 615 (2.6%) 122 (7.5%) Polyhydramnios < 0.001*	Yes	126 (0.5%)	35 (2.2%)	
No 22,595 (97.4%) 1,501 (92.5%) Yes 615 (2.6%) 122 (7.5%) Polyhydramnios <0.001*	Uterine fibroids/adenomyosis			< 0.001*
Yes 615 (2.6%) 122 (7.5%) Polyhydramnios < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < <	No	22,595 (97.4%)	1,501 (92.5%)	
Polyhydramnios < 0.001* No 20,967 (90.3%) 1,399 (86.2%) Yes 2,243 (9.7%) 224 (13.8%) Umbilical cord entanglement 0.973 No 16,254 (70.0%) 1,143 (70.4%)	Yes	615 (2.6%)	122 (7.5%)	
No 20,967 (90.3%) 1,399 (86.2%) Yes 2,243 (9.7%) 224 (13.8%) Umbilical cord entanglement 0.973 No 16,254 (70.0%) 1,143 (70.4%)	Polvhvdramnios			< 0.001*
Yes 2,243 (9.7%) 224 (13.8%) Umbilical cord entanglement 0.973 No 16,254 (70.0%) 1,143 (70.4%)	No	20,967 (90.3%)	1,399 (86.2%)	
Umbilical cord entanglement 0.973 No 16,254 (70.0%) 1,143 (70.4%) Vers 0.973	Yes	2.243 (9.7%)	224 (13.8%)	
No 16,254 (70.0%) 1,143 (70.4%)	Umbilical cord entanglement	· - · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·	0.973
	Νο	16.254 (70.0%)	1.143 (70.4%)	0.27.0
1es 0,950 (30,0%) 480 (79,0%)	Yes	6.956 (30.0%)	480 (29.6%)	

Table 1 (continued)

	Non-PPH(<i>n</i> = 23,210)	PPH(n=1,623)	Р
Premature rupture of membranes			< 0.001*
No	18,247 (78.6%)	1,051 (64.8%)	
Yes	4,963 (21.4%)	572 (35.2%)	
Placental abruption			< 0.001*
No	23,180 (99.9%)	1,605 (98.9%)	
Yes	30 (0.1%)	18 (1.1%)	
Vaginal bleeding during pregnancy			0.867
No	22.030 (94.9%)	1,546 (95,3%)	
Yes	1,180 (5.1%)	77 (4.7%)	
Scarred uterus			< 0.001*
No	22.980 (99.0%)	1.593 (98.2%)	
Yes	230 (1.0%)	30 (1.8%)	
Time of delivery	200 (1.070)	50 (11570)	< 0.001*
0. o 'clock	781 (3.4%)	70 (4 3%)	(0.001
1 o clock	793 (3.4%)	51 (3 1%)	
2 o 'clock	747 (3.2%)	48 (3.0%)	
3 o 'clock	809 (3.5%)	46 (2.8%)	
4 o 'clock	855 (3.7%)	30 (2.0%)	
5 o 'clock	840 (3.6%)	46 (2.8%)	
6 o 'clock	054 (4 1%)	60 (3 7%)	
7 o 'clock	936 (4.0%)	57 (3.5%)	
R a falach	661 (2 804)	26 (2.2%)	
	001 (2.8%)	50 (2.2%) E2 (2.2%)	
9 0 CIOCK	915 (5.9%)	55 (5.5%) E8 (2.6%)	
	904 (4.2%)	56 (5.0%) (0 (4.2%)	
	958 (4.1%)	09 (4.3%) 70 (4.2%)	
	1,112 (4.8%)	/0 (4.3%)	
	1,221 (5.3%)	101 (6.2%)	
	1,286 (5.5%)	91 (5.6%)	
	1,450 (6.2%)	120 (7.4%)	
16 o clock	1,155 (5.0%)	87 (5.4%)	
1/ o'clock	1,0/0 (4.6%)	81 (5.0%)	
18 o clock	1,059 (4.6%)	/1 (4.4%)	
19 o'clock	1,031 (4.4%)	/4 (4.6%)	
20 oʻclock	965 (4.2%)	/9 (4.9%)	
21 oʻclock	951 (4.1%)	63 (3.9%)	
22 oʻclock	871 (3.8%)	72 (4.4%)	
23 oʻclock	826 (3.6%)	81 (5.0%)	
First stage of labor - Latent phase	194.0[124.0, 296.0]	281.0[181.0, 416.0]	< 0.001*
First stage of labor - Active phase	72.0[43.0, 119.0]	108.0[62.0, 205.5]	< 0.001*
Second stage of labor	28.0[15.0, 50.0]	49.0[26.0, 93.0]	< 0.001*
Third stage of labor	5.0[3.0, 7.0]	5.0[4.0, 10.0]	< 0.001*
Newborn weight (grams)	3365.0[3120.0, 3620.0]	3610.0[3400.0, 3930.0]	< 0.001*
Newborn length (centimeters)	51.0[50.0, 52.0]	52.0[50.0, 53.0]	< 0.001*
Analgesia during labor			< 0.001*
No	19,593 (84.4%)	1,177 (72.5%)	
Yes	3,617 (15.6%)	446 (27.5%)	
Placenta accreta spectrum			< 0.001*
No	22,440 (96.7%)	1,458 (89.8%)	
Yes	770 (3.3%)	165 (10.2%)	
Instrumental assistance in delivery			< 0.001*
No	22,704 (97.8%)	1,524 (93.9%)	
Yes	506 (2.2%)	99 (6.1%)	
Lacerations of the cervix, vagina, or perineum			< 0.001*

Table 1 (continued)

	Non-PPH(<i>n</i> = 23,210)	PPH(<i>n</i> = 1,623)	Р
No	21,139 (91.1%)	1,398 (86.1%)	
Yes	2,071 (8.9%)	225 (13.9%)	

Continuous values were presented as median [interquartile range]. Categorical values were presented as number (percentage) PPH: postpartum hemorrhage; BMI: body mass index; GDM: gestational diabetes mellitus; PIH: pregnancy-induced hypertension *: *P* < 0.05



Fig. 1 Patient Selection Criteria Flowchart

scenarios. By entering the actual values of the 15 features required by the model, the application predicts the risk of PPH for individual parturients. It also displays a force plot for each parturient, indicating the features contributing to the decision on PPH: features on the right side in blue indicate factors pushing the prediction towards "non-PPH," while features on the left side in red push the prediction towards "PPH." This web application is accessible online at https://postpartum-hemorrhage-predictio n-model6.streamlit.app/.

Discussion

The era of big data has revolutionized clinical healthcare, and the application of ML in disease prediction and prognosis is increasingly prevalent. This study leverages SHAP values to assist in identifying risk factors and constructs a ML predictive model for PPH to predict the risk in women undergoing vaginal delivery. By utilizing big data to develop a diagnostic system for PPH, we can significantly enhance the accuracy of PPH diagnosis. While artificial intelligence, including ML, is making strides in obstetric disease diagnosis and treatment globally, the application of interpretable ML in clinical practice is still Fig. 2 Comparison of ROC results of different machine learning models. XGBoost: eXtreme Gradient Boosting; LGBM: Light Gradient Boosting Machine; GBDT: Gradient Boosting Decision Tree; GBM: Gradient Boosting Machine; Ada: Adaptive Boosting; BNB: Bernoulli Naive Bayes

Table 2 Performance of the ML models for PPH prediction

	Accuracy	Precision	Recall	F1 score
XGBoost	0.99	1.00	0.78	0.87
LGBM	0.97	0.98	0.50	0.66
GBDT	0.97	0.99	0.58	0.73
Ada	0.94	0.74	0.19	0.30
BNB	0.94	0.73	0.02	0.05

PPH: postpartum hemorrhage; XGBoost: eXtreme Gradient Boosting; LGBM: Light Gradient Boosting Machine; GBDT: Gradient Boosting Decision Tree; Ada: Adaptive Boosting; BNB: Bernoulli Naive Bayes

in its infancy. This study represents an important step towards improving the standardization of obstetric medical care and reducing maternal mortality, particularly by providing a tool for early identification of high-risk patients.

Currently, three widely used postpartum hemorrhage risk assessment tools globally are the California Maternal Quality Care Collaborative (CMQCC) [18] toolkit, the Association of Women's Health, Obstetric and Neonatal Nurses (AWHONN) [19] guidelines, and the New York State Department of Health (NYSBOH) [3] guidelines. These tools have summarized and classified risk factors for PPH into low, medium, and high categories based on expert consensus. However, a comparative study of these three risk assessment scales found that they only have moderate reliability in predicting severe PPH in high-risk cesarean section groups [20]. In these tools, the incidence of PPH is significantly higher in the high-risk group only when a pregnant woman is classified as such [20]. Additionally, in Dilla et al.'s study, the sensitivity of the CMQCC toolkit in predicting PPH requiring transfusion was only 22%, and the probability of severe PPH in the low-risk group was still 0.4–0.6% [21]. Therefore, add-ing more assessment indicators and improving modeling methods may enhance the accuracy of PPH prediction.

In 2021, Venkatesh et al. published a study utilizing ML to predict PPH [22]. This study included 152,279 childbirth cases, of which 7,279 (4.8%) experienced PPH exceeding 1000 milliliters. They included 55 risk factors and used random forest and extreme gradient boosting algorithms to develop ML models. The extreme gradient boosting algorithm achieved the best performance (AUC: 0.93; 95% CI: 0.92–0.93), followed by the random forest, demonstrating the high predictive performance of ML. However, this study primarily focused on PPH cases associated with cesarean sections, with 28% of the patients undergoing cesarean delivery, and 91% of PPH cases occurring in cesarean section patients.

Akazawa's study [23] from 1995 to 2020 at the Tokyo Women's Medical University East Center, involving 9,894 women who underwent vaginal delivery, applied eleven clinical variables to create a ML model predicting PPH, defined as blood loss > 1000mL. The study utilized an ensemble learning approach with five ML classifiers, including logistic regression, support vector machine, random forest, boosting tree, decision tree, and a deep learning model consisting of two-layer neural networks. The deep learning model demonstrated the best performance, achieving an AUC of 0.708 for PPH prediction, with an accuracy of 0.686, false positive rate (FPR) of 0.312, and false negative rate (FNR) of 0.398.

However, previous models lacked interpretability, as they did not explain the results of the ML algorithm predictions due to their black box nature. SHAP, a ML model interpretation method based on Shapley values from game theory, addresses this issue by assigning the contribution of model predictions to each feature, thus explaining the decision-making process of the model [24]. SHAP values quantify the impact of each feature on the model's prediction results, aiding in understanding why the model gives specific predictions. In this study, SHAP was employed in the XGBoost model for its superior predictive performance and interpretability. Personalized explanations constructed through SHAP force analysis help doctors understand why the model makes specific highrisk recommendations, enhancing understanding of the decision-making process.

To further validate the contribution of risk factors to the model, SHAP feature importance and feature effects were calculated. Then 15 key variables that significantly predict PPH were identified. The most important input parameter for PPH was newborn weight, followed by stages of labor, nature of work, premature rupture of membranes, among others. The clinical significance





Fig. 3 Global model explanation of initial XGBoost model SHAP value for all risk factors. (A) SHAP summary bar plot. (B) SHAP summary dot plot

of these variables is consistent with existing literature, emphasizing the importance of these common clinical characteristics in predicting PPH. Previous studies have demonstrated that larger neonatal birth weight is an independent risk factor for PPH [25]. Heavier infants are generally associated with uterine distension, prolonged labor, and difficult placental separation, all of which increase the risk of bleeding [26]. Prolonged labor, which has also been confirmed as a significant predictor of PPH, may indicate insufficient uterine contractions, abnormal fetal position, or difficulty in placental separation, thus increasing the likelihood of postpartum hemorrhage [27-29]. Interestingly, occupational factors were also identified as important predictors of PPH. Individuals engaged in moderate to heavy physical labor occupations had a lower probability of developing PPH, which may reflect the potential influence of lifestyle and socioeconomic factors. Additionally, older maternal age and obesity were found to be associated with a higher risk of PPH. Pregnant women delivering at >40 weeks of gestation require enhanced surveillance, aligning with findings from previous research, especially for those delivering between 41 and 42 weeks [30]. Furthermore, we observed that certain pregnancy complications significantly increased the risk of PPH, highlighting the importance of timely preventive and management measures in clinical practice, such as infection prevention, blood pressure control, and anemia correction, to reduce the risk of PPH. Thus, SHAP analysis in this study not only helped us gain a deeper understanding of the predictive mechanisms of the ML model but also provided a bridge for applying the model in clinical practice, significantly enhancing its clinical feasibility.

However, this study has several limitations. Firstly, the data were derived solely from the Shenyang area, potentially introducing significant selection bias and limiting generalizability. Secondly, while the predictive model incorporated multiple risk factors and demonstrated high overall efficacy, the restrictive selection criteria may have influenced PPH prediction, hindering objective clinical utility assessment. Thirdly, while ML techniques require 'big data' for predictive model construction, there are no established standards for calculating the sample size needed. Therefore, caution should be exercised in interpreting the conclusions, and further evidence is warranted to confirm these findings in diverse populations.

Conclusion

In conclusion, our study has successfully developed an interpretable ML model capable of predicting postpartum hemorrhage (PPH) in patients undergoing vaginal delivery using readily available clinical data extracted from the hospital information system (HIS). his model



Fig. 4 Performance of XGBoost models to predict PPH. (A) AUC of the XGBoost model with varied numbers of features. (B) F1 score of the XGBoost model with varied numbers of features. (C) Pearson correlation plot of 15 features



Fig. 5 Global model explanation of final XGBoost model SHAP value for 15 risk factors. (A) SHAP summary bar plot. (B) SHAP summary dot plot

represents a promising tool for early risk assessment and intervention in clinical practice. The final XGBoost model exhibited outstanding predictive performance for PPH, as validated internally and externally. Moving forward, prospective studies are essential to assess whether implementing individualized and timely treatment measures guided by our predictive model can lead to improved maternity outcomes, particularly in reducing PPH-related morbidity and mortality. This represents a crucial step towards personalized healthcare in obstetrics, potentially enhancing patient care and reducing maternal morbidity and mortality rates.



Fig. 6 SHAP dependence plot. Each dependence plot shows how a single feature affects the output of the prediction model, and each dot represents a single patient. The SHAP values for specific features exceeding zero push the decision towards the "PPH" class. LPL: light physical labor; MPL: moderate physical labor; HPL: heavy physical labor; PROM: premature rupture of membranes; BMI: body mass index



Fig. 7 Local model explanation by the SHAP method. (A) Waterfall plot of risks contributed by each feature for individual patient at low; (B) Waterfall plot of risks contributed by each feature for individual patient at high; (C) Force plot of risks contributed by each feature for individual patient at low; (D) Force plot of risks contributed by each feature for individual patient at high; (E) Evolution of risks contributed by each feature for individual patient at low; (F) Evolution of risks contributed by each feature for individual patient at low; (F)



Fig. 8 Model evaluation. (A) ROC of train cohort; (B) ROC of internal validation cohort; (C) ROC of external validation cohort; (D) calibration curve of train cohort; (E) DCA curve of train cohort; (F) calibration curve of internal validation cohort; (G) DCA curve of internal validation cohort; (H) calibration curve of external validation cohort; (I) DCA curve of external validation cohort. ROC: receiver operating characteristic curve; AUC: area under curve; DCA: decision curve analysis

Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s12884-025-07633-w.

Supplementary Material	1
Supplementary Material 2	2
Supplementary Material	3
Supplementary Material	4

Acknowledgements

We would like to express our gratitude to all those who helped us during the writing of this manuscript. Thanks to all the peer reviewers for their opinions and suggestions.

Author contributions

ZS and DZ designed the study and drafted the manuscript. HL, MS, and XW done the data collection. YZ and XC designed the statistical analysis plan. ZS has participated the training and reviewed and co-authored the manuscript with DZ.

Funding

This study was supported in part by grants from 345 Talent Project of Shengjing Hospital of China Medical University (No. M0946), and Medical Education Research Project of Liaoning Province (No. 2024-N004-03), Liaoning Province Science and Technology Plan Joint Program (2024-MSLH-561).

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Ethics approval and consent to participate: The study was approved by the Ethics Committee of Shengjing Hospital of China Medical University (No. 2016PS344K, Date.17/12/2016). All participants provided informed consent.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Received: 4 June 2024 / Accepted: 21 April 2025 Published online: 03 May 2025

References

- 1. Patek K, Friedman P. Postpartum Hemorrhage-Epidemiology, risk factors, and causes. Clin Obstet Gynecol. 2023;66(2):344–56.
- Federspiel JJ, Eke AC, Eppes CS. Postpartum hemorrhage protocols and benchmarks: improving care through standardization. Am J Obstet Gynecol MFM. 2023;5(2s):100740.
- Zheutlin AB, Vieira L, Shewcraft RA, Li S, Wang Z, Schadt E, Gross S, Dolan SM, Stone J, Schadt E, et al. Improving postpartum hemorrhage risk prediction using longitudinal electronic medical records. J Am Med Inf Assoc. 2022;29(2):296–305.
- Faysal H, Araji T, Ahmadzia HK. Recognizing who is at risk for postpartum hemorrhage: targeting anemic women and scoring systems for clinical use. Am J Obstet Gynecol MFM. 2023;5(2s):100745.
- Gonzalez-Brown V, Schneider P. Prevention of postpartum hemorrhage. Semin Fetal Neonatal Med. 2020;25(5):101129.
- Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning?? A primer for the epidemiologist. Am J Epidemiol. 2019;188(12):2222–39.
- 7. Mitchell TMJCA. Machine learning and data mining. 1999, 42(11).

- Li R, Chen Y, Ritchie MD, Moore JH. Electronic health records and polygenic risk scores for predicting disease risk. Nat Rev Genet. 2020;21(8):493–502.
- Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine Learning-Based model for prediction of outcomes in acute stroke. Stroke. 2019;50(5):1263–5.
- McCoubrey LE, Gaisford S, Orlu M, Basit AW. Predicting drug-microbiome interactions with machine learning. Biotechnol Adv. 2022;54:107797.
- Leong KTG, Wong LY, Aung KCY, Macdonald MR, Cao Y, Lee SSG, Chow WL, Doddamani S. Richards AMJTAjoc: risk stratification model for 30-Day heart failure readmission in a multiethnic South East Asian community. Am J Cardiol. 2017;119 9:1428–32.
- 12. Lindberg DS, Prosperi M, Bjarnadottir RI, Thomas J, Crane M, Chen Z, Shear K, Solberg LM, Snigurska UA, Wu Y, et al. Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: A machine-learning approach. Int J Med Inf. 2020;143:104272.
- Wang H, Fu T, Du Y, Gao W, Huang K, Liu Z, Chandak P, Liu S, Van Katwyk P, Deac A, et al. Scientific discovery in the age of artificial intelligence. Nature. 2023;620(7972):47–60.
- Lewanowicz A, Wiśniewski M, Oronowicz-Jaśkowiak W. The use of machine learning to support the therapeutic process - strengths and weaknesses. Postep Psychiatr Neurol. 2022;31(4):167–73.
- Handelman GS, Kok HK, Chandra RV, Razavi AH, Huang S, Brooks M, Lee MJ, Asadi H. Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. AJR Am J Roentgenol. 2019;212(1):38–43.
- 16. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: Neural Information Processing Systems: 2017; 2017.
- Ali S, Akhlaq F, Imran AS, Kastrati Z, Daudpota SM, Moosa M. The enlightening role of explainable artificial intelligence in medical & healthcare domains: A systematic literature review. Comput Biol Med. 2023;166:107555.
- Main EK, Goffman D, Scavone BM, Low LK, Bingham D, Fontaine PL, Gorlin JB, Lagrew DC, Levy BS. National partnership for maternal safety: consensus bundle on obstetric hemorrhage. Anesth Analg. 2015;121(1):142–8.
- Bingham D, Scheich B, Bateman BT. Structure, process, and outcome data of AWHONN's postpartum hemorrhage quality improvement project. J Obstetric Gynecologic Neonatal Nurs. 2018;47(5):707–18.
- Kawakita T, Mokhtari N, Huang JC, Landy HJ. Evaluation of Risk-Assessment tools for severe postpartum hemorrhage in women undergoing Cesarean delivery. Obstet Gynecol. 2019;134(6):1308–16.
- 21. Dilla AJ, Waters JH, Yazer MH. Clinical validation of risk stratification criteria for peripartum hemorrhage. Obstet Gynecol. 2013;122(1):120–6.
- Venkatesh KK, Strauss RA, Grotegut CA, Heine RP, Chescheir NC, Stringer JSA, Stamilio DM, Menard KM, Jelovsek JE. Machine learning and statistical models to predict postpartum hemorrhage. Obstet Gynecol. 2020;135(4):935–44.
- 23. Akazawa M, Hashimoto K, Katsuhiko N, Kaname Y. Machine learning approach for the prediction of postpartum hemorrhage in vaginal birth. Sci Rep. 2021;11(1):22620.
- Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, Liston DE, Low DK-W, Newman S-F, Kim J, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nat Biomedical Eng. 2018;2(10):749–60.
- Beta J, Khan N, Khalil A, Fiolna M, Ramadan G, Akolekar R. Maternal and neonatal complications of fetal macrosomia: systematic review and metaanalysis. Ultrasound Obstet Gynecol. 2019;54(3):308–18.
- Biguzzi E, Franchi F, Ambrogi F, Ibrahim B, Bucciarelli P, Acaia B, Radaelli T, Biganzoli E, Mannucci PM. Risk factors for postpartum hemorrhage in a cohort of 6011 Italian women. Thromb Res. 2012;129(4):E1–7.
- Nyflot LT, Stray-Pedersen B, Forsen L, Vangen S. Duration of labor and the risk of severe postpartum hemorrhage: A case-control study. PLoS ONE. 2017;12(4):e0175306.
- Ladfors LV, Liu X, Sandstrom A, Lundborg L, Butwick AJ, Muraca GM, Snowden JM, Ahlberg M, Stephansson O. Risk of postpartum hemorrhage with increasing first stage labor duration. Sci Rep. 2024;14(1):22152.
- Frolova AI, Stout MJ, Tuuli MG, Lopez JD, Macones GA, Cahill AG. Duration of the third stage of labor and risk of postpartum hemorrhage. Obstet Gynecol. 2016;127(5):951–6.
- Reale SC, Bateman BT, Farber MK. Exploring new risk factors for postpartum hemorrhage: time to consider gestational age?? Anesthesiology. 2021;134(6):832–4.

Publisher's note

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