

Original Research

Development and Evaluation of a Predictive Tool for Second-Pregnancy Preterm Birth and Cervical Insufficiency in Patients With a History of Cervical Insufficiency

Xin Liu¹, Junhong Jiang¹, Yan Wang^{1,*}¹Department of Obstetrics, Xiamen Humanity Maternity Hospital, 361000 Xiamen, Fujian, China*Correspondence: wyrhzhf@163.com (Yan Wang)

Academic Editor: Christos Iavazzo

Submitted: 13 March 2025 Revised: 30 April 2025 Accepted: 26 May 2025 Published: 25 July 2025

Abstract

Background: Patients with cervical insufficiency are at a higher risk of preterm birth or recurrent cervical dysfunction in a subsequent pregnancy. Exploring key risk factors and developing predictive models is of great clinical significance. **Methods:** Patients diagnosed with cervical insufficiency during their first pregnancy were included and categorized according to whether they experienced preterm birth during their second pregnancy. Lasso regression was used for feature selection, followed by logistic regression (LR) and eXtreme Gradient Boosting (XGBoost) models to analyze the factors influencing the two outcomes. Model performance was evaluated through metrics such as area under the curve (AUC), accuracy, and recall rate. Model interpretation was performed using SHapley Additive exPlanations (SHAP) values, and a risk scoring model was constructed based on the core features to evaluate predictive performance. **Results:** The XGBoost model demonstrated the best performance in predicting preterm birth (AUC = 0.745), whereas the LR model achieved a higher AUC (0.777) for predicting cervical insufficiency. The risk model based on core features demonstrated an AUC of 0.711 for predicting preterm birth and an AUC of 0.686 for predicting cervical insufficiency. **Conclusions:** Cervical length at 16–18 weeks and 22–24 weeks, along with preterm birth and cervical laceration during first pregnancy, are key predictive factors of second-pregnancy preterm birth and cervical insufficiency. The risk model derived from multi-model analysis can effectively identify high-risk pregnancies.

Keywords: cervical insufficiency; premature birth; machine learning; risk prediction

1. Introduction

Cervical insufficiency is a relatively common yet serious obstetric pathological condition and is among the leading causes of mid-pregnancy miscarriage and preterm birth [1,2]. Its pathological features are mainly characterized by abnormal painless cervical dilation during pregnancy in the absence of obvious uterine contractions, often accompanied by membrane rupture or fetal expulsion [3,4]. The occurrence of cervical insufficiency may be closely associated with congenital structural abnormalities of the cervical tissue, a history of trauma (such as cervical surgery or childbirth-related injuries), infections, genetic predisposition, and hormonal changes during pregnancy [5]. Due to its insidious progression, cervical insufficiency is often difficult to diagnose in a timely manner, posing a significant threat to the health of both mother and fetus.

Preterm birth is a leading cause of neonatal mortality and morbidity worldwide, with approximately 15 million babies born prematurely each year, accounting for about 10% of all live births [6,7]. Among them, preterm birth accounts for more than 70% of neonatal deaths, and many surviving preterm infants experience a series of short- and long-term health complications, such as respiratory distress syndrome, necrotizing enterocolitis, cerebral palsy, hearing

impairment, and visual impairment [8–11]. These early and long-term complications not only pose a serious threat to the health of affected newborns but also impose a substantial economic burden on families and society. Given the severe consequences of preterm birth, the early identification and prediction of high-risk populations remain critical areas of research in obstetrics.

Although traditional statistical methods, such as logistic regression (LR), can preliminarily identify relevant factors, they struggle to capture complex nonlinear relationships between variables and cannot fully exploit multidimensional feature data for accurate prediction. Machine learning techniques, especially eXtreme Gradient Boosting (XGBoost) models, provide new opportunities for predicting complex pregnancy outcomes due to their advantages in handling high-dimensional data and intricate variable interactions.

Building on this, the present study focuses on patients diagnosed with cervical insufficiency during their first pregnancy and explores the risk factors and key determinants of premature birth and cervical insufficiency in their subsequent pregnancy. The study employed a combination of LR and XGBoost models to comprehensively evaluate the predictive value of various variables, including patients' baseline characteristics, cervical dynamic



changes, and obstetric history. At the same time, by analysis of SHapley Additive exPlanations (SHAP) values clarified the relative contributions of key risk factors, enabling the construction of a simplified risk model that provides a practical and actionable predictive model for clinical practice. The aim of this study is to systematically evaluate the risk and associated influencing factors of preterm birth during the second pregnancy in women diagnosed with cervical insufficiency during their first pregnancy. The study seeks to identify key risk factors and develop a predictive risk model to provide a scientific basis for clinical practice, support the development of personalized intervention strategies, optimize pregnancy management, reduce the incidence of preterm birth and cervical insufficiency, and improve the overall maternal and neonatal outcomes.

2. Materials and Methods

2.1 Study Cohort

This retrospective study included patients who were diagnosed with cervical insufficiency during their first pregnancy and had a second delivery at Xiamen Humanity Maternity Hospital between January 2019 and January 2024. The inclusion criteria were as follows: age ≥ 18 years; diagnosed with cervical insufficiency during the first pregnancy; completion of a subsequent second pregnancy and delivery; and availability of complete clinical data. The exclusion criteria were as follows: multiple pregnancies; severe pregnancy complications that could significantly affect pregnancy outcomes; presence of severe systemic diseases; major fetal malformations or chromosomal abnormalities; severe hepatic or renal dysfunction; severe mental illness or poor compliance; and missing critical data.

For samples with a small amount of missing data, regression imputation (or multiple imputation) based on other variables was performed. Samples with a substantial amount of missing data, such as those with more than 30% and not missing completely at random, were excluded from the analysis.

According to the inclusion criteria, a total of 550 pregnant women were initially deemed eligible for the study. After applying the exclusion criteria, a total of 480 women were ultimately included in the study. Among them, 86 experienced preterm birth, and 110 were diagnosed with cervical insufficiency during their second pregnancy. All participants were categorized into a preterm birth group and a non-preterm birth group based on the occurrence of preterm birth during their second pregnancy.

2.2 Data Collection

Patient data collection included maternal physiological and health status (age, body mass index [BMI], history of uterine cavity surgery, history of cervical conization surgery, history of uterine surgery, and use of folate and iron supplementation); relevant information from the first pregnancy (whether cervical cerclage was performed

and its type, premature rupture of membranes [PROM], pregnancy-induced hypertension, gestational diabetes, placental abruption, placenta previa, mode of delivery, occurrence of preterm birth, cervical laceration, and the degree of cervical softening or dilation); measurement of cervical length at different gestational weeks during the first pregnancy (16–18 weeks, 20–22 weeks, and 22–24 weeks); the interpregnancy interval; and fetal related characteristics (including birth weight, fetal position, and amniotic fluid index).

Cervical insufficiency: According to the *Clinical Practice Guideline No. 373 on Cervical Insufficiency and Cerclage*, published by the Society of Obstetricians and Gynaecologists of Canada (SOGC) in 2019 [12], cervical insufficiency is defined as a cervical length of less than 25 mm during the mid-trimester in the absence of clinical signs of preterm labor. It is typically characterized by painless, progressive cervical dilation in the second or third trimester, with or without PROM or protrusion of the amniotic sac through the cervix.

2.3 Feature Selection

Lasso regression was used to screen significant variables. The optimal regularization parameter (lambda) was selected through cross-validation. The optimal lambda value was 0.0047 for the preterm birth model and 0.0087 for the cervical insufficiency model.

2.4 Model Construction and Evaluation

The dataset was randomly split into a training set and a test set at a ratio of 7:3. Within the training set, 10-fold cross-validation was used to optimize the parameters of the LR model. For the XGBoost model, hyperparameters were tuned using grid search, which systematically explores combinations of hyperparameters (e.g., learning rate, maximum tree depth, and subsample ratio) to identify the optimal configuration. The final LR and XGBoost models were constructed using the optimal parameter settings. Model performance was evaluated on the test set using area under the curve (AUC), precision, accuracy, recall, and F1 score.

2.5 Key Feature Extraction and Risk Model Construction

Key predictive features with high importance in both the LR and XGBoost models were selected to construct the risk model. LR coefficients and SHAP values from the XGBoost model were standardized, and the overall risk score was calculated as:

$$\text{Risk Model} = \sum (\text{Factor}_i \times \text{Coef}_i)$$

where *Factor* represents the value of each selected feature, and *Coef* represents the standardized weight.

The predictive performance of the risk model was assessed using receiver operating characteristic (ROC) curves, with evaluation metrics such as AUC, sensitivity, specificity, optimal threshold, and the Youden Index. Ad-

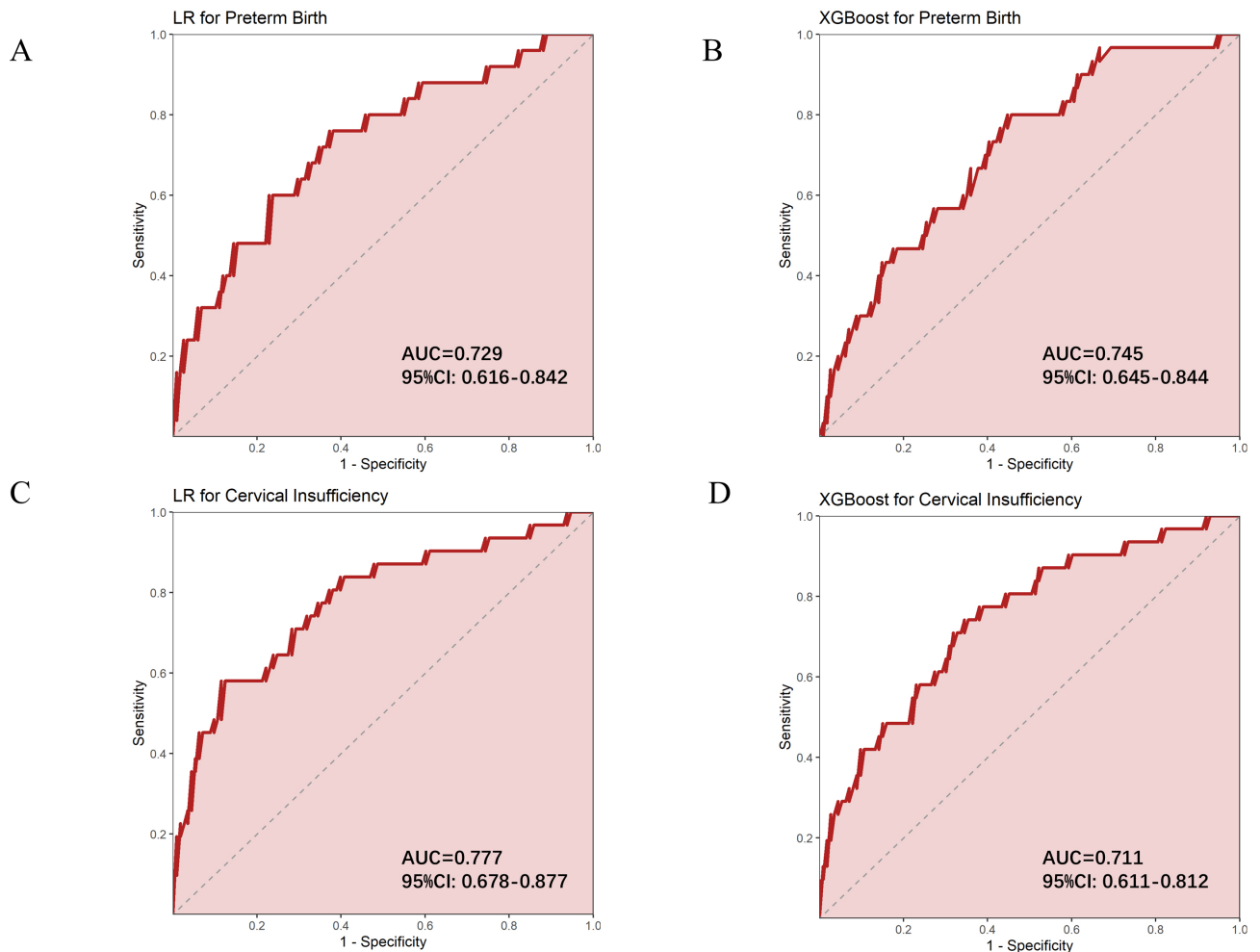


Fig. 1. ROC curves for predicting preterm birth and cervical insufficiency using different models. (A) LR model predicting preterm birth. (B) XGBoost model predicting preterm birth. (C) LR model predicting cervical insufficiency. (D) XGBoost model predicting cervical insufficiency. AUC, area under the curve; ROC, receiver operating characteristic; LR, logistic regression; XGBoost, eXtreme Gradient Boosting.

ditionally, the model's applicability to predicting the risk of both preterm birth and cervical insufficiency was evaluated.

2.6 Statistical Analysis

All analyses in this study were performed using R software version 4.4.0 (R Foundation for Statistical, Vienna, Austria). Continuous data were presented as mean $\pm$ SD or median (minimum–maximum) and compared using either the *t*-test or Mann-Whitney U test, depending on data distribution. Categorical data were expressed as frequency (percentage). For the analysis of categorical variables, Fisher's exact test was used when the theoretical frequency (*T*) was less than 1. When $1 \leq T < 5$ and the total sample size was ≥ 40 , the continuity-corrected chi-square test was applied.

3. Results

3.1 Differences in Baseline Characteristics Between Preterm and Non-Preterm Pregnant Women

Compared with pregnant women who did not experience preterm birth in their second pregnancy, pregnant those who experienced preterm birth exhibited statistically significant differences in several baseline characteristics. The incidence of a history of uterine cavity surgery was higher in the preterm birth group (13.95% vs. 6.85%, $p = 0.049$). The proportion of emergency cervical cerclage was significantly higher in the preterm birth group (6.98% vs. 2.03%, $p = 0.017$). The incidence of PROM was higher in the preterm group (20.93% vs. 10.66%, $p = 0.015$). The proportion of moderate and severe cervical lacerations was significantly higher in the preterm group compared to the non-preterm group (moderate: 25.58% vs. 17.77%; severe: 17.44% vs. 7.87%, $p = 0.002$). The proportion of moderate cervical softening or dilation abnormalities was significantly higher in the preterm birth group compared to the

Table 1. Differences in baseline characteristics between pregnant women with and without preterm birth during their second pregnancy.

	All patients	Non-preterm birth	Preterm birth	<i>p</i> -value
Maternal physiology and health status				
Age (years)	29 (21–40)	28 (21–38)	32 (22–40)	0.099
BMI (kg/m ²)	27.82 (18.55–37.58)	27.56 (18.55–37.58)	29.19 (18.77–37.30)	0.419
History of uterine cavity surgery				0.029
Yes	39 (8.12%)	27 (6.85%)	12 (13.95%)	
No	441 (91.88%)	367 (93.15%)	74 (86.05%)	
History of cervical conization				0.098
Yes	9 (1.88%)	5 (1.27%)	4 (4.65%)	
No	471 (98.12%)	389 (98.73%)	82 (95.35%)	
History of uterine surgery				0.082
Yes	20 (4.17%)	13 (3.30%)	7 (8.14%)	
No	460 (95.83%)	381 (96.70%)	79 (91.86%)	
Folic acid supplementation				0.121
Yes	291 (60.62%)	232 (58.88%)	59 (68.60%)	
No	189 (39.38%)	162 (41.12%)	27 (31.40%)	
Iron supplementation				0.176
Yes	313 (65.21%)	251 (63.71%)	62 (72.09%)	
No	167 (34.79%)	143 (36.29%)	24 (27.91%)	
Information during the first pregnancy				
Cervical cerclage				0.017
No	349 (72.71%)	284 (72.08%)	65 (75.58%)	
Prophylactic cerclage	117 (24.38%)	102 (25.89%)	15 (17.44%)	
Emergency cerclage	14 (2.92%)	8 (2.03%)	6 (6.98%)	
Pregnant women undergoing cerclage surgery				
Gestational age at cerclage placement (weeks)				
Prophylactic cerclage	13 (12–14)	13 (12–14)	13 (12–14)	0.726
Emergency cerclage	20 (16–25)	20 (16–25)	20 (16–24)	0.636
Suture type				0.228
Mersilene tape	89 (67.94%)	77 (65.81%)	12 (85.71%)	
Prolene	42 (32.06%)	40 (34.19%)	2 (14.29%)	
Suturing method				0.092
McDonald	116 (88.55%)	106 (90.60%)	10 (71.43%)	
Shirodkar	15 (11.45%)	11 (9.40%)	4 (28.57%)	
PROM				0.009
Yes	60 (12.50%)	42 (10.66%)	18 (20.93%)	
No	420 (87.50%)	352 (89.34%)	68 (79.07%)	
Gestational hypertension				0.073
Yes	12 (2.50%)	7 (1.78%)	5 (5.81%)	
No	468 (97.50%)	387 (98.22%)	81 (94.19%)	
Gestational diabetes				0.081
Yes	24 (5.00%)	16 (4.06%)	8 (9.30%)	
No	456 (95.00%)	378 (95.94%)	78 (90.70%)	
Placental abruption				0.084
Yes	3 (0.62%)	1 (0.25%)	2 (2.33%)	
No	477 (99.38%)	393 (99.75%)	84 (97.67%)	
Placenta previa				0.483
Yes	14 (2.92%)	10 (2.54%)	4 (4.65%)	
No	466 (97.08%)	384 (97.46%)	82 (95.35%)	

Table 1. Continued.

	All patients	Non-preterm birth	Preterm birth	<i>p</i> -value
Mode of delivery				0.245
Vaginal delivery	177 (36.88%)	150 (38.07%)	27 (31.40%)	
Cesarean section	303 (63.12%)	244 (61.93%)	59 (68.60%)	
Preterm birth				0.001
Yes	105 (21.88%)	74 (18.78%)	31 (36.05%)	
No	375 (78.12%)	320 (81.22%)	55 (63.95%)	
Cervical laceration				0.002
No or mild	342 (71.25%)	293 (74.37%)	49 (56.98%)	
Moderate	92 (19.17%)	70 (17.77%)	22 (25.58%)	
Severe	46 (9.58%)	31 (7.87%)	15 (17.44%)	
Failure to soften or dilate normally				0.002
No or mild	418 (87.08%)	352 (89.34%)	66 (76.74%)	
Moderate	54 (11.25%)	35 (8.88%)	19 (22.09%)	
Severe	8 (1.67%)	7 (1.78%)	1 (1.16%)	
Cervical length at 16–18 weeks (mm)	29.10 (25.80–32.60)	29.40 (25.80–32.60)	27.70 (25.80–31.40)	<0.001
Cervical length at 20–22 weeks (mm)	27.30 (23.20–30.70)	27.60 (23.20–30.70)	26.50 (23.90–30.60)	0.058
Cervical length at 22–24 weeks (mm)	26.00 (22.10–28.60)	26.20 (22.10–28.60)	25.00 (22.50–28.60)	0.004
Interpregnancy interval	32.50 ± 8.10	33.20 ± 7.90	32.25 ± 7.75	0.687
Fetal related information				
Fetal birth weight				0.484
<2500 g	22 (4.58%)	16 (4.060%)	6 (6.98%)	
2500–4000 g	426 (88.75%)	351 (89.09%)	75 (87.21%)	
>4000 g	32 (6.67%)	27 (6.85%)	5 (5.81%)	
Fetal presentation				0.045
Cephalic presentation	469 (97.71%)	388 (98.48%)	81 (94.19%)	
Breech presentation	8 (1.67%)	4 (1.02%)	4 (4.65%)	
Transverse lie	3 (0.62%)	2 (0.51%)	1 (1.160%)	
Amniotic fluid index (cm)	14.60 ± 3.65	14.45 ± 4.05	15.10 ± 3.38	0.369

BMI, body mass index; PROM, premature rupture of membranes.

non-preterm group (22.09% vs. 8.88%, $p = 0.002$). Additionally, cervical length at both 16–18 weeks and 22–24 weeks of gestation was significantly shorter in the preterm group (16–18 weeks: 27.70 vs. 29.40 mm, $p < 0.001$; 22–24 weeks: 25.00 vs. 26.20 mm, $p = 0.004$). In addition, the preterm birth group had a higher proportion of preterm birth history during the first pregnancy (36.05% vs. 18.78%, $p = 0.001$), and the incidence of abnormal fetal position (breech presentation) was significantly increased (4.65% vs. 1.02%, $p = 0.045$) (Table 1). The median gestational age at intervention was 13 weeks (range: 12–14 weeks) in the prophylactic cerclage group and 20 weeks (range: 16–25 weeks) in the emergency cerclage group. The usage rate of Mersilene tape was 67.94%, while Prolene suture was used in 32.06% of cases. The McDonald suturing technique accounted for 88.55% of procedures, whereas the Shirodkar technique was used in 11.45%. In most cases, cervical cerclage was performed independently by an experienced obstetrician. In more complex cases, two operators were involved. No cervical trauma was observed during any of the procedures in this study.

3.2 Lasso Regression Screening Input Features

The optimal lambda value for the Lasso model predicting preterm birth was 0.004735971, while the optimal lambda value for the Lasso model predicting cervical insufficiency was 0.008738651. The features selected for the Lasso model predicting preterm birth included: prophylactic cerclage, emergency cerclage, PROM, preterm birth, cervical laceration, failure of the cervix to soften or dilate normally, cervical length at 16 to 18 weeks, and cervical length at 22 to 24 weeks. The selected features for the Lasso model predicting cervical insufficiency included: history of uterine cavity surgery, prophylactic cerclage, emergency cerclage, PROM, preterm birth, cervical laceration, failure of the cervix to soften or dilate normally, fetal presentation (breech presentation), cervical length at 22 to 24 weeks, and cervical length at 16 to 18 weeks.

3.3 Model Performance Evaluation

The performance of the four models varied. For preterm birth prediction, the XGBoost model achieved an AUC of 0.745, slightly higher than the LR model's AUC of 0.729. Additionally, XGBoost outperformed LR in preci-

Table 2. ROC curve parameters of machine learning models and risk prediction models.

	AUC	AUC_CI_Lower	AUC_CI_Upper	Best_Threshold	Youden	Sensitivity	Specificity	Accuracy	Precision	Recall	F1 score
LR for preterm birth	0.729	0.616	0.842	0.442	0.387	0.760	0.627	0.889	0.706	0.800	0.749
LR for cervical insufficiency	0.777	0.678	0.877	0.388	0.466	0.581	0.885	0.847	0.833	0.595	0.695
XGBoost for cervical insufficiency	0.711	0.611	0.812	0.234	0.353	0.800	0.553	0.861	0.690	0.645	0.667
XGBoost for preterm birth	0.745	0.645	0.844	0.237	0.397	0.742	0.655	0.924	0.885	0.742	0.805
Risk model for preterm birth	0.711	0.653	0.768	40.285	0.407	0.779	0.628	-	-	-	-
Risk model for cervical insufficiency	0.686	0.628	0.745	41.039	0.396	0.733	0.663	-	-	-	-

ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval; LR, logistic regression; XGBoost, extreme gradient boosting.

Table 3. LR analysis of factors influencing preterm birth and cervical insufficiency.

		Estimate	SE	Statistic	<i>p</i> -value	OR	CI_lower	CI_upper
Preterm birth	Prophylactic cerclage	-0.479	0.235	-2.037	0.042	0.619	0.390	0.982
	Emergency cerclage	2.220	0.645	3.445	0.001	9.211	2.604	32.578
	PROM	0.666	0.274	2.427	0.015	1.946	1.137	3.333
	Preterm birth	0.782	0.226	3.456	0.001	2.185	1.403	3.403
	Cervical laceration	0.539	0.138	3.895	0.000	1.714	1.307	2.248
	Failure to soften or dilate normally	0.795	0.224	3.549	0.000	2.215	1.428	3.436
	Cervical Length at 16–18 weeks	-0.835	0.047	-2.890	0.004	0.434	0.395	0.476
	Cervical Length at 22–24 weeks	-0.841	0.053	-4.518	0.000	0.431	0.389	0.479
Cervical insufficiency	History of uterine cavity surgery	1.346	0.513	2.623	0.009	3.842	1.405	10.503
	Prophylactic cerclage	-0.161	0.367	-0.438	0.661	0.851	0.415	1.749
	Emergency cerclage	1.356	0.896	1.513	0.130	3.881	0.670	22.474
	PROM	1.666	0.400	4.161	0.000	5.290	2.414	11.594
	Preterm birth	1.489	0.329	4.528	0.000	4.431	2.326	8.439
	Cervical laceration	0.691	0.221	3.120	0.002	1.996	1.293	3.080
	Failure to soften or dilate normally	0.856	0.335	2.554	0.011	2.354	1.220	4.539
	Breech presentation	3.377	1.320	2.559	0.011	29.285	2.203	389.228
	Cervical length at 22–24 weeks	-2.190	0.083	-2.274	0.023	0.113	0.094	0.131
	Cervical length at 16–18 weeks	-3.212	0.080	-2.633	0.008	0.040	0.034	0.047

OR, odds ratio.

sion, accuracy, and F1 score, demonstrating strong predictive ability (Fig. 1A,B). In comparison, the LR model had a precision of 0.706 and an F1 score of 0.749. Although the LR model performed well overall, it was still inferior to XGBoost. For the prediction of cervical insufficiency, the LR model achieved an AUC of 0.777, demonstrating high specificity (0.885) but a relatively low recall rate (0.595), resulting in an F1 score of 0.695. The overall performance was acceptable but somewhat limited. The XGBoost model achieved an AUC of 0.711, with a moderate recall rate (0.645) but low specificity (0.553) (Fig. 1C,D). XGBoost performs slightly worse than LR in predicting cervical insufficiency; however, its higher recall rate provides a greater number of positive samples for identifying preterm birth (Table 2). Additionally, we found that although breech presentation was identified as an independent risk factor for predicting cervical insufficiency, its odds ratio (OR) was exceptionally high and accompanied by a wide confidence interval (CI). This may be due to the relatively small number of breech presentations in our dataset, leading to an unstable estimation of this variable within the model. Therefore, caution is warranted when interpreting or using this variable for predictive purposes.

3.4 Multivariate LR Analysis of Factors Affecting Preterm Birth and Cervical Insufficiency

The results of LR analysis indicate that prophylactic cervical cerclage is associated with a reduced risk of preterm birth ($OR = 0.619, p = 0.042$), suggesting that patients who undergo this procedure during their first pregnancy have a lower risk of preterm birth during their second pregnancy. Conversely, emergency cervical cerclage is significantly positively correlated with the risk of preterm birth ($OR = 9.211, p = 0.001$), indicating that patients who undergo this procedure during their first pregnancy have a higher likelihood of preterm birth during their second pregnancy. PROM also significantly increases the risk of preterm birth ($OR = 1.946, p = 0.015$). In addition, cervical laceration and failure to soften or dilate normally were also significantly associated with an increased risk of preterm birth ($OR = 1.714, p < 0.001$ and $OR = 2.215, p < 0.001$). Cervical length at 16–18 weeks and 22–24 weeks was significantly negatively correlated with preterm birth ($OR = 0.434, p = 0.004$ and $OR = 0.431, p < 0.001$, respectively), indicating that a shorter cervical length during the first pregnancy may increase the risk of preterm birth in the second pregnancy.

For predictive factors of cervical insufficiency, a history of uterine cavity surgery was significantly associated with an increased risk ($OR = 3.842, p = 0.009$). PROM also significantly increased the risk of cervical insufficiency ($OR = 5.290, p < 0.001$), as did a history of preterm birth ($OR = 4.431, p < 0.001$). Cervical laceration and failure to soften or dilate normally also had significant effects on cervical insufficiency ($OR = 1.996, p = 0.002$ and $OR = 2.354, p = 0.011$). In addition, the breech presentation significantly

increased the risk of cervical insufficiency ($OR = 29.285, p = 0.011$, respectively). The cervical length at 16–18 weeks and 22–24 weeks was negatively associated with cervical insufficiency, especially at 22–24 weeks ($OR = 0.113, p = 0.023$ and $OR = 0.040, p = 0.008$, respectively), indicating that a shorter cervical length during the first pregnancy is closely related to the occurrence of cervical insufficiency in the second pregnancy.

In summary, the occurrence of preterm birth and cervical insufficiency in the second pregnancy is influenced by cervical condition, cervical surgery, and pregnancy-related factors during the first pregnancy, with changes in cervical length playing an important role in both outcomes (Table 3).

3.5 Visual Analysis of the XGBoost Model

The XGBoost model was interpreted using SHAP visualization analysis, which revealed that for preterm birth, key predictive features include cervical length at 16–18 weeks and 22–24 weeks during the first pregnancy, cervical softening or dilation at the first delivery, cervical laceration, and a history of preterm birth in the first pregnancy. These factors significantly influence the risk of preterm birth in the subsequent pregnancy. For cervical insufficiency, key predictive factors included length of the cervix at 16–18 weeks and 22–24 weeks during the first pregnancy, history of preterm birth during the first pregnancy, cervical lacerations, and PROM. These features significantly influence the risk of cervical insufficiency in the subsequent pregnancy (Fig. 2A,B).

3.6 Construction of Risk Model

Four characteristic factors were selected to construct the risk model, namely cervical length at 16–18 weeks of the first pregnancy, cervical length at 22–24 weeks, history of preterm birth, and cervical laceration. Subsequently, the ROC curve was used to evaluate the performance of this model in predicting preterm birth and cervical insufficiency. The results showed that the risk model for preterm birth achieved an AUC of 0.711, indicating good predictive performance, with a 95% CI of 0.653 to 0.768. The optimal threshold is 40.285, Youden index is 0.407, sensitivity is 0.779, and specificity is 0.628, indicating that the model has a good balance of sensitivity and specificity in predicting premature birth. For the risk model of cervical incompetence, the AUC value is 0.686, indicating that the predictive ability of the model is moderate, with a 95% CI of 0.628 to 0.745. The optimal threshold is 41.039, with a Youden index is 0.396, sensitivity of 0.733, and specificity of 0.663. These results indicate that the model has high sensitivity but relatively lower specificity in predicting cervical insufficiency, suggesting a potential risk of false positives (Table 2) (Fig. 3A,B).

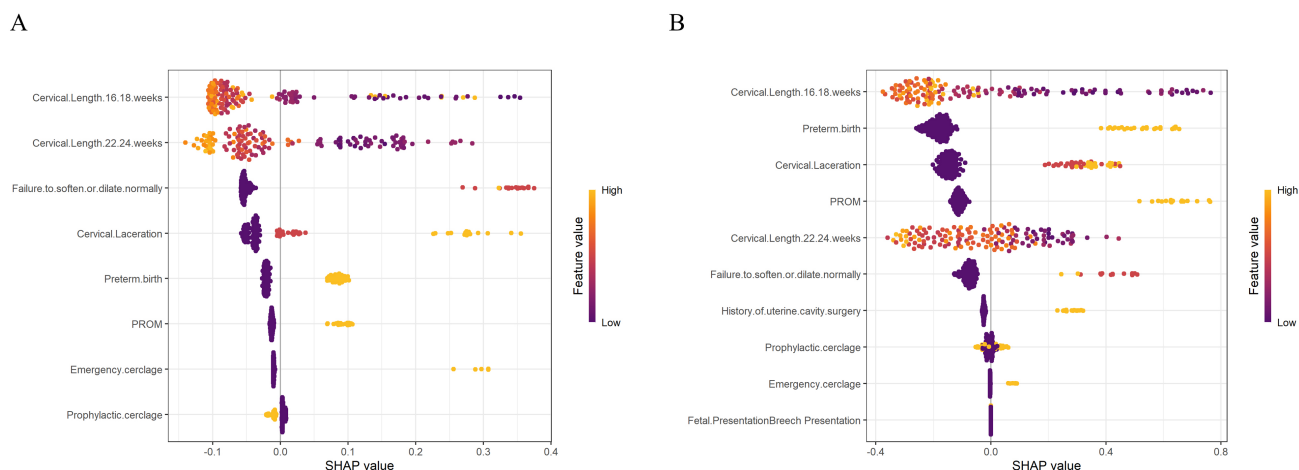


Fig. 2. SHAP visualization analysis of XGBoost models for pregnancy outcomes. (A) Prediction of preterm birth. (B) Prediction of cervical insufficiency. SHAP, SHapley Additive exPlanations.

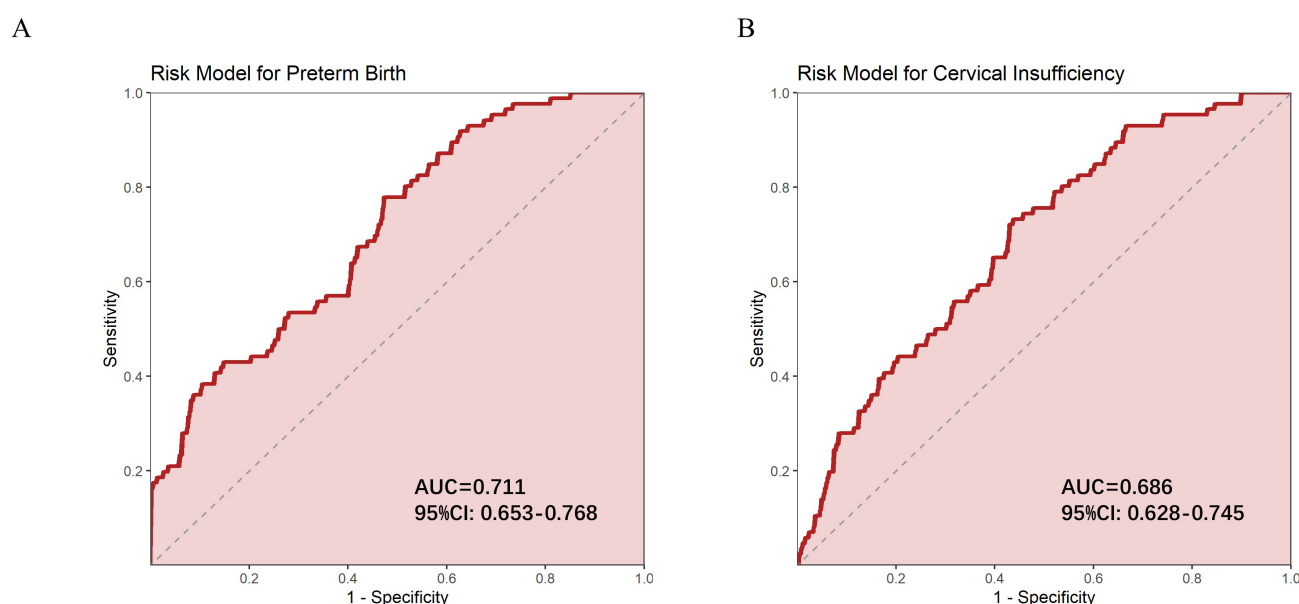


Fig. 3. ROC curves for prediction models. (A) ROC curve of risk model predicting preterm birth. (B) ROC curve for predicting cervical incompetence using a risk model. ROC, receiver operating characteristic.

4. Discussion

This study analyzed the clinical characteristics of patients with cervical insufficiency during their first pregnancy and found that multiple factors have significant predictive value for preterm birth and cervical insufficiency in a subsequent pregnancy. Based on various statistical and machine learning methods, a predictive model was developed to serve as a reference for the management of high-risk pregnancies.

Our research has found that pregnant women with cervical insufficiency who experience preterm birth in their first pregnancy may experience structural damage or functional abnormalities of the cervix, such as cervical softening, shortening, or abnormal dilation. These changes may persist into their second pregnancy, compromising the

cervix's ability to effectively support the weight of the fetus during pregnancy and thereby increasing the risk of preterm birth [13]. In addition, preterm birth in the first pregnancy may be associated with abnormal immune or physiological responses, which may persist into the second pregnancy, further increasing the risk of preterm birth [14]. Preventive cervical cerclage is negatively correlated with the risk of second preterm birth ($OR = 0.619, p = 0.042$). This may be because undergoing cerclage during the first pregnancy induces adaptive changes in the cervical fibrous structure, enhancing its tolerance to pregnancy-related stress [15]. In the second pregnancy, these adaptive changes may enable the cervix to remain closed more effectively, thereby prolonging gestation and reducing the risk of preterm birth. Compared with preventive cervical cerclage, emergency cervi-

cal cerclage is significantly correlated an increased risk of preterm birth ($OR = 9.211$, $p = 0.001$), indicating that patients who undergo this procedure during their first pregnancy have a higher risk of preterm birth during their second pregnancy [16,17]. This finding may reflect that emergency cervical cerclage is usually performed in cases of severe cervical dysfunction, indicating that these patients already had significant cervical issues during the first pregnancy, which in turn increases the risk of preterm birth in the second pregnancy [18].

Cervical length measured at 16–18 weeks and 22–24 weeks is significantly negatively correlated with preterm birth ($OR = 0.434$, $p = 0.004$ and $OR = 0.431$, $p < 0.001$, respectively), especially at 16–18 weeks where cervical length shows a significant impact on cervical insufficiency ($OR = 0.040$, $p = 0.008$). This suggests that a shorter cervical length during the first pregnancy may indicate the occurrence of preterm birth or cervical insufficiency in the second pregnancy [19]. This may be because a shorter cervical length can reflect underlying anatomical or functional abnormalities of the cervix, such as fragile cervical tissue structure, inflammatory infections, hormonal dysregulations, or mechanical damage, which can impair the cervix's ability to withstand the pressure of pregnancy and potentially lead to permanent damage [20]. In addition, a shorter cervix may cause excessive tension on the fetal membranes during pregnancy, further exacerbating cervical damage. These issues may persist into the second pregnancy, rendering the cervix more susceptible to preterm birth or functional impairment.

Cervical lacerations during the first pregnancy and delivery significantly increase the risk of preterm birth and cervical insufficiency. This may be due to structural damage and weakening of the cervix caused by lacerations, reducing its ability to effectively withstand the pressure exerted by the fetus and amniotic fluid in later stages of pregnancy [21]. In addition, the healing process of cervical lacerations may lead to the formation of scar tissue with reduced elasticity, further compromising the ability of the cervix to remain close. Laceration may also cause local inflammation or infection, disrupt the cervical defense barrier, and increase the risk of PROM or infection-related preterm birth [22]. A damaged cervix may be more sensitive to hormonal changes and uterine contractions, making it more prone to premature softening or dilation, ultimately leading to premature birth or cervical insufficiency.

In the performance evaluation of multiple models, this study found that XGBoost outperformed LR models in predicting premature birth, highlighting the potential of machine learning methods in handling multivariate and complex nonlinear relationships. In contrast, LR models perform better in predicting cervical insufficiency, possibly due to a clearer linear relationship between the outcome variable and the input features. The model interpretation based on SHAP values shows that cervical length, preterm

birth in the first pregnancy, and cervical laceration are the most influential features across multiple models, corroborating both clinical experience and statistical analysis.

In this study, we developed a risk prediction model by combining LR and XGBoost approaches. This model comprehensively analyzes both linear and nonlinear relationships among features, ensuring the robustness of its predictive outcomes. Furthermore, the model is capable of predicting both the risk of preterm birth and cervical insufficiency, demonstrating its broad clinical utility. The selected features are readily obtainable in clinical practice. Clinicians can input data from a patient's first pregnancy—including cervical length, history of preterm birth, and presence of cervical laceration—into the risk model to calculate an individual risk score. For patients with higher risk scores, preventive measures can be implemented in subsequent pregnancies to reduce the risk of preterm birth and cervical insufficiency. This model can aid in the early identification and management of high-risk pregnant women. However, given that the AUC value indicates moderate predictive ability, the model should not be solely relied upon or clinical decision-making. In specific clinical practice, the potential for misdiagnosis or missed diagnosis should also be considered. Therefore, clinicians should use the model as a reference tool and make more comprehensive judgments and decisions in conjunction with their clinical experience.

This study has several limitations. First, its retrospective design, which may introduce bias in data selection. Secondly, the sample size is limited, and not all potential factors that may influence preterm birth or cervical insufficiency, such as uterine anomalies, previous full dilation cesarean section, fibroids, and polyhydramnios, were fully considered. Future research should consider including additional factors and conducting larger-scale, prospective randomized controlled trials to broaden and deepen the scope of the study.

5. Conclusions

This study identified cervical length at 16–18 weeks and 22–24 weeks, a history of preterm birth in the first pregnancy, and cervical laceration as key predictive factors for preterm birth and cervical insufficiency in the second pregnancy. The constructed LR and XGBoost models demonstrate good predictive performance, with XGBoost showing superior accuracy in predicting preterm birth. Lastly, the risk scoring model, integrating key predictive factors, demonstrated superior performance in predicting preterm birth and cervical insufficiency.

Availability of Data and Materials

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Author Contributions

XL, JJ, and YW designed the research study. XL and JJ performed the research. YW provided help and advice on data analysis. XL analyzed the data. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The study has been approved by the Ethics Review Committee of Xiamen Humanity Maternity Hospital with the approval number 20024-B088, and strictly adhered to the Declaration of Helsinki. All participants provided written informed consent before participating in the study.

Acknowledgment

This study sincerely thanks all the medical staff of Xiamen Humanity Maternity Hospital for their invaluable support in data collection and patient care.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Roman A, Suhag A, Berghella V. Overview of Cervical Insufficiency: Diagnosis, Etiologies, and Risk Factors. *Clinical Obstetrics and Gynecology*. 2016; 59: 237–240. <https://doi.org/10.1097/GRF.0000000000000184>.
- [2] Han Y, Li M, Ma H, Yang H. Cervical insufficiency: a noteworthy disease with controversies. *Journal of Perinatal Medicine*. 2020; 48: 648–655. <https://doi.org/10.1515/jpm-2020-0255>.
- [3] Liu Y, Ke Z, Liao W, Chen H, Wei S, Lai X, *et al.* Pregnancy outcomes and superiorities of prophylactic cervical cerclage and therapeutic cervical cerclage in cervical insufficiency pregnant women. *Archives of Gynecology and Obstetrics*. 2018; 297: 1503–1508. <https://doi.org/10.1007/s00404-018-4766-9>.
- [4] Lotgering FK. Clinical aspects of cervical insufficiency. *BMC Pregnancy and Childbirth*. 2007; 7: S17. <https://doi.org/10.1186/1471-2393-7-S1-S17>.
- [5] Marchand GJ, Masoud AT, Galitsky A, Sainz K, Azadi A, Ware K, *et al.* Complications of Laparoscopic and Transabdominal Cerclage in Patients with Cervical Insufficiency: A Systematic Review and Meta-analysis. *Journal of Minimally Invasive Gynecology*. 2021; 28: 759–768.e2. <https://doi.org/10.1016/j.jmig.2020.11.014>.
- [6] Vogel JP, Chawanpaiboon S, Moller AB, Watananirun K, Bonet M, Lumbiganon P. The global epidemiology of preterm birth. *Best Practice & Research. Clinical Obstetrics & Gynaecology*. 2018; 52: 3–12. <https://doi.org/10.1016/j.bpobgyn.2018.04.003>.
- [7] Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet (London, England)*. 2008; 371: 75–84. [https://doi.org/10.1016/S0140-6736\(08\)60074-4](https://doi.org/10.1016/S0140-6736(08)60074-4).
- [8] Kachikis A, Walker CL, McAdams RM, Gyamfi-Bannerman C, Adams Waldorf KM. Phenotypic overlap in neonatal respiratory morbidity following preterm premature rupture of membranes versus spontaneous preterm labor. *The Journal of Maternal-fetal & Neonatal Medicine: the Official Journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians*. 2021; 34: 1941–1948. <https://doi.org/10.1080/14767058.2019.1651280>.
- [9] Ream MA, Lehwald L. Neurologic Consequences of Preterm Birth. *Current Neurology and Neuroscience Reports*. 2018; 18: 48. <https://doi.org/10.1007/s11910-018-0862-2>.
- [10] Allen MC. Neurodevelopmental outcomes of preterm infants. *Current Opinion in Neurology*. 2008; 21: 123–128. <https://doi.org/10.1097/WCO.0b013e3282f88bb4>.
- [11] Zhu X, Lei X, Dong W. Change to hearing loss-related risks and screening in preterm infants. *American Journal of Perinatology*. 2022; 39: 501–512. <https://doi.org/10.1055/s-0040-1717071>.
- [12] Brown R, Gagnon R, Delisle M. No. 373-Cervical insufficiency and cervical cerclage. *Journal of Obstetrics and Gynaecology Canada*. 2019; 41: 233–247. <https://doi.org/10.1016/j.jogc.2018.08.009>.
- [13] Abdel-Aleem H, Shaaban OM, Abdel-Aleem MA, Aboelfadle Mohamed A. Cervical pessary for preventing preterm birth in singleton pregnancies. *The Cochrane Database of Systematic Reviews*. 2022; 12: CD014508. <https://doi.org/10.1002/14651858.CD014508>.
- [14] Vink J, Feltoovich H. Cervical etiology of spontaneous preterm birth. *Seminars in Fetal & Neonatal Medicine*. 2016; 21: 106–112. <https://doi.org/10.1016/j.siny.2015.12.009>.
- [15] Giouleka S, Boureka E, Tsakiridis I, Siargkas A, Mamopoulos A, Kalogiannidis I, *et al.* Cervical Cerclage: A Comprehensive Review of Major Guidelines. *Obstetrical & Gynecological Survey*. 2023; 78: 544–553. <https://doi.org/10.1097/OGX.0000000000001182>.
- [16] Hulshoff CC, Bosgraaf RP, Spaanderman MEA, Inthout J, Scholten RR, Van Drongelen J. The efficacy of emergency cervical cerclage in singleton and twin pregnancies: a systematic review with meta-analysis. *American Journal of Obstetrics & Gynecology MFM*. 2023; 5: 100971. <https://doi.org/10.1016/j.ajogmf.2023.100971>.
- [17] Wierzychowska-Opoka M, Kimber-Trojnar Ż, Leszczyńska-Gorzela B. Emergency Cervical Cerclage. *Journal of Clinical Medicine*. 2021; 10: 1270. <https://doi.org/10.3390/jcm10061270>.
- [18] Chatzakis C, Efthymiou A, Sotiriadis A, Makrydimas G. Emergency cerclage in singleton pregnancies with painless cervical dilatation: A meta-analysis. *Acta Obstetrica et Gynecologica Scandinavica*. 2020; 99: 1444–1457. <https://doi.org/10.1111/aogs.13968>.
- [19] Son M, Miller ES. Predicting preterm birth: Cervical length and fetal fibronectin. *Seminars in Perinatology*. 2017; 41: 445–451. <https://doi.org/10.1053/j.semperi.2017.08.002>.
- [20] Palacio M, Cobo T, Bosch J, Filella X, Navarro-Sastre A, Ribes A, *et al.* Cervical length and gestational age at admission as predictors of intra-amniotic inflammation in preterm labor with intact membranes. *Ultrasound in Obstetrics & Gynecology: the Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*. 2009; 34: 441–447. <https://doi.org/10.1002/uog.6437>.
- [21] Wong LF, Wilkes J, Korgenski K, Varner MW, Manuck TA. Intrapartum Cervical Laceration and Subsequent Pregnancy Outcomes. *AJP Reports*. 2016; 6: e318–23. <https://doi.org/10.1055/s-0036-1592198>.
- [22] Hossein-Pour P, Rajasingham M, Muraca GM. Risk of cervical laceration in forceps vs vacuum delivery: A systematic review and meta-analysis. *Acta Obstetrica et Gynecologica Scandinavica*. 2025; 104: 29–38. <https://doi.org/10.1111/aogs.14969>.