Am J Transl Res 2026;18(1):154-166
www.ajtr.org /ISSN:1943-8141/AJTR0168813

Original Article

Establishment and validation of a prediction model for successful trial of labor after cesarean in scarred uteruses

Xueqin Han[1*], Juhua Ding[1*], Gang Li[2], Min Su[3], Dong Yin[1], Gong Chen[1], Huiling Liu[1], Wenli Zhou[1], Xiaoqiong Li[1]

[1]Department of Obstetrics, Huai'an Maternal and Child Health Hospital, Huai'an 223000, Jiangsu, China; [2]Department of Obstetrics, Peking University People's Hospital Qingdao Hospital, Qingdao 266109, Shandong, China; [3]Department of Obstetrics and Gynecology, Affiliated Hospital of Nantong University, Nantong 226000, Jiangsu, China. [*]Equal contributors and co-first authors.

Received September 15, 2025; Accepted December 2, 2025; Epub January 15, 2026; Published January 30, 2026

Abstract: Objective: To develop and validate a prediction model for successful trial of labor after cesarean (TOLAC) in women with a scarred uterus using clinical data from a single center with an external temporal validation set. Methods: This retrospective study developed a prediction model using data from a single center (Huai'an Maternal and Child Health Hospital) with scarred uterus planning TOLAC, collected from January 2016 to December 2023. Assessed variables included maternal characteristics, obstetric history, and prenatal examination findings. Univariate and multivariate logistic regression were used to identify predictors and develop the model. The model was temporally validated using an independent external dataset from two other tertiary hospitals. Its performance was evaluated by the area under the ROC curve (AUC), sensitivity, and specificity, and was compared with the Grobman model. Results: Among 2,386 eligible women, 1,721 (72.1%) had a successful TOLAC. Univariate analysis identified 10 significant variables (P<0.05). Multivariate analysis retained six independent predictors: age, body mass index, history of vaginal delivery, cervical score, estimated fetal weight, and gestational age. The model expression was: Logit(P) = -3.82 + 0.04 × age - 0.12 × BMI + 1.56 × vaginal delivery history + 0.37 × cervical score - 0.02 × fetal weight + 0.18 × gestational age. Internal validation showed an AUC of 0.85 (95% CI: 0.82-0.88), sensitivity of 82.3%, and specificity of 78.5%. Temporal Validation (n=524) yielded an AUC of 0.83 (95% CI: 0.79-0.87), which was significantly higher than the Grobman model (AUC=0.79, P<0.05). Conclusions: The developed prediction model demonstrates good performance and generalizability for predicting TOLAC success, potentially aiding clinical decision-making and improving maternal and neonatal outcomes.

Keywords: Trial of labor after cesarean, scarred uterus, predictive model, vaginal trial of labor, logistic regression

Introduction

With the annual global rise in cesarean section rates, a scarred uterus has become an undeniable challenge in modern obstetric practice [1]. Data from the World Health Organization show that cesarean section rates in some countries are as high as 40% or more [2]. Although China has managed to reduce its cesarean section rate through clinical standard management in recent years, the number of pregnant women with a scarred uterus continues to grow [3, 4]. For these women, trial of labor after cesarean (TOLAC) is recognized as an important strategy to lower the repeat cesarean rate and reduce maternal and neonatal complications [5]. Research indicates that successful TOLAC can significantly decrease the risk of postpartum hemorrhage, infection, and neonatal respiratory complications [6]. However, failed trials may lead to severe consequences such as uterine rupture and emergency cesarean sections, with a low but life-threatening incidence of uterine rupture at approximately 0.5% to 1.0% [7].

Currently, taking into account factors like maternal age, obstetric history, and cervical conditions, clinical decisions for TOLAC mainly rely on physicians' experiential judgment [8]. This subjective evaluation leads to significant differenc-

https://doi.org/10.62347/TRPI8982

es in TOLAC success rates across different medical institutions (ranging from 50% to 80%), highlighting inadequacies in the risk prediction system [9]. Although internationally established TOLAC prediction models, such as the Grobman model, provide quantitative tools for clinical use, they are primarily built on Western population data [10]. Variables, such as indications for previous cesarean section and delivery intervals, differ from the clinical characteristics of Chinese pregnant women. For instance, average BMI, the proportion of vaginal delivery history, and pregnancy management models among Chinese women are notably different from those in Western populations, leading to limited predictive performance of these models in China (AUC mostly between 0.75 and 0.80) [11]. Moreover, existing models often depend on single-center small sample data lacking multicenter validation, and thus their generalization ability requires further verification.

The core of accurate prediction of TOLAC outcomes lies in identifying key influencing factors and establishing stable quantitative models [12]. Previous studies, including those conducted in China, have confirmed that maternal age, BMI, history of vaginal delivery, and cervical maturity are the important factors affecting TOLAC success [13]. However, prediction models that systematically integrate these factors to quantify their synergistic effects and relative weights are primarily derived from Western populations [14]. However, current research on prediction models for TOLAC targeting scarred uteruses in China remains scarce, lacking rigorous internal validation and external independent dataset validation tools [15], and their performance in the Chinese population has been limited [16]. Therefore, a model developed and validated specifically for Chinese pregnant women is warranted. Given this situation, this study aims to utilize data from a single center to construct and validate a TOLAC success prediction model suitable for Chinese pregnant women with a scarred uterus. The model's effectiveness will be assessed through internal cross-validation and external independent sample validation, ultimately providing clinicians with precise and reliable TOLAC risk assessment tools to promote standardized and individualized management of pregnancies with a scarred uterus.

Materials and methods

Study subjects

This retrospective study included 2,386 eligible subjects. Among them 1,862 cases were collected from January 2016 to December 2021 in the training set and 524 cases were collected from January 2022 to December 2023 in the temporal Validation set. All of whom were pregnant women with a scarred uterus and planned to undergo TOLAC. The data for model development (training set) were collected from a single center: Huai'an Maternal and Child Health Hospital. The independent temporal validation set was derived from Peking University People's Hospital Qingdao Hospital and the Affiliated Hospital of Nantong University. All procedures involving human participants in this study were conducted in accordance with the Declaration of Helsinki (revised in 2013). The study was approved by Ethics Committee of Huai'an Maternal and Child Health Hospital.

Inclusion and exclusion criteria

Inclusion criteria: (1) A single prior cesarean section (uterine lower segment transverse incision preferred, with no history of incision extension or infection); (2) Singleton pregnancy with cephalic presentation in the current gestation; (3) Gestational age between 37 and 42 weeks; (4) Planned to undergo TOLAC; (5) Complete clinical data available for analysis.

Exclusion criteria: (1) History of classical cesarean section (corporal incision) or longitudinal uterine incision; (2) Severe pregnancy complications (e.g., preeclampsia, placenta accrete, severe gestational diabetes mellitus); (3) Fetal malformations or abnormal fetal presentations (e.g., breech, transverse lie); (4) Coexisting uterine conditions in the current pregnancy (e.g., uterine fibroids, uterine malformations, uterine scar diverticulum with depth >2 mm); (5) Inability to complete follow-up or incomplete clinical data.

Sample size estimation

The sample size was estimated using the formula for a single rate: $n = Z^2 \times P(1 - P)/d^2$, where Z=1.96 (α=0.05), and d=0.03 (allowable error). The assumed success rate of TOLAC (P) was set at 70%, which was consistent with the

rates reported in previous studies conducted in the Chinese population (ranging from 60% to 80%) [17]. Based on this, the minimum required sample size was 897 subjects. To enhance the robustness of the multivariate logistic regression model and account for a potential 10% loss to follow-up rate, 2,386 cases were finally included. Data Collection Basic information included the pregnant woman's age, body mass index, obstetric history, prenatal examination, placental position etc.

Statistical analysis

All statistical analyses were performed using SPSS 26.0 and R 4.2.0 software. A two-sided *P*-value of <0.05 was considered statistically significant. Continuous variables were presented as mean ± standard deviation, and categorical variables as frequency and percentage (n, %).

Variable selection and model development

Univariate analyses were conducted to identify variables associated with TOLAC success. Continuous variables were compared using the independent samples t-test or Mann-Whitney U test, as appropriate. Categorical variables were compared using the Chi-square test or Fisher's exact test. Variables with a *P*-value <0.05 in the univariate analysis were considered candidates for the multivariate model.

Subsequently, all variables with P<0.05 in the univariate analysis were included as candidates in the multivariate logistic regression model. A backward stepwise selection procedure was then employed, with a removal criterion of P>0.10, to identify independent predictors and build the final parsimonious model. This approach ensured that only variables retaining statistical significance in the presence of other predictors were included in the final model. The model was presented as $Logit(P) = \beta_0 + \beta_1X_1 + ... + \beta_kX_k$. Multicollinearity among the included variables was assessed using the variance inflation factor (VIF), with a VIF<5 indicating no significant collinearity.

The variable selection process followed a two-stage approach: (1) Univariate screening: All clinically relevant variables were first assessed using univariate analysis, with P<0.05 considered statistically significant for inclusion in the multivariate candidate pool. (2) Multivariate refinement: Candidate variables from the univariate analysis were entered into a multivariate logistic regression model, and backward stepwise elimination was performed to remove variables that did not maintain independent predictive significance (P>0.10). This methodology ensured that the final model included only the most clinically relevant and statistically robust predictors while minimizing overfitting.

Model validation and performance assessment

The model's performance was evaluated in terms of discrimination, calibration, and clinical utility. At the same time, the prediction performance of this model was compared with the existing Grobman model to evaluate its relative advantages. To ensure a fair comparison, the Grobman model was applied to our temporal Validation set. The probability of successful TOLAC for each patient in the validation set was calculated using the published Grobman nomogram or regression formula [18]. Subsequently, the AUC, sensitivity, and specificity of the Grobman model were calculated and compared with those of our model using the same statistical methods (e.g., DeLong's test for AUC). These performance metrics (discrimination, calibration, and clinical utility) were evaluated on both the internal training set and the temporal Validation set to comprehensively assess the model's generalizability.

Internal validation: The model's stability was assessed via 10-fold cross-validation on the training set.

Temporal validation: The generalizability of the final model was tested on an independent temporal Validation set (n=524). On the temporal Validation set, DeLong's test was used to compare the AUC between the proposed model and the Grobman model. The training set comprised data from Huai'an Maternal and Child Health Hospital. The independent temporal validation set was derived from Peking University People's Hospital Qingdao Hospital and the Affiliated Hospital of Nantong University to assess the model's generalizability.

Discrimination was quantified by the area under the receiver operating characteristic curve (AUC) with a 95% confidence interval. Sensitivity

Table 1. Comparison of baseline data between the training set and the validation set (n=2,386)

Variable type	Index	Training set (n=1,862)	Validation set (n=524)	χ^2/t	*P* value
Demographic characteristics					
	Age (years, Mean ± SD)	28.8±4.3	28.6±4.2	0.871	0.384
	BMI (kg/m^2, Mean ± SD)	24.1±3.0	24.3±2.9	-1.052	0.293
	History of vaginal delivery (Yes, (n, %))	825 (44.3%)	231 (44.1%)	0.011	0.917
Prenatal examination					
	Cervical score (points, Mean ± SD)	6.8±1.5	6.7±1.4	1.231	0.219
	Estimated fetal weight (g, Mean ± SD)	3205±350	3190±345	0.761	0.447
	Gestational age (weeks, Mean ± SD)	38.2±1.1	38.1±1.0	1.452	0.147
	Late pregnancy systolic BP (mmHg, Mean ± SD)	118.42±9.85	119.10±10.20	1.120	0.263
Placental position, (n, %)				0.062	0.970
	Anterior	745 (40.01%)	210 (40.08%)		
	Posterior	892 (47.90%)	252 (48.09%)		
	Lateral	225 (12.08%)	62 (11.83%)		
Number of previous cesarean sections, (n, %)				0.183	0.671
	1 time	1,780 (95.60%)	502 (95.80%)		
	≥2 times	82 (4.40%)	22 (4.20%)		
Gestational diabetes mellitus (current pregnancy), (n, %)				0.031	0.862
	Yes	298 (16.00%)	85 (16.22%)		
	No	1,564 (84.00%)	439 (83.78%)		
Outcome	TOLAC success (n, %)	1326 (71.2%)	395 (75.4%)	2.891	0.089

Note: Continuous variables are expressed as Mean ± SD, and the t-test is used for comparison between groups; categorical variables are expressed as n (%), and the χ^2 test is used for comparison between groups. The training set is the internal cross-validation sample, and the validation set is a temporal validation sample. Abbreviations: BMI, body mass index; SD, standard deviation; BP, blood pressure; TOLAC, trial of labor after cesarean.

and specificity were calculated based on the optimal cutoff determined by the Youden index. On the temporal Validation set, DeLong's test was used to compare the AUC of our model with the Grobman model.

Calibration was assessed using the Hosmer-Lemeshow test and visualized with a calibration plot.

Clinical utility was evaluated using decision curve analysis (DCA) to estimate the net benefit across a range of threshold probabilities.

Results

Basic characteristics of study subjects

There were no statistically significant differences between the training set and the validation set in baseline data such as age, BMI, history of vaginal delivery, cervical score, estimated fetal weight, and gestational week (all P>0.05), showing good consistency (**Table 1**). Among all subjects, 1,721 cases had successful TOLAC, achieving a success rate of 72.1%; 665 cases failed, with a failure rate of 27.9%.

Univariate analysis results

Univariate analysis showed that ten variables including age, BMI, history of vaginal delivery, cervical score, estimated fetal weight, gestational week, late pregnancy systolic blood pressure, placental position, previous cesarean se ction times, and gestational diabetes mellitus in this pregnancy were related to the success of TOLAC (all P<0.05). Among these, key indicator differences between the success group and the failure group are as follows. The average age in the success group was lower than that in the failure group (28.5±4.1 years vs 29.3±4.5 years), BMI was lower (23.9±2.8 kg/m^2 vs 25.1±3.3 kg/m^2), the proportion of history of vaginal delivery was higher (51.8% vs 24.7%), the cervical score was higher (7.2±1.3 points vs 5.1±1.2 points), the estimated fetal weight was lower (3180±320 g vs 3350±380 g),

Table 2. Definitions and data types of candidate variables screened for univariate analysis

Category	Variable Name	Definition/Unit	Data Type
Demography	Age	Years	Continuous variable
	BMI	Weight/height^2 (kg/m^2)	Continuous variable
Obstetric history	History of vaginal delivery	Yes/No (Yes =1/No =0)	Binary variable
Prenatal examination	Cervical score	0-10 points (higher score indicates more mature cervix)	Continuous variable
	Estimated fetal weight	Grams	Continuous variable
	Gestational age	Weeks	Continuous variable
	Third-trimester blood pressure (systolic blood pressure)	mmHg	Continuous variable
	Placental position	Anterior wall/Posterior wall/Lateral wall	Categorical variable
Medical history	Number of previous cesarean sections	1 time/≥2 times	Categorical variable
	Gestational diabetes in current pregnancy	Yes/No	Binary variable

Abbreviations: BMI, body mass index; TOLAC, trial of labor after cesarean.

Table 3. Results of univariate analysis between the TOLAC success group and the failure group

Variable	Success group (n=1,721)	Failure group (n=665)	χ^2/t	*P* value
Age (years, Mean ± SD)	28.5±4.1	29.3±4.5	t-3.261	<0.001
BMI (kg/m^2, Mean ± SD)	23.9±2.8	25.1±3.3	-6.180	<0.001
History of vaginal delivery (Yes, n,%)	892 (51.8%)	164 (24.7%)	156.321	<0.001
Cervical score (points, Mean ± SD)	7.2±1.3	5.1±1.2	32.571	<0.001
Estimated fetal weight (g, Mean ± SD)	3180±320	3350±380	-8.743	<0.001
Gestational age (weeks, Mean ± SD)	38.3±1.0	37.6±1.2	9.522	<0.001

Abbreviations: BMI, body mass index; SD, standard deviation; TOLAC, trial of labor after cesarean.

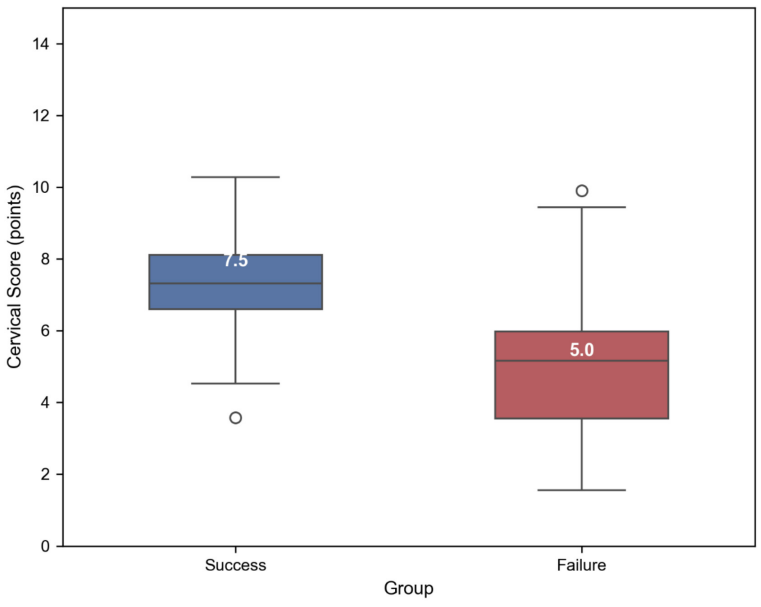

Figure 1. Cervical scores for successful vs failed TOLAC. Abbreviations: TOLAC, trial of labor after cesarean.

and the gestational week was larger (38.3±1.0 weeks vs 37.6±1.2 weeks) (**Tables 2**, **3**). **Figure 1** shows that the median cervical score in the TOLAC success group was 7.5 points, significantly higher than 5.0 points in the failure group, intuitively reflecting the impact of cervical maturity on TOLAC outcomes.

Construction of the multivariate logistic regression model

Based on univariate analysis, after multivariate logistic regression analysis, six independent predictors were ultimately included: age, BMI, history of vaginal delivery, cervical score, estimated fetal weight, and

Table 4. Six key variables and coefficients of multivariate Logistic regression

Variable	Coefficient	Standardized Coefficient	OR Value	95% CI
Age	0.04	0.08	1.04	1.01-1.07
BMI	-0.12	-0.15	0.89	0.85-0.93
History of vaginal delivery (Yes =1)	1.56	0.28	4.76	4.01-5.65
Cervical score	0.37	0.22	1.45	1.36-1.55
Estimated fetal weight (g)	-0.02	-0.11	0.98	0.97-0.99
Gestational age	0.18	0.13	1.20	1.13-1.27

Model expression: Logit(P) = -3.82 + 0.04 × Age - 0.12 × BMI + 1.56 × History of vaginal delivery + 0.37 × Cervical score - 0.02 × Estimated fetal weight + 0.18 × Gestational age. Abbreviations: BMI, body mass index; OR, odds ratio; CI, confidence interval; TOLAC, trial of labor after cesarean.

Table 5. Model performance indicators and comparison with existing models

Model	Dataset	AUC (95% CI)	Sensitivity	Specificity	Youden Index	*P* value
Logistic model of this study	Internal validation set (n=1,862)	0.85 (0.82-0.88)	82.3%	78.5%	0.608	-
	Temporal Validation set (n=524)	0.83 (0.79-0.87)	81.0%	77.2%	0.582	-
Grobman model	Temporal Validation set (n=524)	0.79 (0.75-0.83)	76.5%	72.0%	0.485	<0.05

Abbreviations: AUC, area under the curve; CI, confidence interval; TOLAC, trial of labor after cesarean.

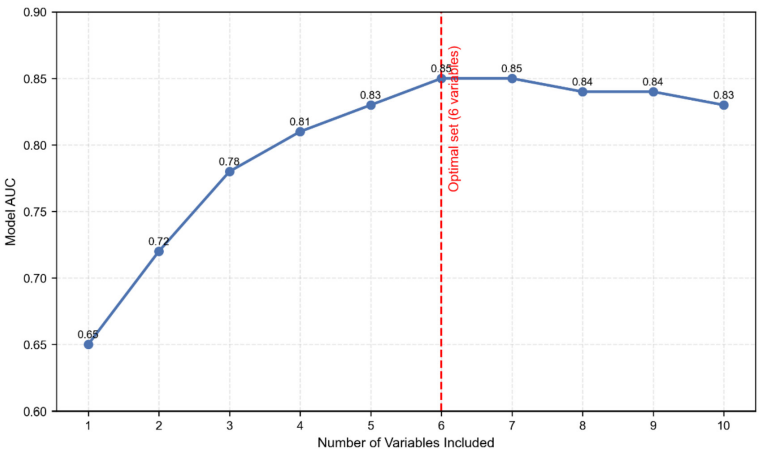

Figure 2. Variable selection trajectory (stepwise regression). Abbreviations: AUC, area under the curve.

gestational week. The model expression is: Logit(P) = -3.82 + 0.04 × age - 0.12 × BMI + 1.56 × history of vaginal delivery (yes =1/no =0) + 0.37 × cervical score - 0.02 × estimated fetal weight + 0.18 × gestational week. The coefficients, standardized coefficients, OR values, and 95% CIs of each variable are detailed in **Table 4**. Among them, the OR value of history of vaginal delivery was the highest (4.76, 95% CI: 4.01-5.65), indicating it as the strongest predictor.

Model validation results

Internal validation: The internal validation of the model was performed using 10-fold cross-validation, showing an AUC of 0.85 (95% CI: 0.82-0.88), sensitivity of 82.3%, specificity of 78.5%, and Youden's index of 0.608 (**Table 5**). **Figure 2** demonstrates that the AUC of the model combining six variables is 0.85, significantly higher than the predictive efficacy of a single variable. The calibration curve (**Figure 3A**) shows good consistency between the predicted probability and actual TOLAC success rate (Hosmer-Lemeshow test, P=0.587). Decision curve analysis (**Figure 3B**) indicated that within the threshold probability range of 0.1-0.8, the net benefit of this model was superior to both the "treat all" and "treat none" strategies. Specifically, at clinically relevant threshold probabilities of 30-60%, the model provided a net benefit of 0.15-0.25, meaning that using the model to guide clinical decisions would yield the equivalent of 15-25 additional appropriate TOLAC decisions per

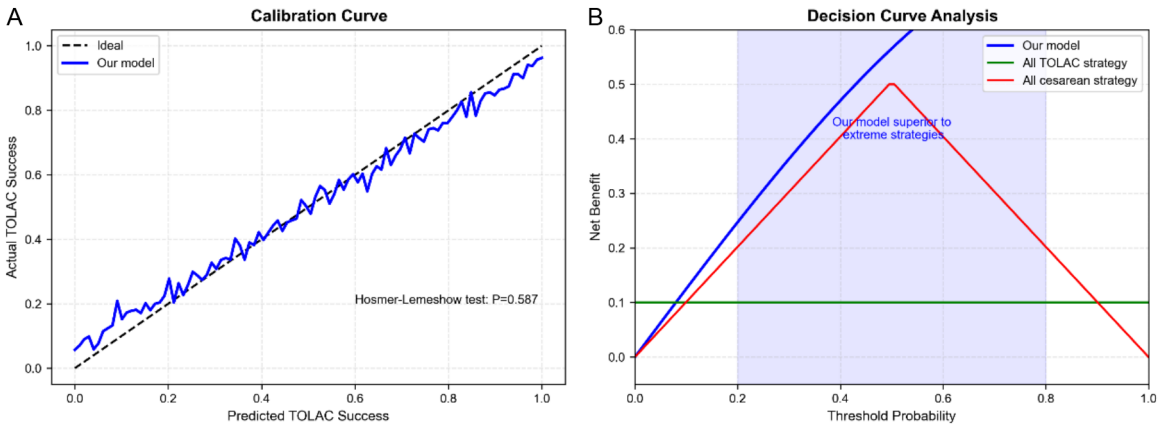

Figure 3. Calibration Curve (A) and Decision Curve (B) for Logistic Model. (A) The calibration curve shows the agreement between the predicted probability of successful TOLAC and the observed actual success rate. The dashed line denotes perfect calibration, and the solid line represents the model's calibration. The Hosmer-Lemeshow test (P=0.587) indicates no significant deviation between predicted and observed outcomes, confirming good model calibration. (B) The decision curve evaluates the model's clinical utility across threshold probabilities. The blue line represents the proposed model, with red and green lines for "all TOLAC" and "all cesarean" strategies, respectively. The model has superior net benefit over the two extreme strategies within 0.1-0.8 threshold probabilities; specifically, it provides a 0.15-0.25 net benefit at the clinically relevant 0.3-0.6 threshold (15-25 additional appropriate decisions per 100 patients). Abbreviations: TOLAC, trial of labor after cesarean.

100 patients compared to the extreme strategies, without increasing the rate of adverse outcomes.

Temporal validation: In 524 independent external samples, the model's AUC was 0.83 (95% CI: 0.79-0.87), sensitivity was 81.0%, specificity was 77.2%, and Youden's index was 0.582 (**Table 5**). The nomogram (**Figure 4**) visually presents the corresponding relationship between scores of each variable and the probability of TOLAC success, facilitating quick estimation of predictions in clinical settings.

The calibration curve for the temporal Validation set (**Figure 5A**) demonstrated good agreement between predicted and observed probabilities (Hosmer-Lemeshow test, P=0.452). Decision curve analysis (**Figure 5B**) showed that the model provided positive net benefit across a wide range of threshold probabilities (0.2-0.7) in the external cohort, outperforming both the "treat all" and "treat none" strategies.

Comparison with existing models

On the temporal Validation set, the AUC of this study's model (0.83) was significantly higher than that of the Grobman model (0.79, P<0.05), with both sensitivity and specificity superior. **Figure 6** visually demonstrates this comparison, showing the ROC curves of our model and the Grobman model applied to the temporal Validation set. The superior performance of our model is evident from its ROC curve lying above that of the Grobman model across most of the specificity range.

Subgroup analysis

To assess model robustness across clinical subgroups, we conducted subgroup analyses by prior vaginal birth after cesarean (VBAC) history and cervical score (≥6 vs. <6), using the Grobman model as a benchmark. As shown in **Figure 7**, our model consistently outperformed the Grobman model in all subgroups. Among those with prior VBAC, AUCs were 0.92 vs. 0.85 (cervical score ≥6) and 0.85 vs. 0.80 (cervical score <6); among those without prior VBAC, AUCs were 0.82 vs. 0.80 (cervical score ≥6) and 0.80 vs. 0.73 (cervical score <6). These results demonstrate that our model maintains strong predictive performance across varying VBAC histories and cervical maturity levels, confirming its robustness in diverse clinical settings.

Discussion

The choice of delivery method for pregnant women with a scarred uterus has always been a focus and clinical challenge in obstetrics [19]. With the implementation of the two-child and

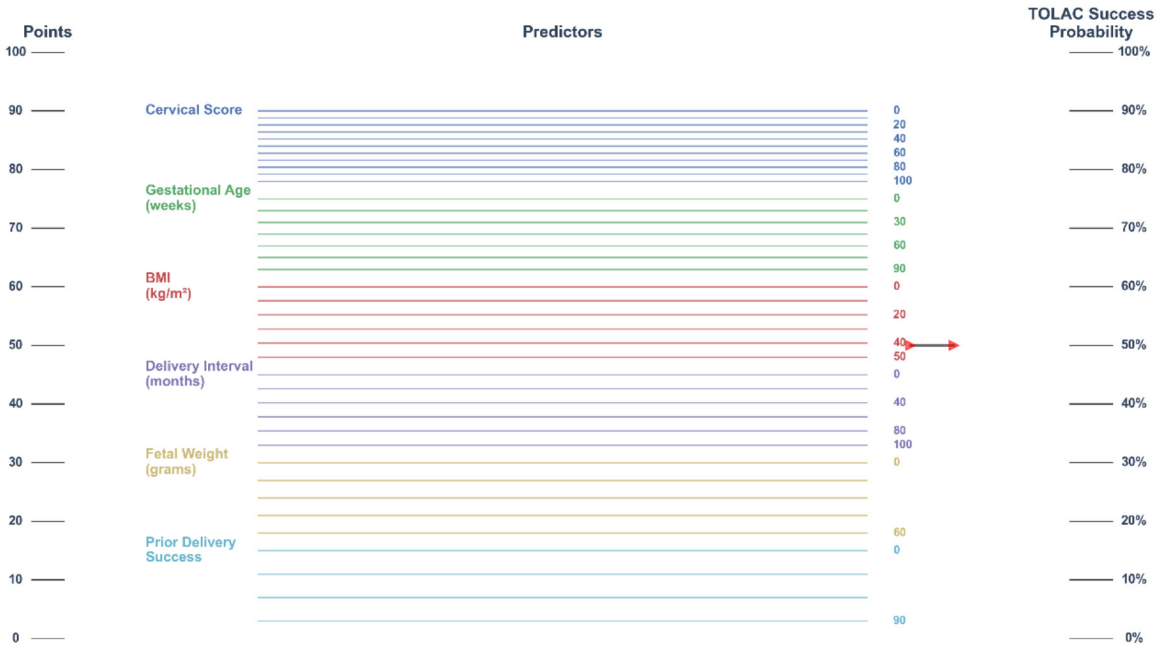

Figure 4. Nomogram for predicting the probability of successful trial of labor after cesarean (TOLAC). Instructions for use: For each patient, locate her value on the axis for each predictor. Draw a vertical line upward to the 'Points' axis to determine the score for that variable. Sum the scores for all six variables to obtain the 'Total Points'. Finally, draw a vertical line downward from the 'Total Points' axis to the 'Predicted Probability' axis to read the estimated probability of TOLAC success. Worked example (as detailed in the Discussion): For a 30-year-old woman (≈ 48 points) with a BMI of 24 kg/m^2 (≈ 38 points), a history of vaginal delivery (≈ 72 points), a cervical score of 7 (≈ 58 points), an estimated fetal weight of 3,200 g (≈ 48 points), and a gestational age of 38 weeks (≈ 62 points), the total points are approximately 326. This corresponds to a predicted probability of TOLAC success of approximately 80%. Abbreviations: TOLAC, trial of labor after cesarean; BMI, body mass index.

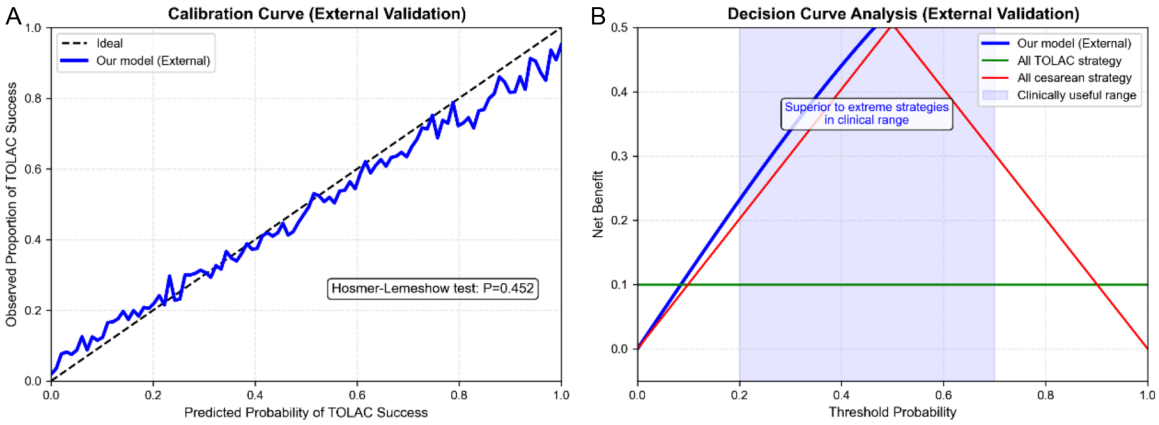

Figure 5. Temporal validation - calibration and decision curve analysis. A. Calibration curve showing the agreement between predicted probabilities and observed outcomes in the temporal Validation set (n=524). The dashed line represents perfect calibration. The Hosmer-Lemeshow test indicates no significant deviation from perfect fit (P=0.452). B. Decision curve analysis evaluating the clinical utility of the model in the temporal Validation set across different threshold probabilities. The model demonstrates positive net benefit in the clinically relevant range (20-70% threshold probability) compared to the strategies of performing TOLAC for all patients or performing cesarean section for all patients. Abbreviations: TOLAC, trial of labor after cesarean.

three-child policies, the number of pregnant women with a scarred uterus in China has seen explosive growth [20]. According to data from this study, the success rate of TOLAC was 72.1%, which, although higher than the 60%-70% reported in some Western countries, still

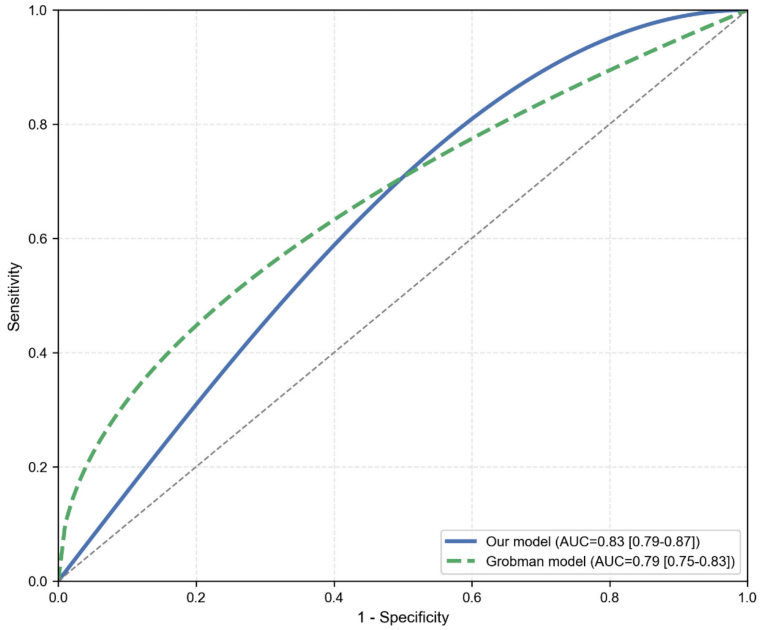

Figure 6. ROC curve comparison (temporal validation). Abbreviations: ROC, receiver operating characteristic; AUC, area under the curve; TOLAC, trial of labor after cesarean.

leaves nearly 30% of pregnant women at risk of trial failure [21, 22]. Failure not only increases the cesarean section rate but may also lead to serious complications such as uterine rupture (incidence approximately 0.5%-1%) and postpartum hemorrhage [22]. Therefore, accurately assessing the feasibility of TOLAC is crucial for maternal and neonatal safety [23]. Current clinical decisions heavily rely on physician judgment, which is subjective and limited in accuracy, while existing prediction models are less effective in Chinese clinical practice due to population differences and limitations in included variables [24]. This study aims to address this clinical need by constructing a predictive model based on a large single-center sample and validated externally, providing individualized TOLAC risk assessment tools for pregnant women with a scarred uterus.

From the results of univariate analysis, among the ten associated variables, the difference in cervical score between groups was most significant, consistent with the physiological mechanism of cervical maturity being a core indicator for the onset of labor [25-27]. Cervical scoring directly reflects the cervix's responses to contractions through evaluating comprehensive indicators such as cervical dilation, position, softness, receptivity, and fetal head position, indicating that the higher the score, the more mature the cervix, and the lower the resistance to natural delivery [28]. Notably, our model identified a positive association between maternal age and TOLAC success, which appears counterintuitive to the conventional view that advanced maternal age is a risk factor for adverse obstetric outcomes. This finding should be interpreted with caution within the specific context of our study population. The age range in our cohort was relatively concentrated (25-35 years), which may not fully represent the risks associated with more advanced maternal age (e.g., >35 years). The observed association could be confounded by unmeasured variables. For instance, in our clinical setting, older women within this range might have been subjected to more stringent selection criteria for TOLAC by their clinicians a priori, leading to a 'healthy candidate' effect where only those with otherwise favorable characteristics were advised to attempt labor. This selective process could result in a higher success rate among the older women who were ultimately included in the TOLAC group. Furthermore, the positive coefficient might also reflect a complex, non-linear relationship that a linear model cannot fully capture. Therefore, while the variable 'age' contributed to the model's discrimination, its independent effect warrants further investigation in larger, prospectively designed studies that can adequately control for such potential confounding and selection biases.

In the multivariate logistic regression model, the OR value for history of vaginal delivery was as high as 4.76 (95% CI: 4.01-5.65), becoming the strongest predictor, a result verified in multiple international studies. Previous vaginal delivery experiences not only cause "memorable" dilation of the birth canal but also enhance the coordination of uterine contractions through neuro-endocrine regulation, reducing the risk of weak contractions. BMI was negatively correlated with TOLAC success, suggesting that obesity may affect trial outcomes by increasing

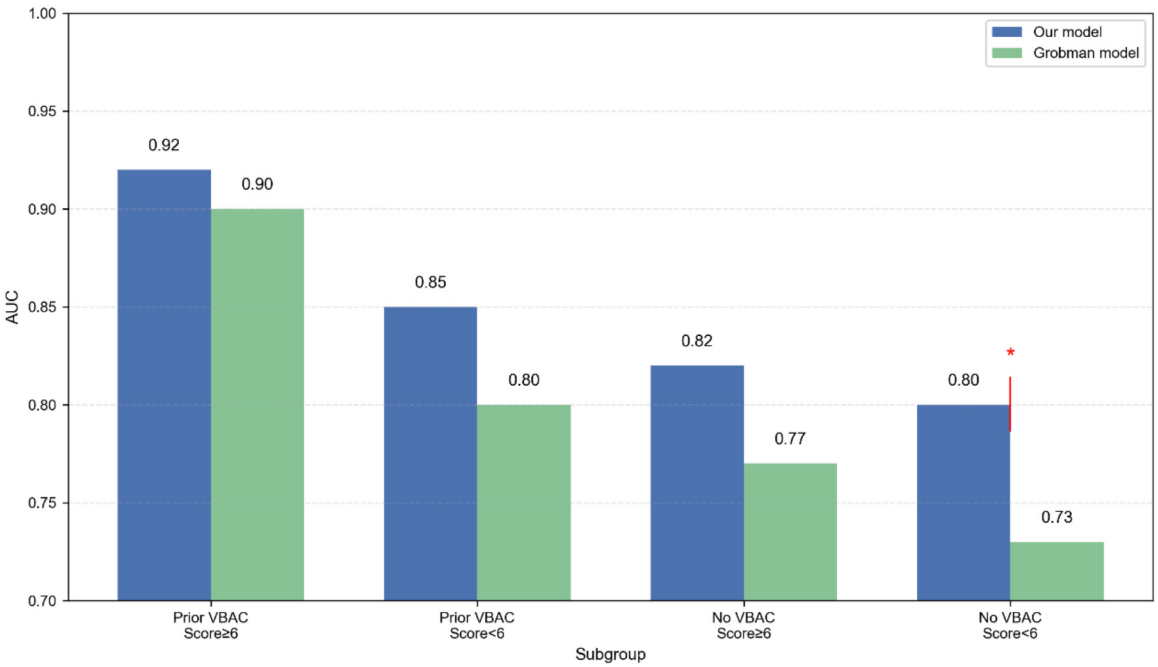

Figure 7. AUC comparison in subgroups. Abbreviations: AUC, area under the curve; VBAC, vaginal birth after cesarean; TOLAC, trial of labor after cesarean.

birth canal resistance and lowering the efficiency of uterine muscle layer contractions. Each increase of 1 kg/m^2 in BMI decreases the probability of TOLAC success by about 11%. The OR value for estimated fetal weight was 0.98, indicating that for every additional 100 g in fetal weight, the success rate of trials decreases by 2%, aligning with the clinical phenomenon where macrosomia easily leads to cephalopelvic disproportion, further supporting the importance of prenatal ultrasound estimation of fetal weight.

Model validation results showed an AUC of 0.85 for internal 10-fold cross-validation and 0.83 for temporal Validation, both at relatively high levels, indicating that the model has stable discriminative ability. Calibration curves revealed good consistency between predicted probabilities and actual outcomes, suggesting no significant calibration bias, allowing accurate quantification of TOLAC success probabilities. Compared with the Grobman model, this model not only exceled in discrimination but also incorporated the dynamic indicator of cervical score, while the Grobman model mainly relied on static variables like obstetric history and gestational week, cervical scores can be updated in real-time before labor, more accurately reflecting physiological conditions during childbirth. Additionally, decision curve analysis confirmed that within the threshold probability range of 0.1-0.8, the net benefit of this model significantly outperforms extreme strategies, especially in the "gray area" of clinical decision-making (0.3-0.6), effectively reducing unnecessary cesarean sections or trial failures. The decision curve analysis provided compelling evidence for the clinical utility of our model. The positive net benefit across the clinically relevant threshold probability range (20-70%) indicated that using this model to guide TOLAC decisions would lead to better patient outcomes compared to either universally recommended TOLAC or repeat cesarean section. This is particularly important in the "gray zone" of clinical decision-making (30-60% predicted probability), where physician uncertainty is highest. The quantitative net benefit values (0.15-0.25 in this range) translate to meaningful clinical impact - potentially avoiding 15-25 inappropriate management decisions per 100 patients while maintaining safety.

From a clinical translation perspective, the nomogram design of this model provided a convenient tool for rapid assessment. For instance, a 30-year-old pregnant woman with a BMI of 24 kg/m^2, a history of vaginal delivery, a cervical score of 7, an estimated fetal weight of 3,200

g, and a gestational age of 38 weeks would score 65 points using the nomogram, corresponding to an approximately 80% probability of TOLAC success, thus encouraging active trial of labor. If her cervical score were 4, the total score would drop to 40, reducing the success probability to 50%, requiring cautious decision-making in conjunction with other indicators. Moreover, since the variables included in the model are routine examination items, there is no need for additional medical costs, offering considerable health economic value [26, 29].

Our model did not include certain variables recognized as influential in international models, such as the 'indication for the previous cesarean section' and the 'time interval since the previous cesarean'. This omission was primarily due to limitations in our retrospective data collection. Specifically, the detailed indications for the prior cesarean (e.g., failure to progress, fetal distress) were often not standardized or comprehensively documented in the medical records available for this study. Furthermore, a significant number of patients were referred from other primary care institutions, and complete obstetric histories, including the exact date of the previous delivery, were frequently missing, making accurate calculation of the inter-delivery interval unfeasible. While these factors are undoubtedly important, our model demonstrated that a robust prediction can be achieved using a parsimonious set of readily available clinical and obstetric variables, which may enhance its practicality in real-world clinical settings where such detailed historical data may be lacking. Future prospective studies should aim to incorporate these variables to further refine predictive accuracy.

This study has several limitations. Firstly, it is a single-center retrospective study, its design may introduce selection bias, such as clinicians tending to opt for direct cesarean sections for pregnant women with poor cervical conditions, leading to a disproportionately high proportion of successful cases in the sample. Secondly, it did not include factors like uterine scar thickness and strength of contractions obtained through ultrasound or labor indicators, which could further enhance model performance. Thirdly, although the model was developed from a single-center cohort and demonstrated good performance in an external temporal validation set from two other hospitals, its generalizability needs to be further confirmed in broader, multi-center, and multi-ethnic populations. Future research could adopt prospective designs, dynamically collect real-time data such as cervical changes and contraction curves during labor, and use machine learning algorithms to build dynamic prediction models while exploring the impact of epigenetic markers (e.g., levels of inflammatory factors in cervical tissue) on TOLAC outcomes to further improve prediction accuracy.

In conclusion, the predictive model for successful TOLAC, based on a single center with external temporal validation, exhibits excellent discrimination, calibration, and generalization capabilities, outperforming existing international models. By quantifying the impacts of critical factors such as age, BMI, and history of vaginal delivery, this model provides precise TOLAC risk assessment for clinical use, aiding in optimizing delivery method choices and reducing the incidence of maternal and neonatal complications, possessing significant clinical application value. As the model is of wide promotion and validation, it is expected to promote standardized and individualized development in managing pregnancies with a scarred uterus, contributing to improving obstetric quality.

Acknowledgements

This work was supported by Huai'an Natural Science Research Program (HAB202219).

Disclosure of conflict of interest

None.

Address correspondence to: Wenli Zhou and Xiaoqiong Li, Department of Obstetrics, Huai'an Maternal and Child Health Hospital, No. 6 Baiyun Road, Qingjiangpu District, Huai'an 223000, Jiangsu, China. Tel: +86-13776746386; E-mail: 137767-46386@163.com (WLZ); Tel: +86-0517-89685623; E-mail: Lixiaoqiong68@163.com (XQL)

References

[1] Levin G, Meyer R, Mor N, Yagel S, David M, Yinon Y and Rottenstreich A. Trial of labor after cesarean in adolescents - a multicenter study. J Pediatr Adolesc Gynecol 2020; 33: 398-402.

[2] Levin G, Tsur A, Tenenbaum L, Mor N, Zamir M and Meyer R. Prediction of successful trial of labor after cesarean among grand-multipa-

rous women. Arch Gynecol Obstet 2022; 306: 373-378.
[3] Levin G, Tsur A, Tenenbaum L, Mor N, Zamir M and Meyer R. Prediction of vaginal birth after cesarean for labor dystocia by sonographic estimated fetal weight. Int J Gynaecol Obstet 2022; 158: 50-56.
[4] Chang WH, Yang ST and Wang PH. Trial of labor after cesarean section. Taiwan J Obstet Gynecol 2022; 61: 739-740.
[5] Lopian M, Kashani-Ligumski L, Cohen R, Assaraf S, Herzlich J, Lessing JB and Perlman S. Twin TOLAC is an independent risk factor for adverse maternal and neonatal outcome. Arch Gynecol Obstet 2021; 304: 1433-1441.
[6] Li H, Sheng W, Cai M, Chen Q, Lin B, Zhang W and Li W. A predictive nomogram for a failed trial of labor after cesarean: A retrospective cohort study. J Obstet Gynaecol Res 2022; 48: 2798-2806.
[7] Lopian M, Kashani-Ligumski L, Cohen R, Herzlich J and Perlman S. A Trial of Labor after Cesarean Section with a Macrosomic Neonate. Is It Safe? Am J Perinatol 2024; 41: e400-e405.
[8] Lopian M, Kashani-Ligumski L, Cohen R, Herzlich J, Vinnikov Y and Perlman S. Grand multiparity, is it a help or a hindrance in a trial of labor after cesarean section (TOLAC)? J Matern Fetal Neonatal Med 2023; 36: 2190835.
[9] Luo ZC, Liu X, Wang A, Li JQ, Zheng ZH, Guiyu S, Lou T, Pang J and Bai XL. Obstetricians' perspectives on trial of labor after cesarean (TOLAC) under the two-child policy in China: a cross-sectional study. BMC Pregnancy Childbirth 2021; 21: 89.
[10] Meyer R, Tsur A, Tenenbaum L, Mor N, Zamir M and Levin G. Sonographic fetal head circumference is associated with trial of labor after cesarean section success. Arch Gynecol Obstet 2022; 306: 1913-1921.
[11] Miazga E, Reed C, Olsthoorn A, Fan I, Zhao E, Shapiro J, Cipolla A, Tunde-Byass M and Shore EM. Investigating the implementation of a Trial of Labour After Caesarean (TOLAC) delivery bundle with respect to decreasing caesarean delivery rates: a multisite quality improvement initiative. BMJ Open 2025; 15: e088734.
[12] Miazga E, Reed C, Shapiro J and Shore EM. Health care practitioners' attitudes and knowledge regarding trial of labour after cesarean delivery (TOLAC): a survey of Canadian obstetrical care providers. J Obstet Gynaecol Can 2022; 44: 658-663.
[13] Miazga E, Reed C, Tunde-Byass M, Cipolla A, Shapiro J and Shore EM. Decreasing cesarean delivery rates using a trial of labour after cesarean (TOLAC) bundle. J Obstet Gynaecol Can 2020; 42: 1111-1115.
[14] Miazga E, Reed C, Tunde-Byass M, Cipolla A, Shapiro J and Shore EM. Corrigendum to 'Decreasing Cesarean Delivery Rates Using a Trial of Labour After Cesarean (TOLAC) Bundle' [Journal of Obstetrics and Gynaecology Canada 42 (2020) 1111-1115]. J Obstet Gynaecol Can 2021; 43: 1120.
[15] Olsthoorn AV, Figueiro-Filho EA, Li YE, Farine D and Sobel ML. Counselling patients for trial of labour after cesarean (TOLAC) and invasive placentation: are we missing the mark? The importance of local data and informed choice. J Obstet Gynaecol Can 2021; 43: 306-312.
[16] Parveen S, Rengaraj S and Chaturvedula L. Factors associated with the outcome of TOLAC after one previous caesarean section: a retrospective cohort study. J Obstet Gynaecol 2022; 42: 430-436.
[17] Qiu L, Zhu J and Lu X. The safety of trial of labor after cesarean section (TOLAC) versus elective repeat cesarean section (ERCS): a systematic review and meta-analysis. J Matern Fetal Neonatal Med 2023; 36: 2214831.
[18] Rao J, Fan D, Ma H, Lin D, Zhang H, Zhou Z, Li P, Chen G, Lu D, Liu Y, Wu Z, He J, Liu X, Peng B, Guo X and Liu Z. Is there an optimal interdelivery interval in women who underwent trial of labor after cesarean delivery (TOLAC)? Reprod Health 2022; 19: 14.
[19] Zhang H, Yu L, Wei S and Li G. Analysis of influencing factors and establishment of prediction model for successful vaginal delivery after cesarean section. Front Glob Womens Health 2025; 6: 1447569.
[20] Shalabna E, Kedar R, Assaf W, Nahshon C, Kugelman N, Lavie O and Sagi-Dain L. The association between obesity and the success of trial of labor after cesarean delivery (TOLAC) in women with past vaginal delivery. J Perinat Med 2024; 52: 158-164.
[21] Grobman WA, Lai Y, Landon MB, Spong CY, Leveno KJ, Rouse DJ, Varner MW, Moawad AH, Caritis SN, Harper M, Wapner RJ, Sorokin Y, Miodovnik M, Carpenter M, O'Sullivan MJ, Sibai BM, Langer O, Thorp JM, Ramin SM and Mercer BM; National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units Network (MFMU). Development of a nomogram for prediction of vaginal birth after cesarean delivery. Obstet Gynecol 2007; 109: 806-812.
[22] Shi H, Li S, Lv J, Wang HHX, Hou Q and Jin Y. Maternal and neonatal characteristics associated with clinical outcomes of TOLAC from 2012-20 in the USA: evidence from a retrospective cohort study. EClinicalMedicine 2022; 54: 101681.
[23] Tegegne GA, Rade BK, Yismaw AE, Taye W and Mengistie BA. Predictors of successful trial of

labor after cesarean section (TOLAC) in women with one prior transverse cesarean section at Tertiary Hospitals in northwest Ethiopia: a multicenter study. BMC Pregnancy Childbirth 2024; 24: 240.

[24] Arkerson BJ, Muraca GM, Thakur N, Javinani A, Khalil A, D'Souza R and Mustafa HJ. Prediction of uterine rupture in singleton pregnancies with one prior cesarean birth undergoing TOLAC: a cross-sectional study. Acta Obstet Gynecol Scand 2025; 104: 185-193.

[25] Dasgupta S and Dasgupta J. Development of a multivariate predictive nomogram among women with antepartum fetal death diagnosed at ≥ 34 weeks of gestation for outcome of TOLAC. Arch Gynecol Obstet 2024; 310: 1131-1139.

[26] Dasgupta S, Dasgupta J, Goswami B and Mondal J. Randomized controlled trial comparing efficacy of a combination regime containing two cervical sensitizers (mifepristone + Foley's catheter) versus single agent mifepristone or Foley's catheter for labor induction in women attempting TOLAC at late third trimester with a dead fetus in utero. J Obstet Gynaecol Res 2023; 49: 2671-2679.

[27] He X, Dai Q, Wu X, Zhou J and Li J. Outcomes and risk factors for failed trial of labor after cesarean delivery (TOLAC) in women with one previous cesarean section: a Chinese population-based study. BMC Pregnancy Childbirth 2022; 22: 676.

[28] Bart Y, Mohr-Sasson A, Mazaki-Tovi S, Kalter A, Shapira M, Sivan E and Hendler I. The effect of polyhydramnios on the success of trial of labor after cesarean delivery (TOLAC): A retrospective cohort. Eur J Obstet Gynecol Reprod Biol 2020; 253: 187-190.

[29] Togioka BM and Tonismae T. Uterine Rupture. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2025, StatPearls Publishing LLC.; 2025.