



OPEN Development and validation of a risk prediction model for caesarean delivery among multiparous women

Nigus Billign Yimer^{1✉} & Eskedar Getie Mekonnen^{2,3}

While caesarean risk prediction models exist for nulliparous and high-risk pregnancies, there is a lack of models that predict the risk of caesarean delivery among multiparous women. This study aimed to develop and validate a risk prediction model for caesarean delivery and assess its clinical utility among multiparous women. Using data from 460 participants, a prediction model was developed to predict the risk of caesarean delivery. The model performance was evaluated using the area under the receiver operating characteristic curve (AUC) and calibration plot, and the model was internally validated using bootstrapping technique. A simplified risk score was calculated, and a nomogram was developed for the individual caesarean delivery risk guide. Additionally, a decision curve analysis was performed to assess the clinical utility of the model. The final model included four predictors: maternal age, previous caesarean delivery, pregnancy-induced hypertension, and antepartum hemorrhage. The model had an AUC of 78.0% (95% CI 71.1–84.8), indicating good discrimination capacity. The model also exhibited good calibration and a low overoptimism coefficient, indicating minimal risk of overfitting. The risk prediction model has good clinical utility for discriminating multiparous women at risk of caesarean delivery. The tool can guide clinicians in estimating the risk of caesarean delivery among multiparous women that could lead to improved maternal and neonatal outcomes, ultimately enhancing the quality of care delivered in low-resource settings.

Keywords Caesarean delivery, Multiparous, Prediction model, Internal validation, Decision curve, Nomogram

Caesarean Delivery (CD) is a widely performed surgical procedure, with approximately 21% of women undergoing CD worldwide¹. The rate of CD continues to rise due to an increase in primary CD and a decline in the number of vaginal births after prior CD². Although CD can reduce maternal-perinatal morbidities and mortality³ when medically indicated, undergoing CD without proper indications can lead to adverse outcomes for both the mother and her baby¹. Long-term complications of CD include miscarriage, antepartum hemorrhage and placenta accreta, stillbirth, subfertility, childhood asthma, and obesity in future pregnancies⁴. However, the likelihood of CD and associated complications varies among women.

Several factors, including advanced maternal age, parity > 1, fetal macrosomia, premature rupture of membranes⁵, and labor induction have been identified to increase the risk of CD in multiparous women⁶. Maternal demographic characteristics, antenatal factors, obstetric and medical complications, prior obstetric history, and current clinical characteristics have also been found to provide relevant information for predicting the risk of CD in multiparous women^{5–7}. In Ethiopia, a recent study revealed that most women with a prior CD underwent a trial of labour after caesarean section (TOLAC)⁸. However, higher rates of morbidity and mortality in women undergoing TOLAC emphasize the importance of meticulous selection of women at higher risk of CD before attempting a trial of labor⁹. CD risk prediction tools have the potential to improve short- and long-term obstetric outcomes by identifying women who are likely to require CD¹⁰.

Over the years, predicting the risk of CD has gained attention from researchers^{7,10–16}. With the increasing complexity and multidimensionality of healthcare data, advanced machine learning algorithms have become popular methods for predicting CD risk^{13,17,18}. Nevertheless, existing CD prediction models primarily focus on nulliparous women, women with high-risk pregnancies, and women undergoing labor induction, often

¹Department of Midwifery, College of Health Sciences, Woldia University, Woldia, Ethiopia. ²School of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia. ³Global Health Institute, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerpen, Belgium. ✉email: nigusb27@gmail.com

considering multiparous women to be at low risk. However, multiparous women with specific risk factors, such as prior caesarean scars and medical complications such as diabetes mellitus and other obstetric risks, i.e., polyhydramnios and premature rupture of membranes (PROM), can have a CD risk as high as that of nulliparous women¹⁹. Therefore, predicting the risk of CD in these women is an essential first step in providing appropriate obstetric interventions and addressing overlooked risks in this population.

Given the lack of prediction tools for assessing the risk of CD in multiparous women, having quantitative knowledge about the risk of CD can enhance provider-patient engagement, facilitate shared decision-making, and optimize resource allocation in clinical care⁷. Identifying women at risk of emergency CD using predictors collected at the onset of labor could reduce the incidence of complications associated with labor¹⁰. Additionally, CD prediction algorithms can help protect physicians against medicolegal issues¹⁷.

Therefore, in this study, we aimed to develop a prediction model for assessing the risk of CD in multiparous women using easily obtainable maternal and obstetric variables. We internally validated the prediction model, calculated a simplified risk score, and developed a nomogram for clinical use. Furthermore, we assessed the clinical utility of the model using decision curve analysis.

Methods

Study design and participants

A prospective cross-sectional study was conducted among multiparous women presenting with labor complaints in two public hospitals in Hawassa city, Ethiopia, between February 1 and June 30, 2018. Multiparous women with singleton pregnancies at or beyond 28 weeks of gestation were included. We excluded women who presented with a second stage of labor or beyond.

Data collection

A total of 460 multiparous women presenting with labor pain at the study hospitals were recruited at admission to the labor ward. The total sample was allocated to the two study hospitals based on the proportion of the average number of childbirths that occurred three months prior to the study. Data were collected using a structured questionnaire from eligible participants and their medical records. A range of sociodemographic characteristics, antenatal factors, obstetric profiles, and pregnancy outcomes were collected. The tool used to collect the data were translated into local languages understandable by the participants (Amharic and Sidamic). The data were collected by six trained midwives. The process of data collection and completeness was followed by the principal investigator (NBY).

Variables

The outcome of interest was CD (yes/no) collected through observation and reviewing the medical charts. The initial predictors for model development were selected based on prior knowledge and clinical relevance²⁰. Variables such as advanced maternal age (categorized as less than 35 (not advanced) and 35 or above (advanced)), parity (2–4 births (multiparous) and 5 and above (grand multiparous)), mode of delivery during last birth (coded as vaginal vs. caesarean delivery), history of medical illness (yes/no), history of abortion (yes/no), antenatal care (ANC) follow-up (yes/no), frequency of ANC (categorized as 1–3 and 4 or more), hypertensive disorders of pregnancy (HDP) (yes/no), antepartum hemorrhage (APH) (yes/no), premature rupture of membranes (PROM) (yes/no), preterm delivery (yes/no), and intrauterine fetal death (yes/no) were included as predictors. The variables were collected through interviewing women (history taking) and extraction from medical charts, which were again validated using physical examination and further clinical investigation results when necessary.

Statistical analysis

The collected data were entered into Epi-data version 4.4.3.1 and exported to R v 4.3.2 (<https://www.r-project.org>) for processing and analysis. The data were cleaned and explored for completeness, structure, and coding consistency before analysis. The distribution of data was checked for continuous variables using plots and Kolmogorov–Smirnov and the Shapiro–Wilk test. Continuous variables were categorized into relevant categories, and categorical coding was also used to determine meaningful factor levels.

The least absolute shrinkage and selection operator (LASSO) regression technique was used to select the predictor variables for our final model. The LASSO approach performs variable selection and shrinkage by adding a penalty term to the sum of residuals and is recommended for small events per predictor²¹. Given that the number of events per predictor in our model was reasonably low, we used LASSO regression for variable selection. The ‘glmnet’ function version 4.1-8 in R was used, and k-fold cross-validation was performed to identify the optimal lambda value, followed by extracting coefficients of the predictors that returned a value greater than the optimal threshold. The predictors identified via the LASSO technique were used to construct the final model.

The LASSO was formulated as:

$$\text{minimize} - \left(\sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \right) + \lambda \sum_{j=1}^p |\beta_j|$$

Where y_i represents the outcome variable, \hat{y}_i is predicted probability based on logistic regression, β_j represent coefficients of predictors, and λ is tuning parameter which controls the degree of shrinkage.

A logistic regression model was built with the selected predictors to calculate the risk of CD. We formulated the logistic regression model as follows:

$$P(Y = 1|X_1, X_2, \dots, X_p) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}}$$

Where $P(Y = 1|X_1, X_2, \dots, X_p)$ is the predicted probability of CD, X_1, X_2, \dots, X_p are the predictors included in the model, β_0 is intercept, $\beta_1, \beta_2, \dots, \beta_p$ are coefficients for each predictor.

The model performance was evaluated using the area under the receiver operating characteristic (AUROC) curve and calibration plot. An AUC value close to 1 was used to indicate good model discriminative power²². The calibration of the model was determined by plotting the predicted probabilities against the observed probabilities and assessing their proximity. Additionally, the p-value from the Hosmer–Lemeshow goodness-of-fit test above 5% was used to indicate the goodness of fit of the model.

The model was internally validated to assess the risk of overfitting using the bootstrapping technique. A total of 1000 bootstrapping samples were created with random resampling with replacement. The over-optimism coefficient was calculated by subtracting the average AUC value of the bootstrapping sample from the original AUC and dividing it by the original AUC. A coefficient of less than 0.1 was used to indicate a low risk of bias or over-fitting²³.

The regression coefficients of the predictors in the final model were used to calculate the simplified risk scores. The scores were calculated by dividing each predictor's coefficient by the smallest coefficient and rounding it to the nearest integer. The total risk score was then estimated as the product of the multiplicative value of predictors with their corresponding risk scores. Furthermore, a nomogram was developed as a visual tool that can be used to guide the individual risk of caesarean section based on the final predictors included in the prediction model. The nomogram was plotted using the 'nomogram' function from the 'rms' package in R. Linear predictors were transformed to a probability scale for intuitive interpretation of the predictions by the model.

Decision curve analysis was used to show the clinical utility of the developed predictive model. The net benefit of the model and other alternative strategies was plotted at different threshold probabilities with a 95% confidence interval (CI). Then, the standardized net benefit was compared for interventions guided by the predictive model, interventions for all patients, and interventions for none of the patients. Additionally, the plot included cost-benefit ratios to show potential benefits and harms associated with different threshold probabilities of the curve. A total of 1000 bootstrap samples were used in the analysis to reduce bias and provide a more accurate estimate of the predictive model's net benefit.

This paper followed the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement²⁴.

Ethical approval

The study was approved by the institutional review board of Hawassa University College of Medicine and Health Sciences (IRB/164/10). Written consent was obtained from the study participants. The confidentiality of the participants was assured by anonymizing personal identifiers in the data. All methods and procedures in the study were performed in accordance with the Declaration of Helsinki.

Results

Characteristics of the sample

Of the total 460 participants, approximately 14% underwent CD. Among the women included in the study, approximately 1 in 10 had a prior caesarean scar. Furthermore, 6.7%, 3.5%, and 4.1% of the participants had hypertensive disorders, antepartum hemorrhage, and premature rupture of membranes during their current pregnancy, respectively (Table 1).

Predictors of caesarean delivery

Predictors of CD in multiparous women were identified through LASSO regression. Variables such as maternal age, abortion, parity, last mode of delivery (prior caesarean delivery), number of antenatal care visits, medical illnesses, hypertensive disorders of pregnancy, and antepartum hemorrhage were included in the LASSO regression. Four of these predictors returned with coefficients shrank to zero, while the remaining four predictors (Table 2) were selected to fit the final model. The optimum lambda value used to select the predictors was 0.0097.

In the logistic regression model, prior caesarean delivery and antepartum haemorrhage during the current pregnancy were found to be significantly associated with the risk of CD. To calculate a simplified risk score, we divided each variable's regression coefficient by the smallest coefficient (0.3 for maternal age) and approximated the result to the nearest integer. The total risk score was then calculated as follows:

$$\begin{aligned} \text{Pr (CD)} = & (1 * \text{advanced maternal age}) \\ & + (11 * \text{caesarean delivery for last birth}) \\ & + (3 * \text{hypertensive disorders of pregnancy}) \\ & + (6 * \text{antepartum haemorrhage}). \end{aligned}$$

Using the coefficients of the predictors in the logistic regression model, a nomogram was developed that provides a graphical aid to estimate the individual probability of caesarean delivery risk (Fig. 1).

Performance of the model

The observed probabilities of the model were plotted against the predicted probabilities and showed a perfect calibration, as evidenced by all dots on the diagonal line (Fig. 2). The p-value from the Hosmer–Lemeshow test also indicated that the model fit the data well.

Characteristics	Caesarean delivery, N= 64 (13.9%)	Vaginal delivery, N= 396 (86.1%)	Overall, N= 460 (100%)
Maternal age			
Less than 35	53 (82.8%)	330 (83.3%)	383 (83.3%)
35 and above	11 (17.2%)	66 (16.7%)	77 (16.7%)
History of medical illness			
No	57 (89.1%)	370 (93.4%)	427 (92.8%)
Yes	7 (10.9%)	26 (6.6%)	33 (7.2%)
History of abortion			
No	52 (81.3%)	331 (83.6%)	383 (83.3%)
Yes	12 (18.8%)	65 (16.4%)	77 (16.7%)
ANC follow-up			
Yes	60 (93.8%)	346 (87.4%)	406 (88.3%)
No	4 (6.3%)	50 (12.6%)	54 (11.7%)
Frequency of ANC			
1–3	58 (90.6%)	345 (87.1%)	403 (87.6%)
4 or more	6 (9.4%)	51 (12.9%)	57 (12.4%)
Last mode of delivery			
Vaginal	32 (50.0%)	378 (95.5%)	410 (89.1%)
Caesarean	32 (50.0%)	18 (4.5%)	50 (10.9%)
Parity			
2–4	45 (70.3%)	258 (65.2%)	303 (65.9%)
5 or more	19 (29.7%)	138 (34.8%)	157 (34.1%)
Hypertensive disorders			
No	56 (87.5%)	373 (94.2%)	429 (93.3%)
Yes	8 (12.5%)	23 (5.8%)	31 (6.7%)
Antepartum hemorrhage			
No	58 (90.6%)	386 (97.5%)	444 (96.5%)
Yes	6 (9.4%)	10 (2.5%)	16 (3.5%)
Premature rupture of membrane			
No	62 (96.9%)	379 (95.7%)	441 (95.9%)
Yes	2 (3.1%)	17 (4.3%)	19 (4.1%)
Preterm delivery			
No	63 (98.4%)	393 (99.2%)	456 (99.1%)
Yes	1 (1.6%)	3 (0.8%)	4 (0.9%)
Intrauterine fetal death			
No	60 (93.8%)	351 (88.6%)	411 (89.3%)
Yes	4 (6.3%)	45 (11.4%)	49 (10.7%)

Table 1. Characteristics of multiparous women.

Variable	OR (95% CI)	Regression coefficient	Simplified risk score
Maternal age	1.35 (0.57, 2.97)	0.30	1
Prior caesarean delivery	23.93 (11.95, 49.80)	3.17	11
Current HDP	2.80 (0.97, 7.32)	1.03	3
Current APH	5.26 (1.47, 17.0)	1.66	6

Table 2. Predictors of caesarean delivery risk in multiparous women.

Additionally, the model exhibited good performance in discriminating true and false positives, with an AUC of 78% (95% CI 71.1–84.8) (Fig. 3). The model built with the simplified risk score as a predictor also exhibited nearly similar performance, indicating a non-significant loss of discriminative power due to simplifying the model.

Internal validation

The mean AUC value of the bootstrapping sample was 74.47%. The over-optimism coefficient was 0.04.

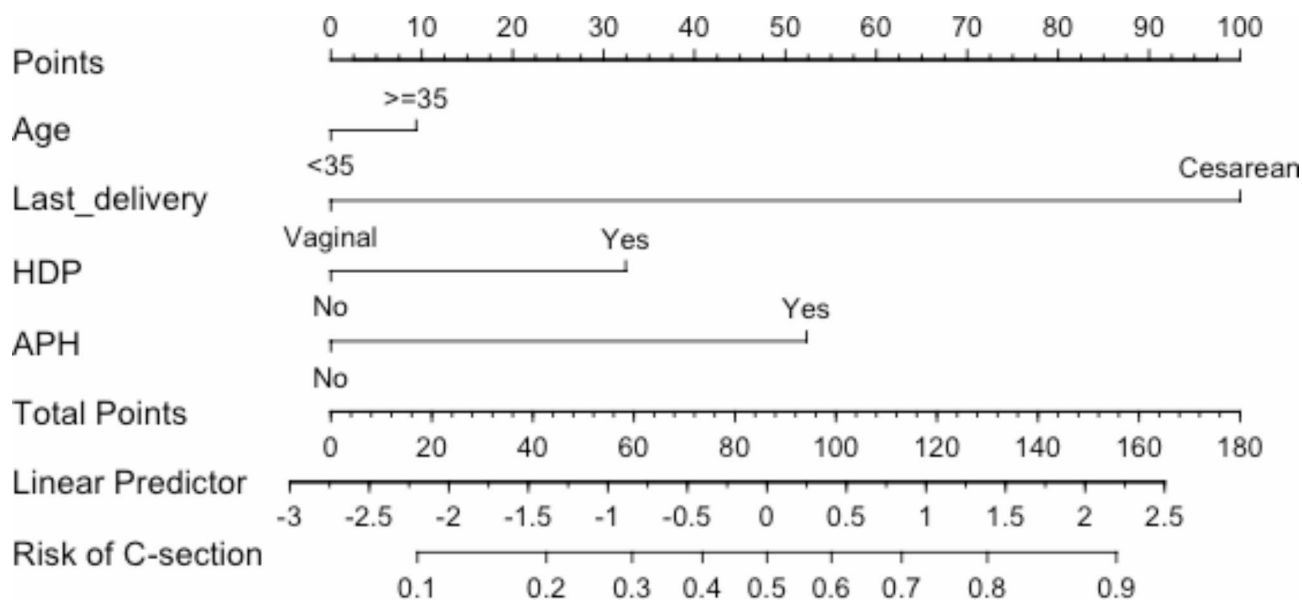


Fig. 1. Nomogram used to predict the risk of caesarean delivery (CD) for multiparous (MP) women at admission to labour. For individual women, points are assigned to each predictor using the point axis at the top of the figure, and a total score is obtained. The total points correspond to the predicted probability of CD for MP women. For example, for a woman aged ≥ 35 years (12 points) who had prior CD (100 points) with HDP (34 points) but without APH (0 points), the calculated total point was 146, which corresponded to 88% risk of CD at admission.

Calibration of the model

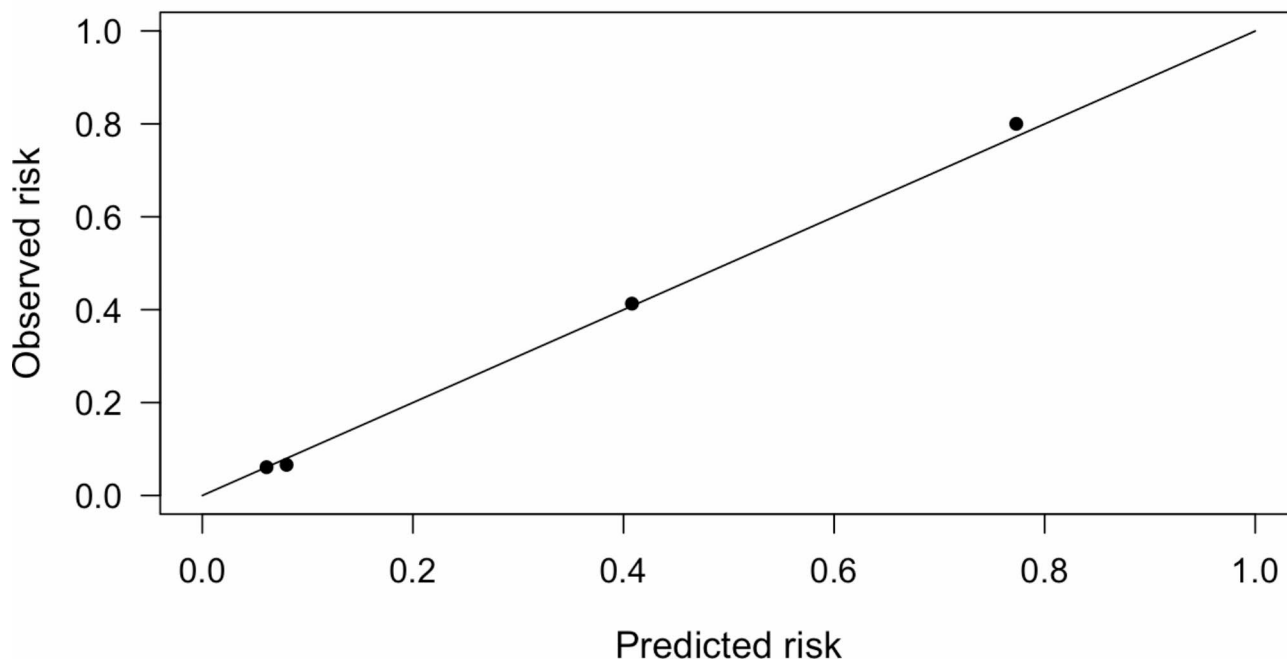


Fig. 2. Calibration plot comparing observed and predicted probabilities of caesarean delivery among multiparous women. The plot demonstrates the predictive accuracy of the model with dots on the diagonal line representing perfect calibration.

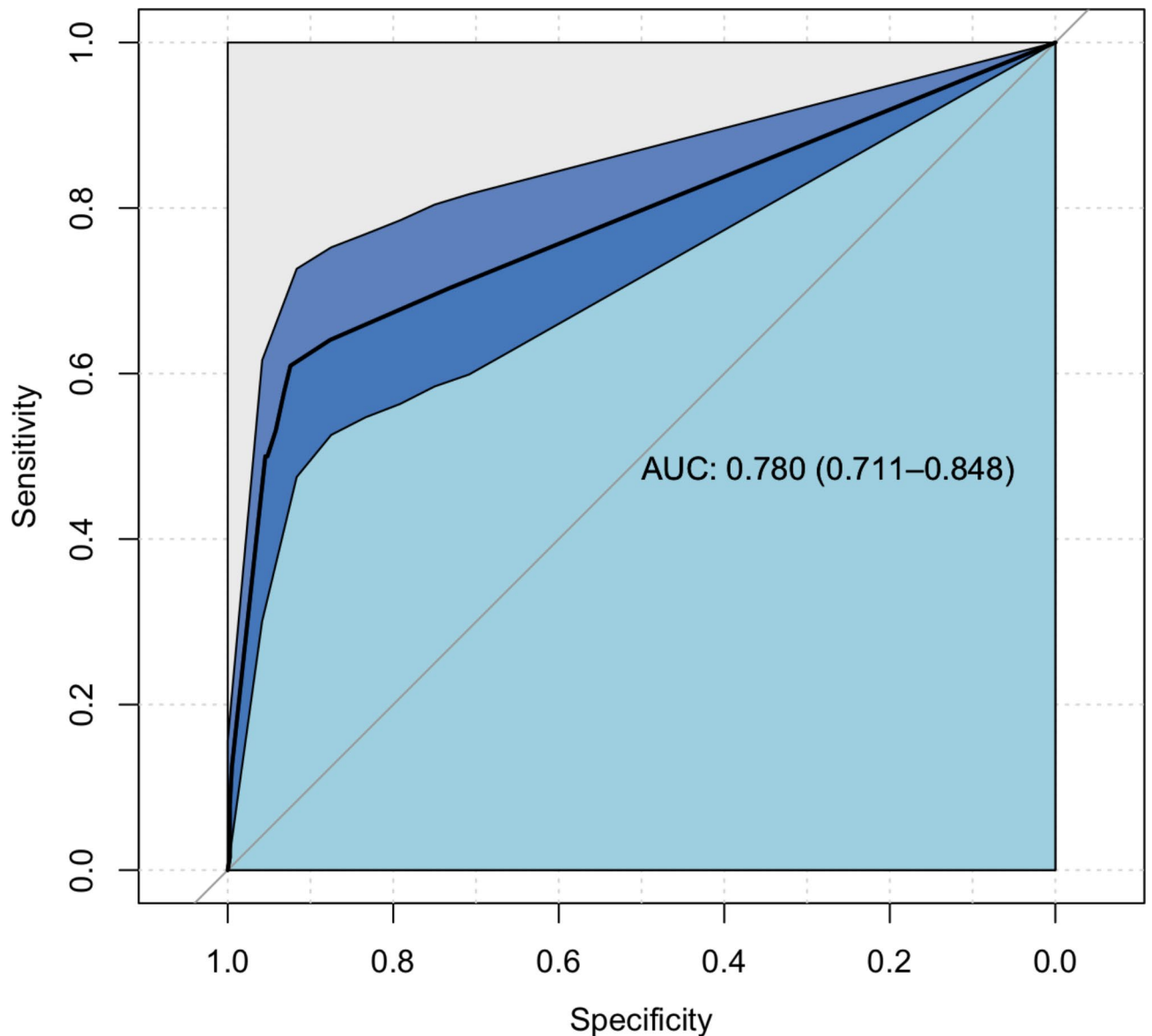


Fig. 3. ROC curve illustrating the performance of the prediction model for caesarean delivery risk in multiparous women. The curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) across different threshold values. The AUC shows the models ability to discriminate between women at risk of caesarean delivery and those not at risk.

Decision curve analysis

The decision curve plot compares the net benefit of using the model to treat women versus treating all or none of the women (Fig. 3). The net benefits of the model were greater than those of the treating all, and none at the different threshold probabilities (Fig. 4).

Discussion

In this study, we developed a prediction model and calculated a simplified risk score for CD among multiparous women using maternal characteristics and obstetric factors at admission to labor. The final model included four predictors: advanced maternal age, prior CD, hypertensive disorders of pregnancy, and antepartum hemorrhage. The model demonstrated good discrimination and calibration ability, indicating its potential usefulness in clinical practice.

The model in the development sample also performed well in the bootstrapping samples, resulting in a very low overoptimism coefficient. This finding is indicative of a lower risk of bias and overfitting, which enhances the potential application of the model in a real-world setting without compromising its performance. The robustness of our model in the bootstrapping analysis further supports its reliability and generalizability, strengthening its value as a predictive tool for CD in multiparous women.

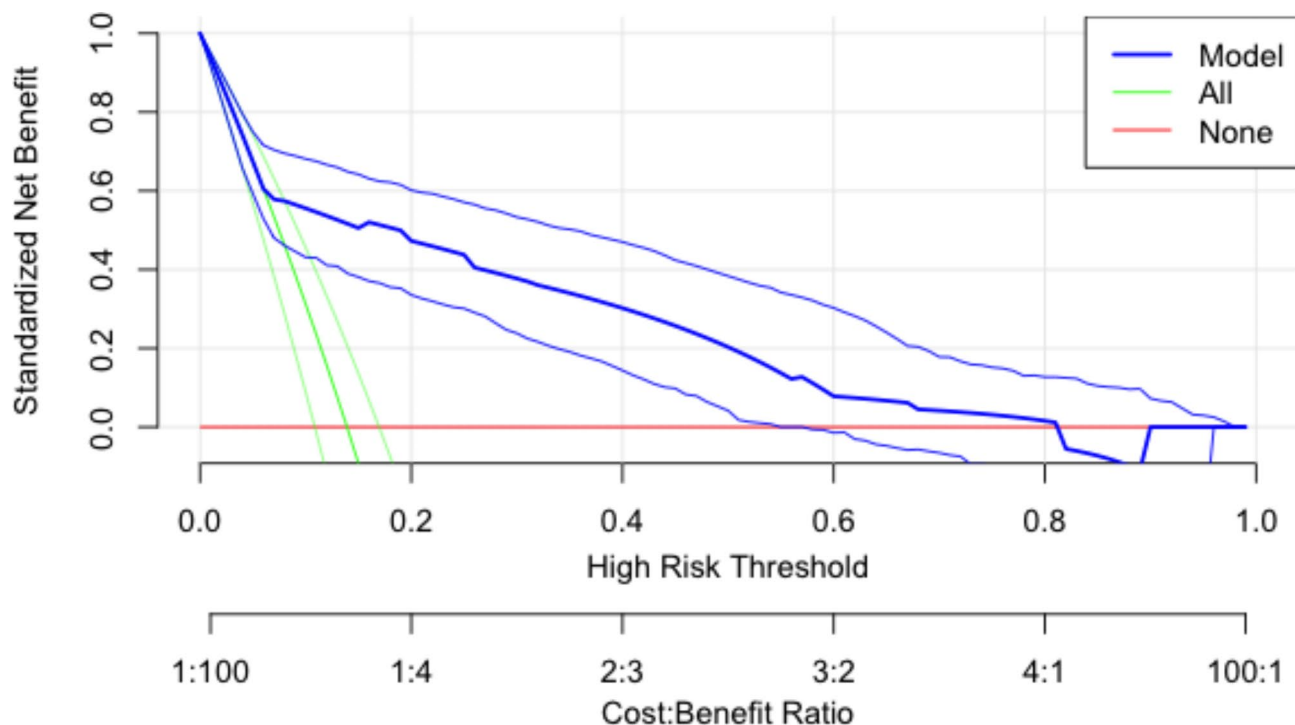


Fig. 4. Decision curve analysis illustrating the prediction model's clinical utility by plotting the net benefit across threshold probabilities. The standardized net benefit includes the trade-offs between true positives and false positives to determine whether the model has greater value compared to 'treat all' or 'treat none'. The light blue and green lines above and below the dark line are confidence intervals for the model's net benefit.

The results of the decision curve analysis also demonstrated a greater net benefit of using our model across a range of threshold probabilities at which considering an alternative intervention is beneficial. Specifically, our model exhibited comparable net benefit to the "treating all" strategy when the threshold probability was between 0 and 0.1. However, the net benefit of our model outperformed that of the default strategies for a threshold probability of 0.1 and above. However, it is important to note that the default strategies of treating all or none of the women may not be the most realistic comparisons. Nevertheless, our model demonstrated promising clinical value in facilitating decision-making.

The simplified risk score and nomogram were also derived from the prediction model, offering a practical tool for assessing the risk of CD. The risk score assigns weights to each predictor based on their regression coefficients, enabling a numerical representation of an individual's likelihood of undergoing CD. This scoring system enhances the interpretability of the model's predictions and provides clinicians with a straightforward risk assessment tool. The risk score ranges from 0 to 21, with higher scores indicating an increased risk of CD. The nomogram serves as a visual and easily interpretable guide for calculating an individual's total points and determining the corresponding risk of CD. This user-friendly tool aids clinicians in risk stratification and facilitates informed decision-making.

Comparing the performance of our model with previous studies, our model outperformed a model developed in Canada (AUC: 74%)¹². Moreover, our model demonstrated comparable performance to similar models developed in Spain (AUC: 74–81%)¹³ and the Netherlands (AUC: 73–81%)¹⁶. However, it is worth noting that our model was constructed using only four easily obtainable predictors, making it a practical and feasible tool for implementation. Although our model performed slightly worse than recent studies conducted in Ethiopia (AUC: 82%)¹⁴ and Indonesia (AUC: 81.3%)¹⁵, it is important to consider that these previous models included numerous antenatal and intrapartum predictors, which may delay diagnosis and risk stratification. Additionally, the inclusion of nulliparous women who underwent labor induction in previous studies introduces different risk factors for CD than the multiparous women included in our study.

The strength of our study lies in its good discrimination and calibration ability, as well as its robust performance in bootstrapping samples with easily obtainable variables. This suggests that our model is not only reliable but also practical for implementation in diverse clinical settings. The simplified risk score and nomogram provide user-friendly tools for risk assessment, facilitating effective risk stratification and enabling healthcare providers to make informed decisions tailored to individual patients.

Moreover, the implications of our model extend beyond individual patient care; it provides a framework for optimizing intervention thresholds. By aiding clinicians in evaluating when to recommend caesarean deliveries, our model implies a more strategic allocation of healthcare resources. This could lead to improved maternal and neonatal outcomes, ultimately enhancing the quality of care delivered in low-resource settings. Thus, the clinical utility of our model is profound, offering both immediate benefits for patient management and broader

implications for healthcare policy and practice. The variables to develop the current model influences the risk of caesarean delivery in several ways. It has been hypothesised that contractility of myometrium and effect of oxytocin decrease with advancing maternal age²⁵. Risk of caesarean delivery tends to be higher in women with prior caesarean scar due to risk of uterine rupture when vaginal delivery is attempted and increased adverse obstetric conditions such as invasive placentation and antepartum haemorrhage which ultimately necessitate repeat caesarean delivery⁹. Similarly, hypertensive disorders during pregnancy, including preeclampsia increases the risk of maternal and neonatal outcomes that necessitates caesarean delivery to prevent severe morbidities and mortality²⁶. The four variables used to build our model can be easily collected from mothers through history and physical examination or non-invasive tests upon presentation to the obstetric ward. Hence, selection of mothers who need caesarean delivery can be aided by this model that in turn helps to prevent adverse outcomes.

However, we acknowledge limitations in our study. Due to limited statistical power, we were unable to include some important clinical variables. Additionally, the specific cause of antepartum hemorrhage (i.e., placental abruption, placenta previa or vasa previa) used for model development was not specified, which may introduce some uncertainty. The wide 95% CI for predictors such as prior CD and antepartum haemorrhage in our model may suggest the degree of uncertainty and its potential to compromise the accuracy of our prediction. Further large-scale studies are imperative to validate the advantages of our predictors and to explore additional clinical variables that could improve the prediction of CD risk. External validation of the model is also important for evaluating its performance in independent datasets and facilitating its implementation in different clinical settings.

Conclusions

The risk prediction model developed to predict CD in multiparous women demonstrated good discrimination and calibration ability, as well as promising clinical utility. The risk score and nomogram provide practical tools for clinicians to estimate individual risk and guide decision-making. While our model performed well compared to previous studies, it should also be noted that it was constructed using easily obtainable predictors, making it practical for implementation. However, the highlighted limitations of this model should be considered when using and interpreting the results of this model.

Data availability

The datasets used to generate the findings of this study are available as supplementary file.

Received: 6 April 2024; Accepted: 7 January 2025

Published online: 13 February 2025

References

1. Angolile, C. M., Max, B. L., Mushemba, J. & Mashauri, H. L. Global increased cesarean section rates and public health implications: a call to action. *Health Sci. Rep.* **6**(5), e1274 (2023).
2. Nagy, S. & Papp, Z. Global approach of the cesarean section rates. *J. Perinat. Med.* **49**(1), 1–4 (2021).
3. Betran, A. P., Torloni, M. R., Zhang, J. J. & Gülmezoglu, A. M. WHO statement on caesarean section rates. *Bjog* **123**(5), 667–670 (2016).
4. Keag, O. E., Norman, J. E. & Stock, S. J. Long-term risks and benefits associated with cesarean delivery for mother, baby, and subsequent pregnancies: systematic review and meta-analysis. *PLoS Med.* **15**(1), e1002494 (2018).
5. Buyuk, G. N. et al. Risk factors for intrapartum cesarean section delivery in low-risk multiparous women following at least a prior vaginal birth (robson classification 3 and 4). *Rev. Bras. Ginecol. Obstet.* **43**(6), 436–441 (2021).
6. Derbent, A. U., Karabulut, A., Yildirim, M., Simavli, S. A. & Turhan, N. Evaluation of risk factors in cesarean delivery among multiparous women with a history of vaginal delivery. *J. Turk. Ger. Gynecol. Assoc.* **13**(1), 15–20 (2012).
7. Ramos, S. Z. et al. Predicting primary cesarean delivery in pregnancies complicated by gestational diabetes mellitus. *Am. J. Obstet. Gynecol.* **229**(5), 549 (2023).
8. Wagaw, A. C., Sendekie, A. K., Nigatu, S. G. & Mihretie, G. S. Preferred mode of delivery and its associated factors in pregnant women with a previous cesarean scar at a tertiary care hospital in Ethiopia: institutional-based cross-sectional study. *BMC Pregn. Childbirth* **23**(1), 585 (2023).
9. Young, C. B. et al. Mode of delivery after a previous cesarean birth, and associated maternal and neonatal morbidity. *Cmaj* **190**(18), E556–e64 (2018).
10. Burke, N. et al. Prediction of cesarean delivery in the term nulliparous woman: results from the prospective, multicenter Genesis study. *Am. J. Obstet. Gynecol.* **216**(6), 598 (2017).
11. Danilack, V. A. et al. Development and validation of a risk prediction model for cesarean delivery after labor induction. *J. Womens Health (Larchmt)* **29**(5), 656–669 (2020).
12. Tun, M. H. et al. Prediction of odds for emergency cesarean section: a secondary analysis of the CHILd term birth cohort study. *PLOS ONE* **17**(10), e0268229 (2022).
13. Campillo-Artero, C., Serra-Burriel, M. & Calvo-Pérez, A. Predictive modeling of emergency cesarean delivery. *PLoS One* **13**(1), e0191248 (2018).
14. Fente, B. M., Asaye, M. M., Gudayu, T. W., Mihret, M. S. & Tesema, G. A. Prediction of unplanned cesarean section using measurable maternal and fetal characteristics, Ethiopia, a retrospective cohort study. *BMC Pregn. Childbirth* **24**(1), 161 (2024).
15. Irwinda, R., Hiksas, R., Lokeswara, A. W. & Wibowo, N. Maternal and fetal characteristics to predict c-section delivery: a scoring system for pregnant women. *Womens Health (Lond.)* **17**, 17455065211061969 (2021).
16. Schuit, E. et al. A clinical prediction model to assess the risk of operative delivery. *Bjog* **119**(8), 915–923 (2012).
17. Beksac, M. S., Tanacan, A., Bacak, H. O. & Leblebicioglu, K. Computerized prediction system for the route of delivery (vaginal birth versus cesarean section). *J. Perinat. Med.* **46**(8), 881–884 (2018).
18. Rahman, S., Khan, M. I., Satu, M. S. & Abedin, M. Z. Risk prediction with machine learning in cesarean section: optimizing healthcare operational decisions. In *Signal Processing Techniques for Computational Health Informatics* (eds. Ahad, M. A. R. & Ahmed, M. U.) 293–314 (Springer International Publishing, 2021).
19. Luo, J. et al. Pregnancy complications among nulliparous and multiparous women with advanced maternal age: a community-based prospective cohort study in China. *BMC Pregnancy Childbirth* **20**(1), 581 (2020).

20. Harrell, F. E. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis* (Springer, 2001).
21. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B: Stat. Methodol.* **58**(1), 267–288 (1996).
22. Steyerberg, E. W. & Vergouwe, Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur. Heart J.* **35**(29), 1925–1931 (2014).
23. Pavlou, M. et al. How to develop a more accurate risk prediction model when there are few events. *Bmj* **351**, h3868 (2015).
24. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann. Intern. Med.* **162**(1), 55–63 (2015).
25. Arrowsmith, S., Robinson, H., Noble, K. & Wray, S. What do we know about what happens to myometrial function as women age? *J. Muscle Res. Cell. Motil.* **33**(3–4), 209–217 (2012).
26. Caughey, A. B., Cahill, A. G., Guise, J.-M. & Rouse, D. J. Safe prevention of the primary cesarean delivery. *Am. J. Obstet. Gynecol.* **210**(3), 179–193 (2014).

Acknowledgements

The authors are thankful to the data collectors and study participants.

Author contributions

NBY conceived the study. NBY and EGM designed the study, cleaned the data, and performed the analysis. NBY wrote the original draft of the manuscript. NBY and EGM wrote, reviewed, and edited the final version of the manuscript. Both authors read and approved the final version of the manuscript.

Funding

The authors did not receive any funds for this work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-86015-w>.

Correspondence and requests for materials should be addressed to N.B.Y.

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

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

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

© The Author(s) 2025