Developing and validating a risk prediction model for preterm birth at Felege Hiwot Comprehensive Specialized Hospital, North-West Ethiopia: a retrospective follow-up study

Sefineh Fenta Feleke,1 Zelalem Alamrew Anteneh,2 Gizachew Tadesse Wassie,2,3 Anteneh Kassa Yalaw3, Anteneh Mengist Dessie4

ABSTRACT

Objective To develop and validate a risk prediction model for the prediction of preterm birth using maternal characteristics.

Design This was a retrospective follow-up study. Data were coded and entered into EpiData V.3.02, and were analysed using R statistical programming language V.4.0.4 for further processing and analysis. Bivariable logistic regression was used to identify the relationship between each predictor and preterm birth. Variables with p<0.25 from the bivariable analysis were entered into a backward stepwise multivariable logistic regression model, and significant variables (p<0.05) were retained in the multivariable model. Model accuracy and goodness of fit were assessed by computing the area under the curve of calibration plot (discrimination) and calibration plot (calibration), respectively.

Setting and participants This retrospective study was conducted among 1260 pregnant women who did prenatal care and finally delivered at Felege Hiwot Comprehensive Specialised Hospital, Bahir Dar city, north-west Ethiopia, from 30 January 2019 to 30 January 2021.

Results Residence, gravidity, haemoglobin <11 mg/dL, early rupture of membranes, antepartum haemorrhage and pregnancy-induced hypertension remained in the final multivariable prediction model. The area under the curve of the model was 0.816 (95% CI 0.779 to 0.856).

Conclusion This study showed the possibility of predicting preterm birth using maternal characteristics during pregnancy. Thus, use of this model could help identify pregnant women at a higher risk of having a preterm birth to be linked to a centre.

INTRODUCTION

Preterm birth is described as babies that are born alive before the end of 37 weeks of pregnancy.1 Preterm birth can be accidental (due to spontaneous preterm labour and/or preterm membrane rupture) or induced by the provider (by caesarean or labour induction).2 Most preterm births happen spontaneously.3

An estimated 15 million babies worldwide are born too early per year. That is more than 1 in 10 infants. About 1 million newborns die per year because of preterm birth complications.4

Across 184 countries, the rate of preterm birth ranges from 5% to 18% of babies born.5 However, there are stark disparities in survival rates around the world. Half of the babies born at or below 32 weeks die in low-income settings due to lack of practical, cost-effective and critical care, such as comfort, breastfeeding assistance, basic infection care and trouble breathing.6

Furthermore, the effect of preterm birth is also prolonged beyond the neonatal
phases and throughout life.7 Hence, the largest risk of severe health issues, including cerebral palsy, intellectual disability, chronic lung disease, and vision and hearing loss, is faced by babies born before maturity. This introduces a lifelong disability dimension. At some point in their lives, most people will face the struggles and potential disasters of preterm birth either directly in their families or indirectly through events for the nations.7,8

To alleviate this burden, in the past few decades, numerous methods have been attempted internationally, including in Ethiopia, to prevent and enhance the treatment of preterm births.9–11 As part of the strategy, it is essential to diagnose or predict preterm birth earlier in pregnancy to take appropriate measures for high-risk groups. However, in most nations, predicting preterm birth is still largely based on subjective clinical experience. This approach may increase unnecessary hospital admissions and unnecessary but potentially harmful treatments, such as the use of steroids for the maturation of the fetal lung and tocolysis.12 13

There were clinical prediction models that aim to estimate the likelihood of preterm birth that include laboratory tests that are typically inaccessible in low-resource settings, such as fetal fibronectin, insulin-like growth factor binding protein-1, interleukin 6 (IL-6) and placental α-macroglobulin-1.14–19 Most current research on Preterm Birth (PTB) prediction focuses on finding PTB risk factors using a hypothesis-testing methodology in highly controlled environments. PTB has been linked to a number of risk factors, including previous preterm labour, multiple gestation (carrying several children), diabetes, problems with the cervix, uterus or placenta, smoking, and infections.20–22 However, women who have preterm delivery often have no known risk factors.23 In addition, some of the predictors (such as prior PTB) do not apply for first-time mothers.

Predicting the risk of PTB in pregnant women has been the subject of numerous studies,24 but no model exists that is accurate enough to be used in clinical settings. Most research (eg, cervical length or fetal fibronectin) has concentrated on predictors during the second or third trimester.25 These predictors, however, can only forecast PTB at intermediate risk and have only been shown to be reliable in high-risk populations. Unfortunately, the majority of women who give birth early have no evident risk factors, and more than half of PTBs happen in low-risk pregnancies, indicating the limited usefulness of using fetal fibronectin or cervical length in the general population.26

Due to scarce resources, using readily available data to predict PTB seems appealing in low-income and middle-income communities. But relatively few models have been made public. The considerable range in PTB occurrence across the globe, which suggests differences in exposure to psychosocial, sociodemographic and medical risk factors as well as genetic variations, is also significant.27–29 As a result, it is necessary to develop and evaluate PTB prediction models in various populations.

Therefore, developing and validating a risk prediction model for the prediction of preterm birth using maternal (clinical and nonclinical) characteristics based on the available measurements is paramount to allow early preterm birth interventions such as in utero transfer to tertiary care centres, appropriate corticosteroid administration while preventing excessive use, neuroprotective magnesium sulfate therapy and antibiotic treatment in the event of infection.14 30

METHODS AND MATERIALS

Study setting

This retrospective study was conducted among 1260 pregnant women who used prenatal care and finally delivered at the Felege Hiwot Comprehensive Specialized Hospital, Bahir Dar city, North-West Ethiopia, from 30 January 2019 to 30 January 2021. Bahir Dar is the capital city of Amhara national regional state and is 575 km north-west of Addis Ababa.

The hospital currently has a total of 1431 employees (5 obstetricians and gynaecologists and 63 midwives among others) in different disciplines. It has a total of 500 formal beds, 11 wards (emergency ward and inpatient wards such as gynaecological and obstetric, surgical, orthopaedic, medical, paediatric, Laison and Director (L&D), eye unit, Neonatal Intensive Care Unit (NICU), psychiatric, oncology, and 22 Outpatient Departements (OPDS)), 39 clinical and non-clinical departments/service units/ providing laboratories, diagnostic, curative and rehabilitation services at outpatient and inpatient bases as well as disease prevention and health promotion services.

Sample size determination

The sample size required for model development was determined based on the minimum standard of 10 events per candidate predictor considered, according to the formula N = (n*10)/I, where N is the sample size, n is the number of candidate predictor variables and I is the estimated event rate in the population.31 Since there were 17 candidate predictors considered and 10 events per candidate predictor, the estimated number of events for the study was 170. According to a study done at Debre Tabor hospital the prevalence of preterm 13%,32 so taking into account this the required sample size was calculated as follows, n=170*100/13=1308.

Study design and participants

The theoretical design of the present study was; the incidence of preterm birth as a function of multiple predictors during pregnancy. The source population of the study was all pregnant mothers who gave birth at Felege Hiwot Comprehensive Specialized Hospital (FHCSH). To be included in this study, mothers must meet all of the following eligibility criteria; all medical records of Addis Ababa.

Study setting

- Felege Hiwot Comprehensive Specialized Hospital, Bahir Dar city, North-West Ethiopia, from 30 January 2019 to 30 January 2021.
Sampling method and procedures
A simple random sampling technique was employed to select participants using the medical registration number of a delivered mother from the delivery registration book. First, all mothers who delivered at FHCSH in the last 2 years were identified from the delivery registration book. After that, records of mothers who met the inclusion criteria were included in the study. Subsequently, a sampling frame was prepared. Finally, the study unit was selected by using a computer-generated random number.

Data collection
Outcome assessment: The outcome variable was attributed to women whose medical records indicated a physician or midwife diagnosis of preterm birth and delivery between 28 and 36 completed weeks of gestation. The gestational age (GA) was measured using either Last Normal Menstrual Period (LNMP), which is found to be a more reliable measure of GA in a low-resource setting,33 34 or an early ultrasound result (12 weeks).

Predictor assessment data were collected using a structured checklist through chart review. Checklists were developed after reviewing various relevant literatures.35–39 Checklists consisted of sociodemographic (maternal age, residence) and maternal obstetric characteristics: (history of preterm birth, history of abortion, history of stillbirth, gravidity, parity, multiple pregnancy, antepartum haemorrhage (APH), premature rupture of membranes (PROM), gestational diabetes mellitus, and pregnancy-induced hypertension (PIH)), and maternal medical condition: (haemoglobin (HGB) level, diabetic mellitus, chronic hypertension, Urethral Tract Infection (UTI) and HIV).

Quality assurance mechanisms
To maintain the quality of data, the data collectors and supervisors were trained for a day on the objective of the study, the content of the checklists and how to fill the checklists. Reviewing of 15 medical chart records of mothers who gave birth at Felege Hiwot Comprehensive Specialized Hospital, which is found in North-West Ethiopia, were done. After that, some adjustments (removing variables that were not available in the medical records of mothers) were made accordingly. The checklist was developed in English.

Data processing and analysis
Data were entered into a software application (EpiData, V.3.02) and were analysed by using R statistical programming language V.4.0.4 for further processing and analysis. There were 13 (1%), 2 (0.2 %), 11 (0.9 %), 15 (2.5%), 21 (1.7%), 29 (2.3%), 20 (1.6%) and 20 (1.6%) missing values for PROM, residence, chronic hypertension, multiple pregnancies, gestational diabetes mellitus, PIH, APH and HGB, respectively.

We assumed the data were missing at random, and we, therefore, performed multivariate imputation by chained equations for all variables evaluated in the prediction model.40 Sensitivity analysis was performed to assess whether the assumption of missing at random is valid or not, and the results were reasonably comparable (table 1). Descriptive statistics including median, IQR and percentages, were carried out.

Model development and validation
For model development, bivariable logistic regression was done to obtain insight into the association between each potential predictor and preterm birth. Variables with 𝑝≤0.25 from the bivariable analysis were entered into a backward stepwise multivariable logistic regression model, and significant variables (𝑝<0.05) were retained in the multivariable model. The results of significant predictors were reported as coefficients with 95% CIs.

To check for the model accuracy and goodness of fit, we computed the area under the receiver operating characteristic (ROC) curve (discrimination) and calibration

Table 1  Sensitivity analysis of the model to predict preterm birth: comparison of the regression coefficients, SEs and p values for complete case analysis and multiple imputed data

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>Complete case analysis</th>
<th>Multiple imputations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td>Chronic hypertension (yes)</td>
<td>0.7313</td>
<td>0.6297</td>
</tr>
<tr>
<td>Residence (rural)</td>
<td>0.815</td>
<td>0.1946</td>
</tr>
<tr>
<td>GDM (yes)</td>
<td>0.709</td>
<td>0.4028</td>
</tr>
<tr>
<td>HGB (&lt;11 g/dL)</td>
<td>0.497</td>
<td>0.2185</td>
</tr>
<tr>
<td>PROM (yes)</td>
<td>1.898</td>
<td>0.2080</td>
</tr>
<tr>
<td>APH (yes)</td>
<td>1.194</td>
<td>0.2858</td>
</tr>
<tr>
<td>PIH (yes)</td>
<td>1.353</td>
<td>0.2600</td>
</tr>
<tr>
<td>Multiple pregnancy (yes)</td>
<td>0.539</td>
<td>0.3173</td>
</tr>
<tr>
<td>Gravidity (primigravida)</td>
<td>0.426</td>
<td>0.1944</td>
</tr>
</tbody>
</table>

APH, antepartum haemorrhage; GDM, gestational diabetes mellitus; HGB, haemoglobin; PIH, pregnancy-induced hypertension; PROM, premature rupture of membranes.
plot (calibration) using ‘classifierplots’ and ‘givitiR’ packages of R, respectively. The area under the curve (AUC) ranged from 0.5 (no predictive ability) to 1 (perfect discrimination). The regression coefficients and their 95% confidence levels, and the AUC were adjusted for overfitting or overoptimism using the bootstrapping technique. To make internal validation, we computed 1000 random bootstrap samples with the replacement of all predictors in the data. The model’s predictive performance after bootstrapping is considered as the performance that can be expected when the model is applied to future similar populations. To evaluate the clinical and public health impact of the model, we performed a decision curve analysis (DCA) of standardised net benefits across a range of threshold probabilities (0 to 1). In the DCA, the model was compared with two extreme scenarios: ‘intervention for all’ and ‘no intervention’. In our case, the intervention considered is the referral of high-risk pregnant women to facilities where appropriate, corticosteroid administration, and antibiotic treatment is given.

**Risk score development**

To construct an easily applicable preterm birth prediction score, we transformed each coefficient of the model into a rounded number by dividing it by the lowest coefficient. The number of points was subsequently rounded to the nearest integer. We determined the total score for each individual by assigning points for each variable present and adding them up. The score was transformed to dichotomous, allowing each pregnant woman to be classified as having a high risk or low risk of preterm birth. The ROC curve was plotted and the AUC was calculated to measure the discriminatory power of the scoring system.

**Patient and public involvement**

There was no direct interaction with patients in this study and no direct patient involvement in the design or conduct of this study.

**RESULT**

**Demographic, obstetric and clinical characteristics of mothers**

A total of 1260 study cards were reviewed from a sample of 1308; about 48 cards were not reviewed due to the outcome of intrauterine fetal death and abortion. Table 2 shows the demographic, obstetric and clinical characteristics of mothers who gave birth, included in the analysis. The median age of the study participants was 26 years with IQR 24–30 years; the majority of the participants (1086, 86.2%) were in the age group of 20–34 years.

More than three-fourths of the participants 926 (73.49%) were urban residents. Of the total of mothers who delivered at FHCSH, more than two-thirds of 841 (66.7%) were multigravida. About parity, above, half of them (713, 56.6%) were multipara. Concerning past obstetric history, 55 (6.5%) of them had a history of previous preterm birth, 76 (9%) of them had a history of stillbirth and 162 (19.3%) of them had a history of abortion.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Category</th>
<th>Frequency</th>
<th>Per cent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravidity</td>
<td>Primigravida</td>
<td>419</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>Multigravida</td>
<td>841</td>
<td>66.7</td>
</tr>
<tr>
<td>Residence</td>
<td>Urban</td>
<td>926</td>
<td>73.5</td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>334</td>
<td>26.5</td>
</tr>
<tr>
<td>GDM</td>
<td>Yes</td>
<td>44</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1216</td>
<td>96.5</td>
</tr>
<tr>
<td>APH</td>
<td>Yes</td>
<td>84</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1176</td>
<td>93.3</td>
</tr>
<tr>
<td>PIH</td>
<td>Yes</td>
<td>110</td>
<td>8.73</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1150</td>
<td>91.27</td>
</tr>
<tr>
<td>HGB level</td>
<td>&lt;11g/dL</td>
<td>236</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td>≥11g/dL</td>
<td>1024</td>
<td>81.3</td>
</tr>
<tr>
<td>Chronic hypertension</td>
<td>Yes</td>
<td>21</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1239</td>
<td>98.3</td>
</tr>
<tr>
<td>PROM</td>
<td>Yes</td>
<td>195</td>
<td>15.5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1065</td>
<td>84.5</td>
</tr>
<tr>
<td>Multiple pregnancies</td>
<td>Yes</td>
<td>90</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1170</td>
<td>92.8</td>
</tr>
</tbody>
</table>

APH, antepartum haemorrhage; GDM, gestational diabetes mellitus; HGB, haemoglobin; PIH, pregnancy-induced hypertension; PROM, premature rupture of membrane.

Development of a prediction model for preterm birth

Out of 1260 delivered neonates, 109 (13.4%, 95% CI 11.6% to 15.4%) were preterm infants.

The bivariable logistic regression analysis found several factors were eligible to be included in the prediction model. These variables were HGB level, gravidity, residence, gestational diabetes mellitus, APH, PIH, chronic hypertension, PROM and multiple pregnancies. Using the results, a prediction model was developed, and the equation for the prediction model was obtained (table 3).

Both backward and forward selection showed the same results; β after internal validation with bootstrapping is shown. Simplified risk score: we divided the coefficient of predictors included in the reduced model by the smallest (0.666). The probability or risk of preterm birth=1/(1-exp – (-3.517+1.148 * Residence (rural)+0.666 *gravidity (primigravida)+2.051*PROM (yes)+1.348 *APH (yes)+1.387*PIH+0.677*HGB (<11 g/dL).

The AUC of the final reduced model was 0.816 (95% CI 0.779 to 0.856) (figure 1A). The calibration test had a p value of 0.492, indicating that the model does not misrepresent the data or the calibration of the model was visually

---

**Table 2** Demographic, obstetric and clinical characteristics of mothers who gave birth at FHCSH, North-West Ethiopia, 2021

- Gravidity: Primigravida 419 (33.3%), Multigravida 841 (66.7%)
- Residence: Urban 926 (73.5%), Rural 334 (26.5%)
- GDM: Yes 44 (3.5%), No 1216 (96.5%)
- APH: Yes 84 (6.7%), No 1176 (93.3%)
- PIH: Yes 110 (8.73%), No 1150 (91.27%)
- HGB level: <11g/dL 236 (18.7%), ≥11g/dL 1024 (81.3%)
- Chronic hypertension: Yes 21 (1.7%), No 1239 (98.3%)
- PROM: Yes 195 (15.5%), No 1065 (84.5%)
- Multiple pregnancies: Yes 90 (7.2%), No 1170 (92.8%)

APH, antepartum haemorrhage; GDM, gestational diabetes mellitus; HGB, haemoglobin; PIH, pregnancy-induced hypertension; PROM, premature rupture of membrane.
accurate since the observed and predicted probabilities were similar (figure 1B).

In addition, to verify whether any maternal characteristics were used as a specific predictor of preterm birth we performed an ROC analysis. The analysis indicated that residence (AUC=0.604, 95% CI 0.564 to 0.643), gravidity (AUC=0.59, 95% CI 0.571 to 0.628), PROM (AUC=0.580, 95% CI 0.544 to 0.616), APH (AUC=0.695, 95% CI 0.661 to 0.729), PIH (AUC=0.721, 95% CI 0.685 to 0.757) and HGB (<11 mg/dL (AUC=0.630, 95% CI 0.591 to 0.668) emerged as better predictors of preterm birth (figure 2).

Validation of the model with the bootstrap technique showed hardly any indication of undue influence by particular observations, with an optimism coefficient of 0.085 and resulting AUC of 0.789 (corrected 95% CI 0.729 to 0.827) (figure 1A).

Using ‘SpEqualSe’, the suggested threshold score to predict preterm birth using risk scores is ≥3 with a sensitivity of 75.14% and specificity of 67.46% (table 4).

When dichotomised into low risk (<3) and high risk (≥3) based on the risk score, 278 (14.36%) were categorised as high risk and 982 (77.9%) as low risk for preterm birth.

**DISCUSSION**

In this study, the incidence of preterm birth was found to be 13.4%. Maternal characteristics were identified in this retrospective study to build a preterm birth prediction risk score. We intended to employ maternal features that are easily accessible and pertinent to clinical practice in countries with constrained resources, including Ethiopia. These nations may not have the financial resources to pay for ultrasound exams and laboratory tests. The optimal combination of maternal features to predict preterm birth includes residency, gravidity and HGB <11 mg/dL, early rupture of membranes, APH, and PIH, according to the prediction model. The model has an AUC of 0.816 (95% CI 0.776 to 0.856). Predicting the probability of preterm birth in pregnant women is essential to take appropriate measures accordingly. Identifying women at risk of preterm birth is an important task for clinical care providers. However, in low-income and middle-income countries, there are only a few methods available for reliably predicting actual preterm labour in women. Previously, the focus of the research was to explain the maternal and fetal determinants of preterm birth. In recent years, the focus shifted to predicting preterm birth optimally using a combined set of characteristics.

Without any advanced laboratory or imaging testing, this study measured the predicted performance of a model based on maternal features during pregnancy. Furthermore, we discovered that using SpEqualSe as an optimal cut-off point, the sensitivity and specificity of this prediction model reached 75.14% and 67.46%, respectively, at the score threshold of 3.

In our study, a combination (residency, gravidity, HGB <11 mg/dL, early rupture of membranes, APH and PIH) of maternal characteristics was used as a predictor of preterm birth, achieving a sensitivity of 75.74%, a specificity of 72.87%, a positive predictive value of 30.2%, and a negative predictive value of 95.1%.

<table>
<thead>
<tr>
<th>Predictor variables*</th>
<th>Original β (95% CI)</th>
<th>Bootstrap β</th>
<th>P value</th>
<th>Risk score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence (rural)</td>
<td>1.161 (0.780 to 1.545)</td>
<td>1.148</td>
<td>&lt;0.001</td>
<td>2</td>
</tr>
<tr>
<td>Gravidity (primigravida)</td>
<td>0.675 (0.291 to 1.061)</td>
<td>0.666</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>PROM (yes)</td>
<td>2.081 (1.669 to 2.50)</td>
<td>2.051</td>
<td>&lt;0.001</td>
<td>3</td>
</tr>
<tr>
<td>APH (yes)</td>
<td>1.364 (0.806 to 1.915)</td>
<td>1.348</td>
<td>&lt;0.001</td>
<td>2</td>
</tr>
<tr>
<td>PIH (yes)</td>
<td>1.387 (0.887 to 1.879)</td>
<td>1.368</td>
<td>&lt;0.001</td>
<td>2</td>
</tr>
<tr>
<td>HGB &lt;11 g/dL</td>
<td>0.676 (0.255 to 1.09)</td>
<td>0.677</td>
<td>&lt;0.001</td>
<td>1</td>
</tr>
</tbody>
</table>

*Variables retained in the reduced model are; residence, APH, HGB, PIH, gravidity and PROM.

APH, antepartum haemorrhage; HGB, haemoglobin; PROM, premature rupture of membrane.

**Table 3** Coefficients and risk scores of each predictor included in the model to predict preterm birth (n=1260)
PIH) of maternal characteristics resulted in an AUC of 0.816 (95% CI 0.776 to 0.856), with excellent accuracy according to diagnostic accuracy classification.44

We found that early rupture of membranes is a strong predictor of preterm birth. Similar evidence was found in different studies.36 37 45 46 The effect of a burst membrane on uterine contraction could explain this. Existing scientific evidence confirms that when a membrane ruptures, natural uterotonic chemicals are released, and these uterotonic chemicals drive uterine contraction, resulting in PTB. This finding suggested that due attention should be given to women with PROM.

In our study, PIH is a strong predictor of preterm birth. Similar studies have demonstrated that PIH was predictive of subsequent preterm birth.47 48 This could be related to vascular injury to the placenta caused by PIH issues or iatrogenesis caused by the severity of hypertension or its complications. As a result, the oxytocin receptors are activated, resulting in preterm labour and delivery. Or else this conclusion could be explained by current scientific evidence suggesting that PIH is linked to vascular and placental injury, which causes oxytocin receptors to be activated, resulting in PTB. Therefore, it is imperative to identify populations at risk of PIH and introduce risk-lowering interventions.

Another strong predictor of preterm birth is the place of residence. Current evidence shows that there is an association between preterm birth and rural residence.49–53
This gap may be explained by the greater accessibility and availability of maternal health services in metropolitan regions. It has long been understood that social deprivation and the nuanced interactions between them affect prenatal outcomes, including premature birth.54 Hence, accessing maternal health services targeted to rural women could improve prenatal outcomes including the risk of pretermbirth.

APH is the predictor of preterm birth which is supported by different studies.55 Identification of groups at risk for APH and the introduction of risk-reducing measures are therefore essential. Other predictors of preterm birth are gravidity and HGB <11 g/dL (anaemia) which is in line with different studies.52,56 Molecular factors could explain how anaemia, iron deficiency, or both, could result in preterm delivery. In reality, a number of plausible molecular processes have linked anaemia to a higher risk of premature birth. Accordingly, maternal and fetal stress can be caused by anaemia (by resulting in hypoxia) and iron deficiency (by increasing serum norepinephrine concentrations), which in turn induces the production of corticotrophin-releasing hormone (CRH). Additionally, iron deficiency may raise the risk of maternal infections, which can again boost the synthesis of CRH. High levels of CRH are known to be a risk factor for PTB since they increase the likelihood of PTB. Thus, we can conclude

Figure 2  Receiver operating characteristic (ROC) curve of maternal parameters for prediction of postpartum glucose intolerance: residence, premature rupture of membranes (PROM), antepartum haemorrhage (APH), pregnancy-induced hypertension (PIH), haemoglobin (HGB) and gravidity.

Figure 3  A decision curve plotting the net benefit of the model against threshold probability.
that, in order to prevent PTB, routine ANC services need to place a greater emphasis on anaemia prevention.

A study conducted in China showed that a model developed using advanced maternal age, lower maternal height, history of preterm delivery, amount of vaginal bleeding during pregnancy, and lack of folic acid intake before pregnancy for the prediction of overall preterm birth had an AUC of (0.6). It had lower discriminatory performance than the present study, this difference may be due to some of the predictors they used such as lower maternal height, lack of folic acid intake before pregnancy and advanced maternal age. However, the predictors they used such as lack of folic acid intake before pregnancy are not easily obtainable information in routine clinical practice, which makes their model less practical in our setting. This prediction model constitutes variables that are easily obtainable and have reasonable accuracy to be used by both mid-level and lower-level health professionals in primary care settings. Among the maternal characteristics included in our model, five can be easily found by history taking and one by test for HGB.

The model’s accuracy is consistent with a retrospective study done in China that established a preterm birth prediction model based on maternal characteristics, including demographics and clinical characteristics, and a model with predictors (gravidity, educational status, residency, history of preterm birth, twin pregnancy, gestational diabetes mellitus (type I or II), chronic hypertension, and place of birth) with AUC of 0.749 (95% CI 0.732 to 0.767).

On the other hand, a model incorporating four predictors (cervical length at admission, GA, amniotic fluid, glucose and IL-6) has an AUC of 0.86 and similarly, the combination of biophysical, biochemical, immunological, microbiological, fetal cell, exosomal, or cell-free RNA at different GAs, integrated as part of a multivariable predictor model may be necessary to advance our attempts to predict spontaneous preterm labor (sPTL) and preterm birth. In the prediction of spontaneous preterm birth within 48 hours, a prognostic model including quantitative fetal fibronectin (qfFN) and clinical risk factors showed excellent results. Both models have higher discriminatory performance. Our study has lower discriminatory performance than the studies described above, this could be because we used secondary data available from the register and as this data set is limited and some variables that require advanced laboratory tests were not included in the model.

Hence, predictors necessitate laboratory testing, which is often unavailable in low-resource settings. As a result, such predictors are difficult to come by in ordinary clinical and public health practice, making the model less useful.

A study conducted in the UK found that data on maternal characteristics and obstetric history at 11–13

Figure 4  Area under the receiver operating characteristics (ROC) curve for the simplified risk score to predict the risk of preterm birth among mothers who gave birth at FHCSH, 2021. AUC, area under the curve.

<table>
<thead>
<tr>
<th>Score* (risk category)</th>
<th>Prediction model based on maternal characteristics</th>
<th>Incidence of preterm birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 (low)</td>
<td>982 (77.9%)</td>
<td>72 (7.9%)</td>
</tr>
<tr>
<td>≥3 (high)</td>
<td>278 (14.36%)</td>
<td>97 (53.59%)</td>
</tr>
<tr>
<td>Total</td>
<td>1260 (100%)</td>
<td>169 (13.4%)</td>
</tr>
</tbody>
</table>

*Score = (2*PIH) + (3*PROM) + (HGB <11 mg/dL) + 2*residence + (2*APH) + gravidity.
APH, antepartum haemorrhage; HGB, haemoglobin; PIH, pregnancy-induced hypertension; PROM, premature rupture of membranes.
weeks of gestation were predictive of spontaneous early preterm delivery; this model had an AUC of 0.67\(^3\) which had lower discriminatory performance than the present study. This difference may be the difference in the study population.

A model that predicts a risk of preterm delivery in women with multiple pregnancy incorporating previous preterm delivery, monochorionicity, smoking, educational level, and triplet pregnancy for preterm and very preterm delivery had c-indexes of 0.68 (95% CI 0.63 to 0.72) and 0.68 (95% CI 0.62 to 0.75), respectively.\(^2\) It had lower discriminatory performance than the present study. This might be due to the difference in the study populations. In the present study, the study population included both women who had multiple pregnancies and singleton pregnancy.

In our prediction score, using 3 as a cut-off point has an acceptable level of specificity, sensitivity, PPV and NPV to predict preterm birth. It is also possible to shift the cut-off point to increase either of the accuracy measures depending on the aim of the programme and availability of resources.

The strength of the study was using an adequate number of participants with the outcome, which helped us to construct the model using a sufficient number of predictor variables. In addition, our prediction model was constructed from easily obtainable maternal characteristics that make it applicable in primary care settings and multiple imputations were used to address missing data, which has been shown to be a valid technique for dealing with missing data within logistic regression models, resulting in less bias than excluding all women with missing data.

However, the findings from this study should be interpreted with the perspective of the following limitations. As a single-site study, it is confined to a single area, which needs external validation before using it in another context. Furthermore, data were collected from each mother’s card; due to this, some important variables were missed, such as previously highlighted factors with preterm birth in different studies.

**Conclusions and recommendations**

This study shows the possibility of predicting preterm birth using a simple prediction model constructed from maternal characteristics. Thus, the optimal combination of maternal characteristics such as residence, gravity, HGB <11 mg/dL, PROM, APH and PIH shows the possibility of predicting preterm birth using a simple prediction model constructed from maternal characteristics. In addition, risk score calculation based on a combination of predictors was effective and had comparable accuracy with the model-based approach of the original β coefficients. This score may assist in clinical decision making. In addition, incorporating this convenient and easily applicable score in the healthcare system can be used by clinicians to inform pregnant mothers about the future course of their outcome after external validation. Doing further research is needed to validate the prediction tool using prospective follow-up studies in another context before introducing it to clinical and public health practices.

**Contributors**

SFF conceived the study and wrote the manuscript. ZAA, SFT, GTW, AKY and AMD contributed to data analysis, study design and supervision of data collection. All authors participated in manuscript revision for intellectual content and approval of the final version. All authors have read and agreed to the published version of the manuscript. SFF is acting as guarantor.

**Competing interests**

None declared.

**Patient and public involvement**

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication**

Not applicable.

**Ethics approval**

This study involves human participants and was approved by the Institutional Review Boards (IRB) of Bahir Dar University. College of Medicine and Health Sciences with protocol number 083/2021. The study does not involve participants directly. We used secondary data.

**Provenance and peer review**

Not commissioned; externally peer reviewed.

**Data availability statement**

Data are available upon reasonable request.

**Open access**

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use

**ORCID iDs**

Sefineh Fenta Feleke http://orcid.org/0000-0002-3051-1296

Zelalem Alamrew Anteneh http://orcid.org/0000-0003-4879-675X

Gizachew Tadesse Wassie http://orcid.org/0000-0001-6213-0191

Anteneh Mengist Desse http://orcid.org/0000-0002-5919-9670

**REFERENCES**


