

BMJ Open 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, Anteneh Kassa Yalew,³
Anteneh Mengist Dessie ⁴

To cite: Feleke SF, Anteneh ZA, Wassie GT, *et al.* Developing and validating a risk prediction model for preterm birth at Felege Hiwot Comprehensive Specialized Hospital, North-West Ethiopia: a retrospective follow-up study. *BMJ Open* 2022;**12**:e061061. doi:10.1136/bmjopen-2022-061061

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-061061>).

Received 01 February 2022
Accepted 31 August 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Public health, Woldia University, Woldia, Ethiopia

²Department of Epidemiology and Biostatistics, Bahir Dar University College of Medical and Health Sciences, Bahir Dar, Ethiopia

³Department of Public Health, Wolkite University, Wolkite, Ethiopia

⁴Department of Public Health, Debre Tabor University, Debre Tabor, Ethiopia

Correspondence to

Sefineh Fenta Feleke;
fentasefineh21@gmail.com

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 $p \leq 0.25$ 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 receiver operating characteristic curve (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

STRENGTHS AND LIMITATIONS OF THE STUDY

- ⇒ An adequate number of participants with the outcome helped us to construct the model using a sufficient number of predictor variables and the inclusion of sensitivity analyses.
- ⇒ 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.
- ⇒ The prediction model is constructed from easily obtainable maternal characteristics that make it applicable in primary care settings.
- ⇒ 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 of preterm birth in different studies.

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

phase 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, Liaison 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 $N = (n \times 10) / I$, 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, $n = 170 \times 100 / 13 = 1308$.

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 mothers who gave live birth and had at least one Antenatal Care (ANC) follow-up in FHCSH 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, Urinary 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 $p \leq 0.25$ 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

Predictor variables	Complete case analysis			Multiple imputations		
	B	SE	P value	B	SE	P value
Chronic hypertension (yes)	0.7313	0.6297	0.24	0.581	0.6285	0.92
Residence (rural)	0.815	0.1946	<0.001	1.154	0.1958	<0.001
GDM (yes)	0.709	0.4028	0.07	0.472	0.4236	0.26
HGB (<11 g/dL)	0.497	0.2185	0.02	0.642	0.2153	0.001
PROM (yes)	1.898	0.2080	<0.001	2.097	0.2129	<0.001
APH (yes)	1.194	0.2858	<0.001	1.298	0.2874	<0.001
PIH (yes)	1.353	0.2600	<0.001	1.368	0.2523	<0.001
Multiple pregnancy (yes)	0.539	0.3173	0.08	0.446	0.3257	0.17
Gravidity (primigravida)	0.426	0.1944	0.02	0.711	0.1976	<0.001

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).⁴¹ 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 bootstrap⁴² 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)⁴³ 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

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

Characteristics	Category	Frequency	Per cent (%)
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	<11d/dl	236	18.7
	≥11 g/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.

of stillbirth and 162 (19.3%) of them had a history of abortion.

Development of a prediction model for preterm birth

Out of 1260 delivered neonates, 169 (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 = $1 / (1 + \exp^{-(-3.517 + 1.148 * \text{Residence (rural)} + 0.666 * \text{gravidity (primigravida)} + 2.051 * \text{PROM (yes)} + 1.348 * \text{APH (yes)} + 1.387 * \text{PIH} + 0.677 * \text{HGB (<11 g/dL)})}$.

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 3 Coefficients and risk scores of each predictor included in the model to predict preterm birth (n=1260)

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

*Variables retained in the reduced model are; residence, APH, HGB, PIH, gravidity and PROM.
 APH, antepartum haemorrhage; HGB, haemoglobin ; 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 (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 0.748 to 0.83).

Using 'SpEqualSe', the predicted risk cutoff point for the coefficient (β) was a probability of >0.1320, and the model has a sensitivity of 75.74%, specificity of 72.87%, a positive predictive value of 30.2%, and a negative predictive value of 95.1%.

When applying DCA, we first evaluate whether our model understudy has a higher net benefit than the default strategies (referring all and none). This model outperforms the default strategies across the relevant threshold range. The model has the highest net benefit across the entire range of threshold probabilities, which indicates that the model has the highest clinical and public health value. Hence, the referral decision made using the model has a higher net benefit than not referring at all or referring all regardless of their risk threshold as shown in figure 3.

Risk classification using a simplified risk score

We created a simplified risk score from the model for practical use. The reduced model's prediction score was simplified by rounding all regression coefficients. The simplified score had a considerably comparable prediction accuracy to the original β coefficients, with an AUC of 0.786 (95% CI 0.729 to 0.827) (figure 4). The possible minimum and maximum scores a mother can have are 0 and 11, respectively.

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 factors 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

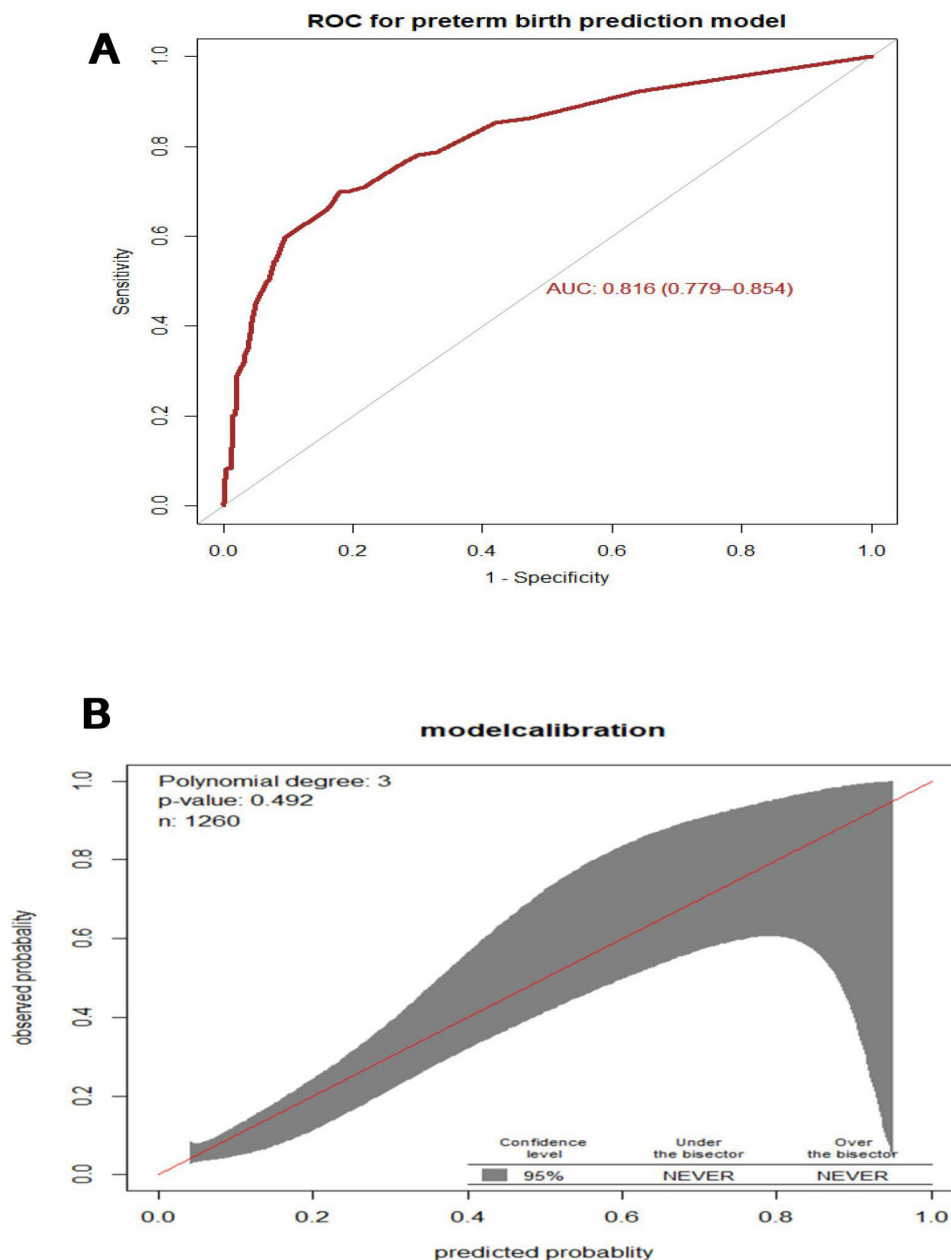


Figure 1 (A) Area under the receiver operating characteristics (ROC) curve for the prediction model, and (B) Predicted versus observed preterm birth probability in the sample. This analysis includes mothers who gave birth at FHCSH, 2021 (n=1260). Calibration plot created using 'givitiCalibrationBelt'f in R programming. AUC, area under the curve.

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}

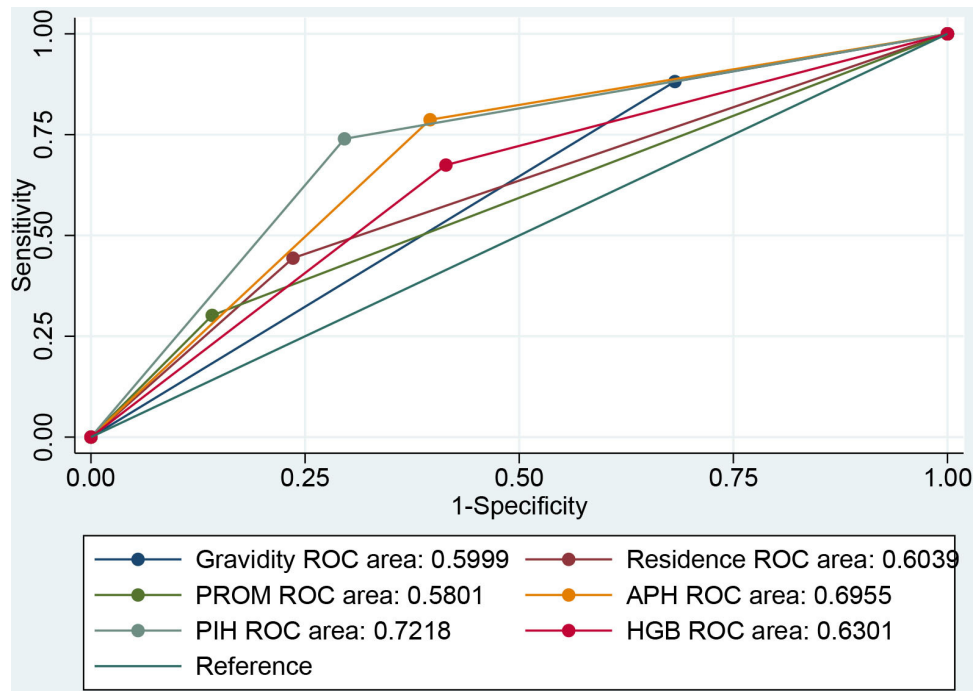


Figure 2 Receiver operating characteristic (ROC) curve of maternal parameters for prediction of postpartum glucose intolerance: residence, premature rupture of membranes (PROM), antepartum haemorrhage (APH), pregnancy-induced hypertension (PIH), haemoglobin (HGB) and gravity.

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 preterm birth.

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 gravity and HGB <11g/dL (anaemia) which is in line

with different studies.^{32,56} 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

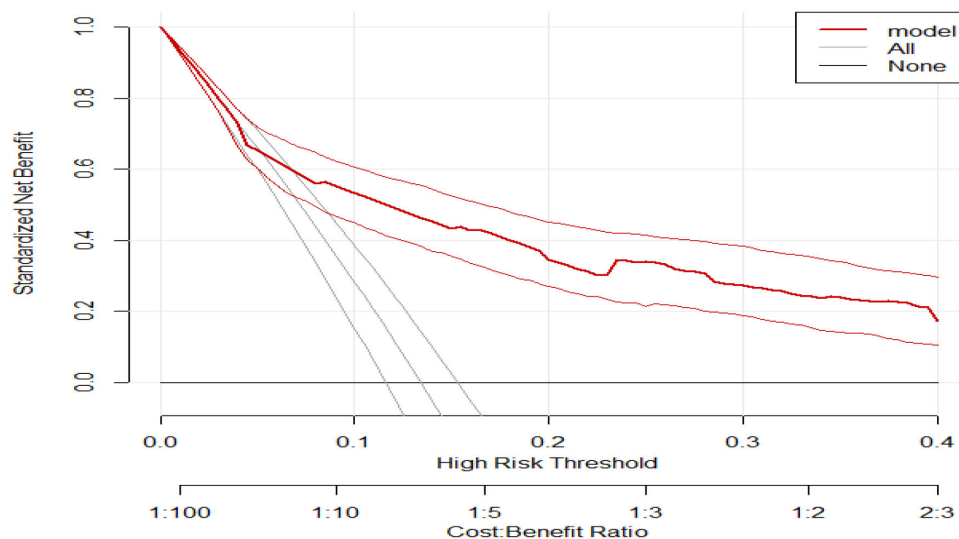


Figure 3 A decision curve plotting the net benefit of the model against threshold probability.

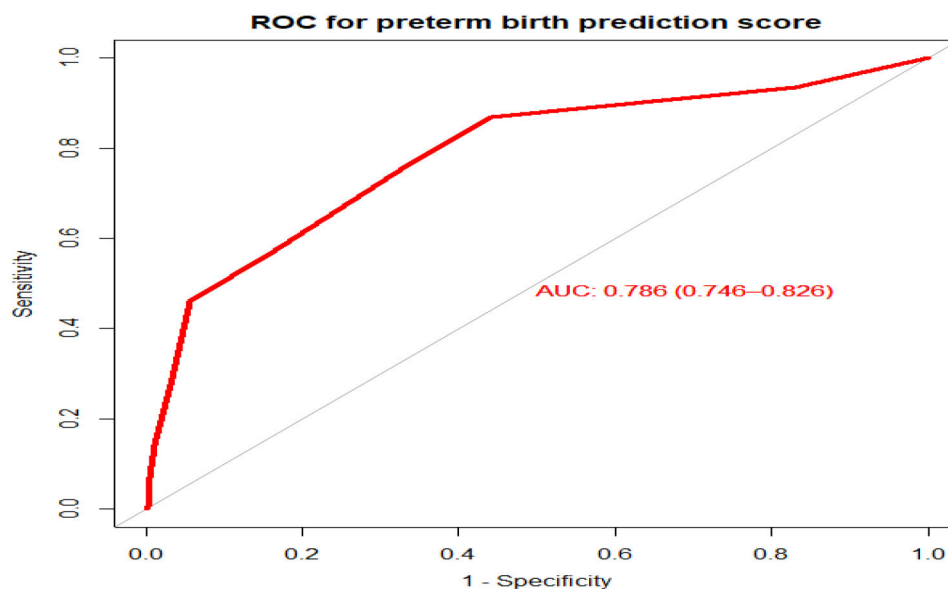


Figure 4 Area under the receiver operating characteristics (ROC) curve for the simplified risk score to predict the risk of preterm birth among mothers who gave birth at FHCSH, 2021. AUC, area under the curve.

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 a AUC of (0.6).⁵⁷ 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, pregestational diabetes mellitus (type I or II), chronic hypertension, and place of birth) with AUC of 0.749 (95% CI 0.732 to 0.767).⁴⁸

On the other hand, a model incorporating four predictors (cervical length at admission, GA, amniotic fluid, glucose and IL-6) has an AUC of 0.86⁵⁸ 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.^{59,60} 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

Table 4 Risk classification of preterm birth using the simplified prediction score (n=1260)

Score* (risk category)	Prediction model based on maternal characteristics	
	Number of mothers	Incidence of preterm birth
<3 (low)	982 (77.9%)	72 (7.9%)
≥3 (high)	278 (14.36%)	97 (53.59%)
Total	1260 (100%)	169 (13.4%)

*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 0.67⁶¹ 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.⁶² 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, gravidity, 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, SFF, 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 is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

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 Dessie <http://orcid.org/0000-0002-5919-9670>

REFERENCES

- World Health Organization. *Preterm birth and low birth weight*, 2020.
- Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. *Lancet* 2008;371:75e84:75–84.
- Althabe F. *Born too soon: the global action report on preterm birth*. World Health Organization, 2012.
- Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the sustainable development goals. *Lancet* 2016;388:3027–35.
- World Health Organization. WHO fact sheet on preterm birth. Available: <http://www.who.int/mediacentre/factsheets/fs363/en/>
- World Health Organization. WHO fact sheet: preterm birth. World Health Organization, Geneva, Switzerland. Available: <http://www.who.int/mediacentre/factsheets/fs363/en/>
- Li S, Xi B, Xi B. Preterm birth is associated with risk of essential hypertension in later life. *Int J Cardiol* 2014;172:e361–3.
- Blencowe H, Cousens S, Chou D, et al. Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health* 2013;10:S2.
- Soon BT. *The global action report on preterm birth*. Geneva: World Health Organization, 2012.
- Griffin JB, Jobe AH, Rouse D, et al. Evaluating WHO-Recommended interventions for preterm birth: a mathematical model of the potential reduction of preterm mortality in sub-Saharan Africa. *Glob Health Sci Pract* 2019;7:215–27.
- Victora CG, Rubens CE, GAPPS Review Group. Global report on preterm birth and stillbirth (4 of 7): delivery of interventions. *BMC Pregnancy Childbirth* 2010;10:S4.
- Kemp MW, Newnham JP, Challis JG, et al. The clinical use of corticosteroids in pregnancy. *Hum Reprod Update* 2016;22:240–59.
- Lorthe E, Goffinet F, Marret S, et al. Tocolysis after preterm premature rupture of membranes and neonatal outcome: a propensity-score analysis. *Am J Obstet Gynecol* 2017;217:212.e1–212.e12.
- Oskovi Kaplan ZA, Ozgu-Erdinc as: prediction of preterm birth: maternal characteristics, ultrasound markers and biomarkers: an updated overview 2018;2018:8367571.
- Lucaroni F, Morciano L, Rizzo G, et al. Biomarkers for predicting spontaneous preterm birth: an umbrella systematic review. *J Matern-Fet & Neo Med* 2018;31:726–34.

- 16 Lee KA, Chang MH, Park M-H, *et al.* A model for prediction of spontaneous preterm birth in asymptomatic women. *J Womens Health* 2011;20:1825–31.
- 17 Georgiou HM, Di Quinzio MKW, Permezel M, *et al.* Predicting preterm labour: current status and future prospects. *Dis Markers* 2015;2015:435014.
- 18 Shennan AH. Prediction and prevention of preterm birth: a quagmire of evidence. *Ultrasound Obstet Gynecol* 2018;51:569–70.
- 19 Son M, Miller ES. Predicting preterm birth: cervical length and fetal fibronectin. *Semin Perinatol* 2017;41:445–51.
- 20 Cobo T, Kacerovsky M, Jacobsson B. Risk factors for spontaneous preterm delivery. *Int J Gynaecol Obstet* 2020;150:17–23.
- 21 Ren H, Du M. Role of maternal periodontitis in preterm birth. *Front Immunol* 2017;8:139.
- 22 Oskovi Kaplan ZA, Ozgu-Erdinc AS. Prediction of preterm birth: maternal characteristics, ultrasound markers, and biomarkers: an updated overview. *J Pregnancy* 2018;2018:8367571.
- 23 Blencowe H, Cousens S, Oestergaard MZ, *et al.* National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012;379:2162–72.
- 24 Kleinrouweler CE, Cheong-See FM, Collins GS, *et al.* Prognostic models in obstetrics: available, but far from applicable. *Am J Obstet Gynecol* 2016;214:79–90.
- 25 Sananès N, Langer B, Gaudineau A, *et al.* Prediction of spontaneous preterm delivery in singleton pregnancies: where are we and where are we going? A review of literature. *J Obstet Gynaecol* 2014;34:457–61.
- 26 Catley C, Frize M, Walker CR, *et al.* Predicting high-risk preterm birth using artificial neural networks. *IEEE Trans Inf Technol Biomed* 2006;10:540–9.
- 27 Raglan GB, Lannon SM, Jones KM, *et al.* Racial and ethnic disparities in preterm birth among American Indian and Alaska native women. *Matern Child Health J* 2016;20:16–24.
- 28 Culhane JF, Goldenberg RL. Racial disparities in preterm birth. *Semin Perinatol* 2011;35:234–9.
- 29 York TP, Strauss JF, Neale MC, *et al.* Racial differences in genetic and environmental risk to preterm birth. *PLoS One* 2010;5:e12391.
- 30 Koullali B, Oudijk MA, Nijman TAJ, *et al.* Risk assessment and management to prevent preterm birth. *Semin Fetal Neonatal Med* 2016;21:80–8.
- 31 Peduzzi P, Concato J, Kemper E, *et al.* A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373–9.
- 32 Mekonen DG, Yismaw AE, Nigusie TS, *et al.* Proportion of preterm birth and associated factors among mothers who gave birth in Debreabor town health institutions, northwest, Ethiopia. *BMC Res Notes* 2019;12:2.
- 33 Tigist B, Abdela A, Zenebe G. Preterm birth and associated factors among mothers who gave birth in Debre Markos town health institutions. In: *Institutional based cross sectional study*, 2013.
- 34 Rosenberg RE, Ahmed ASMNU, Ahmed S, *et al.* Determining gestational age in a low-resource setting: validity of last menstrual period. *J Health Popul Nutr* 2009;27:332.
- 35 Wassie M, Manaye Y, Abeje G, *et al.* Determinants of preterm birth among newborns delivered in Bahir Dar City public hospitals, North West Ethiopia 2020.
- 36 Woday A, Muluneh MD, Sherif S. Determinants of preterm birth among mothers who gave birth at public hospitals in the Amhara region, Ethiopia: a case-control study. *PLoS One* 2019;14:e0225060.
- 37 Wudie FT, Tesfamichael FA, Fisseha HZ, *et al.* Determinants of preterm delivery in the central zone of Tigray, Northern Ethiopia: a case-control study. *South African J Child Health* 2019;13:108–14.
- 38 Woldeyohannes D, Kene C, Gomora D, *et al.* Factors associated with preterm birth among mothers who gave birth in Dodola town hospitals, Southeast Ethiopia: institutional based cross sectional study. *Clinics Mother Child Health* 2019;16:2.
- 39 Muchie KF, Lakew AM, Teshome DF, *et al.* Epidemiology of preterm birth in Ethiopia: systematic review and meta-analysis. *BMC Pregnancy Childbirth* 2020;20:1–12.
- 40 Kwak SK, Kim JH. Statistical data preparation: management of missing values and outliers. *Korean J Anesthesiol* 2017;70:407.
- 41 Grobbee DE, Hoes AW. *Clinical epidemiology: principles, methods, and applications for clinical research*. Jones & Bartlett Publishers, 2014.
- 42 Moons KGM, Kengne AP, Woodward M, *et al.* Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012;98:683–90.
- 43 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
- 44 Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010;5:1315–6.
- 45 Brhane M, Hagos B, Abrha MW. Does short inter-pregnancy interval predicts the risk of preterm birth in Northern Ethiopia? 2019;12:405.
- 46 Sifer S, Kedir B, Demisse G, *et al.* Determinants of preterm birth in neonatal intensive care units at public hospitals in Sidama zone, South East Ethiopia; case control study. *J Pediatr Neonatal Care* 2019;9:180–6.
- 47 Xu X, Tan H, Zhou S, *et al.* [Study on the application of back-propagation artificial neural network used the model in predicting preterm birth]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2014;35:1028–31.
- 48 Chen M, Xie N, Liang Z, *et al.* Early prediction model for preterm birth combining demographic characteristics and clinical characteristics 2020.
- 49 Kebede EB, Terfa YB, Geleta BA, Akuma AO: predictors of preterm birth in Jimma town public hospitals, Jimma, Ethiopia. *Journal of Pediatric and Neonatal Individualized Medicine* 2021;10:e100125.
- 50 Belaynew W, Teumay A, Getachew G, *et al.* Effects of inter pregnancy interval on preterm birth and associated factors among postpartum mothers who gave birth at Felege Hiwot referral hospital. *World J Pharm Pharm Sci* 2015;4:12–25.
- 51 Muhumed II, Kebira JY, Mabalhin MO. Preterm birth and associated factors among mothers who gave birth in Fafen zone public hospitals, Somali regional state, eastern Ethiopia. *Res Report Neonatal* 2021;11:23–33.
- 52 Zhang Y-P, Liu X-H, Gao S-H, *et al.* Risk factors for preterm birth in five maternal and child health hospitals in Beijing. *PLoS One* 2012;7:e52780.
- 53 Wagura P, Wasunna A, Laving A, *et al.* Prevalence and factors associated with preterm birth at kenyatta national Hospital. *BMC Pregnancy Childbirth* 2018;18:107.
- 54 Panaretto K, Lee H, Mitchell M, *et al.* Risk factors for preterm, low birth weight and small for gestational age birth in urban Aboriginal and Torres Strait Islander women in Townsville. *Aust N Z J Public Health* 2006;30:163–70.
- 55 Olusanya BO, Ofovwe GE. Predictors of preterm births and low birthweight in an inner-city hospital in sub-Saharan Africa. *Matern Child Health J* 2010;14:978–86.
- 56 Mekuriyaw AM, Mihret MS, Yismaw AE. Determinants of preterm birth among women who gave birth in Amhara region referral hospitals, Northern Ethiopia, 2018: institutional based case control study. *Int J Pediatr* 2020;2020:1854073.
- 57 He J-R, Ramakrishnan R, Lai Y-M, *et al.* Predictions of preterm birth from early pregnancy characteristics: born in Guangzhou cohort study. *J Clin Med* 2018;7:185.
- 58 Cobo T, Aldecoa V, Figueras F, *et al.* Development and validation of a multivariable prediction model of spontaneous preterm delivery and microbial invasion of the amniotic cavity in women with preterm labor. *Am J Obstet Gynecol* 2020;223:421.e1–e14.
- 59 Lamont R, Richardson L, Boniface J, *et al.* Commentary on a combined approach to the problem of developing biomarkers for the prediction of spontaneous preterm labor that leads to preterm birth. *Placenta* 2020.
- 60 Stock SJ, Horne M, Bruijn M, *et al.* 793: a new prediction model for birth within 48 hours in women with preterm labour symptoms. *Am J Obstet Gynecol* 2020;222:S502.
- 61 Beta J, Akolekar R, Ventura W, *et al.* Prediction of spontaneous preterm delivery from maternal factors, obstetric history and placental perfusion and function at 11–13 weeks. *Prenat Diagn* 2011;31:75–83.
- 62 van de Mheen L, Schuit E, Lim AC, *et al.* Prediction of preterm birth in multiple pregnancies: development of a multivariable model including cervical length measurement at 16 to 21 weeks' gestation. *J Obstet Gynaecol Can* 2014;36:309–19.