BMJ Open  

Maternal near-miss prediction model development among pregnant women in Bahir Dar City administration, northwest Ethiopia: a study protocol

Yinager Workineh 1,2,3  

ABSTRACT

Introduction Maternal near-miss is a condition when a woman nearly died but survived from complications that happened during pregnancy, childbirth or within 42 days after delivery. Maternal near-miss is more prevalent among women in developing nations. Previous studies have identified the impact of different predictor variables on maternal near-miss but shared prognostic predictors are not adequately explored in Ethiopia. It is therefore necessary to build a clinical prediction model for maternal near-misses in Ethiopia. Hence, the aim of this study is to develop and validate a prognostic prediction model, and generate a risk score for maternal near-miss among pregnant women in Bahir Dar City Administration.

Methods and analysis A prospective follow-up study design will be employed among 2110 selected pregnant women in the Bahir Dar City administration from 1 May 2023 to 1 April 2024. At the initial antenatal visit, pregnant women will be systematically selected. Then, they will be followed until 42 days following birth. Data will be collected using structured questionnaires and data extraction sheet. The model will be created using Cox proportional hazard regression analysis. The performance of the model will be assessed based on its capacity for discrimination using c-index and calibration using calibration plot, intercept and slope. The model's internal validity will be evaluated through the bootstrapping method. Ultimately, the model will be illustrated through a nomogram and decision tree, which will be made available to prospective users.

Ethics and dissemination Ethical approval has been obtained from the Institutional Review Board of the College of Medicine and Health Sciences, Bahir Dar University (protocol number 704/2023). Findings will be published in peer-reviewed journals and local and international seminars, conferences, symposiums and workshops. Manuscripts will be prepared and published in scientifically reputable journals. In addition, policy briefs will be prepared.

INTRODUCTION

Maternal near-miss refers to a woman who almost lost her life but survived from a severe obstetric complication that occurred during pregnancy, childbirth or within 42 days following delivery.1,2 Severe acute maternal morbidity (SMM) is a synonymous term for maternal near-miss.3

The notion of maternal near-miss was devised by the WHO to pinpoint life-threatening situations during pregnancy, childbirth and postpartum.4 This concept allows for interventions focusing on the sequence of events that led to a woman with near to death or actual death.5,6 By using maternal near-miss statistics instead of maternal mortality rates (MMR), maternal healthcare system deficiencies and health priorities can be more swiftly identified.4,7 To apply this concept globally, WHO created diagnostic tools for maternal near-miss that encompass clinical, laboratory and management-based criteria.8,9 This tool was adapted and validated for sub-Saharan Africa (SSA) countries.10

Maternal mortality and maternal near-miss are prominent health concerns on a global scale, especially in underprivileged countries. Roughly 303 000 women die annually due to complications during pregnancy and childbirth across the world.11 The 2017 United Nations Population Fund reported that every 2 min a woman died due to pregnancy or childbirth-related complications.12 Low-resource countries account for 99% of all maternal mortalities.13 According to the 2019 Mini-Ethiopian Demographic Health Survey (EDHS) report, Ethiopia’s MMR stood at 412
per 100,000 live births. This number is considerably higher than the average MMR worldwide (211 per 100,000 live births), but lower than the MMR in SSA (553 per 100,000 live births).

Maternal near-miss ranged from 0.80% to 8.23% based on disease-specific criteria, and 0.01%–2.99% based on management-related criteria. The maternal near-miss ratio was 18.57 per 1000 live births globally. The smallest maternal near-miss ratio was found in Europe (3.10 per 1000 live birth), whereas the highest burden of maternal near-miss was found in African and Asian countries. In SSA, the maternal near-miss ratio was 24.2 per 1000 live births. The prevalence of maternal near-miss in Ethiopia was 12.57% with the highest (26.5%) in the Amhara region.

The high burden of maternal near-miss is influenced by a multitude of complex risk factors. Delays in seeking care, reaching to care, receiving adequate and appropriate care are some of the risk factors for maternal near-miss. The delay in seeking care is linked to failure to recognize signs of complications, failure to perceive severity of illness, cost consideration, negative experience with health system, transportation difficulties and need of permission from family members. The reasons for the delay in accessing healthcare services include a considerable distance to the medical facility, poor road conditions and a shortage of transportation options. The uncompassionate demeanour of healthcare providers, inadequacy of supplies and essential equipments, unavailability or inadequate proficiency of medical staff and absence of urgency or comprehension of emergency situations are the reasons for delay in receiving adequate and appropriate care. A community’s delay to take responsibility can also contribute to maternal mortality as a result of the absence of a community-based and community-driven comprehensive approach to maternal health/well-being.

Several measures have been implemented to reduce the load, complications, causes and risk factors linked to maternal mortality and maternal near-miss. In 2012, WHO defined, conceptualised and evaluated severe maternal morbidity. The objective of this effort was to compile numerous definitions of maternal morbidity. In 2013, WHO established the Maternal Mortality Surveillance and Response unit, which concentrates on continuous evaluation of the reasons and factors that lead to maternal mortality.

In addition to the aforementioned global initiatives, individual risk assessment studies like clinical prediction models are important in enhancing maternal health. These prediction models can guide (1) clinical researchers to select appropriate study subjects; (2) patients to choose more beneficial steps for themselves; (3) doctors to make better clinical decisions and (4) health management departments to monitor and manage the quality of medical services and to allocate medical resources more rationally.

The effects of clinical prediction models can be nearly mirrored in any of the three-level prevention system of diseases including the primary (health promotion, prevention and control), secondary (early screening, early diagnosis and prompt treatment) and tertiary (rehabilitation programmes, preventing disease relapse, reducing mortality and disability and promoting functional recovery and quality of life). Prognostic predictive models can provide patients and doctors with a numerical risk value (probability) of identifying specific illness in the future based on current health condition.

The prediction model we are going to develop will be used in daily clinical practice and obstetric patients. It will help pregnant women to identify themselves as a risk group or not. In daily clinical practice, this model will provide clues for obstetricians to select high-risk obstetric patients for further screening, diagnosis and management. In general, the development of numerical and visual maternal near-miss model will assist obstetricians or other healthcare professionals, obstetric patients and their relatives to facilitate shared medical decision-making for diagnostic testing, initiating or discontinuing treatments or making lifestyle changes throughout the perinatal period.

Different researchers investigated the burden and determinants of maternal near-miss in Ethiopia. These studies focused on the individual predictor effects on maternal near-miss. They did not identify shared characteristics of prognostic predictors as a whole, and did not directly indicate the risk stratification of obstetric patients.

The objective of this study was the development and validation of a prediction model for maternal death or severe obstetric morbidity among admitted obstetric patients. The diagnostic prediction model was developed for severe maternal outcome in SSA among admitted obstetric patients. But there are few prognostic prediction models for maternal near-miss. Hence, the development and validation of a prediction model for maternal near-miss among pregnant women is required in Ethiopia. Therefore, the aims of this study are the development and (internal) validation of a prognostic prediction model, as well as the generation of a risk score for maternal near-miss among pregnant women. The specific objectives are:

1. To develop a prognostic model for maternal near-miss among pregnant women in Bahir Dar City administration, Northwest Ethiopia 2023–2024.
2. To validate (internal) the prognostic model of maternal near-miss among pregnant women in Bahir Dar City administration, Northwest Ethiopia, 2023–2024.
3. To create risk score for maternal near-miss among pregnant women in Bahir Dar City administration, Northwest Ethiopia, 2023–2024.

METHODS AND ANALYSIS

The components of this protocol were reported based on the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guideline or checklists (online supplemental file 1).
Study design and period

Prospective follow-up study will be conducted to develop prognostic prediction model of maternal near-miss. The focus of the model is to predict a future occurrence of maternal near-miss using individual level factors at the entrance of antenatal contact (ANC).

Maternal near-miss in the follow-up period \( (O) = f (D1, D2, D3, \ldots, Dn) \)

Where,

\( O \) = occurrence of maternal near-miss

\( D1, D2, D3, \ldots, Dn \) = the predictors

Individual-level variables will be used as prognostic predictors \( (Dn) \) to forecast the incidence of maternal near-miss \( (O) \) during the follow-up period. The occurrence of maternal near-miss among pregnant women as a function of individual level predictors is expressed as \( f \) (age, height, weight…\( Dn \)).

The study will be conducted from 1 May 2023 to 1 April 2024. The requirement of participants and data collection on prognostic predictors will be performed from 1 May 2023 to 5 August 2023 (base line period). The follow-up assessment of the outcome will be at any time after the enrolment of the participants. Hence, identification of the outcome will be carried out from 2 May 2023 to April 2024 (end line period).

Study setting and participants

The study will be carried out in Bahir Dar City administration in Northwest Ethiopia. Bahir Dar is 450 km away from Addis Ababa. Both urban and rural populations inhabit this city. There are three public hospitals, 11 healthcare centres, 15 health posts, 4 private hospitals, 56 private specialty clinics and 13 private medium clinics in the city.4

First, health facilities in Bahir Dar City administration that provide ANC service will be identified. Second, the health facilities to be included in the study will be selected randomly. Third, the total sample size will be proportionally allocated to each health facility based on last year’s reports of first antenatal care visit. Fourth, pregnant women who come for their first ANC visit will be selected using a systematic sampling method. In this regard, in each data collection day, the woman who will arrive first in the ANC clinic will be selected as a starting participant. Then, every other ANC visitor will be selected. Finally, the selected pregnant women will be followed until they either experience the event or reach a state of censoring (figure 1). The event group will consist of women who experience maternal near-miss, while the censoring group will include women who withdraw from ANC visits, transfer out from the selected healthcare facility, are lost to follow-up or pass away during the follow-up period.

Eligibility criteria

Participants who have no plans to relocate from the study area and are considered well enough to be interviewed by the interviewer will be included in the study. Pregnant women who experienced maternal near-miss at the beginning of the cohort and after 42 days of delivery, and who do not remember their last normal menstrual period will be excluded from the study participants.

Predictors and their measurements

Individual-level data on all predictors will be collected by trained midwives using an interview-administered questionnaire and an extraction sheet. These variables include sociodemographic characteristics such as age (in years), residence (coded as ‘0’ for rural and ‘1’ for urban), decision-making for healthcare (coded as ‘1’ for self, ‘2’ for spouse, ‘3’ for relatives and ‘4’ for jointly), height (in centimetres), weight (in kilograms) and mid-upper arm circumference (measured in centimetres). Additionally, obstetric factors such as parity (measured in number), plurality (coded as ‘0’ for single and ‘1’ for multiple), pregnancy intention (coded as ‘0’ for unplanned and ‘1’ for planned), gestational age (in weeks), inter-pregnancy interval (in months) and history of caesarean section (C/S) (coded as ‘0’ for no and ‘1’ for yes) will be recorded. Other factors to be considered include the history of pre-eclampsia (coded as ‘0’ for no and ‘1’ for yes), eclampsia (coded as ‘0’ for no and ‘1’ for yes), sepsis (coded as ‘0’ for no and ‘1’ for yes), haemorrhage (coded as ‘0’ for no and ‘1’ for yes), obstructed labour (coded as ‘0’ for no and ‘1’ for yes), medical morbidity (coded as ‘0’ for no and ‘1’ for yes), history of stillbirth (coded as ‘0’ for no and ‘1’ for yes), history of abortion (coded as ‘0’ for no and ‘1’ for yes), distance from the health facility (measured in kilometres), timing of initial ANC (expressed in weeks) and birth preparedness and complication readiness (coded as ‘0’ for no and ‘1’ for yes). Additionally, baseline clinical indices such as blood pressure (measured in mm Hg) and haematocrit measurement (measured in percentage) will be recorded.

Outcome

Maternal near-miss will be diagnosed during the follow-up phase by trained health professionals using the WHO screening criteria. Women who meet at least one of the clinical, laboratory or management-based criteria will be classified as the event group (maternal near-miss), while the remaining women will be classified as the censoring group. In addition to identifying maternal near-miss cases, pregnancy danger signs, referrals, withdrawals against medical advice, transfers out, deaths and any complications will be extracted from the maternal card during the follow-up period. The survival time will be measured in week(s) from the last date of normal menstrual period to the occurrence of maternal near-miss or censorship.

Blinding assessment of predictors and outcome

Blinding reduces the risk of bias that may be introduced in the model development. This will be done by measuring the predictors at the base line and outcome at the follow-up period. The data collectors for the baseline and end line surveys will be different.
Sample size determination

The sample was determined based on the minimum sample size calculation criteria for a time to event study. The minimum criteria for this calculation are as follows: (1) a minimum heuristic shrinkage factor, $S$, greater than 0.9 (targeting less than 10% overfitting), (2) a small difference between Nagelkerke’s $R^2_{\text{app}}$ and $R^2_{\text{adj}}$ (targeting less than 0.05 absolute difference) and (3) a small margin of error in the overall risk estimate (targeting less than 0.05 absolute error). These criteria, including the number of parameters (P), heuristic shrinkage factor ($S$), overall risk in the population and the model’s anticipated Cox-Snell $R^2$ (or C-statistics) were reviewed from previous studies. Approximately 25 candidate parameters, 26.6% of maternal near-miss and C-statistics of 0.1142 were utilised to calculate the sample size. The sample size was then calculated using the Stata command ‘pmsampsize, type (b) rsquared (0.11) parameter25 prev (0.266)’. The resulting sample size was 1918 (table 1).

Data quality assurance

Input from study participants and subject-matter experts will be used to construct and validate the questionnaire in terms of face and content validity. After tool validation, data collectors and supervisors will receive 2 days of training to become familiar with the questionnaires, the data collection processes, the ethical considerations and a purpose of the study. Similarly, health professionals who work at antenatal, delivery and postnatal departments will be trained on the screening criteria of maternal near-miss. Then, a pretest will be conducted by four data collectors (Bachelor degree in Midwifery) and one supervisor (Epidemiologist). This pretest will be undertaken to ensure the accuracy of the data and to check for ambiguities in language after switching from English to
Amharic. Data collection tools will be changed in light of the pretest’s results. Daily supervision of data collectors and daily verification of all collected data will be done by the supervisor. The data will be regularly checked for completeness, and any problems during data collection will be addressed appropriately.

**Statistical analysis methods**

**Data processing**

The data from Epi-Collect5 will be downloaded and transferred to Microsoft Excel. The illogical values and steps will be checked during the design of the questionnaire/abstraction sheet in Epi-Collect5. The collected data will be examined for consistency and completeness. Then the data will be exported to R 4.2.2 software for analysis.

Data cleaning will be performed to check the completeness of the data, remove or correct noise, outliers and missing values in order to prepare the data for the subsequent steps of model development and validation. The cleaned data will be further processed using feature selection and extraction algorithms, which will involve deriving new attributes and summarising the data. Data transformations, such as normalisation, will be applied to remove noise and correct inconsistencies in the data.

For continuous variables, dichotomisation or categorisation will be performed based on widely accepted clinical cut-off values. If there are no clinical cut-off points, restricted cubic splines or fractional polynomials will be applied to model non-linear relationships.

The distribution of continuous variables will be examined graphically using a histogram and statistically using the Shapiro-Wilk test. If the data do not follow a normal distribution, data transformation and standardisation will be performed. Multicollinearity between each independent predictor will be checked using the variance inflation factor (VIF). If the VIF is greater than 10, there is no multicollinearity.

**Missing data handling**

Once the data are prepared for analysis, a thorough evaluation will be conducted to identify missing data. Missing data is a common problem that can impact the accuracy of classification and the models generated from data mining algorithms.

The first step in dealing with missing data is to understand the patterns of missing values. The Hmisc library’s ‘naclus’ and ‘naplot’ functions, as well as the recursive partitioning library of Atkinson and Therneau, will be applied for this purpose. The ‘naclus’ function identifies variables that tend to be missing for the same participants and computes the proportion of missing values for each variable. The ‘naplot’ function builds a tree to predict which types of participants tend to have missing values.

After understanding the patterns of missing values, statistical methods will be employed to handle the missing values. One common method is complete case analysis, which involves excluding all subjects with missing values for any potential predictor or outcome. Complete case analysis considers subjects with complete data for a specific predictor, even if they have missing values for other covariates not included in the specific model. This method discards information from subjects who have information on some predictors but not all, making it statistically inefficient. Therefore, methods that replace missing values with substituted values based on various criteria are preferred. These methods include: (1) replacing the missing value with a constant, (2) replacing the missing value with the mean of the field and (3) replacing the missing values with randomly generated values from the observed variable distribution.

In the current model development, missing values will be handled by replacing them with randomly generated values from the observed variable distribution. Specifically, multiple imputation techniques will be used when the amount of missing data is less than 10%. The steps of multiple imputation involve: (1) replacing missing values with randomly selected values from specific distributions to create complete case data sets, (2) conducting the same analysis on each of these data sets and (3) pooling the results together. This imputation process will be repeated five times.

**Model building**

**Predictor selection**

The prognostic predictors will be selected by considering existing knowledge of previously established predictors; applicability and costs of predictor measurement relevant to the targeted setting; and statistical power.

**Model estimation and specification**

Cox proportional hazard regression model will be used. The strength of the association will be measured in terms of HR at 95% confidence limits. The Cox proportional hazard model assumption will be checked using scaled
Schoenfeld residual test and graphically with log–log Cox adjusted survival estimate. The model fitness will be checked using Cox-nell residuals test. Finally, the multivariable cox proportional hazard model will be declared at a p<0.05 with 95% CI as statistical significance.

Model performance
Calibration and discrimination are two crucial factors in assessing the prognostic accuracy of the model. At a group level, calibration assesses how accurately the absolute predicted risks align with the real hazards. Discrimination refers to the model’s capacity to differentiate between patients who experience the relevant event and those who do not.

By utilising the concordance index, the model’s capacity for differentiation can be assessed (C-index). To compute the C-index, all possible pairs of patients with and without the outcome are examined. If the patient with the result has a higher expected risk than the patient without the result, the pair is considered to be in agreement. The C-index is equivalent to the area under the curve (AUC). The AUC values ranging from 0.9 to 1.0, 0.8 to 0.9, 0.7 to 0.8, 0.6 to 0.7 and 0.5 to 0.6 are classified as excellent (A), good (B), fair (C), poor (D) and fail (F), respectively, in order to evaluate the model’s ability to differentiate.

A time-to-event C-statistic, such as Harrel’s C-statistic or Uno’s C-statistic, can be calculated for Cox proportional hazards prediction models. Two patients who both have the event can be matched together in these calculations, and if the patient who has the event first has the higher predicted risk, they are considered to be a concordant pair. Because this project is focused on survival, the time-to-event C-statistic method will be used to assess the model’s ability to discriminate.

Calibration assesses if the predicted hazards are comparable to the actual hazards. Calibration can be evaluated through a calibration plot, intercept and slope. By computing a Cox proportional hazards model utilising the prognostic indicator, the calibration slope is determined. The calibration slope represents the coefficient of regression applied to the prognostic indicator.

Internal validation
Different techniques exist for internal verification of prognostic forecasting models, including cross-validation, bootstrapping and split-sample validation. The bootstrapping approach surpasses the split-sample and cross-validation methods in addressing optimism. Therefore, this study will utilise 10 000 random bootstrap samples with replacement on all predictors in the data set to validate the model. The C-index with a 95% CI, as well as sensitivity and specificity, will be employed in conjunction with the Hosmer-Lemeshow statistic.

Model accuracy
Using a Brier score, the accuracy of a set of probabilistic forecasts will be evaluated. The Brier score is the average of the squared disparities between the probabilistic predictions and the actual event outcomes, with the probabilistic forecasts being provided for those specific events.

Clinical and public health impact assessment
Decision curve analysis (DCA) is a commonly employed method to evaluate the efficacy of clinical prediction models. Traditional measures of diagnostic performance, such as sensitivity, specificity and area under the receiver operating characteristic curve do not consider the clinical value of a specific model. Instead, these measures only compare the diagnostic accuracy of different prediction models. Therefore, in this study, DCA will be utilised to assess the clinical and public health impacts of the maternal near-miss model.

Model presentation
Once the model is developed and validated, a risk score will be generated for maternal near-misses. This score will be based on predictors at the individual level and will be user-friendly. The coefficients of each predictor that are statistically significant in the multivariable cox proportional hazard regression model will be adjusted to calculate the risk scores. The cut-off point for the risk score and the development of a nomogram will be determined using the Youden index value (sensitivity+specificity−1) for each risk category. Additionally, a decision tree will be presented to potential users as part of the clinical prediction model.

Patient and public involvement
No patient or public has been involved while developing this study protocol.

ETHICS AND DISSEMINATION
The Institutional Review Board of the College of Medicine and Health Sciences, Bahir Dar University has granted ethical clearance (protocol number 704/2023) for this study. Results will be shared via scientific publications, conference presentations, community meetings, and policy briefs.

Acknowledgements We thank Bahir Dar University for the facilities we have used while preparing this protocol.

Contributors YW, GDA and GAF contributed equally to the design of the study. YW drafted the manuscript and all the authors revised and approved the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

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

Patient consent for publication Consent obtained directly from patient(s).

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability
of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

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 iD
Yinager Workineh http://orcid.org/0000-0002-5175-1045

REFERENCES


