**BMJ Open**

Development and internal validation of a clinical risk score for in-hospital mortality after stroke: a single-centre retrospective cohort study in Northwest Ethiopia

Tiruayehu Getinet Abebe, Sefineh Fenta Feleke, Anteneh Mengist Dessie, Rahel Mulatie Anteneh, Zelalem Alamrew Anteneh

**ABSTRACT**

**Objective** To develop and validate a clinical risk score for in-hospital stroke mortality.

**Design** The study used a retrospective cohort study design.

**Setting** The study was carried out in a tertiary hospital in the Northwest Ethiopian region.

**Participants** The study included 912 patients who had a stroke admitted to a tertiary hospital between 11 September 2018 and 7 March 2021.

**Main outcome measures** Clinical risk score for in-hospital stroke mortality.

**Methods** We used EpiData V.3.1 and R V.4.0.4 for data entry and analysis, respectively. Predictors of mortality were identified by multivariable logistic regression. A bootstrapping technique was performed to internally validate the model. Simplified risk scores were established from the beta coefficients of predictors of the final reduced model. Model performance was evaluated using the area under the receiver operating characteristic curve and calibration plot.

**Results** From the total stroke cases, 132 (14.5%) patients died during the hospital stay. We developed a risk prediction model from eight prognostic determinants (age, sex, type of stroke, diabetes mellitus, temperature, Glasgow Coma Scale, pneumonia and creatinine). The area under the curve (AUC) of the model was 0.895 (95% CI: 0.859–0.932) for the original model and was the same for the bootstrapped model. The AUC of the simplified risk score model was 0.893 (95% CI: 0.856–0.929) with a calibration test p value of 0.225.

**Conclusions** The prediction model was developed from eight easy-to-collect predictors. The model has excellent discrimination and calibration performance, similar to that of the risk score model. It is simple, easily remembered, and helps clinicians identify the risk of patients and manage it properly. Prospective studies in different healthcare settings are required to externally validate our risk score.

**INTRODUCTION**

Stroke is the second leading cause of death and the third leading cause of disability-adjusted life years (DALYs) worldwide. It is a disease of immense public health importance with serious economic and social consequences, despite the advances in understanding its epidemiology, effects on quality of life and pathophysiology. The public health burden of stroke is set to rise over the next few decades because of demographic transitions in populations, particularly in low/middle-income countries. The prognosis of stroke depends on the quick diagnosis of the type, followed by appropriate and fast management.

Although stroke mortality rates declined from 1990 to 2017, the absolute number of people who developed a new stroke, died or became disabled from a stroke has almost doubled. In 2017, there were 6.2 million fatal cases of stroke and 132.1 million stroke-related DALYs worldwide, where 80% of all incident strokes, 87% of all stroke deaths and 89% of all stroke-related DALYs occurred in low/middle-income countries. Another evidence from the Global Burden of Disease report showed that mortality from stroke in sub-Saharan African countries is nearly doubled from 1990 to 2019.

Several studies revealed that mortality from stroke varies across the African countries, and the estimate ranges from 6.2% in Benin to 34.8% in Sierra Leone. This might reflect the variations in the access to quality healthcare, absence of national health insurance...
schemes, lack of trained medical workforce, lack of diagnostic imaging and lack of appropriate treatment due to the absence of in-hospital stroke units.

In Ethiopia, stroke was the second leading cause of death, causing an estimated 32859.5 deaths among all ages and gender groups in 2016. Deaths from stroke also contributed to 109499 (30%) of the total cardiovascular deaths in 2016. Reports also showed that death from stroke was estimated to have increased by 23.8% between 2000 and 2016. A review study conducted in Ethiopia showed that nearly one-fifth (18%) of patients who had a stroke have died during hospitalisation. According to this study, there was a regional difference in the in-hospital mortality rate of stroke, which ranges from 10.7% in the Tigray region to 35.2% in Southern Nations, Nationalities and People’s Region. Moreover, the in-hospital mortality rate is increasing (15.1%) until 2016, then it becomes 19.6% after 2016.

Several factors have been identified as predictors for in-hospital mortality, including age, localisation of ischaemia, the Glasgow Coma Scale (GCS), stroke severity measured by the National Institutes of Health Stroke Scale (NIHSS), posterior circulation stroke syndrome, non-lacunar stroke cause, pre-stroke functional disability (modified Rankin Scale >0), pre-existing heart disease, being into a coma, having cerebral oedema, atrial fibrillation, diabetes mellitus and hypertension, tobacco smoking, alcohol use and family history of stroke.

Previous studies conducted so far have placed emphasis on the causes of in-hospital mortality from stroke; however, stroke is a complex condition occurring in heterogeneous populations with a variety of mechanisms, each having a different prognosis. Nowadays, the focus has shifted to prediction research to determine the risk of an individual patient dying during a hospital stay by combining multivariable prognostic predictors. World-wide, several risk prediction models have been developed previously to predict the in-hospital mortality of stroke, but the Get With The Guidelines-Stroke (GWTG-Stroke) is the only valid model that includes both ischaemic and haemorrhagic stroke types. The GWTG-Stroke risk score predicted in-hospital stroke mortality using prognostic determinants such as age, stroke type, method of arrival at the hospital, history of atrial fibrillation, previous stroke, coronary artery disease, carotid stenosis, diabetes, peripheral vascular disease, hypertension, dyslipidaemia, smoking, and weekend or night-time admission. However, it is complex and applicable with the aid of web-based computational support, which cannot be used in a resource-limited set-up like in Ethiopia. In addition, previous studies conducted on mortality due to stroke in Ethiopia focused on determining the causal role of determinants of mortality from stroke. However, there are no clinical prediction models and scores available to assist the clinical judgement process by healthcare workers at a healthcare facility. Therefore, this study aimed to develop and validate a prediction model and risk score for in-hospital stroke mortality using information routinely available to clinicians at hospital admission, such as demographic characteristics, clinical presentations, risk factors and complications, and biochemical characteristics.

METHODS

Study setting

This study was conducted at Felege Hiwot Comprehensive Specialized Hospital (FHCSH) found in Bahir Dar City, the capital of Amhara National Regional State, located at a distance of 560 km from Addis Ababa, the capital city of Ethiopia.

Study design and participants

We used a retrospective cohort data collection technique among patients who had a stroke in the medical ward. This study used the data of patients admitted to the ward from 11 September 2018 to 7 March 2021. The theoretical design of this current study is to examine the occurrence of in-hospital stroke mortality as a function of multiple prognostic determinants among patients admitted to the medical ward of FHCSH.

Inclusion and exclusion criteria

All medical records of patients who had a stroke admitted to the medical ward at FHCSH from 11 September 2018 to 7 March 2021 were included in the current study. However, medical records of patients with an unknown outcome who changed their diagnosis of stroke after admission, lost cards or replaced medical records that had no admission data were excluded.

Sample size determination and sampling technique

The sample size was determined based on the minimum standard of 10 events per candidate predictors considered. Using the rule of thumb, the assumption of minimum standard 10 events per predictor variable was used to determine required stroke cases in this study. Thus, the number of patients who had a stroke (N)=(n×10)/I, where N is the sample size, n is the number of candidate predictors and I is the estimated event rate in the population based on the previous study conducted. We have 20 predictors for in-hospital mortality, and the incidence of in-hospital mortality of stroke was 21.6%. Therefore, the required sample size was 926. The flow chart for sampling procedure was illustrated below (figure 1).

Definition of terms

**Diabetes mellitus:** when a patient had a medical history of diabetes mellitus or had admission random blood sugar of greater than 200mg/dL or fasting blood sugar of greater than or equal to 126mg/dL.

**Hypertension:** when a patient had a history of antihypertensive treatment before admission or systolic blood pressure greater than or equal to 140mm Hg or diastolic blood pressure greater than or equal to 90 mm Hg during admission.
Data collection tool and technique
Data were extracted using structured checklists after the identification of medical registration numbers from the inpatient registration abstract. Four trained health professionals collected the data, which contained sociodemographic characteristics, risk factors, complications, clinical presentations, biochemical characteristics and outcomes.

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.

Data processing and analysis
Data were collected using EpiData V.3.02, and were exported to R statistical programming language V.4.0.4 for analysis. Data were assumed missing at random; missing data of predictors were replaced by a multiple imputation chained equations approach. Descriptive statistics were performed based on the type of variable. Categorical variables were described by using percentage and frequency, and continuous variables were summarised using median and IQR.

Bivariate logistic regression analysis was done to show unadjusted associations between each prognostic determinant and in-hospital mortality due to stroke to select candidate determinants for multivariate prediction modelling. The variables with a p value of <0.25 in the bivariate regression were entered into the multivariable analysis. We performed multicollinearity tests between each determinant using variance inflation factor (VIF); VIF >10 was used as a cut-off point to exclude a variable for multivariate modelling.17

A backward stepwise logistic regression analysis was performed. The regression coefficients with their 95% confidence levels and p values were reported for the models. Determinants with a p value of <0.05 in the reduced model were used to develop a prediction model and risk score for in-hospital stroke mortality. To evaluate the accuracy of the model, the area under the receiver operating characteristic (ROC) curve (AUC) (discrimination) and calibration plot (calibration) were used. Calibration is a goodness-of-fit or reliability test that compares observed versus predicted event rates for groups of patients. It is described by a calibration plot and the Hosmer-Lemeshow test and scores with nonsignificant p values, suggesting good calibration. An AUC less than or equal to 0.5 is considered to represent a non-discriminative model, whereas an AUC equal to or greater than 0.7 is considered as fair discriminative ability; an AUC equal to or greater than 0.8 is considered to represent good discriminative ability, and an AUC equal to or greater than 0.9 means excellent discriminative ability.18

To internally validate our model, a bootstrapping technique was used. Bootstrapping is one of the most widely used internal validating methods, where a model is fitted to the full sample and multiple samples are then drawn (with replacement). On each bootstrap sample, the model is redeveloped. For the model derivation, we had drawn 1000 random bootstrap samples with replacement on all predictors. The model’s predictive performance after bootstrapping was considered as the performance that can be expected when the model is applied to future patients who would have a stroke. The regression coefficients and AUC of the validated model were reported and compared with the original model.

To develop a simplified and easily applicable prediction score for the in-hospital stroke mortality, each regression coefficient in the validated model was divided by the smallest coefficient and rounded to the nearest integer. We determined the total score for each individual by assigning the points for each variable and adding them up. For simple interpretation in a clinical setting, we categorised the total risk score into two based on the Youden index (optimal cut-off point). Then, patients were categorised into high-risk or low-risk groups based on the summation of individualised risk scores. As development of a simplified risk score might result in the loss of information that can result in a loss in prediction accuracy, a risk score model was developed to compare its accuracy with the original model. The ROC was plotted and an AUC with its 95% confidence level was computed to evaluate the discriminatory power of the scoring system.

RESULTS
Baseline patient characteristics
A total of 912 patients who had a stroke were enrolled in the study. The median age of patients was 65 years (IQR 55–65). Over half (58.2%) of patients were male and 53.4% of patients were urban dwellers. The incidence of in-hospital mortality in this study was 132 (14.5%). Of the total in-hospital deaths, 94 (71.2%) were aged ≥65 years, and women accounted for 67 (50.8%) of in-hospital deaths. More than half of deaths (55.3%) occurred among urban residents.
Ischaemic stroke was the most common type of stroke, accounted for 617 (67.7%) of admitted patients who had a stroke. One hundred forty (15.4%) patients were comatose at admission, and the median time to admission was 24 hours (IQR 12–24). Hemiplegia was one of the clinical presentations in 271 (29.7%) of admitted patients who had a stroke. Five hundred fifty-nine (71.2%) in-hospital deaths were from patients who were comatose during admission.

Hypertension was the most common risk factor found in 462 (50.7%) of patients who had a stroke. Sixty-three (6.9%) patients had diabetes mellitus, and seizure was found only in 42 (4.6%) of patients who had a stroke.

Fifteen per cent of patients had a history of stroke or transient ischaemic attack.

Biochemical characteristics were not normally distributed when distribution was checked by the Shapiro-Wilk normality test. The median haemoglobin level was 138 g/L (IQR 125–138). The median blood urea nitrogen (BUN) of patients who had a stroke was 36 mg/dL (IQR 20–36) and the median creatinine was 0.854 mg/dL (IQR 0.68–0.85).

### Development and validation of a prediction model

Out of 912 admitted patients who had a stroke finally enrolled in the study, 132 (14.5%) died during their hospital stay.
hospital stay. There were 223 (24.5%) BUN, 155 (17%) creatinine, 120 (13.2%) haemoglobin, 12 (1.32%) temperature and 1 (0.1%) GCS missing values. To predict in-hospital mortality, 20 sociodemographic, risk factors/complications, clinical and biochemical characteristics predictors were included in this study. The bivariable logistic regression analysis identified 15 candidate predictors for the multivariable prediction model. To be liberal, a p value of 0.25 as a cut-off value was used to enter into the multivariable regression. Accordingly, variables were age, sex, type of stroke, heart rate, respiratory rate, temperature, hemiplegia, level of consciousness/GCS, hypertension, diabetes mellitus, pneumonia, seizure, urinary incontinence, BUN and creatinine. Finally, eight variables remained in the final reduced model with a p value of <0.05 (table 1).

The AUC of the final reduced model was 0.895 (95% CI: 0.859–0.932) (figure 2). The model fitness test had a p value of 0.225; the calibration curve is nearly 45°.

Table 2  Prognostic predictors and prognostic index scores derived from 912 patients who had a stroke admitted to the medical ward of Felege Hiwot Comprehensive Specialized Hospital from 11 September 2018 to 7 March 2021

<table>
<thead>
<tr>
<th>Predictors</th>
<th>β coefficients</th>
<th>Risk score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt;65</td>
<td>0.722</td>
<td>1</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>0.642</td>
<td>1</td>
</tr>
<tr>
<td>Creatinine (&gt;1.2 mg/dL)</td>
<td>0.790</td>
<td>1</td>
</tr>
<tr>
<td>Temperature &gt;37.5°C</td>
<td>1.128</td>
<td>2</td>
</tr>
<tr>
<td>Haemorrhagic stroke</td>
<td>0.817</td>
<td>1</td>
</tr>
<tr>
<td>GCS (3–8)</td>
<td>2.908</td>
<td>5</td>
</tr>
<tr>
<td>Diabetes mellitus (yes)</td>
<td>1.061</td>
<td>2</td>
</tr>
<tr>
<td>Pneumonia (yes)</td>
<td>0.914</td>
<td>1</td>
</tr>
</tbody>
</table>

GCS, Glasgow Coma Scale.

showing that there is no difference between predicted and observed probabilities (figure 3).

To minimise too optimistic results from the original model, we used a bootstrapping technique using mrs package to validate our model. This study used 1000 bootstrap samples with replacement; the corrected AUC was the same as the original model and the optimism coefficient for the validated model was 0.079. The \( \beta \) coefficients from the bootstrapped model produced the same results as the original \( \beta \) coefficients. Based on the optimal cut-off point (Youden index), the predicted risk cut-off point of \( \beta \) coefficients of predicted probability is 0.1792; the model has accuracy of 0.905 (95% CI: 0.884, 0.923), with a sensitivity of 0.7955, specificity of 0.9231, positive predictive value of 0.6364 and negative predictive value of 0.9639.

The linear relationship of the original model is shown as:

\[
\log \text{odds of } y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \\
\log \text{odds of in-hospital mortality of stroke} = -4.649 + 0.722 \text{ age} + 0.642 \text{ sex} + 1.128 \text{ temperature} + 0.817 \text{ type of stroke} + 2.908 \text{ GCS} + 0.914 \text{ pneumonia} + 1.061 \text{ diabetes mellitus} + 0.790 \text{ creatinine}.
\]

**Clinical risk score development**

For practical utility, a simplified risk score was developed from eight variables (age in years, sex, temperature, type of stroke, diabetes mellitus, pneumonia, level of consciousness and creatinine). Each variable score was assigned by dividing its corrected \( \beta \) coefficient with the least coefficient in the model (beta coefficient of sex) and rounding to the nearest integer score (table 2).

One point was assigned each for age ≤65 years, female sex, haemorrhagic type of stroke, creatinine >1.2mg/dL and pneumonia; 2 points were assigned for temperature ≥37.5°C and diabetes mellitus; and 6 points were assigned for coma level of consciousness (GCS ≤8). Age <65 years, male sex, ischaemic stroke, temperature ≤37.5°C, having no comorbidity of diabetes mellitus and pneumonia, and patients having admission GCS 9–15 were given each a score of 0. The simplified score had an AUC of 0.893 (95% CI: 0.856–0.929). The possible minimum and maximum scores a patient who had a stroke can have are 0 and 14, respectively. Bias of \( \beta \) coefficients was corrected after internal validation.

For practical utility, we developed a simplified risk score. The risk score produced prediction accuracy of an AUC of 0.893 (95% CI: 0.856–0.929), which is nearly a comparable prediction accuracy with the original model (see figure 4). In this study, the maximum total risk score is 14; for simple interpretation in the clinical settings, we categorised risk scores into less than or equal to 5 points (low-risk group), and greater or equal to 5 points (high-risk group) based on the Youden index (optimal cut-off point). Therefore, a patient can have minimum and maximum risk scores of 0 and 14, respectively. Out of the total 132 (14.5%) cases of in-hospital deaths, 33 (25%) were in the low-risk group and 99 (75%) were in the high-risk group (table 3).

Based on the Youden index (cut-off point >5), the risk score model performance has a sensitivity of 78.79% and a specificity of 91.67%.

Score of individual patient = age + sex + pneumonia + 2 × diabetes mellitus + 5 × GCS + 2 × temperature + type of stroke + creatinine.

**DISCUSSION**

Stroke is one of the leading causes of death and disability from cardiovascular diseases; it is the main driver of increasing healthcare costs, which affect the economic stability of communities. The United Nations Sustainable Development Goals aim to reduce the risk of premature death among people aged 30–69 years from cardiovascular disease, cancer, diabetes and chronic lung disease by one-third by 2030. As part of this target, accurate risk stratification tools and models to guide clinical decision-making for patients who had a stroke in the healthcare settings are very essential. Therefore, this study aimed to develop and validate a clinical risk score for in-hospital stroke mortality. Therefore, a prediction model was developed using sociodemographic, clinical presentation, risk factors/complications and biochemical characteristic variables.

In the present study, the model was developed with eight predictors (age, sex, stroke type, temperature, level of consciousness, creatinine, and presence of comorbidity of pneumonia and diabetes mellitus).

Our model has an excellent discrimination performance of AUC 0.895 (95% CI 0.859–0.932), suggesting that the sensitivity and specificity of the model appear robust enough to be an aid of the clinical judgement of patients who had a stroke. In addition, the model has a p value of 0.225 and the calibration curve is nearly 45°, showing a magnificent agreement between predicted and observed probabilities. The model helps to identify patients who had a stroke having a higher risk of in-hospital mortality at admission and helps clinicians to focus early on potentially modifiable risk factors to prevent further progression of the high-risk patient population.

Based on the researchers’ search previously, many prediction models were developed either for haemorrhagic or ischaemic stroke, but there was only one

---

**Table 3** Risk stratifications for in-hospital stroke mortality using a simplified prediction score

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Number of patients who had a stroke (%)</th>
<th>Number of patients who died in hospital (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk (≤5)</td>
<td>766 (84.0)</td>
<td>33 (25.0)</td>
</tr>
<tr>
<td>High risk (&gt;5)</td>
<td>146 (16.0)</td>
<td>99 (75)</td>
</tr>
<tr>
<td>Total</td>
<td>912 (100)</td>
<td>132 (100)</td>
</tr>
</tbody>
</table>
inclusive model. Different models were developed for in-hospital mortality of ischaemic stroke including the PREMISE score,11 the PLAN score21 and the model developed for in-hospital mortality of elderly patients who had an acute stroke.22 23 Other models were also developed for haemorrhagic stroke including the max-intracerebral haemorrhage score24 and prediction models for aneurysmal subarachnoid haemorrhage,25 but there was only one developed model for all types of stroke.

Previously, the GWTG-Stroke score was the first developed and validated model for all types of stroke using predictor variables: age, stroke type, method of arrival at the hospital, history of atrial fibrillation, previous stroke, coronary artery disease, carotid stenosis, diabetes, peripheral vascular disease, hypertension, dyslipidaemia, smoking, and weekend or night-time admission. The GWTG-Stroke mortality score had fair discrimination ability without including the NIHSS variable. The score had good discrimination only when it included the NIHSS variable. It also has limited clinical practicability because the score is difficult to calculate with points ranging from 0 to 204, and its application needs the aid of web-based computational support.14

A biomarker study conducted to assist clinical decision-making for in-hospital mortality due to stroke revealed that biomarkers such as endostatin, tumour necrosis factor receptor-1 and interleukin 6 were predictors of in-hospital stroke mortality. According to the findings of this study, the risk of mortality increased up to 69% when these biomarkers were combined. The study stressed that the addition of these biomarkers to clinical predictors improved the discrimination performance of the model.25 However, the biomarker mechanism of intervention is less likely to be implemented and to be used in routine patient care practice in low-income settings.

Our prediction model and risk score are the only ones developed for predicting in-hospital stroke mortality in Ethiopia. Although the model is simple and easily applicable in clinical settings and was developed from a smaller number of predictors with good discrimination ability, it was developed in a single hospital. An important limitation of our work might be the fact that in-hospital stroke mortality could be affected by treatment effects, which in the current study were not considered. Furthermore, because this is a single-centre study, it cannot be replicated in other hospitals without external validation.

CONCLUSIONS

The risk score has good discrimination ability and was developed from predictors that are routinely measured, simple and easily remembered. Researchers’ further investigation is required to externally validate the risk score in different populations and settings. Health professionals’ acceptance and application of a developed risk score may be necessary at the bedside, and health policymakers may incorporate the developed risk score in stroke management, follow up its application and judge it.

Acknowledgements We acknowledge Bahir Dar University for offering us ethical approval to commence our study and the Felege Hiwot Comprehensive Specialized Hospital (FHCSh) for granting permission to conduct this study. In particular, we are thankful to the medical ward and card room workers for their kind cooperation and patience throughout the data collection period.

Contributors TGA is the guarantor who conceptualised and designed the study, supervised the data collection, performed data analysis, interpreted the findings, drafted the manuscript, and generally takes responsibility for the overall contents. SFF, AMD and RMA were responsible for data entry, analysis, interpretation and manuscript review. ZA conceptualised the study, supervised every component of the study process, performed data analysis and interpretations of the findings, and revised the manuscript. All authors read and approved the final version of 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, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The research proposal was approved by the Ethical Review Board of Bahir Dar University, College of Medicine and Health Sciences, with a protocol number 088/2021, and ethical clearance was obtained from the board. Ethical clearance and supporting letter were submitted and permission was obtained from the hospital management. To ensure confidentiality of the patients’ information, the name and address of the patients were not recorded during the data collection.

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, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ids

Tiruayehu Getnet Abebe http://orcid.org/0000-0003-1512-1668
Selfinhe Fenta Feleke http://orcid.org/0000-0002-3051-1296
Anteneh Mengist Dessie http://orcid.org/0000-0002-5919-9670
Rahel Muluati Anteneh http://orcid.org/0000-0002-5960-893X
Zalemam Almaw Anteneh http://orcid.org/0000-0003-4879-675X

REFERENCES
