Ethiopian Journal of Health and Biomedical Sciences

Original article Open access

Modeling Determinants of Time to Death of Stroke Patients in Harari Regional State, Ethiopia: An Application of Shared Frailty Models

***Alebachew Abebe*[1*], *Kumela Ayansa*[2], *Kasahun Takele*[1]**

[1]*Department of Statistics, College of Computing and Informatics, Haramaya University, P.O.Box 138, Dire Dawa, Ethiopia,* [2]*Ethiopian Statistical Services, Oromia Regional State, Ethiopia.*

***Corresponding author:** Alebachew Abebe; **Email:** aleb.abebe19@gmail.com.*

Citation: Abebe A, Ayansa K, Takele K. Modeling Determinants of Time to Death of Stroke Patients at Harar Regional State, Ethiopia: Application of Shared Frailty Models. Ethiop J Health Biomed Sci. 2025;15(1):45-58.

DOI: https://doi.org/10.20372/ejhbs.v15i1.979

Article History

Submitted: January 1, 2025
Revised: July 18, 2025
Accepted: July 31, 2025

Keywords: Stroke Patients, Time-to-Death, Retrospective Study Design, Frailty Model, Shared Frailty Model.

Publisher: University of Gondar

Abstract

Background: *Stroke, a condition caused by interrupted blood flow to the brain, is the second leading cause of death worldwide and a major contributor to morbidity and disability. Despite its global impact, there remains a need for more comprehensive and flexible statistical models to better understand the timing of stroke-related mortality. This study addressed this gap by applying robust shared frailty models to stroke patient data.*

Objective: *The main purpose of this study was to identify the factors influencing time to death among stroke patients using shared frailty models, accounting for hospital-level clustering effects.*

Methods: *A retrospective study was conducted at Harar Regional State, Ethiopia, across three hospitals: Harar General Hospital, Jegol hospital, and Hiwot Fana Specialized University Hospital. A total of 224 stroke patients admitted to medical wards between September 1, 2020, and November 1, 2023, were included. Data was coded, cleaned, and entered in to SPSS version 25 and further analyzed using R-Studio. The presence of clustering (frailty) effects among hospitals was evaluated, and different shared frailty models were compared to identify the best-fitting model.*

Results: *Of the total 224 stroke patients, 51(22.77%) died during the study period, while 173 (77.23%) were censored. The median survival time was estimated at 14 days, highlighting the acute nature of the disease. The Weibull-inverse Gaussian shared frailty model provided the best fit for the data, accounting for unobserved heterogeneity across hospitals. Significant predictors of shorter time to death included hypertension (φ = 2.118; 95% CI: 1.145-3.917), cardiac disease (φ = 2.667; 95% CI: 1.343-5.296), diabetes mellitus (φ = 3.035; 95% CI: 1.1.560-5.906), atrial fibrillation (φ = 3.247; 95% CI: 1.619-6.511), and presence of basic complications (φ = 2.983; 95% CI: 1.477-6.023).*

Conclusion: *This study highlights the importance of hospital-level clustering effects in survival analysis of stroke patients. The significant frailty effect suggests variability in outcomes across hospitals, underlining the need for tailored interventions. Clinicians and hospital administrators should consider these differences when managing stroke patients, emphasizing timely follow-up, individualized care, and resource allocation to improve survival outcomes.*

Introduction

A potentially deadly stroke may arise from an inadequate blood supply to the brain. Two types of strokes signify a halt to blood flow or bleeding: hemorrhagic and ischemic strokes. Dementia and depression can also result from strokes [1]. Stroke rates have already risen to epidemic levels. Worldwide, strokes have claimed the lives of nearly 110 million people. A stroke will occur in the lifetime of one in four adults over the age of 25. According to WHO data from 2022, around 12.2 million people worldwide will have a stroke for the first time, and the illness will be the cause of 6.5 million deaths [2].

Stroke is the world's second leading cause of death, after ischemic heart disease. The global burden of stroke, including mortality, morbidity, and disability, is rising [3]. In high-income countries, in-hospital mortality for stroke patients ranges from 3% to 11%, whereas in low-and middle-income countries it increase to 7%-15%. Hemorrhagic strokes carry a substantially higher fatality rate (nearly 38%) compared with ischemic strokes (8%-12%), with outcomes influenced by factors such as stroke severity, patient age, comorbidities, treatment efficacy, and complications [3]. Given the rising global burden of stroke, approximately one person worldwide suffers a stroke every two seconds [4]. According to the 2022 Global Stroke Fact Sheet, the lifetime risk of experiencing a stroke has increased by 50 percent over the past 17 years. Between 1990 and 2019, stroke incidence rose by 70 percent, mortality by 43 percent, prevalence by 102 percent, and disability-adjusted life years (DALYs) by 143 percent. Notably, low- and lower-middle-income countries shoulder the greatest portion of this burden, accounting for 86 percent of stroke-related deaths and 89 percent of stroke-attributable DALYs [5].

In Sub-Saharan African (SSA) countries, the range was 11.1% to 43.4%, three to four times higher than in developed nations [6]. While stroke rates have dropped by 42% in high-income nations, they have risen by over 100% in low- to middle-income countries. Research suggests that stroke incidence and prevalence show minimal geographic variation [7]. According to WHO data released in 2020, stroke caused 39,362 deaths in Ethiopia, accounting for approximately 6.98% of all deaths in the country. Ethiopia has an age-adjusted stroke mortality rate of 83.71 per 100,000 population, ranking 90thglobally. A study conducted at St. Paul's Millennium Medical College in Ethiopia reported an average hospital stay of nine days for stroke patients, with an in-hospital mortality rate of 14.7% (23.5% for hemorrhagic strokes and 6.1% for ischemic strokes) [8]. The 2013 Global Burden of Disease report indicates that strokes are responsible for 80% of mortality in low- and middle-income countries. Despite stroke consistently ranking among the top three causes of morbidity and mortality in Ethiopia in recent years, there is limited knowledge about the outcomes of stroke treatment when adequate resources are available [9].

In a frailty model, the acceleration factor (φ) quantifies the effect of a covariate on the time to an event. That means φ is equivalent value with a risk ratio of stroke patients and interpreted as 100 people per φ (risk ratio). It essentially represents how much faster or slower an individual experiences the event compared to a baseline, based on the value of the covariate. A value of φ greater than 1 indicates a faster time to event, while a value less than 1 indicates a slower time to event. These models are used in survival analysis, particularly when dealing with clustered or correlated data. They account for unobserved heterogeneity (frailty) that might influence the survival times of individuals within the same cluster [15]. This study aims to evaluate the heterogeneity of unobserved factors including a random effect (frailty) in the model. It focused on the risk factors of mortality among stroke patients treated at Jegol Hospital, Harar General Hospital, and Hiwot Fana Specialized University Hospital in the Harari regional state, Ethiopia. The main objective of this study was to model the determinants of the time to death among stroke patients using a shared frailty model.

Method

Study setting, design, and period

This study utilized a retrospective study design from September 1, 2020 to November 1, 2023 in the three hospitals of Harar City: Harar General Hospital, Jegol Hospital, and Hiwot Fana Specialized University Hospital. The city is located in eastern Ethiopia approximately 526 kilometers from Addis Ababa, the capital city of Ethiopia. Those hospitals are well-equipped to provide a quality healthcare services give with an infrastructure and ample supplies for the patients. Offering a comprehensive range of services from inpatient and outpatient care to emergency and specialized treatments

such as surgery, they play a crucial role in meeting the diverse healthcare needs of the population. Furthermore, the subsidies for healthcare services at these hospitals make them affordable and accessible to nearly all residents in Harar City.

Source and study population: A secondary source of a stroke patients' data were collected for this study. The population study was considered a patient who received a treatment from medical ward outpatient in the hospitals from September 1, 2020 to November 1, 2023.

Inclusion and exclusion criteria's

Inclusion criteria: Stroke patients who had been admitted to the intensive care unit ward at those hospitals were included in this study from September 1, 2020 to November 1, 2023. All patient charts with a complete data in the variables of interest were included in this study.

Exclusion criteria: In hospitals, stroke patients' charts were excluded from certain treatments or clinical trials due to specific criteria like the severity of the stroke, the patient's medical history, or time elapsed since the stroke's onset. Additionally, patients with certain medical conditions or those who don't meet the specific time windows for treatment was also be excluded.

Data collection methods and tools

A data collection was reviewed a stroke patients medical records, using tools structured data collection forms. The data collection tool was included patient demographics (sex, age and place of residence), disease types (hypertension, cardiac disease and diabetes mellitus), atrial fibrillation, baseline complication, stroke types, drug types, and patients' treatment from time-to-death. The following activities were done for a data collection methods and tools in the study. **Pre-testing the data collection tool:** Before the main study, the data collection tool was pre-tested on a small sample of records to identify any potential issues with clarity, applicability, or consistency of the variables. **Trained data collectors:** Trained medical professionals, such as nurses or residents, were involved in the data extraction process to ensure accurate and consistent data collection. **Retrospective review of medical records:** This involves carefully examined the existing patient charts and files were extracted a relevant information for stroke patients datasets. **Structured data collection forms:** These forms were designed to systematically gather specific data points from the medical records, ensuring consistency, and completeness.

Study variables

Dependent variable: Time to death measured (in days) from the start of anti-stroke treatment to the date of the patient's death or censoring. **Independent variables:** Sex, age, residence, hypertension, cardiac disease, diabetes mellitus, atrial fibrillation, baseline complication, stroke types, and drug types were considered under this study. The details of those variables, categories (codes), and data nature were described in the table below **(Table 1).**

Table 1: Descriptions of the variables, categories (codes) and data nature of stroke patients in Harari regional state from September 1, 2020 to November 1, 2023.

Variable	Levels and Codes	Data Nature
Sex	Female=0, Male=1	Categorical
Age	year	continuous
Place of residence	Urban=0, Rural=1	categorical
Hypertension	No=0, yes=1	categorical
Cardiac disease	No=0, yes=1	categorical
Diabetes mellitus	No=0, yes=1	categorical
Atrial fibrillation	No=0, yes=1	categorical
Baseline complication	No=0, yes=1	categorical
Stroke types	Ischemic=1, Hemorrhagic=2	categorical
Drug types	Anti-coagulants and Anti-platelet=1, Anti-coagulants, Anti-platelet and Anti-hypertensive=2, Anti-coagulants, Anti-platelet, and Statin=3, Anti-coagulants, Anti-platelet, Statin and Anti-hypertensive=4, Anti-coagulants, Anti-platelet, Statin, Anti-hypertensive and Antibiotics=5, Anti-platelet, Statin and Anti-hypertensive=6, Anti-hypertensive and Antibiotics=7, Anti-coagulants and Anti-hypertensive=8 and Other=9	categorical

Data management and analysis

A stroke patients' datasets was coded, cleaned, and entered in to SPSS version-25 for data management and R-studio version -4.3 for analyzed. The first step of the analysis was made a descriptive statistics that shown a frequencies and percentages of the variables. At second stage: a non-parametric test was used for compare groups of variables such as Kaplan Meier estimates, log-rank test and Wilcoxon test, whose P<0.05 were considered for multivariate analysis. In third step, fit a Cox-proportional hazard model to check a relationship between the dependent and independent variables. Finally, a marginal log-likelihood approach was used a parameter estimation to fit a robust model for shared frailty model was considered different frailty distributions and predicting.

The Survival models

The time until an event occurs is the outcome variable of interest in survival analysis, which is a collection of statistical processes for data analysis. The time variable in a survival analysis is commonly referred to as "survival time (ST)", since it indicates how long an individual has survived over a given period of time [10]. Let T be a random variable associated with the STs, t is the value of T, and $f(t)$ be the probability density function (PDF) of ST at a time value, t. A cumulative distribution function (CDF), $F(t), t \geq 0.$

$$F(t) = P(T \leq t) = \int_0^t f(u)du \qquad (1)$$

Survival and hazard function (SF & HF), and commutative hazard function (CHF) in the equation number 2, 3 and 4 were presented, respectively.

$$S(t) = P(T \geq t) = \int_t^{\infty} f(u)du \qquad (2)$$

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}\ln S(t) \qquad (3)$$

$$H(t) = \int_0^t h(u)du \qquad (4)$$

Non-Parametric Survival Model

Kaplan-Meier Estimator (KME): Suppose that r individuals have failures in a group of individuals and $0 \leq t(1) \leq t(2) \leq ... \leq t(r) < \infty$ be the observed ordered death times [11]. Assume that $r(j)$ be size of the risk at $t(j),$ where the risk set encompasses individuals alive and uncensored before $t(j).$ Let $d(j)$ be the number of observed events at $t(j), j = 1,2,...,r.$ Then, the KMEs for both survival and CHF had a probability of developing a disease at any time in equation 5 and 6 were presented, respectively.

$$\hat{S}(t) = \prod_{t(j)<t} \frac{[r(j) - d(j)]}{r(j)} \qquad (5)$$

$$H(t) = -\ln \hat{S}(t) \qquad (6)$$

Median Survival Time (MST): The smallest observed ST for which the value of the estimated SF is less than 0.5[12]. Where, $t(i)$ is an observed ST for the i^{th} individual, $i = 1,2,...,n.$

$$t(50) = \min \frac{t(i)}{S(t_i)} \leq 0.5 \qquad (7)$$

Frailty Model (FM): The random effects (frailty) element was added to standard models of analysis to account for unmeasured variables or linked survival data [13].

Shared Frailty Model (SFM): Conditional on the random term as a frailty denoted by $u_i,$ the STs in cluster-i $(1 \leq i \leq n)$ was assumed to be independent and the proportional hazard frailty model (PHFM) [14].

$$h_{ij}(t/x_{ij}, u_i) = h_0 \exp(\beta x_{ij} + u_i) \qquad (8)$$

Where, as an alternative, if the PHs assumption does not hold, then an AFTFM was applied.

$$h_{ij}(t/x_{ij}, u_i) = h_0 \exp(\beta x_{ij} + u_i)[\exp(\beta x_{ij} + u_i)t] \qquad (9)$$

Where, i is indicates the i^{th} cluster and j is indicates the j^{th} individual for the i^{th} cluster, $h_0(.)$ is the baseline hazard, u_i is the random term for all the subjects in cluster i, X_{ij} is a vector of the covariates for subject j in cluster i, and β is a vector of the regression coefficients.

Test of Unobserved Heterogeneity: The variance of θ is both large and statistically different from 0, indicates a cluster's heterogeneity, and the individuals' had a strong association to each other. Let θ=0, frailties=1, means that the cluster effects are not existent and the occurrences are independent for both within and between clusters [15].

Frailty Distribution

Gamma Frailty Distribution (GFD): The distribution of frailty Z is one of a parameter for GD [16]. The density of a GD of a random variable with parameter, $\theta > 0.$

$$f_Z(Z_i) = \frac{Z_i^{\frac{1}{\theta}} \exp(-Z_i/\theta)}{\Gamma(1/\theta)^{1/\theta}} \quad (10)$$

Where, $\Gamma(.)$ is GF and GD $(\mu, \theta),$ with μ is fixed to be one for identifiable and its variance becoming $\theta.$ Where, $Z_i > 1,$ individuals in the group-i are more frailty, $Z_i < 1,$ individuals are less frailty and have lower risk.

Conditional SF and HF of the GFD were given, respectively [17].

$$S_\theta(t) = [1 - \theta \ln S(t)]^{\frac{-1}{\theta}} \quad (11)$$

$$h_\theta(t) = h(t)[1 - \theta \ln S(t)]^{-1} \quad (12)$$

GD measures the association between any two event times from the same cluster in the multivariate case.

$$\tau = \frac{\theta}{\theta + 2}, \ where\ \tau \in (0,1) \quad (13)$$

Inverse Gaussian Frailty Distribution (IGFD): The PDF of an IND of a random variable with a mean of 1 and variance, $\sigma^2 = \theta$ [18].

$$f(Z) = \frac{1}{\sqrt{2\prod \theta Z^3}} \exp[\frac{-1}{2\theta Z}(Z-1)^2] \quad (14)$$

Where, $\theta > 0\ \&\ z > 0.$ Consequently, the Laplace transform (LT) of the IND.

$$L(s) = \exp[\frac{1}{\theta}(1 - \sqrt{1 + 2\theta s})];\ \theta\ and\ s > 0 \quad (15)$$

Conditional SF & HF were presented, respectively.

$$S_\theta(t) = \exp[\frac{1}{\theta}(1 - \sqrt{1 - 2\theta \ln S(t)})];\ \theta > 0 \quad (16)$$

$$h_\theta(t) = h(t)[1 - 2\theta \ln S(t)]^{-1/2};\ \theta > 0 \quad (17)$$

IGD frailty yields a Kendall's Tau.

$$\tau = \frac{1}{2} - \frac{1}{\theta} + 2\frac{\exp(2/\theta)}{\theta^2}\int_{\frac{2}{\theta}}^{\infty} \frac{\exp(-u)}{u} du;\ where, \tau \in (0, \frac{1}{2}) \quad (18)$$

Parameter Estimation

The ST of the random variables was given the covariate information from the marginal log-likelihood of the observed data [19].

$$Lm\arg(\varphi, \beta, \theta, Z, X) = \sum_{i=1}^{s}[\sum_{j=1}^{n_i} \delta_{ij}(\log(h_0(y_{ij}) + X_{ij}^T\beta)] + \log[(-1)^{d_i} L^{d_i}(\sum_{i=1}^{n_i} H_0(y_{ij})\exp(X_{ij}^T\beta))] \quad (19)$$

Where, $d_i = \sum_{j=1}^{n_i} \delta_{ij}$ is the number of events in the i^{th} clusters & $L_q(.)$ is the q^{th} derivative of LT for the FD of Z.

$$L^q(s) = -1^q \int_0^\infty Z^q \exp(-Zs) f(Z) dz;\ Where,\ q > 0. \quad (20)$$

Where, φ =vector of the parameters for a baseline HF, β =vector of the regression coefficients & θ =variance of a random effect. Then, $\varphi, \beta\ \&\ \theta$ were obtained by maximizing a marginal log-likelihood. $L^q(.)$ =LT up-to q=max $(d_1, ..., d_s).$

Prediction of Frailties

The frailty term Z_i was predicted as $Z_i = E[Z/z_i, \varphi, \beta, \theta],$ with z_i is the data of the i^{th} cluster. Conditional expectation [21]:

$$Z_i = E[Z/z_i, \varphi, \beta, \theta] = -\frac{L^{(d_i+1)}[\sum_{j=1}^{n_i} H_0(y_{ij})\exp(X_{ij}^T\beta)]}{L^{d_i}[\sum_{j=1}^{n_i} H_0(y_{ij})\exp(X_{ij}^T\beta)]} \quad (21)$$

Models Selection

Akaike Information Criterion (AIC): The model with the smallest AIC value was considered as a better fit [21].

Model Diagnostics

Likelihood Ratio Test (LRT): To test an association of covariates with an outcome in a frailty model.

Cox Snell Residuals: In semi-parametric residual plots were made with a redefinition of the various residuals to incorporate the parametric form of the baseline hazard rates [22].

Data quality control

Before data collection, an ethical approval was obtained from relevant ethics committees of the respective hospitals, was accessed and used a stroke patient datasets. Finally, data quality was ensured by implemented rigorous data collection procedures, including training for data collectors and regular data quality checks for stroke patients.

Result

Descriptive statistics

In this study, 224 charts were assessed for followed a stroke treatment in Harar General Hospital, Jegol Hospital, and Hiwot Fana Specialized University from September 1, 2020 to November 1, 2023. The main objective of this finding was to model the determinants of the time to death of stroke patients. From a total of 224 stroke patients, 51(22.78%) were dead and the rest, 173(77.23%), were censored. This is indicated by the fact that most off the stroke patients were cured after getting a different treatment in the hospitals. The hypertension disease in developing a stroke patients out of 224, 140 (62.5%) were absent of hypertension and the remaining 84 (37.5%) were present with hypertension. The death rate appears to be highest in stroke patients who had hypertension, 52.94%, compared to 47.06% in non-hypertensive stroke patients **(Table 2).**

When accounting for the 224 hospitals that received medical care, the percentages were 35.27%, 22.32%, and 42.41%, the hospitals that were treated in that order being Harar General Hospital, Jegol Hospital and Hiwot Fana Specialized University Hospital, respectively. According to the hospital mortality total, the rates at Harar General Hospital, Jegol Hospital, and Hiwot Fana Specialized University were 27.45%, 39.21%, and 33.34%, respectively. Similarly, fashion other variables were presented as detailed in table 2.

In table 3 shows that the mean age of the stroke patients was 55.7 years, with the oldest and youngest being 93 and 20 years old, respectively and with a standard deviation of 16.4. Out of 224 total stroke patents of the median survival time was 14 days for this study.

Survival of Significantly Different Groups

The log-rank test results in table 4 indicate that a significant difference in death events across the groups of hypertension, diabetes mellitus, atrial fibrillation, cardiac disease, and basic complications at the 5% level of significance.

The Kaplan-Meier (KM) estimator survival curve was used to estimate the survival function among different covariates so that one can make a comparison. Separate graphs of the estimates of the KM survivor functions were constructed for different categorical covariates. In general, the survivorship pattern of one was laid above another, which means that the group was defined by the upper curve has a better survival rate than the group was defined by the lower curve (Figure 1).

Table 2: Descriptive statistics for categorical variables of stroke patients from September 1, 2020 to November 1, 2023 in Harari regional state.

Covariate Variables	Category	Patients Status		
		Censored	Dead	Total
Sex	Female	60(75%)	20(25%)	80(35.71%)
	Male	113(78.47%)	31(21.53%)	144(64.29%)
Residence	Urban	81(75%)	27(25%)	108(48.21%)
	Rural	92(79.31)	24(20.69%)	116(51.79%)
Hypertension	No	116(82.86)	24(17.14%)	140(62.50%)
	Yes	57(67.86%)	27(32.14%)	84(37.50%)
Cardiac Disease	No	129(84.87%)	23(15.13%)	152(67.86%)
	Yes	44(61.11%)	28(38.89%)	72(32.14%)
Diabetes Mellitus	No	148(86.05%)	24(13.95%)	172(76.79%)
	Yes	25(48.08)	27(51.92%)	52(23.21%)
Atrial fibrillation	No	151(86.29%)	24(13.71%)	175(78.12%)
	Yes	22(44.90)	27(51.10%)	49(21.88)
Basic complication	No	108(83.72%)	21(16.28%)	129(57.59%)
	Yes	65(68.42%)	30(31.58%)	95(42.41%)
Stroke types	Ischemic	114(77.03%)	34(22.97%)	148(66.07%)
	Hemorrhagic	59(77.63%)	17(22.37%)	76(33.93%)
Drug types	Anti-coagulants and Anti-platelet	12(66.67%)	6(33.33%)	18(8.04%)
	Anti-coagulants, anti-platelet and anti-hypertensive	19(82.61%)	4(17.39)	23(10.27%)
	Anti-coagulants, anti-platelet, and statin	28(84.85%)	5(15.15%)	33(14.73%)
	Anti-coagulants, anti-platelet, statin and anti-hypertensive	25(73.53%)	9(26.47%)	34(15.18%)
	Anti-coagulants, anti-platelet, statin, anti-hypertensive and antibiotics	17(70.83%)	7(29.17)	24(10.71%)
	Anti-platelet, statin and anti-hypertensive	7(77.78%)	2(22.22%)	9(4.02%)
	Anti-hypertensive and antibiotics	18(78.26%)	5(21.74)	23(10.27%)
	Anti-coagulants and anti-hypertensive	15(88.24%)	2(11.76%)	17(7.59%)
	Other	32(74.42%)	11(25.58%)	43(19.20%)
Hospitals	Harar General Hospital	65(82.28%)	14(17.72%)	79(35.27%)
	Jegol Hospital	30(60%)	20(40%)	50(22.32%)
	Hiwot Fana Specialized University Hospital	78(82.11%)	17(17.89%)	95(42.41%)

Table 3: Descriptive statistics for continuous variables of stroke patients

Age	**Minimum**	**Maximum**	**Mean**	**Standard deviation**
	20	93	55.71	16.39
Survival time	**n**	**Events**	**Median**	**95% CI**
	224	51	14	[10, NA]

Table 4: Log rank test for equality of survival function of different groups

Covariates	Chi-square value	df	Pr > Chi-Square
Sex	0	1	0.8
Residence	0.1	1	0.7
Hypertension	3.5	1	0.045
Cardiac Disease	12.9	1	<0.001
Diabetes Mellitus	19.9	1	<0.001
Atrial fibrillation	37.5	1	<0.001
Basic complication	4.3	1	0.04
Stroke types	0.1	1	0.8
Drug types	6.4	8	0.6

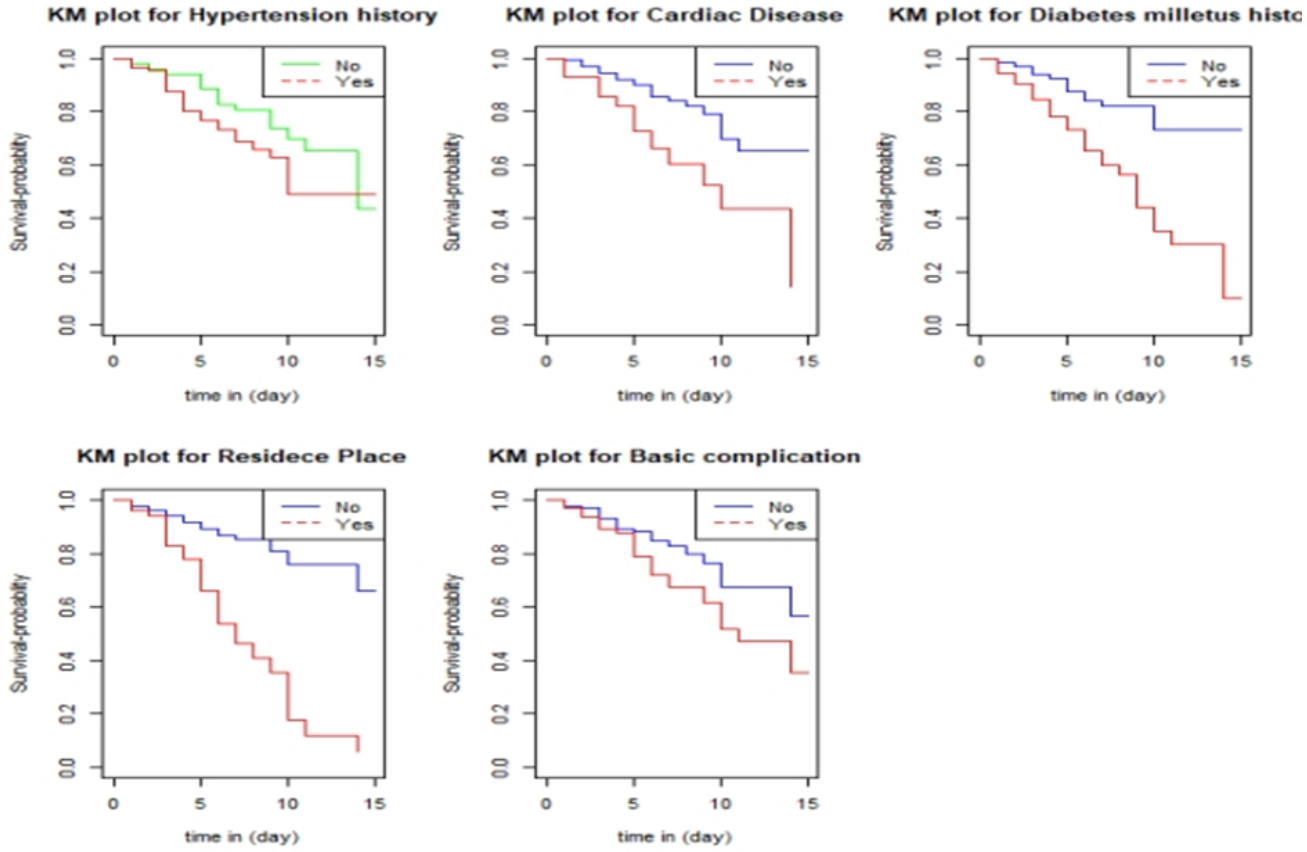

Figure 1: *Kaplan-Meier survivor curves for significantly difference groups*

Test Unobserved Heterogeneity

To predict the random effect of θ to get an idea of heterogeneity among clusters. When θ is large and significant, indicates that a heterogeneity among clusters and a strong correlation among individuals in the same cluster. Conversely, when θ=0, the frailties=1, suggesting that there were no-clustered effects and that events were occurred independently for both in-side and between clusters [15]. The outcomes given in Table 3.4 shows that the likelihood ratio tests of variance of random term(θ) for exponential gamma, exponential inverse Gaussian, Weibull gamma, Weibull inverse Gaussian, log-logistic gamma, and log-logistic inverse Gaussian shared frailty models whose p<0.001 for all shared frailty models. Thus, refers that an indications of unobservable heterogeneity was significant effect (P<0.05) for all models in the stroke patient datasets.

The variance of random effect was highest (θ=0.82) for the Weibull inverse Gaussian shared frailty model and the least (θ=0.23) for the exponential gamma shared frailty model with an exponential baseline hazard. A Kendall's tau(τ) was used to measure the dependence with-in the hospitals (clusters). The values of τ for exponential gamma, exponential inverse

Gaussian, Weibull gamma, Weibull inverse Gaussian, log-logistic gamma, log-logistic inverse Gaussian, log-normal gamma, and log-normal inverse Gaussian shared frailty models were 0.101, 0.097, 0.151, 0.201, 0.149, 0.189, 0.135, and 0.123, respectively. This reveled that, on average had a positive correlation between time-to-death and stroke patients with-in the hospitals **(Table 5).**

Table 5: Test of unobserved heterogeneity using LRT

Shared Frailty Model	LRT	θ	τ	P- value
Exponential Gamma	38.02087	0.225	0.101	<0.001
Exponential Inverse Gaussian	38.25056	0.264	0.097	<0.001
Weibull Gamma	58.07973	0.356	0.151	<0.001
Weibull Inverse Gaussian	58.43818	0.822	0.201	<0.001
Log-logistic Gamma	57.68213	0.349	0.149	<0.001
Log-logistic Inverse Gaussian	58.03188	0.734	0.189	<0.001
Log-normal Gamma	55.79835	0.312	0.135	<0.001
Log-normal Inverse Gaussian	56.11336	0.367	0.123	<0.001

LRT=likelihood ratio test, θ=variance of random terms, τ=Kendall's tau

Model Comparison

In table 6 summarized that all the outcomes of the four baseline hazard functions with two frailty models. Among those models, the Inverse Gaussian frailty model with the Weibull baseline hazard function had the smallest AIC of 383.46. It was the most appropriate model to describe time-to-death for stroke patient datasets.

Table 6: AIC values for the parametric frailty models

Baseline Hazard Function	Frailty Distribution	AIC
Exponential	Gamma	401.877
	Inverse Gaussian	401.6473
Weibull	Gamma	383.8181
	Inverse Gaussian	383.4597
Loglogistic	Gamma	384.2157
	Inverse Gaussian	383.866
Lognormal	Gamma	384.2157
	Inverse Gaussian	385.7845

Multivariable Analysis

A findings of a multivariable frailty models were revealed that the covariates: hypertension(φ=2.118; 95%CI:1.145-3.917), cardiac disease(φ=2.667; 95%CI:1.343-5.296), diabetes mellitus(φ=3.035; 95%CI:1.1.560-5.906), atrial fibrillation(φ=3.247; 95%CI:1.619-6.511), and basic complications (φ=2.983; 95% CI:1.477-6.023) had a significant frailty effect(P<0.05) on the time to death of stroke patients. However, sex was the only insignificant effect(P>0.05) in the model. A variable sex had contained one in the 95% CI of an acceleration effect(φ) in the models in table 7 The interpretation for accelerated factor (φ) under the final multivariate shared frailty model for a stroke patients datasets were as followed: The value φ=2.12 represents a risk ratio. A risk ratio of 2.12 means that for every 100 people with hypertension experiencing the outcome, there would be approximately $47(100/2.12)=47$ people without hypertension experiencing the same outcome. In health terms: this finding highlights hypertension as a serious health concern. It underscores the importance of managing blood pressure to reduce the risk of these related complications. Effective management strategies, such as lifestyle changes and medication, can help mitigate the increased risk associated with hypertension.

A variable cardiac disease had a value of φ=2.67 represents a risk ratio. A risk ratio of 2.67 means that for every 100 people with cardiac disease experiencing in the stroke patients, there would be approximately 38 people without cardiac disease experiencing in the stroke patients. A risk ratio(φ)=2.67, indicates that for every 100 people with car-

diac disease had experienced in the stroke patients, approximately 38 people without cardiac disease had experienced in the stroke patients. A value of φ=3.04(risk ratio) means that for every 100 people with diabetes mellitus were experienced in the stroke patients, approximately 33 people without diabetes mellitus were experienced in the stroke patients. A risk ratio of 3.25(φ=3.25) revealed that for every 100 people with atrial fibrillation was experienced in the stroke patients, there would be approximately 31 people without atrial fibrillation was experienced in the stroke patients. Similarly, A basic complications had φ=2.98 =risk ratio, means that for every 1000 people with basic complications were experienced in the stroke patients, that leads approximately 34 people without basic complications experiencing in the stroke patients.

Table 7: Weibull versus inverse Gaussian multivariable analysis shared frailty model

Covariates	Category	β	φ	St. err	P-value	95% CI
Sex [Female(Ref.)]	Male	0.167	1.181252	0.326	0.726	[0.624, 2.237]
Hypertension [No(Ref.)]	Yes	0.75	2.117636	0.314	0.013 *	[1.145, 3.917]*
Cardiac Disease [No(Ref.)]	Yes	0.981	2.667516	0.35	0.004 **	[1.343, 5.296]*
Diabetes Mellitus [No(Ref.)]	Yes	1.11	3.035117	0.34	0.001 **	[1.560, 5.906]*
Atrial fibrillation [No(Ref.)]	Yes	1.178	3.246609	0.355	0.002 **	[1.619, 6.511]*
Basic complication [No(Ref.)]	Yes	1.093	2.983047	0.359	0.003 **	[1.477, 6.023]*

τ=0.201; θ=0.822; λ=0.014; ρ=1.667; AIC= 383.4597.

φ=acceleration factor, τ=Kendall's tau, θ=variance of random effect, λ=scale, ρ=shape, CI=Confidence Interval & Ref.= reference.

Checking for Overall Goodness of Fit

Diagnostic Plots of the Parametric Baselines

To ascertain if a fitted parametric model accurately represents the data, this choice needs to be made. Out of the four parametric baseline graphs, the Weibull curve exhibits greater linearity than others. This suggests that a Weibull baseline hazard is a better choice for the stroke disease datasets **(Figure 2).**

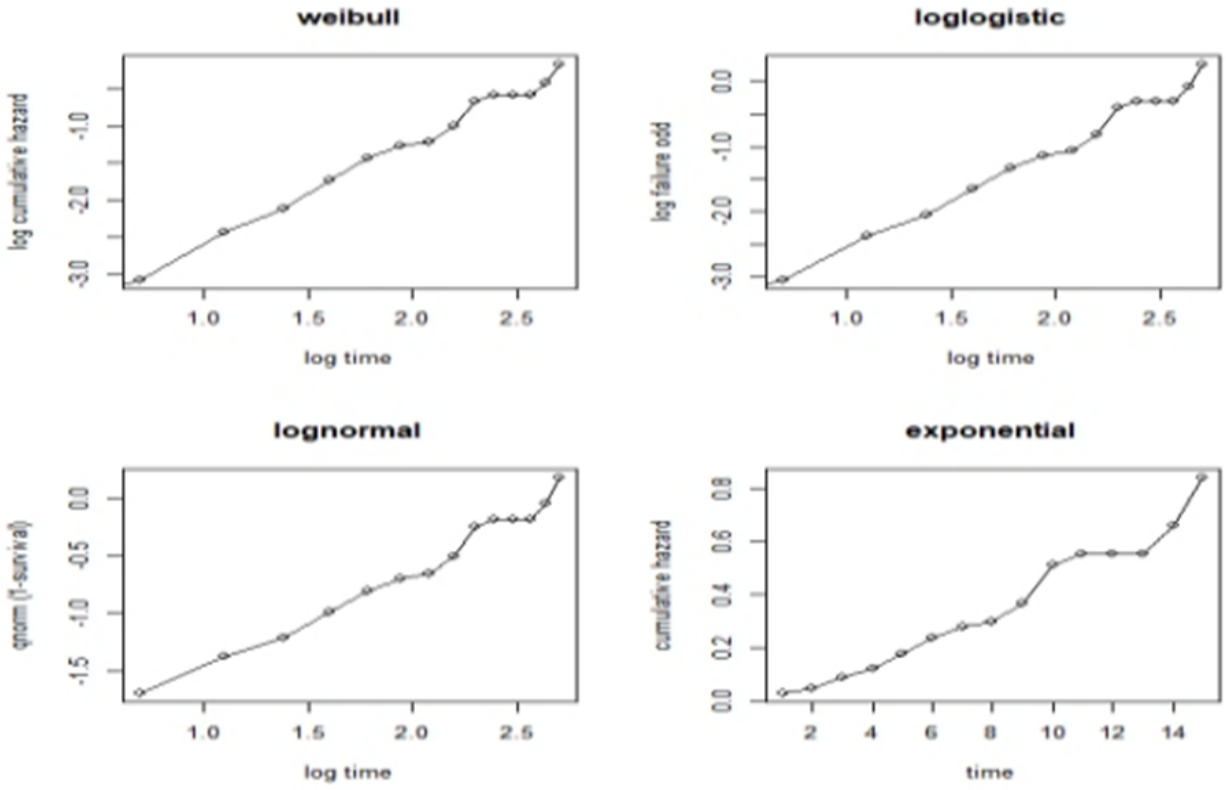

Figure 2: *Diagnostic plot for baselines hazards*

Cox Snell Residual Plots

The residual plot for the Weibull hazard function is rather near the 45-degree straight line through the origin. The **figure 3** revealed that out of all models, the Weibull model that was fitted well to the data was accepted.

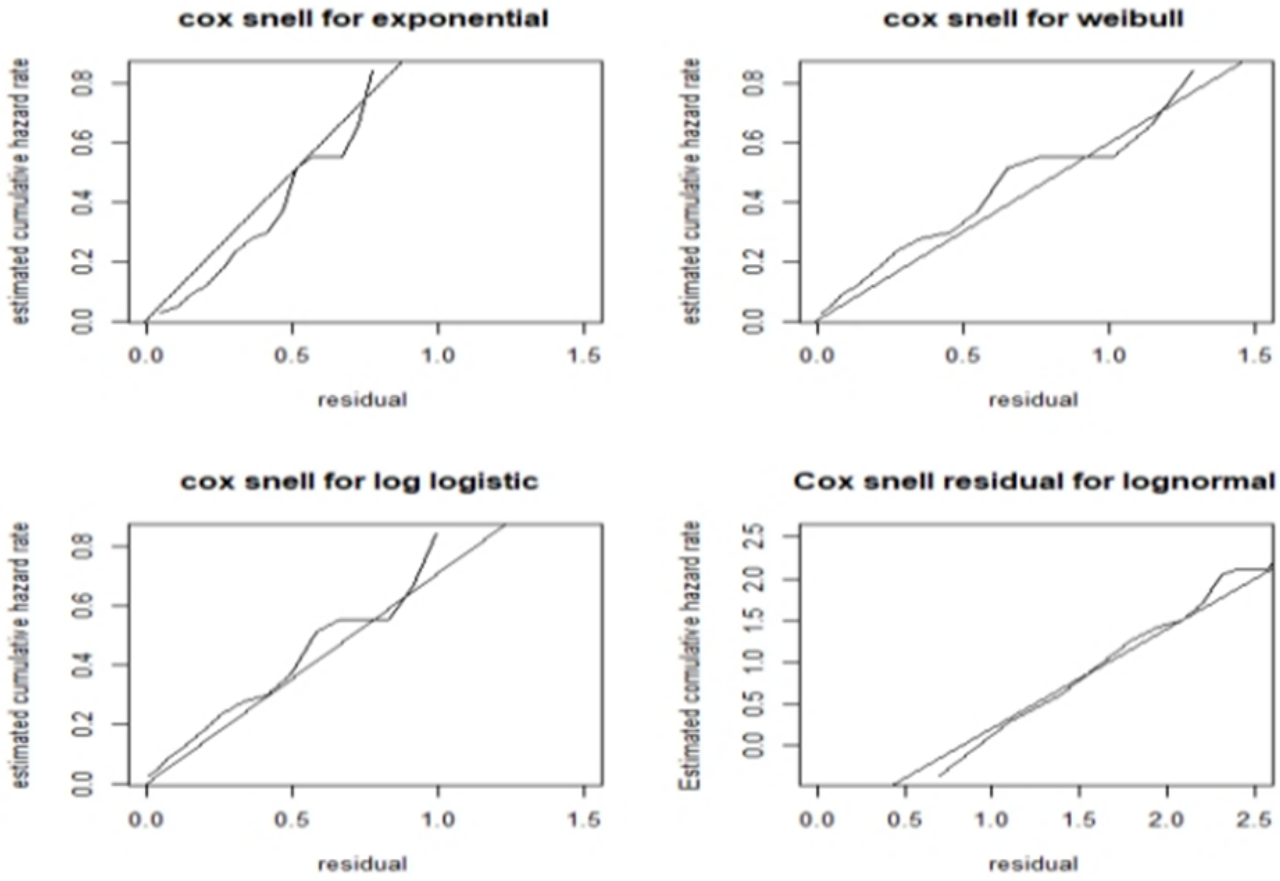

Figure 3: *Cox snell residual plots*

Discussion

This study aimed to model a parametric shared frailty model. The time to death of stroke patients data were used for gamma and inverse-Gaussian frailty distributions among the various baseline hazard functions, including exponential, Weibull, log-logistic, and lognormal. In this study, the population in the same hospital relatively shared some factors, such as the skill of doctors, bedrooms, environment, health facilities, and resources in determining the time to death of stroke patients. A findings of a multivariate shared frailty model revealed that hypertension, cardiac disease, diabetes mellitus, atrial fibrillation, and basic complications had a significant frailty effect (P<0/05) on the stroke patients. However, the variable sex was the only insignificant frailty effect (P>0.05) on the stroke patients. A Weibull inverse Gaussian frailty model had the smallest AIC of 383.46, which makes it the most suitable for explaining time-to-death in the stroke patient datasets. These study findings were supported with stroke patients in Africa by Akinyemi R. et. al., 2019[23].

An outcomes of the models showed that hypertension had a significant impact on the time to death of stroke patients. The prognostic factor of hypertension has increased the risk by a factor of φ=2.12 was compared to patients with non-hypertension; other covariates have remained constant. The value φ=2.12 represents a risk ratio. A risk ratio of 2.12 means that for every 100 people with hypertension experiencing the outcome, there would be approximately $47(100/2.12)=47$ people without hypertension experiencing the same outcome. In health terms: this finding highlights hypertension as a serious health concern. It underscores the importance of managing blood pressure to reduce the risk of these related complications. Effective

management strategies, such as lifestyle changes and medication, can help mitigate the increased risk associated with hypertension.

The results were compared with a related study was conducted at Felege Hiwot Referral Hospital using a retrospective cohort design by Abay Kassie et. al., 2019[24]. The time to death for stroke patients was significantly impacted by cardiac disease, depending on the prognostic factors. For patients without cardiac disease, the model's output shows that the prognostic factor for cardiac disease has increased by a factor of φ=2.67, which was in line with study conducted in Saudi Arabia by Alhazzani A. et. al., 2018[25]. A risk ratio (φ)=2.67, indicates that for every 100 people with cardiac disease had experienced in the stroke patients, approximately 38 people without cardiac disease had experienced in the stroke patients.

According to this study a diabetes mellitus to be a significant risk factor for stroke patients' death, and other studies have continuously corroborated with the study was conducted by Amanual G. et. al., 2019 [26]. The results of the Weibull-inverse Gaussian frailty model showed that diabetes mellitus, which increases the risk of death by a factor of φ=3.034, compared to the reference groups. A value of φ=3.04(risk ratio) means that for every 100 people with diabetes mellitus were experienced in the stroke patients, approximately 33 people without diabetes mellitus were experienced in the stroke patients. Atrial fibrillation was a significant on the time to death of stroke patients at a 5% level of significance. A risk ratio of 3.25(φ=3.25) revealed that for every 100 people with atrial fibrillation was experienced in the stroke patients, there would be approximately 31 people without atrial fibrillation was experienced in the stroke patients. The similarity study was done at Addis Ababa, Ethiopia using a retrospective study design by Ayehu K., et. al., 2020 [27].

When compared to the reference groups, the basic complication increased the risk of mortality by a factor of φ=2.98, which had a significant impact on the survival time of stroke patients at a 5% level of significance. A basic complications had φ=2.98 =risk ratio, means that for every 1000 people with basic complications were experienced in the stroke patients, that leads approximately 34 people without basic complications experiencing in the stroke patients. A related study was carried out at Mettu Karal Referal Hosipital by Dereje G. and Azmeraw G., 2022[28].

The Weibull baseline was the best fit for the stroke datasets, when compared to the exponential, log-logistic, and log-normal hazard functions. Diagnostic graphs were generated to evaluate the model's suitability. The Weibull plot of log cumulative hazard versus log time was more linear. This findings were further supported by the cumulative hazard plot for the Cox snell residuals of the log-normal, Weibull, exponential, and log-logistic models. In this case the plots were closer to the line, shows that the Weibull model performed the best, the synonyms study was conducted at New York by Klein J. and Moeschberger M., 2003[29].

Limitations: Due to the lack of proper stroke patients' data management at those hospitals, some of the vital factors were not included, such as family history, heavy alcohol consumption, physical inactivity, and smoking status.

Conclusion

This study highlights the importance of the hospital-level cluster effects in the survival analysis of stroke patients. The significant frailty effect suggests variability in outcomes across hospitals, underlining the need for tailored interventions. Clinicians and hospital administrators should consider these differences when managing stroke patients, emphasizing timely follow-up, individualized care, and resource allocation to improve the survival outcomes.

List of Abbreviations: SSA: Sub-Saharan African; WHO: World Health Organization; ST: Survival Time; PDF: Probability Distribution Function; CDF: Cumulative Distribution Function; SF: Survival Function; HF: Hazard Function; KME: Kaplan-Meier Estimator; MST: Median Survival Time; MF: Modeling Frailty; SFM: Shared Frailty Model; CHF: Cumulative Hazard Function; PHFM: Proportional Hazard Frailty Model; AFTFM: Accelerated Failure Time Frailty Model; GFD: Gamma Frailty Distribution; IGFD: Inverse Gaussian Frailty Distribution; AIC: Akaike Information Criterion; LRT: Likelihood Ratio Test.

Declarations

Ethical Approval and Consent to Participate: Ethical clearance had been obtained from Department of Statistics, Haramaya University, Ethiopia #5487, issued on October 03, 2023.

Availability of Data and Materials: This work is basically considered the Harari health office to collect stroke patient data from three hospitals such as Harar General Hospital, Jegol Hospital and Hiwot Fana Specialized University Hospital determinates were included under this study.

Acknowledgement

The authors acknowledge Harari health office was permitted to collect stroke patient data from these three hospitals.

Author's Information: Alebachew Abebe is an Assistant Professor of the department of Statistics at Haramaya University, Ethiopia; He had previous ten manuscripts and one proceeding. Kumela Ayansa (MSc) is working in the position of Statistician, Oromia regional state, Ethiopia. Kashun Takele (Dr.), PhD in Data Science at department of Statistics at Haramaya University, Ethiopia; he had more than twelve manuscripts under his research profile.

Author Contributions: KA is developed the original draft preparation, conceptualization, data collection, analysis & interpretation; AA prepared a manuscript and report writing and KT reviewed and edited the overall document.

Consent for Publication: This manuscript has not been published elsewhere and is not under consideration by another journal. Authors had approved the final manuscript and agreed with its submission to this journal. We agreed about authorship for this manuscript.

Competing Interests: The authors declare that no competing interests.

Conflict of Interest: No conflict of interests is found.

Disclosure of Potential Conflicts of Interest: Not applicable!

Informed Consent: Not applicable!

Funding: No fund for this manuscript.

References

1. Kuriakose, D., and Xiao, Z. 2020. Pathophysiology and treatment of stroke: present status and future perspectives. International journal of molecular sciences, 21 (20), 7609.
2. WHO. 2022. Global Stroke Fact Sheet 2022 Purpose : Data sources: 1-14. https://www.world-stroke.org/assets/downloads/WSO Global Stroke Fact Sheet.pdf
3. Abere Wondimu Kassie. 2022. Determinants of survival in stroke patients: application of Cox proportional hazards regression model.
4. Cloud, G., Markus, H., Pereira, A. 2017. Stroke Medicine. United Kingdom: Oxford University Press.
5. Feigin, V. L., Brainin, M., Norrving, B., Martins, S., Sacco, R. L., Hacke, W. and Lindsay, P.2022. World Stroke Organization (WSO): global stroke fact sheet 2022. International Journal of Stroke, *17*(1), 18-29.
6. Samson Getachew Erkabu,Yinager Agedie , Dereje Desta Mihretu, Akiberet Semere, and Yihun Mulugeta Alemu. 2018. Ischemic and Hemorrhagic Stroke in Bahir Dar, Ethiopia: A Retrospective Hospital-Based Study. Journal of Stroke and Cerebrovascular Diseases, 27(6), 1533-1538.
7. Destaye, Alemnew. 2021. Time to Death and Determinant Factors of Stroke Patients at Debre Berhan Comprehensive Specialized Hospital.
8. Getnet Mersha, A., Abegaz, T. M., Seid, M. A., Gebreyohannes, E. A., Bhagavathula, A. S., Gebresillassie, B. M., and Ayalew, M. B. 2019. Mortality rate and predictors of Stroke: A Meta-Analysis and Systematic Review. Research Square, 1-14.
9. Temesgen, T. G., Teshome, B., and Njogu, P. 2018. Treatment Outcomes and Associated Factors among Hospitalized Stroke Patients at Shashemene Referral Hospital, Ethiopia.
10. Kleinbaum, D. G., and Klein, M. 2012. Survival analysis a self-learning text. Springer.
11. Kaplan, E. L., and Meier, P. 1958. Nonparametric esti-

mation from incomplete observations. Journal of the American statistical association, 53(282), 457-481.

12. Collett, D. 2023. Modelling survival data in medical research. CRC press.
13. Vaupel, J. W., Manton, K. G., and Stallard, E. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16(3), 439-454.
14. Huang, X., and Wolfe, R. A. 2002. A frailty model for informative censoring. Biometrics, 58(3), 510-520.
15. Wienke, A., Arbeev, K., Locatelli, I., and Yashin, a. I. 2003. A simulation study of different correlated frailty models and estimation strategies. 49(0), 1-17.
16. Glidden, D. V., and Vittinghoff, E. 2004. Modelling clustered survival data from multicentre clinical trials. Statistics in Medicine, 23(3), 369-388.
17. Pipper, C. B., and Martinussen, T. 2004. An estimating equation for parametric shared frailty models with marginal additive hazards. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 66(1), 207-220.
18. Munda, M., Rotolo, F., and Legrand, C. 2012. Parfm: Parametric frailty models in R. Journal of Statistical Software, 51(11).
19. Jiang, X., Liu, W., and Zhang, B. 2021. A note on the prediction of frailties with misspecified shared frailty models. Journal of Statistical Computation and Simulation, 91(2), 219-241.
20. Akaike, H. 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6), 716-723.
21. Klein, J. 1992. Survival analysis: techniques for censored and truncated data. Medical College of Wisconsin.
22. George, B., Seals, S., and Aban, I. 2014. Survival analysis and regression models. Journal of nuclear cardiology : Official publication of the American Society of Nuclear Cardiology, 21(4), 686. https://doi.org/10.1007/s12350-014-9908-2
23. Akinyemi, R. O., Owolabi, M. O., Ihara, M., Damasceno, A., Ogunniyi, A., Dotchin, C., and Kalaria, R. N. 2019. Stroke, cerebrovascular diseases and vascular cognitive impairment in Africa. *Brain research bulletin*, *145*, 97-108.
24. Abay Kassie, Salie Ayalew, and Mandefro Abere. 2019. Section: Medicine Survival Time of Adult Ischemic Stroke Patients and Associated Risk Factors: A Retrospective Cohort Study at Felege Hiwot Referral Hospital. In Asian Journal of Medical Research (Vol. 8).
25. Alhazzani, A. , Mahfouz, A., Abolyazid, A. Y., Awadalla, N. J., Katramiz, K., Faraheen, A., and Aftab, R. 2018. In hospital stroke mortality: rates and determinants in Southwestern Saudi Arabia. International journal of environmental research and public health, 15(5), 927.
26. Amanual Getnet Mersha, Tadesse Melaku Abegaz, Mohammed Assen Seid, Eyob Alemayehu Gebreyohannes, Bhagavathula, A. S., Begashaw Melaku Gebresillassie, and Mohammed Biset Ayalew. 2019. Mortality rate and predictors of stroke: a meta-analysis and systematic review.
27. Ayehu Kassaw Asres, Amsale Cherie, Tadesse Bedada, and Hailemikeal Gebrekidan .2020. Frequency, nursing managements and stroke patients' outcomes among patients admitted to Tikur Anbessa specialized hospital, Addis Ababa, Ethiopia a retrospective, institution based cross-sectional study. International Journal of Africa Nursing Sciences, 13, 100228.
28. Dereje Gebeyehu Ababu, and Azmeraw Misganaw Getahun. 2022. Determinants of Stroke Mortality through Survival Models: The Case of Mettu Karl Referral Hospital, Mettu, Ethiopia.
29. Klein, J. P., and Moeschberger, M. L. 2003. Survival analysis: techniques for censored and truncated data (Vol. 1230). New York: Springer.