

RESEARCH

Open Access



Assessing survival time of heart failure patients: using Bayesian approach

Tafese Ashine^{1*}, Geremew Muleta² and Kenenisa Tadesse²

*Correspondence:

Tafese145@gmail.com

¹ Department of Statistics,
College of Natural Science,
Assosa University, Assosa,
Ethiopia

Full list of author information
is available at the end of the
article

Abstract

Heart failure is a failure of the heart to pump blood with normal efficiency and a globally growing public health issue with a high death rate all over the world, including Ethiopia. The goal of this study was to identify factors affecting the survival time of heart failure patients. To achieve the aim, 409 heart failure patients were included in the study based on data taken from medical records of patients enrolled from January 2016 to January 2019 at Jimma University Medical Center, Jimma, Ethiopia. The Kaplan Meier plots and log-rank test were used for comparison of survival functions; the Cox-PH model and the Bayesian parametric survival models were used to analyze the survival time of heart failure patients using R-software. Integrated nested Laplace approximation methods have been applied. Out of the total heart failure patients in the study, 40.1% died, and 59.9% were censored. The estimated median survival time of patients was 31 months. Using model selection criteria, the Bayesian log-normal accelerated failure time model was found to be appropriate. The results of this model show that age, chronic kidney disease, diabetes mellitus, etiology of heart failure, hypertension, anemia, smoking cigarettes, and stages of heart failure all have a significant impact on the survival time of heart failure patients. The Bayesian log-normal accelerated failure time model described the survival time of heart failure patient's data-set well. The findings of this study suggested that the age group (49 to 65 years, and greater than or equal to 65 years); etiology of heart failure (rheumatic valvular heart disease, hypertensive heart disease, and other diseases); the presence of hypertension; the presence of anemia; the presence of chronic kidney disease; smokers; diabetes mellitus (type I, and type II); and stages of heart failure (II, III, and IV) shortened their survival time of heart failure patients.

Keywords: Heart failure, Survival time, Bayesian, Survival data analysis, INLA

Introduction

Background of the study

Heart failure (HF) is defined as a clinical syndrome; specifically, failure of the heart to pump blood with normal efficiency, characterized by typical symptoms (shortness of breath, persistent coughing or wheezing, ankle swelling, and fatigue) that may be accompanied by the following signs (jugular venous pressure, pulmonary crackles, increased heart rate, and peripheral oedema) caused by a structural and functional cardiac abnormality, resulting in reduced cardiac output and elevated intracardiac pressures at rest or

during stress. In addition, HF is a syndrome and not a disease, its diagnosis relies on a clinical examination and can be challenging [31, 38].

Heart failure is a global major cause of death and is a rapidly growing public health issue affecting approximately 40 million individuals worldwide and an estimated 287,000 deaths occur a year, making it the most quickly growing cardiovascular disorder. Its ever-increasing prevalence across developed and developing countries has resulted in complications from an increasingly aging population [36]. In the United State of America, the prevalence of HF is nearly 6.5 million; approximately 960,000 new cases of HF are diagnosed each year, the incidence of HF approaching 21 per 1000 population and also an estimated 1 in 8 deaths in 2017 [8]. The prevalence of symptomatic HF is estimated at 5% of the population, and the mortality is estimated at 13% in Europe [21].

In Africa, HF has emerged as a major public health problem, imposing enormous pressure on the health care systems. The sub-Saharan Africa Survey of HF, a prospective multi-center study of HF across the continent, showed that HF is predominantly non-ischemic, most commonly hypertension; HF strikes individuals in sub-Saharan Africa at a much younger age than in the United States and Europe [14]. Similarly, HF is reported to have caused 2.5% of deaths among all age groups in a sampled hospital-based mortality in Ethiopia [29]. In this article, we have used a parametric survival models of Bayesian approach. However, most of those studies have been conducted on HF patients in hospitals based using descriptive statistics [19], and logistic regression analysis [5]. These statistical methodologies are not capable of considering the survival rate of the patients in the hospital and also multivariable logistic regression does not account for the censoring of observations, that is, it does not hold for time-to-event data.

Most medical studies have used cox regression model for assessing the survival distribution of HF patients, while alternative parametric models including exponential, Weibull, log-normal, and log-logistic models have been used to identify the prognostic factors [18, 20]. The parametric survival models could provide a more suitable description of the survival data if one can identify the distribution of the survival time [25]. The Accelerated Failure Time (AFT) models (i.e., exponential, Weibull, log-normal and log-logistic) have a more realistic interpretation and provide more informative results than the Cox Proportional Hazard (Cox-PH) model [32]. Epidemiologists have documented several risk factors for the development of HF, such as age, hypertension [34], and anemia [4], which were associated with an increased risk of mortality among HF patients.

Parametric survival models play an important role in Bayesian survival analysis since many Bayesian analyses in practice are carried out using parametric AFT models and provide computational advantages via the implementation of the Markov Chain Monte Carlo (MCMC) method. The Bayesian approach assumes that the observed data is fixed and those model parameters are random. The prior probability distributions represent a powerful mechanism for incorporating information from previous studies and for controlling confounding [22, 23]. The Bayesian methods combine objective prior knowledge with the information acquired from the data by using the Bayes theorem [17]. MCMC methods have some limitations, like the burden of time in approximating the posterior and convergence problem [9, 11]. Bayesian approach with Integrated Nested Laplace Approximation (INLA) method of estimation provides fast and accurate approximations to the posterior marginal distributions of the parameters in the model over other

methods of estimations through clever use of Laplace approximations and advanced numerical methods taking computational advantage of sparse matrices [33].

Thus, considering the advantages of Bayesian application is the key for the motivation to apply it for the HF data-set under this article. So, we chose the Bayesian parametric survival models using the INLA method to analyze HF data-set because HF is a growing problem in countries with hospital-based and gaps found in different studies. Therefore, this study aims to answer the basic questions on:-which factors significantly affect the survival time of heart failure patients, what is the estimated survival time of heart failure patients, and which parametric survival models are the most appropriate for analyzing the heart failure data set.

The main aim of this study was to assess the survival time of heart failure patients using Bayesian approach in Jimma University Medical Center, Jimma, Ethiopia. It seeks to identify prognostic factors in HF patients, determine the best parametric survival models for a HF data-set, estimate HF patient survival time, and investigate the Bayesian accelerated failure time models using the INLA method.

Significance of the study

Studying the survival time of HF patients is a mechanism of overcoming the problem of health in society by identifying factors associated with death. On top of this, the result of this study might be used to improve awareness of the factors that trigger the death of HF patients. It also enables us to provide scientific information about the finding to the ministry of health in Ethiopia that helps policymakers to enhance the awareness of society about factors that increase the probability of death due to HF, which is protect-able and curable if it is screened and treated in its earlier stage with appropriate treatment.

Methodology

Data description

Study area

The study has been conducted on data taken from Jimma University Medical Center (JUMC) which is located in Oromia National Regional State, Jimma town 350 km southwest of Addis Ababa, Ethiopia. JUMC is the only medical center in the Jimma zone serving the majority of people living in Jimma city and its surrounding areas.

Study design and population

A retrospective study has been applied to obtain data on HF patients were recorded in JUMC. The population of this study was all HF patients who had been registered at JUMC for 3 years starting from first January 2016 up to first January 2019. The data has been carefully reviewed from the registration log book and patient's registration card; any inadequate information encountered was checked from the file and excluded from analysis if proven to be inadequate. Thus, the data has been collected from patient follow-up records based on the variables in the study.

Inclusion criteria: All persons registered with full information including study variables of interest in the registration book or the chart were considered to be eligible for the study. The patients were to be included in the study they must take treatment at least one time from the hospital.

Exclusion criteria: The patients with insufficient information regarding study variables on the registration book or the card were not eligible. Thus, the HF patients lost from the study without starting any treatment were not included.

Data collection methods

Ethical permission has been obtained from the Jimma University Medical Center, Jimma, Ethiopia. Then secondary data was taken based on data existing in the hospital by a trained enumerator and the principal investigator using a checklist (data extraction form).

Starting time: the start time of the interval (in months). Time origin or the beginning of the study, the entry of the survival data would be considered from the day that the heart failure patients start diagnosis; when the patient first received the treatment.

Ending time: the time (in months) at which the event occurred, when the heart failure patients died or were lost to follow-up at first January 2019 (at the end of study). This means that the type of survival data is right-censored.

Variables in the study

The response variable was survival time of HF patients (in months), which defined as the difference between the time of diagnosis and time to one of the events "death", "lost to follow up", "dropped out", "stopped", "transferred out to other health centers or hospitals" occurred. Death was considered to be an event of interest. The status variable was coded as 0 for censored and 1 for death. The factors considered for their influence on survival time of HF patients were as follows: sex, age, smoking cigarettes, diabetes mellitus, hypertension, residence, alcohol consumption, history of HF, anemia, treatments taken, etiology of HF, chronic kidney disease, and stages of HF.

Method of data analysis

Descriptive statistics

The description of survival data utilizes non-parametric methods to compare the survival functions of two or more groups. The use of Laplace or lapse needs to be consistent and Kaplan Meier plot (s) would be employed for this purpose [24]. The frequency distribution table was used to summarize the data obtained from the registration book of patients based on the study variables in Jimma University Medical Center.

Survival data analysis

Survival data are censored in the sense that they did not provide complete information since subjects of the study may not have experienced the event of interest [1].

Survival analysis is well suited for heart failure data-set which are very common in medical research since studies in medical areas have a special feature that follow-up studies could start at a certain observation time and could end before all experimental units had experienced an event.

Right censoring occurs to the right of the last known survival time and the observation of the patient is terminated before the event occurs. This type of censoring is commonly recognized in survival analysis and also considered in this study [26].

Comparison of survival function

The Kaplan–Meier plots are used to determine whether or not there is a difference in survival times between groups of covariates under consideration. But, the plot cannot be used to decide whether the survival time of heart failure patients in each covariate is different or not, and the log-rank test was used for this purpose [27]. The hypotheses to be tested are:

H_0 : There is no difference between the survival curves.

H_1 : There is a difference between the survival curves.

Bayesian survival analysis

The Bayesian approach is preferred over the frequentist approach in survival analysis in that the power of information obtained from the approach is much better as it is the combination of likelihood data and prior information about the distribution of the parameter. The Bayesian approach is more useful in clinical data analysis than the frequentist approach and is a more appropriate data analysis technique for clinical researchers [10]. The main reasons why one might choose to use Bayesian statistics: some complex models simply cannot be estimated using conventional statistics, one might prefer the definition of probability, background knowledge can be incorporated into the analysis, and Bayesian statistics are not based on large samples (i.e., the central limit theorem) and hence large samples are not required to make the math work. Furthermore, Bayesian statistics allow for the incorporation of uncertainty about a parameter and the updating of this knowledge via the prior distribution [15].

The Bayesian approach considers the parameters of the model as random variables and requires that prior distributions be specified for them and data are considered as fixed. The Bayesian approach is well known that survival models are notoriously difficult to fit, particularly in the presence of complex censoring schemes. With the use of the Gibbs sampler and other (Markov Chain Monte Carlo) MCMC techniques, fitting complex survival models are fairly straightforward, and the availability of software eases the implementation greatly [22]. MCMC methods have some limitations, like the burden of time in approximating the posterior and convergence problem [9, 11]. As of 2009, the other news was welcomed with very flexible and fast approximation techniques called Integrated Nested Laplace Approximation (INLA). The Bayesian approach with the INLA method is focused on providing a good approximation to the posterior marginal distributions of the parameters in the model [33].

Prior Distribution $\pi(\theta)$, Its probability distribution is used to express uncertainty about unknown quantities parameter before the data are taken into account. It is prior distribution, which is a probability distribution that represents the prior information associated with the parameter of interest [22].

Likelihood $L(\theta/\text{data})$, It is a likelihood function, which is a function that gives the probability of observing the sample data given the current parameters. For a set of unknown parameters in the presence of right censoring it can be written as:

$$L(\theta/\text{data}) = \prod_{i=1}^n \left[f(t_i/x_i; \theta)^{\sigma_i} * S(t_i/x_i; \theta)^{1-\sigma_i} \right]$$

where σ_i is censoring indicator (0 = censored and 1 = death) and $f(t_i/x_i; \theta)$ & $S(t_i/x_i; \theta)$ are the probability density and survival distributions respectively [16].

Posterior distribution, It is a combination of the prior distribution and likelihood using the Bayes rule, a likelihood which includes information about model parameters based on the observed data, and a prior, which includes prior information (before observing the data) about model parameters. It is obtained by multiplying the prior distribution over all parameters, θ , by the full likelihood function, $L(\theta/\text{data})$ [13]. Given by

$$\text{Posterior} = \frac{\text{Likelihood} * \text{prior}}{\int \text{Likelihood} * \text{prior} d\theta}$$

Assuming that θ is a random variable and has a prior distribution denoted by $\pi(\theta)$, then posterior distribution, $\pi(\theta/X)$, of θ is given by:

$$\pi(\theta/X) = \frac{L(X/\theta) * \pi(\theta)}{\int L(X/\theta) * \pi(\theta) d\theta}$$

It is clear that $\pi(\theta/X)$ is proportional to the likelihood multiply by the prior, $\pi(\theta/X) \sim L(X/\theta) * \pi(\theta)$ and thus it involves a contribution from the observed data through $L(X/\theta)$ and contribution from prior information quantified through $\pi(\theta)$. The quantity $m(x) = \int L(X/\theta) * \pi(\theta) d\theta$ is the normalizing constant of $\pi(\theta/X)$, and is often called the marginal distribution of the data or the prior predictive distribution. Parametric survival models play an important role in Bayesian survival analysis since many Bayesian analyses in practice are carried out using parametric survival models (Exponential, Weibull, Log-Normal, and Log-Logistic). Parametric modeling offers straightforward modeling and analysis techniques [22].

Integrated Nested Laplace Approximation method

The Integrated Nested Laplace Approximation (INLA) method was used to estimate the parameters in the Bayesian parametric survival models. Survival analysis consists of a great body of work using latent Gaussian models. According to [33], INLA computes posterior marginal for each component in the model and it is from these that the posterior expectations and standard deviations can be found. The survival models can be expressed as a latent Gaussian model on which the integrated nested Laplace approximations can be applied. In addition, INLA provides both extremely fast and very accurate approximations to the posterior marginal through clever use of Laplace approximations and advanced numerical methods and it can be adapted to fit survival models [6]. An R package called R-INLA works as an interface for INLA and it is used just like the other R functions. The INLA programme and the R package for INLA are freely available from (<http://www.r-inla.org>).

Bayesian model selection criterion

For Bayesian models, we might prefer the Deviance Information Criteria (DIC) was used for Bayesian parametric survival models comparison. The preferable model is the one with the lowest value of the DIC [35]. An alternative is the Watanabe Akaike information criterion (WAIC) [37], which follows a more fully Bayesian approach to construct a criterion [17]. Claims the WAIC is preferable to the DIC.

Bayesian model diagnostics

The most common ways of checking the goodness of fit are the Bayesian Cox-Snell residual plot and Predictive Distribution. Model-checking and adequacy play an important role in models for survival data. In Bayesian analysis [12], defined the Bayesian version of the residuals.

Statistical software

The secondary data were entered into SPSS software version 21 and then exported to R software version 3.6.1 for analysis.

Results and discussions

Descriptive summaries

The data for this study was collected from 409 patients who received treatments for HF at least once at Jimma University Medical Center in Jimma, Ethiopia, between first January 2016, and first January 2019. The minimum and maximum event time observed from HF patient's follow-up were 6 and 36 months respectively. Among those HF patients, about 59.90% were censored (right censored), and the remaining 40.10% have died. Fifty percent of HF patients who received treatment survived 31 months or above (Table 4).

Almost half, 52.81%, of the HF patients were female and the remaining were male during the follow-up study. However, the survival time of male patients seems lower. The majority of HF patients, approximately 64.79%, live in rural areas, with the remainder living in urban areas. The survival time of HF patients seems less as they get older. About 20.05%, 22.25%, 23.72%, 25.43%, and 8.55% of HF patients were ischemic heart disease, rheumatic valvular heart diseases, cardiomyopathy heart disease, hypertensive heart disease, and other diseases respectively.

By observing the smoking status of HF patients, most HF patients were 74.82% non-smokers and the death proportion seems highest for those HF patients who were smokers, which was 54.88% compared to non-smokers which were 45.12%. About 64.55% of HF patients were not alcohol users and 35.45% were alcohol users.

Moreover, about 19.08% of HF patients are treated in the hospital with a combination of two or more treatments, and 19.32% of HF patients take digoxin. In addition, the remaining 24.2%, 25.18%, and 11.49% of HF patients were treated with spironolactone, atorvastatin, and other treatments respectively. About 58.19%, 13.69%, and 28.12% of HF patients were non-diabetic, type I diabetes mellitus, and type II diabetes mellitus respectively.

By observing the chronic kidney disease of HF patients, about 30.32% and 69.68% were HF patients with chronic kidney disease and without chronic kidney disease respectively, in which HF patients with chronic kidney disease seem to have lower survival time. Most HF patients have no hypertension, 60.64%, and the remaining have hypertension.

Looking at the stage at which the HF patients go to the hospital for treatment, about 36.92%, 28.61%, 19.07%, and 15.4% were in stage IV, in stage III, in stage II, and in stage I respectively. Most of, about 54.87% death, HF patients go for treatment in the hospital at a later stage and their survival time seems low at this stage (Table 5).

The Kaplan Meier estimate for some covariate

Figure 1(a) below, the overall survival rate at the end of the first year was almost 93.1%, and the overall survival rate at the end of 34 months in this study was 31%, 95% confidence interval was (23.9%, 40.2%).

Figure 1(b) below, indicated that HF patients whose age was below 49 years were at a higher probability of surviving than patients whose age was 49 to 65 years and also patients whose age was greater than or equal to 65 years. The probability of surviving becomes less for patients whose age was greater than or equal to 65 years.

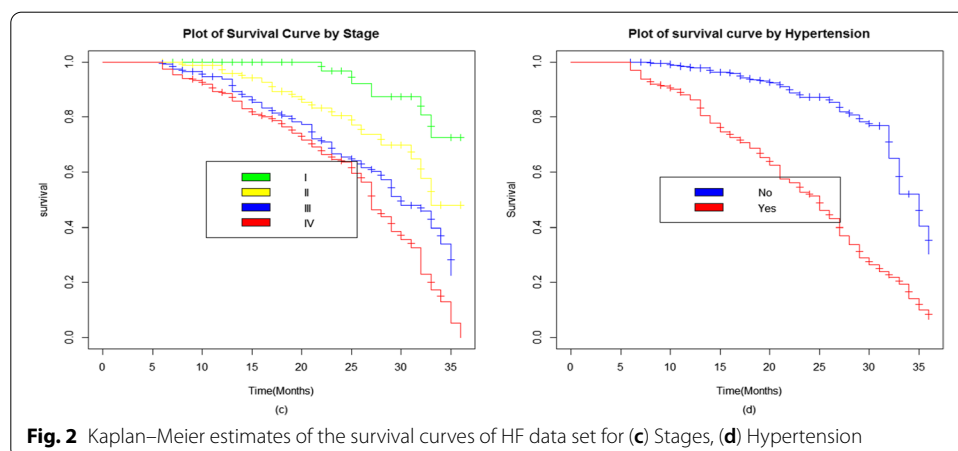
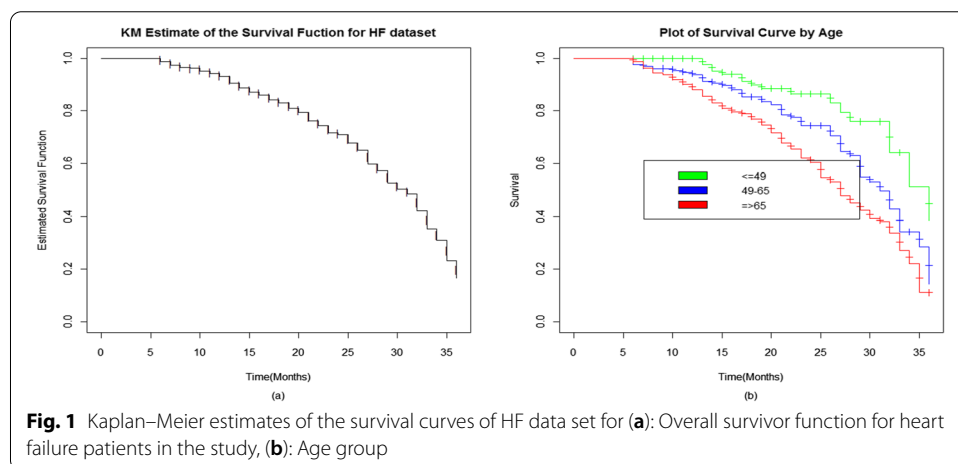


Figure 2(c) below, shows that HF patients with stage I had a higher chance of surviving than other stages. The survival curve for patients with stage II was above the survival curve of those patients with stage III and stage IV. The probability of surviving becomes less for HF patients with stage IV.

Figure 2(d) shows that HF patients without hypertension had a better chance of survival than HF patients with hypertension.

Checking the assumption of Cox-PH

As shown in Table 1, the p-values for alcohol consumption, chronic kidney disease, and anemia are less than the common (5%) level of significance using the correlation test (rho), indicating that the Cox-PH model assumption was not valid for the HF data set. Furthermore, by looking for a global test, the Cox-PH assumption fails because the test result was significant.

Bayesian survival analysis

As it can be shown in Table 1, the assumption of the Cox-PH model was not valid for the HF data set; in this case, parametric AFT models were used for the HF data set. For the HF data set, the time t_i where $i = 1, 2, \dots, 409$ of HF patients. Given that $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ is the vector of coefficients of the covariates considered for analysis β_0 is the intercept and p the number of covariates ($p = 13$), we assume that all these coefficients have a normal prior with mean 0 and variance of 1000. We assume that a gamma prior was applied to the Weibull, Log-normal, and Log-logistic distributions with shape parameter 1 and inverse scale parameter 0.001 [6]. Table 2 below, shows the analysis of the HF data set for model comparison using the INLA method. To compare the efficiency of these various models, DIC and WAIC were used, and the one with the smallest value and the best fit was chosen. Accordingly, the Bayesian lognormal AFT model (DIC = 1297.84; WAIC = 1297.47) was found to be the best for survival time of HF patients data-set from the given alternative because the bold values are smallest.

Table 3 shows the final results for the Bayesian log-normal AFT model, and as this result shows, the survival time of HF patients is statistically significantly affected by age, chronic kidney disease, diabetes mellitus, etiology of HF, hypertension, anemia, smoking, and stages of HF.

Table 1 Shows test of assumption in Cox-PH model

Covariates	Rho	Chi-square	P-value
Age	0.0125	0.0269	0.86981
Sex	-0.0349	0.2344	0.62829
History of HF	0.0367	0.2201	0.63896
Alcohol consumption	0.1639	4.5498	0.03292
Hypertension	-0.0639	0.7364	0.39081
Chronic kidney disease	-0.1970	7.9647	0.00477
Etiology of HF	-0.1205	2.4705	0.11600
Smoking cigarette	0.0907	1.6229	0.20269
Diabetes mellitus	-0.0179	0.0573	0.81077
Anemia	0.2030	8.2169	0.00415
GLOBAL TEST	NA	23.4782	0.00911

Table 2 The comparisons of Bayesian AFT model using INLA methods

Distributions	Pd	DIC	WAIC
Exponential	12.24	1400.62	1522.88
Log-Normal	17.06	1297.84	1297.47
Weibull	11.63	1389.20	1383.43
Log-logistic	7.59	1326.88	1326.39

Bold values indicate better results than other filtering methods

Source: JUMC, Jimma, Ethiopia; from first January 2016 to first January 2019

From Table 3, the final model was interpreted using acceleration factor, 95% credible interval of Bayesian accelerated failure time estimated values. The estimated acceleration factor is defined as $\gamma = \left[\exp(\hat{\beta}) \right] = [\exp(\text{posterior mean})]$.

Under the Bayesian log-normal AFT model, keeping the effect of other factors constant, the estimated acceleration factor for the age group of HF patients were 49 to 65 and greater than or equal to 65 years old is estimated to be 0.7726 with [95% CrI 0.6138, 0.9646] and 0.7146 with [95% CrI 0.5729, 0.9875] respectively. Thus,

Table 3 Indicating the results for Bayesian log-normal AFT model using INLA method

Covariates	Categories	Pmean	Sd	Median	Credible Interval	Mode	KId
Age	Intercept	4.953	0.221	4.945	[4.541, 5.409]*	4.929	0
	≤ 49–65	Ref					
	49–65	−0.258	0.115	−0.256	[−0.488, −0.036]*	−0.253	0
	≥ 65	−0.336	0.110	−0.335	[−0.557, −0.125]*	−0.331	0
Hypertension	No	Ref					
	Yes	−0.301	0.076	−0.300	[−0.452, −0.153]*	−0.299	0
CKD	No	Ref					
	Yes	−0.389	0.075	−0.388	[−0.537, −0.244]*	−0.387	0
Etiology of Heart failure	IHD	Ref					
	RVHD	−0.302	0.116	−0.302	[−0.533, −0.076]*	−0.300	0
	Cardiomyopathy	−0.158	0.113	−0.158	[−0.382, 0.063]	−0.157	0
	HDD	−0.258	0.115	−0.257	[−0.486, −0.035]*	−0.255	0
	Others	−0.381	0.160	−0.381	[−0.693, −0.066]*	−0.382	0
Smoking cigarette	No	Ref					
	Yes	−0.156	0.073	−0.155	[−0.300, −0.014]*	−0.154	0
Stages of Heart failure	I	Ref					
	II	−0.400	0.190	−0.397	[−0.782, −0.038]*	−0.389	0
	III	−0.423	0.176	−0.419	[−0.781, −0.090]*	−0.410	0
	IV	−0.506	0.173	−0.501	[−0.857, −0.180]*	−0.492	0
Diabetes mellitus	Not	Ref					
	Type I	−0.232	0.10	0.231	[−0.431, −0.036]*	−0.230	0
	Type II	−0.422	0.086	−0.421	[−0.593, −0.255]*	−0.419	0
Anemia	No	Ref					
	Yes	−0.154	0.072	−0.153	[−0.298, −0.013]*	−0.152	0
Tau parameter	For log-normal	4.30	0.497	4.28	[3.38, 5.33]*	4.24	−

Source: JUMC, Jimma, Ethiopia; from first January 2016 to first January 2019

* Is indicated statistically significant. Pmean, Posterior Mean; CKD, Chronic kidney disease; IHD, Ischemic heart disease; RVHD, Rheumatic valvular heart disease; HDD, Hypertensive heart disease, Sd, Standard deviation, KLD, Kullback–Leibler Divergence

the expected survival time of HF patients decreased by 22.74% and 28.54% for HF patients aged group 49 to 65 and 65 or above 65 years older respectively as compared to HF patients aged group 49 or below 49 years (Reference).

The 95% CrI for acceleration factor for both age groups did not include one, implying that both age groups have a significant effect on HF patients' survival time. Looking for chronic kidney disease and controlling for other factors, the estimated acceleration factor of HF patients with chronic kidney disease is 0.6777 with [95% CrI 0.5844, 0.7835], implying that the expected survival time decreases by 32.23% compared to HF patients without chronic kidney disease. The 95% CrI for acceleration factor of HF patients with chronic kidney disease did not include one, implying that HF patients with chronic kidney disease have a significant effect on HF patients' survival time.

By observing hypertension, keeping the effect of other factors constant, the estimated acceleration factor for HF patients with hypertension is estimated to be 0.74 with [95% CrI 0.5844, 0.7834] in which the expected survival time is a 26% decrease as compared to HF patients without hypertension (Reference). The 95% credible interval for acceleration factor of HF patients with hypertension did not include one which implies that HF patients with hypertension have a significant (in the Bayesian sense) effect on the survival time of HF patients.

On the other hand, keeping the effect of other factors constant, the estimated acceleration factor for HF patients who were smoking cigarettes is estimated to be 0.8555 with [95% CrI 0.7408, 0.986]. The 95% credible interval for the acceleration factor of HF patients who were smoking cigarettes did not include one. Thus, HF patients who smoked cigarettes had a significant effect on patient survival time, and the expected survival time of HF patients who smoked cigarettes was 14.45% shorter than that of HF patients who did not smoke cigarettes.

Regarding the etiologies of HF, keeping the effect of other factors constant, the estimated acceleration factor for etiologies of HF were rheumatic valvular heart disease, hypertensive heart disease, and other heart diseases are estimated to be 0.7393 with [95% CrI 0.5868, 0.9268], 0.772 with [95% CrI 0.615, 0.965] and 0.683 with [95% CrI 0.5, 0.936] respectively. Thus, the expected survival time of HF patients decreased by 27.07% were rheumatic valvular heart disease, 22.8% were hypertensive heart disease and 31.7% were other heart diseases as compared to ischemic heart disease of HF patients. The 95% CrI for acceleration factor of HF patients for etiology of HF were rheumatic valvular heart disease, hypertensive heart disease, and other heart disease did not include one, implying that HF patient with rheumatic valvular heart disease, hypertensive heart disease, and other heart disease has a significant effect on HF patients' survival time, whereas cardiomyopathy does not affect.

Moreover, for diabetes mellitus, keeping the effect of other factors constant, the estimated acceleration factor for HF patients with type I diabetic and type II diabetic is estimated to be 0.793 with [95% CrI 0.649, 0.964] and 0.655 with [95% CrI 0.552, 0.774] respectively. Thus, the expected survival time of HF patients decreases by

20.7% for type I diabetics and 34.5% for type II diabetics as compared to HF patient's non-diabetic (Reference). The 95% credible interval for acceleration factor of HF patients with both types of diabetes did not include one, implying that HF patients with both types of diabetes have a significant effect on HF patients' survival time.

Looking for anemia while controlling for other factors, the estimated acceleration factor of HF patients with anemia is 0.857 with [95% CrI 0.742, 0.987], implying that the expected survival time decreases by 14.3% compared to HF patients without anemia. The 95% CrI for acceleration factor in HF patients with anemia did not include one, implying that HF patients with anemia have a significant effect on HF patients' survival time.

Finally, observing stages of HF, keeping the effect of other factors constant, the estimated acceleration factor for stage II, III and IV of HF patients is estimated to be 0.67 with [95% CrI 0.457, 0.962], 0.655 with [95% CrI 0.457, 0.913] and 0.602 with [95% CrI 0.424, 0.835] respectively. Thus, the expected survival time of HF patients decreases by 33%, 34.5%, and 39.8% for stage II, III, and IV of HF patients respectively as compared to stage I. The 95% credible interval for acceleration factor of HF patients with stage II, III, and IV did not include one, indicating that stage II, III, and IV have a significant effect on heart failure patients' survival time.

From Table 3, the Kullback–Leibler divergence values for all significant parameters in the Bayesian log-normal AFT model were 0, and thus, small values indicate that the posterior distribution was well approximated by a normal distribution. The most efficient algorithm was a simplified Laplace approximation, which improved efficiency and resulted in faster computation speed.

Bayesian model diagnostics

By observing the Bayesian cox-snell residual plots figure below, shows that the Bayesian log-normal AFT model best fit HF data-set among the five models, since the plot of Cox-Snell residuals against cumulative hazard function of residuals was approximately a straight line with slope one and Bayesian cox-snell residual plot for Bayesian log-normal AFT model was nearest to the line through the origin. In addition, the plot also indicated that the Bayesian log-normal model describes the HF data-set well.

The histograms of the cross-validated probability integral transform values show that the posterior predictive p-values are to some extent closer to uniformly distributed with some observations outliers in the HF data-set. The Conditional predictive ordinate values are significantly smaller (order of magnitude smaller) than the others, so the observed values would be considered surprising concerning the Bayesian log-normal model because the sum of the observations associated with failure flags is equal to zero in the HF data-set. The plots include a 95% credibility interval by observing the posterior density for the parameters that were normally distributed in the HF data-set. The kullback–leibler divergence (kld) is a diagnostic that measures the accuracy of the INLA approximation since Table 3 shows that the kld values for all significant parameters in the Bayesian log-normal AFT model were 0.

Discussions

The main aim of this study was to assess the determinant of survival time of the HF data set, which was obtained from Jimma University Medical Center. Heart failure is a growing problem in the world and the overall prevalence of HF in the adult population in developing countries is 7%-10% with an exponential rise with age [3]. The descriptive results of the study indicated that a total of 409 HF patients were included in this study. The minimum and maximum event time observed from HF patient's follow-up were 6 and 36 months respectively.

In addition, fifty percent of HF patients who received treatment survived 31 months or above. In this study, among those HF patients, about 59.90% were censored (right censored) and the remaining 40.10% died. This finding was similar to a study conducted by [20] which found that 31.3% of the HF patients were dead, while the remaining 68.7% were censored.

In the study, the main etiologies of HF were ischemic heart disease 20.05%, rheumatic valvular heart disease 22.25%, cardiomyopathy heart disease 23.72%, hypertensive heart disease 25.43%, and the remaining were 8.55%. This finding was consistent with a study conducted by [2], which found that the main causes of HF were ischemic heart disease (15.8%), rheumatic valvular heart disease (40.1%), cardiomyopathy (12.5%), hypertensive heart disease (16.0%), and other causes constituted the majority of all admissions due to HF.

The Bayesian parametric survival models were applied to this data set. But, the assumption of the Cox-PH model was violated. The Bayesian approach was applied to parametric AFT models and to compare the efficiency of different AFT models, DIC and WAIC were used [35, 37]. The Bayesian log-normal AFT model was the best model to describe the HF data set from the given alternative. Similar results were obtained from a previous study [7].

However, the results of the Bayesian log-normal AFT model using the INLA method in this study show that age, chronic kidney disease, diabetes mellitus, etiology of HF, hypertension, anemia, smoking cigarettes, and stage of HF all have a significant effect on the survival time of HF patients.

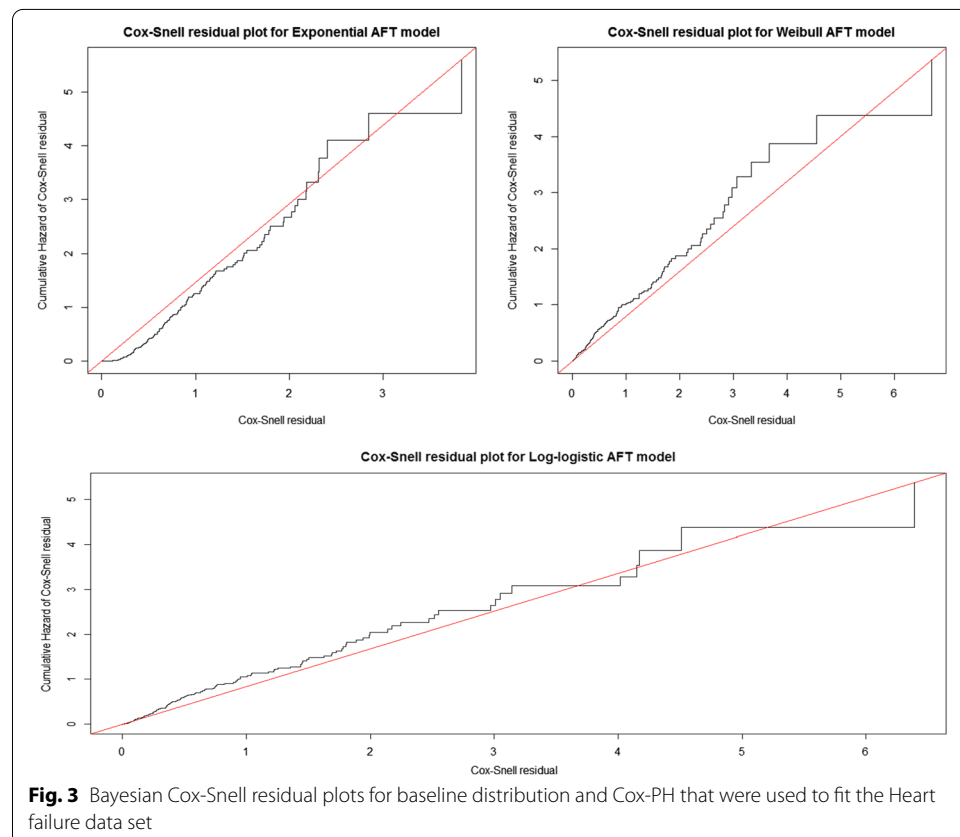
According to the findings of this study, age has a significant impact on the survival time of HF patients. Furthermore, the survival time of HF patients seems less as they get older (greater than or equal to 65 years) and different studies also persisted with these results, [3, 34], and [39]. The survival time of HF patients without hypertension was higher than that of HF patients with hypertension, indicating that hypertension had a significant effect on HF patients. The studies were done by [4], and [34] show the same results.

On the other hand, the survival time of smoker HF patients decreased as compared to nonsmokers, which is similar to the study done by [4]. The survival time of HF patients was significantly affected by both types of diabetes mellitus, and the expected survival time of HF with both types of diabetes mellitus was shorter than that of HF patients

without diabetes. These results are consistent with studies done by [4, 30, 39]. Furthermore, chronic kidney disease had a significant impact on the survival time of HF patients, and the survival time was higher for HF patients who did not have chronic kidney disease compared to HF patients who did have chronic kidney disease. This result was confirmed with the study done by [39].

According to the findings of [4] and [39], anemia has a significant impact on the survival time of HF patients, and the expected survival time of HF with anemia was shorter than that of HF patients who were not anemic. This study was consistent with the current study. The stages of HF patients have a significant impact on their survival time. The study done by [39], shows that the stages of HF patients have a significant impact on their survival time. From the results of this study, the survival time of HF patients decreased as the stage increased.

To check the adequacy of the model, the cumulative hazard plots for the Bayesian Cox Snell residuals of the Cox-PH, Exponential, Weibull, Log-normal, and the Log-logistic models were plotted as in Fig. 3. The plots were more approached to the line in the case of the Bayesian log-normal model that indicates the Bayesian log-normal was best in the HF data-set which is similar to the previous study conducted by [7]. The conditional predictive ordinates and probability integral transforms were also used to check the model in this study. Before adequacy checking using graphical methods, it can be important to check whether the usual numerical problem occurred during the computation of conditional predictive ordinate. Thus, since the sum of the number of failures in conditional



predictive ordinate was zero, no failure was detected, meaning that no numerical problem had occurred in the HF data-set. The histogram and scatter plot of probability integral transform indicated that the plots of predictive residual-based values were to some extent uniformly distributed with some deviated outlier and there is a reasonable predictive distribution that matches the actual data. This finding was supported by other studies conducted by [6] and [28].

The Bayesian log-normal AFT model diagnostic plots, including a 95% credibility interval, shows that the plot of posterior density for the parameters was normally distributed. Similarly, the kullback–leibler divergence is a diagnostic that measures the accuracy of the INLA approximation. In this study, the values of kld for all significant parameters in the Bayesian log-normal AFT model were 0. This indicates that the Bayesian log-normal AFT model was faster and has higher accuracy. These results are also confirmed by other studies done by [28], and [6]. However, this study was not done without limitations. The study was conducted based on secondary data gathered from the registration log book and patients' registration cards, which might have incomplete and biased information. As different literature pointed out, there are different prognostic factors (Body mass index and weight) that are assumed to have an impact on the survival time of HF. However, data on those variables could not be available in hospital records.

Conclusions

This study used the survival time of heart failure patient's data set, for those patients who were receiving treatment for at least one time at Jimma University Medical Center. The Bayesian lognormal AFT model performed better than various parametric models with baseline distribution (Exponential, Weibull, Log-logistic, and Log-normal) for this study. Fifty percent of heart failure patients who received treatment survived 31 months or above. The survival time of HF patients has been significantly affected by age, chronic kidney disease, diabetes mellitus, etiology of heart failure, hypertension, anemia, smoking cigarettes, and stages of heart failure. The findings of this study suggested that the age group (49 to 65 years, and greater than or equal to 65 years); etiology of heart failure (rheumatic valvular heart disease, hypertensive heart disease, and other diseases); the presence of hypertension; the presence of anemic; the presence of chronic kidney disease; smokers; diabetes mellitus (type I, and type II diabetic); and stages of heart failure (II, III, and IV) shortened their survival time of heart failure patients.

Appendix

See Tables 4, 5.

Table 4 Descriptive summaries of patient's for HF data set

Patients status				
Covariates	Categories	Number of censored (%)	Number of death (%)	Total
Sex	Female	148 (60.41)	68 (41.46)	216 (52.81)
	Male	97 (39.59)	96 (58.54)	193 (47.19)
Age	≤ 49	98 (40.0)	17 (10.36)	115 (28.12)
	49–65	78 (31.84)	53 (32.32)	131 (32.03)
	≥ 65	69 (28.16)	94 (57.32)	163 (39.85)
Alcohol	No	172 (70.20)	92 (56.10)	264 (64.55)
	Yes	73 (29.80)	72 (43.90)	145 (35.45)
Place of residence	Urban	97 (39.59)	47 (28.66)	144 (35.21)
	Rural	148 (60.41)	117 (71.34)	265 (64.79)
History of heart failure	New	74 (30.20)	46 (28.05)	120 (29.34)
	HF patient before	84 (34.29)	79 (48.17)	163 (39.85)
	Medical OPD	87 (35.51)	39 (23.78)	126 (30.81)
Chronic kidney disease	No	224 (91.43)	61 (37.20)	285 (69.68)
	Yes	21 (8.57)	103 (62.80)	124 (30.32)
Hypertension	No	203 (82.86)	45 (27.44)	248(60.64)
	Yes	42 (17.14)	119 (72.56)	161 (39.36)
Anemia	No	205 (83.67)	62 (37.80)	267 (65.28)
	Yes	40 (16.33)	102 (62.20)	142 (34.72)
Diabetes mellitus	Not	204 (83.26)	34 (20.73)	238 (58.19)
	Type I	18 (7.35)	38 (23.17)	56 (13.69)
	Type II	23 (9.39)	92 (57.0)	115 (28.12)
Etiology of heart failure	IHD	62 (25.31)	20 (12.19)	82 (20.05)
	RVHD	47 (19.18)	44 (26.83)	91 (22.25)
	Cardiomyopathy	55 (22.45)	42 (25.61)	97 (23.72)
	HHD	55 (22.45)	49 (29.88)	104 (25.43)
	Others	26 (10.61)	9 (5.49)	35 (8.55)
Smoking cigarette	No	232 (94.69)	74 (45.12)	306 (74.82)
	Yes	13 (5.31)	90 (54.88)	103 (25.18)
Treatments taken	Digoxin	48 (19.59)	31 (18.90)	79 (19.32)
	Spironolactone	65 (26.53)	34 (20.73)	99 (24.20)
	Atorvastatin	62 (25.31)	41 (25.0)	103 (25.18)
	Others	27 (11.02)	20 (12.20)	47 (11.49)
	Combination ≥ 2	43 (17.55)	38 (23.17)	81 (19.80)
Stages of heart failure	I	58 (23.67)	5 (3.05)	63 (15.40)
	II	61 (24.90)	17 (10.37)	78 (19.07)
	III	65 (26.53)	52 (31.71)	117 (28.61)
	IV	61 (24.90)	90 (54.87)	151 (36.92)

IHD, Ischemic heart disease; RVHD, Rheumatic valvular heart disease; and HHD, Hypertensive heart disease. Source: JUMC, Jimma, Ethiopia; from first January 2016 to first January 2019

Table 5 Counts for patient status

Patients status	Number of patients (%)
Censoring	245 (59.90)
Death	164 (40.10)

Abbreviations

HF: Heart Failure; AFT: Accelerated Failure Time; Cox PH: Cox Proportional Hazard; MCMC: Markov chain Monte Carlo; INLA: Integrated Nested Laplace Approximation; JUMC: Jimma University Medical Center; DIC: Deviance Information Criteria; WAIC: Watanabe Akaike Information Criterion; KLD: Kullback–Leibler Divergence; OPD: Out-Patient Department.

Acknowledgements

We thank Jimma University and the Ethiopian ministry of science and higher education for financing this study and covering all expenses of the paper. We sincerely thank Dr. Birhanu Teshome (Ph.D.) for his advice and guidance. Finally, thanks to all the staff members of Jimma University medical center for allowing us to collect the data set from the hospital. I am also very thankful to all my beloved parents and friends whose presence and support have always been important to me.

Authors' contributions

This study was designed and compiled by TA (MSc in Biostatistics) as the principal investigator. The development of the basic research questions, identifying the problems and selecting appropriate statistical models, data collection, data analysis, interpretation, and critical review of the paper have been done by him. The edition of the overall progress of the work was supported by GM, KT. All authors read and approved the final manuscript.

Funding

The study was fully funded by Jimma University and the Ministry of science and higher education of Ethiopia. The collaborative fund obtained from the two institutions covered all costs associated with data access and related activities.

Availability of data and materials

The data set was taken from Jimma University Medical Center, Jimma, Ethiopia, from first January 2016 to first January 2019. There are no restrictions on the availability of data and R-code, and the investigators are willing to provide the code as well.

Declarations

Ethics approval and consent to participate

The Research Ethics Review Board of Jimma University would provide an ethical clearance for the study. The data has been collected after written permeation was obtained from Jimma University Medical Center and department of statistics write an official co-operation letter to the Hospital for the permeation. The data has been carefully reviewed from the registration log book and patients registration card. Confidentiality of any information related to the patients and their clinical history has been maintained by keeping both the hardcopy and softcopy of every collected data in a locked cabinet and password secured computer. Only the researcher would access to the de-identified data that has been kept in a secure place. All data has been coded with numbers and without personal identifiers. All analysis has been on de-identified and coded data. During the study, there is no contact between the patients and the researcher. The study is non-invasive and without any harm to the patients. Then, the data obtained from the hospital has been secured.

Consent for publication

Not applicable.

Competing interests

The authors declare that they do have no competing interests available.

Author details

¹Department of Statistics, College of Natural Science, Assosa University, Assosa, Ethiopia. ²Department of Statistics, College of Natural Science, Jimma University, Jimma, Ethiopia.

Received: 20 March 2021 Accepted: 6 November 2021

Published online: 11 December 2021

References

1. Aalen O, Borgan O, Gjessing H. Survival and event history analysis: a process point of view. Berlin: Springer Science & Business Media; 2008.
2. Abebe TB, Gebreyohannes EA, Tefera YG, Abegaz TM. Patients with hfpwf and hfref have different clinical characteristics but similar prognosis: a retrospective cohort study. *BMC Cardiovasc Disord*. 2016;16(1):232.
3. Adebayo SO, Olunuga TO, Durodola A, Ogah OS, et al. Heart failure: definition, classification, and pathophysiology—a mini-review. *Nigerian J Cardiol*. 2017;14(1):9.
4. Ahmad T, Munir A, Bhatti SH, Aftab M, Raza MA. Survival analysis of heart failure patients: a case study. *PLoS ONE*. 2017;12(7):e0181001.
5. Amare H, Hamza L, Asefa H. Malnutrition and associated factors among heart failure patients on follow-up at Jimma university specialized hospital, Ethiopia. *BMC Cardiovasc Disord*. 2015;15(1):128.
6. Akerkar R, Martino S, Rue H. Implementing approximate Bayesian inference for survival analysis using integrated nested Laplace approximations. *Preprint Stat Norwegian Univ Sci Technol*. 2010;1:1–38.
7. Avi E. Bayesian survival analysis: comparison of survival probability of hormone receptor status for breast cancer data. *Int J Data Anal Tech Strategies*. 2017;9(1):63–74.
8. Benjamin EJ, Muntner P, Bittencourt MS. Heart disease and stroke statistics 2019 update: a report from the American heart association. *Circulation*. 2019;139(10):e56–528.

9. Berger JO. Statistical decision theory and Bayesian analysis. Berlin: Springer Science & Business Media; 2013.
10. Bhattacharjee A. Application of Bayesian approach in cancer clinical trial. *World J Oncol*. 2014;5(3):109.
11. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat*. 1998;7(4):434–55.
12. Chaloner K. Bayesian residual analysis in the presence of censoring. *Biometrika*. 1991;78(3):637–44.
13. Christensen R, Johnson W, Branscum A, Hanson TE. Bayesian ideas and data analysis: an introduction for scientists and statisticians. CRC Press; 2011.
14. Damasceno A, Mayosi BM, Sani M, Ogah OS, Mondo C, Ojji D, Dzudie A, Kouam CK, Suliman A, Schrueder N, et al. The causes, treatment, and outcome of acute heart failure in 1006 Africans from 9 countries: results of the Sub-Saharan Africa survey of heart failure. *Arch Intern Med*. 2012;172(18):1386–94.
15. Depaoli S. The impact of inaccurate informative priors for growth parameters in Bayesian growth mixture modeling. *Struct Equ Modeling*. 2014;21(2):239–52.
16. Ganjali M, Baghfalaki T. Bayesian analysis of unemployment duration data in presence of right and interval censoring. *JRSS*. 2012;5(1):17–32.
17. Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. *Stat Comput*. 2014;24(6):997–1016.
18. Giolo SR, Krieger JE, Mansur AJ, Pereira AC. Survival analysis of patients with heart failure: implications of time-varying regression effects in modeling mortality. *PLoS ONE*. 2012;7(6):e37392.
19. Habte B, Alemseged F, Tesfaye D. The pattern of cardiac diseases at the cardiac clinic of Jimma university specialised hospital, south west Ethiopia. *Ethiop J Health Sci*. 2010;20(2):99.
20. Hailay A, Kebede E, Mohammed K. Survival during treatment period of patients with severe heart failure admitted to intensive care unit (ICU) at Gondar university hospital (GUH), gondar, ethiopia. *Am J Health Res*. 2015;3(5):257–69.
21. Huffman MD, Berry JD, Ning H, Dyer AR, Garside DB, Cai X, Daviglius ML, Lloyd-Jones DM. Lifetime risk for heart failure among white and black Americans: cardiovascular lifetime risk pooling project. *J Am College Cardiol*. 2013;61(14):1510.
22. Ibrahim JG, Chen M-H, Sinha D. Bayesian survival analysis. Berlin: Springer Science & Business Media; 2001.
23. Ibrahim JG, Zhu H, Tang N. Bayesian local influence for survival models. *Lifetime Data Anal*. 2011;17(1):43–70.
24. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457–81.
25. Khanal SP, Sreenivas V, Acharya SK. Accelerated failure time models: an application in the survival of acute liver failure patients in India. *Int J Sci Res*. 2014;3(6):161–6.
26. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. Berlin: Springer Science & Business Media; 2006.
27. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22(4):719–48.
28. Martino S, Akerkar R, Rue H. Approximate Bayesian inference for survival models. *Scand J Stat*. 2011;38(3):514–28.
29. Misganaw A, Mariam DH, Ali A, Araya T. Epidemiology of major non-communicable diseases in Ethiopia: a systematic review. *J Health Popul Nutr*. 2014;32(1):1.
30. Miyagawa S, Pak K, Hikoso S, Ohtani T, Amiya E, Sakata Y, Ueda S, Takeuchi M, Komuro I, Sawa Y. Japan heart failure model derivation and accuracy of survival prediction in Japanese heart failure patients. *Circulation Rep*. 2019;1(1):29–34.
31. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, Falk V, Gonzalez-Juanatey JR, Harjola V-P, Jankowska EA, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the task force for the diagnosis and treatment of acute and chronic heart failure of the European society of cardiology (ESC). Developed with the special contribution of the heart failure association (HFA) of the ESC. *Eur J Heart Failure*. 2016;18(8):891–975.
32. Qi J. Comparison of proportional hazards and accelerated failure time models. Ph. D. thesis; 2009.
33. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Series b (Stat Methodology)*. 2009;71(2):319–92.
34. Sheng J, Qian X, Ruan T. Analysis of influencing factors on survival time of patients with heart failure. *Open J Stat*. 2018;8(04):651.
35. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation, vol. 13. Wiley; 2004.
36. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A, Degenhardt L, Dicker D, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. *Lancet*. 2015;386(9995):743–800.
37. Watanabe S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J Mach Learn Res*. 2010;11:3571–94.
38. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American college of cardiology foundation/American heart association task force on practice guidelines. *J Am Coll Cardiol*. 2013;62(16):e147–239.
39. Zeru MA. Assessment of major causes of heart failure and its pharmacologic management among patients at Felege Hiwot referral hospital in Bahir dar, Ethiopia. *J Public Health Epidemiol*. 2018;10(9):326–33.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.