

Bayesian Survival Analysis of Heart Failure Patients: A Case Study in Jimma University Medical Center, Jimma, Ethiopia

Tafese Ashine Tefera (✉ Tafese145@gmail.com)

Assosa University

Geremew Muleta

Jimma University

Kenenisa Tadesse

Jimma University

Research

Keywords: Heart failure, Bayesian, Survival Analysis, INLA

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-339250/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Bayesian Survival Analysis of Heart Failure Patients: A Case Study in Jimma University Medical Center, Jimma, Ethiopia

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

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

²Department of Statistics, College of Natural Science, Jimma University, Jimma, Ethiopia

Email address: Tafese145@gmail.com (Tafese Ashine), Yohager@yahoo.com (Geremew Muleta), kenenisatadesse@gmail.com (Kenenisa Tadesse)

*Corresponding Author

Abstract: Heart failure is failure of the heart to pump blood with normal efficiency and globally growing public health issue with high death rate over the world including Ethiopia. The aim of this study was to identify factors affecting the survival time of heart failure patients in Jimma University Medical Center. To reach the aim, 409 heart failure patients were including in the study based on data taken from medical record of patients enrolled during January, 2016 to January, 2019. Kaplan Meier plots and log rank test were used for comparison of survival function; Bayesian survival models was used to identify factors affecting the survival time heart failure patients. Of the total patients in the study 164 (40.1%) were died. The estimated median survival time of patients was 31 months. Bayesian log-normal accelerated failure time model fit heart failure data-set better than other Bayesian accelerated failure time models used in this study. From the results of this model shows that the survival time of heart failure patients significantly affected by age, chronic kidney disease, diabetes mellitus, etiology of heart failure, hypertension, anemia, smoking cigarette and stages of heart failure. Bayesian log-normal accelerated failure time model describes the heart failure data-set well. Age group (49 to 65 years and greater than 65 years); etiology of heart failure (rheumatic valvular heart disease, hypertensive heart disease and Other diseases); presence of hypertension; presence of anemic; presence of chronic kidney disease; smokers; diabetes mellitus (type I and type II diabetic); and stages of heart failure (II, III and IV) were prolong the timing death of heart failure patients. The hospital, Jimma University medical center, need to improve public awareness for early detection of heart failure.

Keywords: Heart failure, Bayesian, Survival Analysis, INLA

1. Background of the Study

Heart failure 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 a 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 (Ponikowski *et al.*, 2016; Yancy *et al.*, 2013).

Heart failure is global major cause of death and is a rapidly grown public health issue affecting approximately 40 million individuals worldwide and an estimated 287,000 deaths occurred a year, making it the most quickly growing cardiovascular disorder. It is ever increasing prevalence across developed and developing countries resulted as a complications from an increasing aging population (Vos *et al.*, 2015). 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 1,000 population and also an estimated 1 in 8 deaths in 2017 (Benjamin *et al.*, 2019). The prevalence of symptomatic HF is estimated to 5% of the population, and the mortality is estimated at about 13% in Europe (Huffman *et al.*, 2013).

In Africa, HF has emerged as a major public health problem, imposing enormous pressure on the health care systems; HF is not a disease by itself patients with HF have other causes of death. 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 (Damasceno *et al.*, 2012). Similarly, HF is reported to have caused 2.5% of deaths among all age groups in a sampled hospital based mortality in Ethiopia (Misganaw *et al.*, 2014). In this study, the researchers applied survival analysis since it addresses the limitation of classical regressions like logistic and linear regressions. Most medical studies has been used cox regression model for assessing the survival distribution of heart failure patients, while alternative parametric models including exponential, weibull, log-normal, and, log-logistic model has been used to identify the prognostic factors (Giolo *et al.*, 2012; Hailay *et al.*, 2015). The parametric survival models could provide a more suitable description of the survival data if one is able to identify the

distribution of the survival time (Khanal *et al.*, 2014). The parametric AFT models (i.e exponential, weibull, log-normal and log-logistic) has a more realistic interpretation and provides more informative results than Cox-PH model (Qi, 2009). Epidemiologists have documented several risk factors for the development of HF like as age, hypertension (Sheng *et al.*, 2018) and anemia (Ahmad *et al.*, 2017) were increased risk of mortality among HF patients. Factors such as age, sex, stages of HF, hypertension, anemia, and diabetes mellitus has statistically significant effect on the survival of HF patients (Zeru, 2018).

The parametric survival models play an important role in Bayesian survival analysis, since many Bayesian analysis in practice are carried out using parametric AFT models and provide computational advantages. It generates conclusions based on the synthesis of new information from an observed data and historical knowledge or expert opinion. Historical knowledge from past similar studies can be very helpful in interpreting the results of the current study Ibrahim *et al.* (2001). The Bayesian approach assumes that the observed data is fixed and that model parameters are random. The prior probability distributions represents a powerful mechanism for incorporating information from previous studies and for controlling confounding (Ibrahim *et al.*, 2011). The Bayesian methods combine objective prior knowledge with the information acquired from the data by using Bayes theorem (Gelman *et al.*, 2014).

In this study, Bayesian Survival Models would be used to identify the factors that affecting the survival time of heart failure patients in JUMC, Jimma, Ethiopia so, the interesting application of INLA method of estimation with Bayesian survival models are the most key for the motivation to apply it for the HF data-set under this study. The main aim of this study was to identify the factors that affecting survival time of HF patients. It quests to identify the prognostic factors of HF patients, to determine the best parametric survival models for heart failure data-set, to estimate the survival time of HF patients and to explore the Bayesian AFT models using INLA method.

Statement of the Problem

Heart failure is a serious condition during which the guts is unable to pump enough blood to satisfy the requirements of the body (Lloyd-Jones *et al.*, 2002). In approximately all regions of the world, HF is both common and increasing; it is predicted that the number of patients with HF to increase in countries with aging populations and the leading cause of HF death (Benjamin *et al.*, 2019). Studies show that HF is extremely increasing in African countries, including Ethiopia (Damasceno

et al., 2012; Misganaw *et al.*, 2014). Bayesian approach is the possibility of improving the precision of the results by introducing external information in terms of the prior distribution.

Thus, considering the advantages of Bayesian application is the most key for the motivation to apply it for the heart failure data-set under this study. So, chosen Bayesian Survival Analysis using INLA method to analysis heart failure data-set. Considering HF is a growing problem in the countries hospital based and gaps found with different studies, the researchers was explore the Bayesian Survival Analysis of HF patients in JUMC, Jimma, Ethiopia.

Therefore, this study aims to answer the following scientific questions:-

- Which factors significantly affect the survival time of heart failure patients?
- What is the estimated survival time of heart failure patients?
- Which parametric survival models is the most appropriate for analyzing the heart failure data set?

Objectives of the Study

Generalized Objective

The aim of this study was to identify factors affecting the survival time of heart failure patients in Jimma University Medical Center, Jimma, Ethiopia using Bayesian Survival Models.

Specific Objectives

The specific objectives of this study were:-

- To identify the prognostic factors of heart failure patients.
- To estimate the survival time of heart failure patients.
- To determine the best parametric survival models for heart failure data-set.
- To explore the bayesian accelerating failure time models for the heart failure data-set using INLA method.

Significance of the Study

Studying the survival time of HF patients is a mechanism of overcoming the problem of healthy in the society by identifying factors associated with death. On top of this, the result of this study might be used to improve awareness on the factors that trigger the death of HF patients. It also enables to provide scientific information about the finding to ministry of health in Ethiopia that helps policymakers to enhance the awareness of the 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.

2. Methodology

2.1. Data Description

2.1.1. Study Area

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

2.2.2. Study Design and Population

A retrospective study has been applied to obtain data on HF patients that recorded in JUMC, Jimma, Ethiopia. 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 patients 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 and Exclusion Criteria

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

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

Data Collection Methods

Ethical permission has been obtained from the JUMC, Jimma, Ethiopia. Then secondary data were taken based on data existing in the hospital by trained enumerator and the principal investigator using check list (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 starts diagnosis; when the patient first received the treatment.

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

Variables in the Study

The response variable was survival time of heart failure patients (in months), which defined as the difference between 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 the purpose of Bayesian survival analysis of HF patients were as follows:-

Sex	Age
Residence	History of Heart Failure
Etiology of Heart	Chronic kidney disease
Smoking Cigarette	Hypertension
Diabetes Mellitus	Treatments
Anemia	Alcohol Consumption
Stages of Heart failure	

2.2. Method of Data Analysis

2.2.1. Descriptive Statistics

The description of survival data utilizes non-parametric methods to compare the survival functions of two or more groups and kaplan-meier plot(s) would be employed for this purpose (Kaplan and Meier, 1958). The frequency distribution table was used to summarize the data obtained from registration book of patients based on the study variables in JUMC, Jimma, Ethiopia.

2.2.2. Survival Data Analysis

Survival analysis is the statistical analytic method used for modeling and analyzing the data that have a principal end point the time until an event occurs. 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 (Aalen *et al.*, 2008).

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 patient is terminated before the event occurs. This type of censoring is commonly recognized survival analysis and also considered in this study (Klein and Moeschberger, 2006).

Estimation of Survival Function

The Kaplan-Meier estimator, non-parametric estimator used to estimate the survival function with censoring, which is not based on the actual observed event and censoring times, but rather on the order in which events and censored observations occur (Kaplan and Meier, 1958). Therefore, the Kaplan Meier estimate of the survival function at time t , ($S(t)$) is given by:

$$\widehat{S}(t) = \prod_{j:r_j \leq t} \left[1 - \frac{d_j}{r_j}\right]$$

Where: - r_j denote the set of k distinct death time in the observed in the sample, r_j is the number of subjects alive (at risk) just before time t_j (the j^{th} ordered survival time) and d_j denotes the number who died at time.

Comparison of Survival Function

The Kaplan-Meier plots are used to see whether there is difference in survival times or not between groups of covariate under investigation. But, the plot cannot be used to decide whether the survival time of heart failure patients in each covariate is different or not and log-rank test was used for this purpose (Mantel and Haenszel, 1959).

The test statistic for log rank test is given by:

$$\chi^2_{\text{logrank}} = \frac{[\sum (d_{0j} - r_{0j} \frac{d_j}{r_j})]^2}{\sum \frac{r_{0j} r_{1j} (r_j - d_j)}{r_j^2 (1 - r_j)}} \sim \chi^2_{(1)}$$

Where: d_{0j} is the number of failure in j^{th} time of 1st group, d_{1j} is the number of failure in j^{th} time of 2nd group, d_j is the number of failure in j^{th} time ($d_{0j} + d_{1j}$), r_{0j} is the number at risk at j^{th} time of 1st group, r_{1j} is the number at risk at j^{th} time of 2nd group and r_j is the number at risk at j^{th} time ($r_{0j} + r_{1j}$).

The hypotheses to be tested are:

H_0 : There is no difference between the survival curves.

H_1 : There is difference between the survival curves.

2.2.3. Bayesian Survival Analysis

Bayesian approach is preferred over the frequentist approach in survival analysis is 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. In addition, Bayesian approach has several advantages over classical methods, it is well known that survival models are generally quite hard to fit, especially in the presence of complex censoring schemes. With the use of the Gibbs sampler and other MCMC techniques, fitting complex survival models is fairly straightforward, and the availability of software like BUGS eases the implementation greatly (Ibrahim *et al.*, 2001). MCMC methods, has some limitation like the burden of time in approximating the posterior and convergence problem (Brooks and Gelman, 1998; Berger, 2013). As of 2009, the other news was welcomed with very flexible and fast approximation techniques called Integrated Nested Laplace Approximation. Bayesian approach with INLA method is focused on providing a good approximation to the posterior marginal distributions of the parameters in the model (Rue *et al.*, 2009).

The main reasons why one might choose to use Bayesian statistics to produce more accurate parameter estimates. In addition, Bayesian statistics one can incorporate uncertainty about a parameter and update this knowledge through the prior distribution (Depaoli, 2014). Bayesian approach is more useful in clinical data analysis over frequentist and suitable data analysis technique for clinical researchers (Bhattacharjee, 2014). 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.

Components of Bayesian inference:

Prior Distribution: - $\pi(\theta)$, It probability distribution used to expresses uncertainty about unknown quantities parameter, θ 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 (Ibrahim *et al.*, 2001).

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

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

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 (Ganjali and Baghfalaki, 2012).

Posterior Distribution:- Posterior Distribution is a combination of prior distribution and likelihood using the Bayes rule, 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, $L(\theta/\text{data})$ (Christensen *et al.*, 2011). 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 (Ibrahim *et al.*, 2001).

Parametric models play an important role in Bayesian survival analysis, since many Bayesian analyses in practice are carried out using parametric models (Exponential Model, Weibull Model, Log-Normal Model, and Log-Logistic Model). Parametric modeling offers straightforward modeling and analysis techniques (Ibrahim *et al.*, 2001).

2.2.4. Estimation of Parameter in Bayesian Survival Models

Integrated Nested Laplace Approximation (INLA) Method

The Integrated Nested Laplace Approximation Method was used to estimate the parameters in Bayesian survival model. Survival analysis consists of a great body of work using latent Gaussian models. According to Rue *et al.* (2009), 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 provide both extremely fast and very accurate approximations to the posterior marginal through a clever use of laplace approximations and advanced numerical methods and it can be adapted to fit survival models (Akerkar *et al.*, 2010). An R package called R-INLA works as an interface for INLA and it is used just as 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 was used for Bayesian survival model comparison. The preferable model is the one with the lowest value of the DIC (Spiegelhalter *et al.*, 2004). An alternative is the Watanabe Akaike information criterion, (Watanabe, 2010) which follows a more fully Bayesian approach to construct a criterion. (Gelman *et al.*, 2014) claims the WAIC is preferable to the DIC.

Bayesian Model Diagnostics

The most common ways of checking goodness of fit are: Bayesian Cox-Snell residual plot and Predictive Distribution. Model checking and adequacy play an important role in models for survival data. In Bayesian analysis, Chaloner (1991) defined the Bayesian version of the residuals. The idea on how to classify the sample data and techniques undergone for the application of criticism for predictive distribution (Piiironen and Vehtari, 2017).

3. Results and discussions

Descriptive Summaries

The data for this study has been taken from 409 patients receive treatments for HF, at least one time, at Jimma University Medical Center, Jimma, Ethiopia in the year first January, 2016 to first January, 2019. The minimum and maximum event time observed from HF patients follow up where 6 and 36 months respectively. Among those HF patients, about 59.90% were censored (right censored) and remaining 40.10% were died. Fifty percent of HF patients who receive treatments, survived 31 months or above it. Almost half, 52.81%, of the HF patients were female and the remaining were male during the follow up study. However, their survival time of male patients seems lower.

Most of HF patients, about 64.79% live in rural area and the remaining were in urban area. The survival time of HF patients seems less as they gets 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 disease 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 smoker which was 54.88% compared to non-smokers which was 45.12%. About, 64.55% of HF patients were not alcohol users and 35.45% were alcohol users.

Moreover, about 19.08% HF patients treated in the hospital with a combination of two or more treatments and 19.32% HF patients take digoxin. In addition, the remaining 24.2%, 25.18% and 11.49% of HF patients treated with spironolactone, atorvastatin and other treatments respectively. About 58.19%, 13.69% and 28.12% 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 seems lower survival time. Most of HF patients has no hypertension, 60.64%, and the remaining has hypertension.

Looking the stage at which the HF patients goes 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 into the hospital at later stage and their survival time seems low at this stage.

Table 1: Descriptive summaries of patient's for HF data set

Covariates	Categories	Patients Status		
		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)
	No	224(91.43)	61(37.20)	285(69.68)

Chronic kidney disease	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

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 the patients whose age was greater than or equal to 65 years.

Figure 2 (c) below, shows that HF patients with stage I were 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 the HF patients with stage IV.

Figure 2 (d) below, one sees that the HF patients without hypertension was at a higher chance of survival than the HF patients with hypertension.

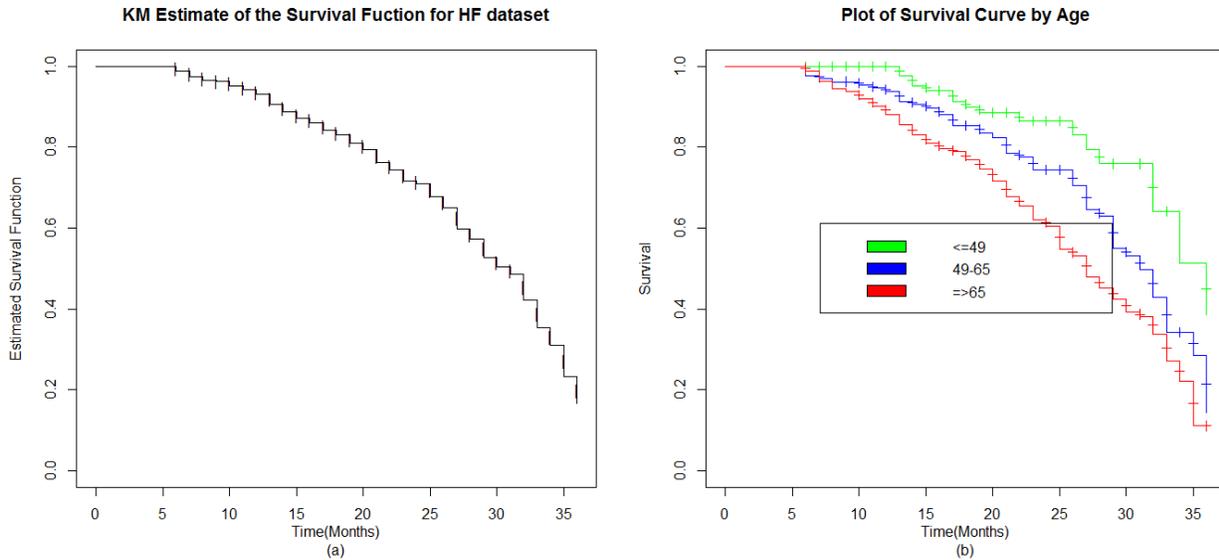


Figure 1: Kaplan-Meier estimates of the survival curves of HF data set for (a): Overall survivor function for heart failure patients in the study, (b): Age group

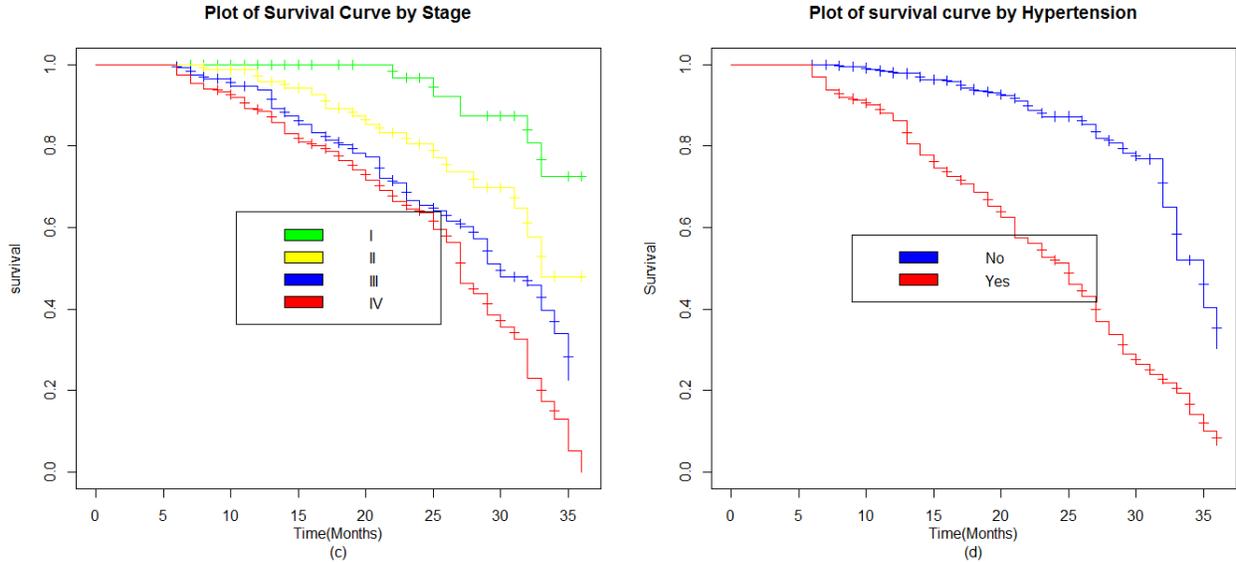


Figure 2: Kaplan-Meier estimates of the survival curves of HF data set for (c): Stages, (d): Hypertension

Checking the Assumption of Cox-PH

As it can be shown in the Table 2 below, the p-values for alcohol consumption, chronic kidney disease and anemia are less than common (5%) level of significance using correlation test (rho)

and these shows as the assumption of Cox-PH model was not valid for HF data set. In addition, by looking for global test the assumption of Cox-PH fails since the test result was significant.

Table 2: 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

Bayesian Survival Data Analysis

As it can be shown in the Table 2, the assumption of Cox-PH model was not valid for heart failure data set; in this case parametric AFT models was used for HF data set. For the HF data set, the time t_i where $i = 1, 2, \dots, 409$ of heart failure 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 1000. We assume that scale parameter have a gamma prior with shape parameter 1 and inverse scale parameter 0.001, this prior was used for Weibull, Log-normal and Log-logistic distribution for INLA (Akerkar *et al.*, 2010). From Table 3 below, shows that analysis of HF data set for model comparison using INLA method. To compare the efficiency of these different models DIC and WAIC were used and the one with smallest value and seems best fit. Accordingly, Bayesian lognormal AFT model using INLA (DIC= 1297.84; WAIC= 1297.47) found to be the best for survival time of HF patients data-set from the give alternative.

Table 3: 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

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

Bayesian Log-Normal AFT Model using INLA method

The final results for the Bayesian log-normal AFT model using INLA method was shown as in Table 4 and as this result shows the survival time of HF patients statistically significantly affected by age, chronic kidney disease, diabetes mellitus, etiology of heart failure, hypertension, anemia, smoking cigarette and stages of heart failure.

Table 4: Indicating the results for Bayesian log-normal AFT model using INLA method

Covariates	Categories	Pmean	Sd	Median	Credible Interval	Mode	Kld
	Intercept	4.953	0.221	4.945	[4.541, 5.409]*	4.929	0
Age	≤ 49	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	
	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	-

*is indicated statistically significant. Pmean: Posterior Mean; Kld: Kullback-leibler divergence; CKD: Chronic kidney disease; IHD: Ischemic heart disease; RVHD: Rheumatic valvular heart disease; HDD: Hypertensive heart disease, Sd: Standard deviation.

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

Interpretation of Bayesian log-normal AFT model using INLA method

From Table 4, the final model were interpreted using acceleration factor, 95% credible interval of Bayesian accelerated failure time estimated values. The estimated acceleration factor is defined as $\gamma = [\exp(\hat{\beta})] = [\exp(\text{posterior mean})]$.

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

The 95% Credible Interval for acceleration factor of both age group did not include one which implies that both age group has significant effect on the survival time of HF patients. Looking for chronic kidney disease, keeping the effect of other factors constant, the estimated acceleration factor of HF patients with chronic kidney disease is estimated to be 0.6777 with [95% CrI: 0.5844, 0.7835] which implies the expected survival time decreases by 32.23% than HF patients without chronic kidney disease. The 95% CrI for acceleration factor of HF patients with chronic kidney disease did not include one which implies that HF patients with chronic kidney disease has significant effect on the survival time of HF patients.

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 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 has significant (in the Bayesian sense) effect on the survival time of HF patients.

On other hand, keeping the effect of other factors constant, the estimated acceleration factor for HF patients who were smoking cigarette 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 cigarette did not include one. Thus, HF patients who were smoking cigarette has significant effect on the survival time of patients and the expected survival time of HF patients who were smoking cigarette decreases by 14.45% than not smoking cigarette. Regarding to 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 disease 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 decreases by 27.07% were rheumatic valvular heart disease, 22.8% were hypertensive heart disease and 31.7% were other heart disease 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 which implies that etiology of HF were rheumatic valvular heart disease, hypertensive heart disease and other heart disease has significant effect on the survival time of HF patients, while the etiology of HF were cardiomyopathy heart disease has not significant effect on the survival time of HF patients.

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 are 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 diabetic and 34.5% for type II diabetic as compared to HF patients non-diabetic (Reference). The 95% credible interval for acceleration factor of HF patients with both type of diabetes did not include one which implies that HF patients with both type of diabetes has significant effect on the survival time of HF patients.

Looking for anemia, keeping the effect of other factors constant, the estimated acceleration factor of HF patients with anemia is estimated to be 0.857 with [95% CrI: 0.742, 0.987] which implies the expected survival time decreases by 14.3% than HF patients without anemia. The 95% CrI for acceleration factor of HF patients with anemia did not include one which implies that HF patients with anemia has significant effect on the survival time of HF patients.

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 are 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 having stage II, III and IV did not include one this indicates that stage II, III and IV has significant effect on the survival time of heart failure patients.

From Table 4, the Kullback-leibler divergence values for all significant parameters in Bayesian log-normal AFT model were 0, and thus, small values indicate that the posterior distribution was well approximated by a normal distribution. A simplified laplace approximation was the most efficient algorithm with improved efficiency and results to higher computation speed. In addition, there is no need to perform the more computationally intense full Laplace approximation. Therefore, simplified laplace approximation was appropriate.

Bayesian Model Diagnostics

Bayesian Cox Snell Residual Plots

By observing Bayesian cox-snell residual plots figure below, 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 were nearest to the line through the origin. In addition, the plot also indicated that Bayesian log-normal model describes the HF data-set well.

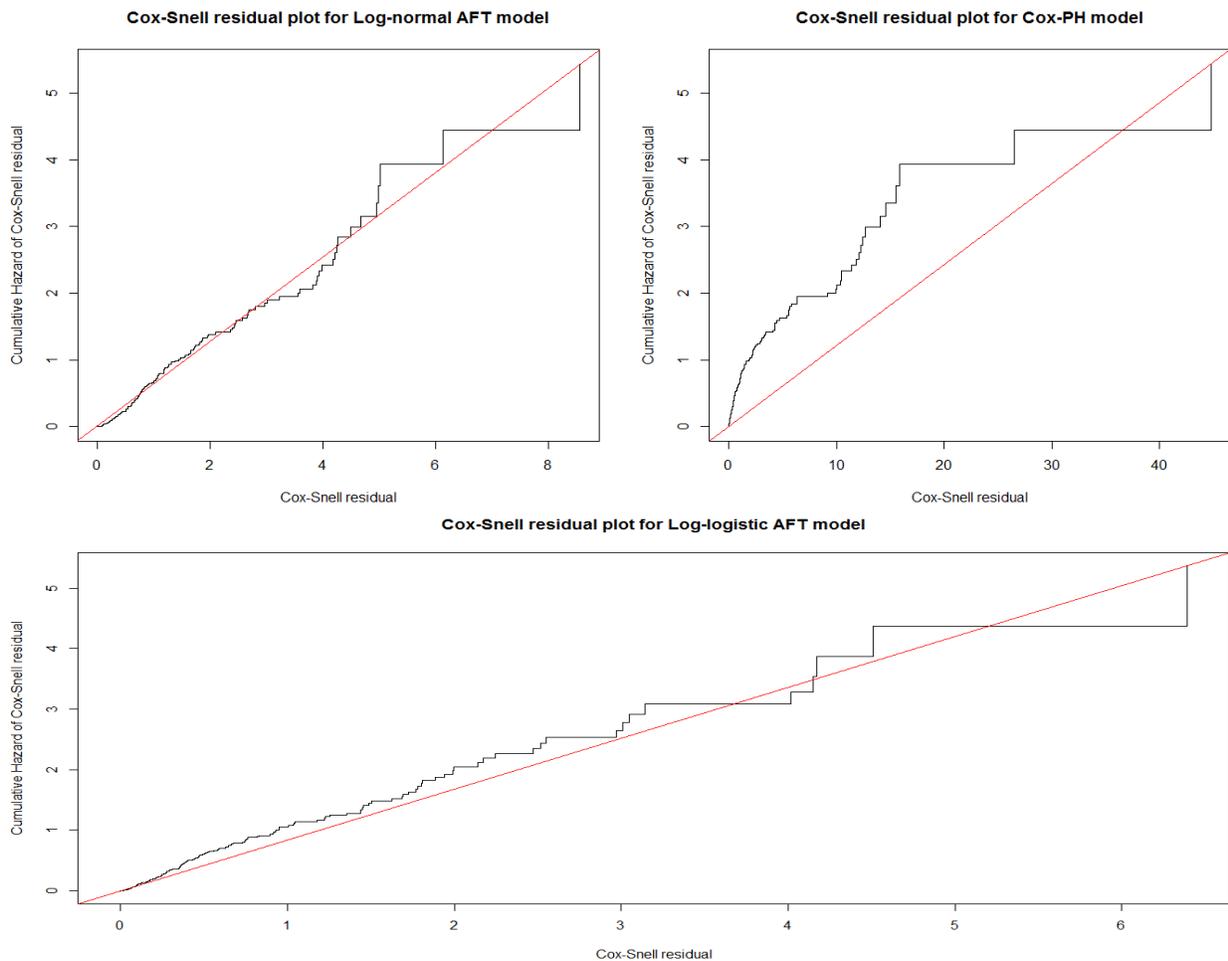


Figure 3: Bayesian Cox-Snell residual plots for baseline distribution and Cox-PH that were used to fit the Heart failure data set.

The histograms of the cross-validated probability integral transform values shows that the posterior predictive p-values are some extent closer to a uniformly distributed with some observations outliers in HF data-set. The Conditional predictive ordinate values that are

considerably smaller (order of magnitude smaller) than the others, so with respect to the Bayesian log-normal model, the observed values would be considered 'surprising', because the sum of the observations associated with failure flags are equal to zero in HF data-set. The plots including 95% credibility interval by observing the posterior density for the parameters was normally distributed in HF data-set. The kullback-leibler divergence is a diagnostic that measures the accuracy of the INLA approximation, since from Table 4 shows that the kld for all significant parameters in the Bayesian log-normal AFT model were 0.

Discussions

The main aim of this study was to identify factors affecting the survival time of HF data set, which was obtained from JUMC, Jimma, Ethiopia. Heart failure is a growth problem in the world and the overall prevalence of HF in the adult population in developing countries is 7% - 10% with exponential rise with age (Adebayo *et al.*, 2017). 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 patients follow up where 6 and 36 months respectively.

In addition, fifty percent of HF patients who receive treatments, survived 31 months or above it. In this study, among those HF patients, about 59.90% were censored (right censored) and remaining 40.10% were died. This finding was similar to a study conducted by Hailay *et al.* (2015) shows among those HF patients, 31.3% of them were dead while the rest 68.7% were censored. The Survival model was applied for this data set. But, the assumption of Cox-PH model was violated. Bayesian approach was applied on parametric AFT models and to compare the efficiency of different AFT models DIC and WAIC were used (Spiegelhalter *et al.*, 2004; Watanabe, 2010). Bayesian log-normal AFT model was the best model to describe HF data set from the given alternative. This result was similar with study done by Avi (2017).

However in this study, the results of Bayesian log-normal AFT model using INLA method shows that the survival time of HF patients significantly affected by age, chronic kidney disease, diabetes mellitus, etiology of HF, hypertension, anemia, smoking cigarette and stage of HF.

From the result of this study the age group has a significant effect on the survival time of HF patients. In addition, the survival time of HF patients seems less as they gets older (greater than or equal to 65 years) and different studies were also persisted with this results Adebayo *et al.* (2017), Zeru (2018), and Sheng *et al.* (2018). The survival time of HF patients has no hypertension was higher than that of with hypertension and thus, hypertension had a significant effect on HF patients, the studies done by Ahmad *et al.* (2017), and Sheng *et al.* (2018) shows the same results.

On other hand, the survival time of smoker HF patients were decreases as compared to nonsmoker which is similar to study done by Ahmad *et al.* (2017). The survival time of HF patients significantly affected by both type of diabetes mellitus and the expected survival time of HF with both type of diabetes mellitus was less as compared to HF patients without diabetic, this results consistent with studies done by Ahmad *et al.* (2017), and Zeru (2018). In addition, chronic kidney disease was significantly affected the survival time of HF patients and the survival time was high for HF patients do not have chronic kidney disease as compared to HF patients having chronic kidney disease, this results, also confirmed with study by Zeru (2018).

The studies done by Ahmad *et al.* (2017), and Zeru (2018) shows that the survival time of HF patients significantly affected by anemia and the expected survival time of HF with anemia was less as compared to HF patients without anemic. This studies was consistent with the current study. The stages of HF patients has significant effect on the survival time of HF patients. The study done by Zeru (2018) shows that the stages of HF patients has been significantly affected the survival time of HF patients. From the results of these study the survival time of HF patients was smaller as the stage increases as in result of this study.

For checking 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 Figure 3. The plots were more approached to the line in case of the Bayesian log-normal model that indicates the Bayesian log-normal was best in HF data-set. This result was consistent with other study done by Avi (2017). The conditional predictive ordinate and probability integral transform were also used for model checking. 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 failure in conditional predictive ordinate was zero, no failure was detected and meaning that no numerical problem has occurred in HF data-set. The histogram and scatter plot of probability integral transform were indicated that the plots of predictive residual based values were to some extent uniformly distributed with some deviated outlier and there is reasonable predictive distribution matches the actual data. This result was persisted with other studies done by Akerkar *et al.* (2010), and Martino *et al.* (2011).

The Bayesian log-normal AFT model diagnostic plots including 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 indicate that Bayesian log-normal AFT model using INLA method was fast and higher accuracy. This results, also confirmed by other studies done by Martino *et al.* (2011), and Akerkar *et al.* (2010).

However, this study was not done without limitation. The study was conducted based on secondary data gathered from registration log book and patients registration card, which might have incomplete and biased information; Lack of published literature's on the countries hospital based related to the survival time of heart failure patients using Bayesian survival models; As different literature pointed out, there are different prognostic factors (Body mass index and Weight) that are assumed to have impacts on the survival time of HF. However, data on those variables could not be available in hospital records.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

This study used survival time of heart failure patients data set, for those patients who were receiving treatments for at-least one time in Jimma University Medical Center. Bayesian lognormal AFT model performing 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 receive treatments, survived months or above it.

The survival time of HF patients significantly affected by age, chronic kidney disease, diabetes mellitus, etiology of heart failure, hypertension, anemia, smoking cigarette and stages of heart failure. Of all this statistically significant covariates; age (49 to 65 years and greater than 65 years); etiology of heart failure (rheumatic valvular heart disease, hypertensive heart disease and Other diseases); presence of hypertension; presence of anemic; presence of chronic kidney disease; smokers; diabetes mellitus (type I and type II diabetic); and stages of heart failure (II, III and IV) were prolong the timing death of heart failure patients.

Recommendations

Based on the finding of the study it is recommended as follows:-

- The ministry of health and policy makers should work on awareness by letting to know the risk factors for heart failure.

- The hospital, JUMC, need to improve public awareness for early detection of HF.
- Awareness has to be given for the society regarding smoking cigarette. The mass media can play an effective role in this regard.

Declaration

During conducting the study, the investigators have included the following declarations.

ACRONYMS

AFT:	Accelerated Failure Time
AHA:	America Health Association
Cox PH:	Cox Proportional Hazard
CPO:	Conditional Predictive Ordinate
DIC:	Deviance Information Criteria
HF:	Heart Failure
INLA:	Integrated Nested Laplace Approximation
JUMC:	Jimma University Medical Center
JUSH:	Jimma University Specialized Hospital
KLD:	Kullback-Leibler Divergence
NYHA:	New York Heart Association
OPD:	Out Patient Department
PIT:	Probability Integral Transform
SSA:	Sub-Saharan Africa
TASH:	Tikur Anbessa Specialized Hospital
USA:	United State of America
WAIC:	Watanabe Akaike Information Criterion

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.

Competing Interests

The authors declare that they have not competing interest were available.

Funding: The study was fully funded by Jimma University and ministry of science and higher education of Ethiopia. All the expenses for the accessing of the data and related were covered by the collaborative fund obtained from the two institutions.

Author's Contributions

This study was designed and compiled by Tafese Ashine (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. Edition of the overall progress of the work was supported by Geremew Muleta and Kenenisa Tadesse.

Authors' Information

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

²Department of Statistics, College of Natural Science, Jimma University, Jimma, Ethiopia

Consent for publication

Not applicable

Acknowledgments

We thank Jimma University and Ethiopian ministry of science and higher education for financing this study and cover all expenses of the paper. We sincere thanks to Dr. Birhanu Teshome (PhD) for his advice and guidance. Finally thanks to all staff members of Jimma University medical center for allowing us to collect the data set from the hospital.

References

Aalen, O., O. Borgan, and H. Gjessing (2008). *Survival and event history analysis: a process point of view*. Springer Science & Business Media.

Adebayo, S. O., T. O. Olunuga, A. Durodola, O. S. Ogah, et al. (2017). Heart failure: Definition, classification, and pathophysiology—a mini-review. *Nigerian Journal of Cardiology* 14(1), 9.

Ahmad, T., A. Munir, S. H. Bhatti, M. Aftab, and M. A. Raza (2017). Survival analysis of heart failure patients: A case study. *PloS one* 12(7), e0181001.

Akerkar, R., S. Martino, and H. Rue (2010). Implementing approximate bayesian inference for survival analysis using integrated nested laplace approximations. Preprint Statistics, Norwegian University of Science and Technology 1, 1–38.

Avi, E. (2017). Bayesian survival analysis: comparison of survival probability of hormone receptor status for breast cancer data. *International Journal of Data Analysis Techniques and Strategies* 9(1), 63–74.

Benjamin, E. J., P. Muntner, and M. S. Bittencourt (2019). Heart disease and stroke statistics 2019 update: a report from the american heart association. *Circulation* 139(10), e56–e528.

Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.

Bhattacharjee, A. (2014). Application of bayesian approach in cancer clinical trial. *World journal of oncology* 5(3), 109.

Brooks, S. P. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics* 7(4), 434–455.

Chaloner, K. (1991). Bayesian residual analysis in the presence of censoring. *Biometrika* 78(3), 637–644.

Christensen, R., W. Johnson, A. Branscum, and T. E. Hanson (2011). *Bayesian ideas and data analysis: an introduction for scientists and statisticians*. CRC Press.

Damasceno, A., B. M. Mayosi, M. Sani, O. S. Ogah, C. Mondo, D. Ojji, A. Dzudie, C. K. Kouam, A. Suliman, N. Schrueder, et al. (2012). 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. *Archives of internal medicine* 172(18), 1386–1394.

Depaoli, S. (2014). The impact of inaccurate informative priors for growth parameters in bayesian growth mixture modeling. *Structural Equation Modeling: A Multidisciplinary Journal* 21(2), 239–252.

Ganjali, M. and T. Baghfalaki (2012). Bayesian analysis of unemployment duration data in presence of right and interval censoring. *JRSS* 5(1), 17–32.

Gelman, A., J. Hwang, and A. Vehtari (2014). Understanding predictive information criteria for bayesian models. *Statistics and computing* 24(6), 997–1016.

Giolo, S. R., J. E. Krieger, A. J. Mansur, and A. C. Pereira (2012). Survival analysis of patients with heart failure: implications of time-varying regression effects in modeling mortality. *PloS one* 7(6), e37392.

Hailay, A., E. Kebede, and K. Mohammed (2015). Survival during treatment period of patients with severe heart failure admitted to intensive care unit (icu) at gondar university hospital (guh), gondar, ethiopia. *American Journal of Health Research* 3(5), 257–269.

Huffman, M. D., J. D. Berry, H. Ning, A. R. Dyer, D. B. Garside, X. Cai, M. L. Daviglus, and D. M. Lloyd-Jones (2013). Lifetime risk for heart failure among white and black americans: cardiovascular lifetime risk pooling project. *Journal of the American College of Cardiology* 61(14).

Ibrahim, J. G., M.-H. Chen, and D. Sinha (2001). *Bayesian Survival Analysis*. Springer Science & Business Media.

Ibrahim, J. G., H. Zhu, and N. Tang (2011). Bayesian local influence for survival models. *Lifetime data analysis* 17(1), 43–70.

Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association* 53(282), 457–481.

Khanal, S. P., V. Sreenivas, and S. K. Acharya (2014). Accelerated failure time models: an application in the survival of acute liver failure patients in india. *Int. J. Sci. Res* 3(6), 161–166.

Klein, J. P. and M. L. Moeschberger (2006). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.

Lloyd-Jones, D. M., M. G. Larson, E. P. Leip, A. Beiser, R. B. D'Agostino, W. B. Kannel, J. M. Murabito, R. S. Vasan, E. J. Benjamin, and D. Levy (2002). Lifetime risk for developing congestive heart failure: the framingham heart study. *Circulation* 106(24), 3068–3072.

Mantel, N. and W. Haenszel (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the national cancer institute* 22(4), 719–748.

Martino, S., R. Akerkar, and H. Rue (2011). Approximate bayesian inference for survival models. *Scandinavian Journal of Statistics* 38(3), 514–528.

Misganaw, A., D. H. Mariam, A. Ali, and T. Araya (2014). Epidemiology of major noncommunicable diseases in ethiopia: a systematic review. *Journal of health, population, and nutrition* 32(1), 1.

Miyagawa, S., K. Pak, S. Hikoso, T. Ohtani, E. Amiya, Y. Sakata, S. Ueda, M. Takeuchi, I. Komuro, and Y. Sawa (2019). Japan heart failure model derivation and accuracy of survival prediction in japanese heart failure patients. *Circulation Reports* 1(1), 29–34.

Piironen, J. and A. Vehtari (2017). Comparison of bayesian predictive methods for model selection. *Statistics and Computing* 27(3), 711–735.

Ponikowski, P., A. A. Voors, S. D. Anker, H. Bueno, J. G. Cleland, A. J. Coats, V. Falk, J. R. González-Juanatey, V.-P. Harjola, E. A. Jankowska, et al. (2016). 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. *European journal of heart failure* 18(8), 891–975.

Qi, J. (2009). Comparison of proportional hazards and accelerated failure time models. Ph. D. thesis.

Rue, H., S. Martino, and N. Chopin (2009). Approximate bayesian inference for latent Gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)* 71(2), 319–392.

Sheng, J., X. Qian, and T. Ruan (2018). Analysis of influencing factors on survival time of patients with heart failure. *Open Journal of Statistics* 8(04), 651.

Spiegelhalter, D. J., K. R. Abrams, and J. P. Myles (2004). Bayesian approaches to clinical trials and health-care evaluation, Volume 13. John Wiley & Sons.

Vos, T., R. M. Barber, B. Bell, A. Bertozzi-Villa, S. Biryukov, I. Bolliger, F. Charlson, A. Davis, L. Degenhardt, D. Dicker, et al. (2015). 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. *The Lancet* 386(9995), 743–800.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* 11(Dec), 3571–3594.

Yancy, C. W., M. Jessup, B. Bozkurt, J. Butler, D. E. Casey, M. H. Drazner, G. C. Fonarow, S. A. Geraci, T. Horwich, J. L. Januzzi, et al. (2013). 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. *Journal of the American College of Cardiology* 62(16), e147–e239.

Zeru, M. A. (2018). Assessment of major causes of heart failure and its pharmacologic management among patients at felege hiwot referral hospital in bahir dar, ethiopia. *Journal of Public Health and Epidemiology* 10(9), 326–33

Figures

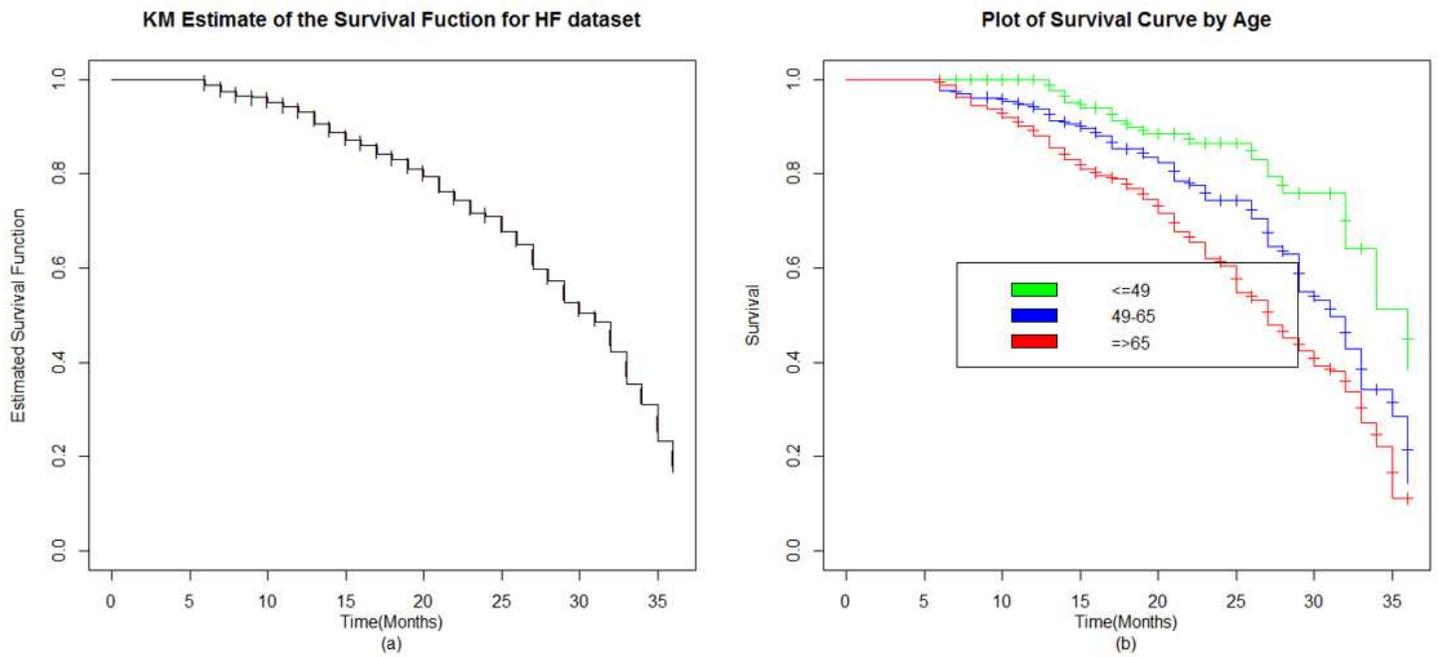


Figure 1

Kaplan-Meier estimates of the survival curves of HF data set for (a): Overall survivor function for heart failure patients in the study, (b): Age group

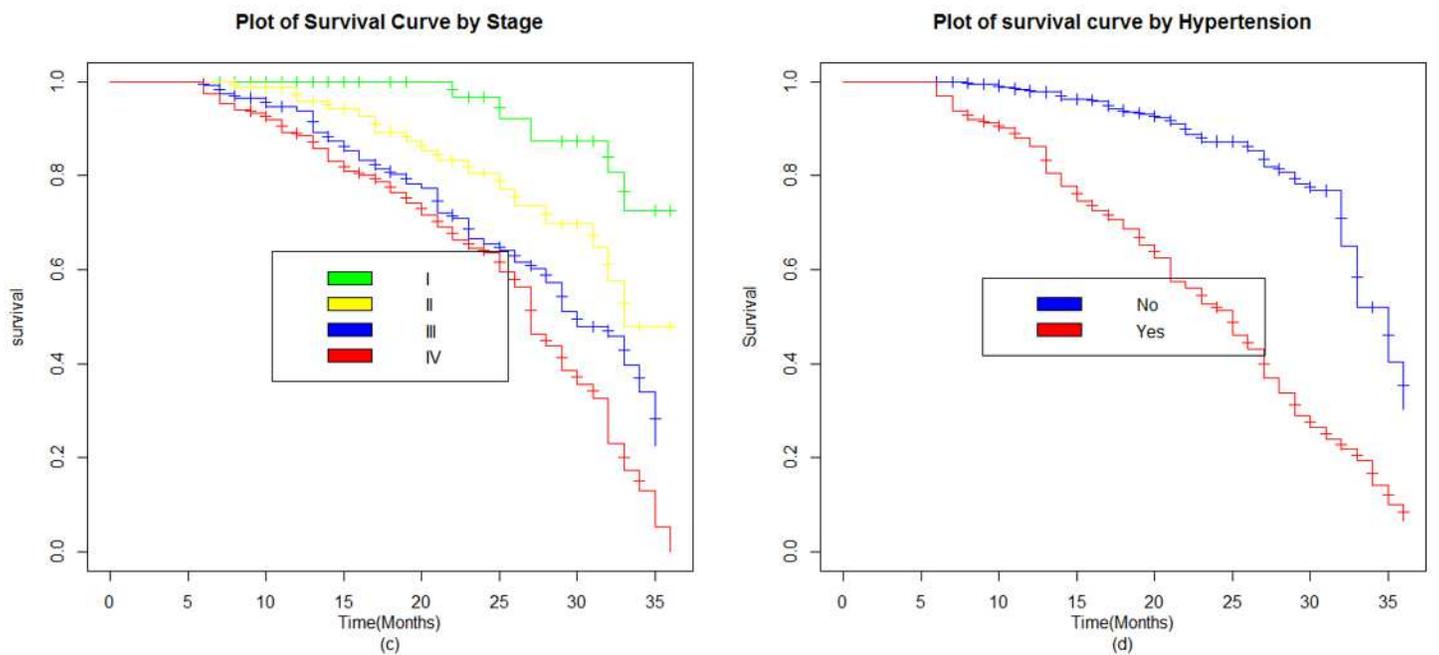


Figure 2

Kaplan-Meier estimates of the survival curves of HF data set for (c): Stages, (d): Hypertension

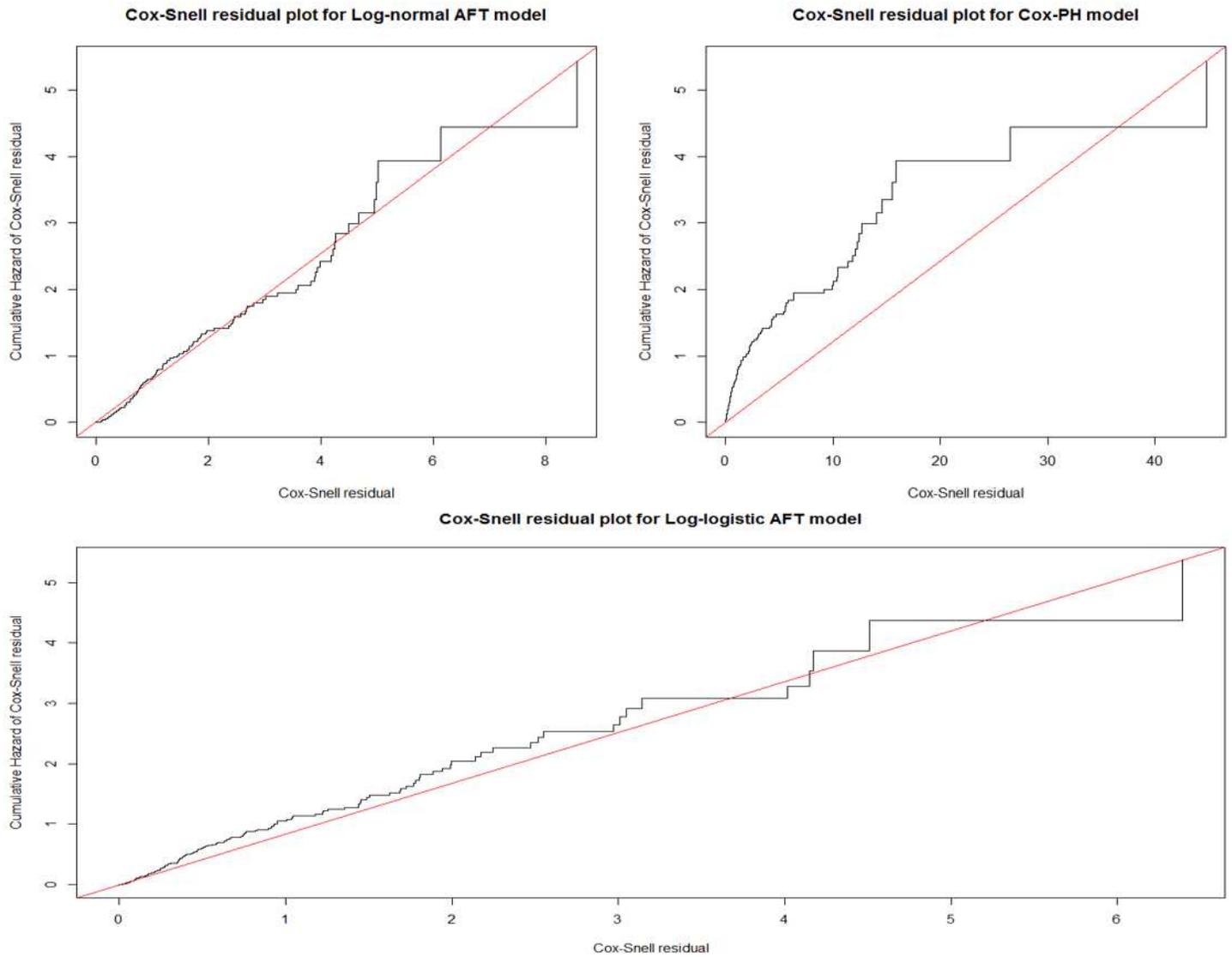


Figure 3

Bayesian Cox-Snell residual plots for baseline distribution and Cox-PH that were used to fit the Heart failure data set.