Survival Time Analysis of Hypertension Patients Using Parametric Models

Markos Abiso Erango¹*, Kabtamu Tolosie Gergiso¹ and Sultan Hussen Hebo²

¹Department of Statistics, College of Natural Science, Arba Minch University, Ethiopia.
²Department of Reproductive Health, Arba, Minch University, Arba Minch, Ethiopia.

Authors’ contributions

This work was carried out in collaboration among all authors. Author MAE conceived and designed the study, developed data collection instruments and supervised data collection. Authors MAE and KTG participated in the testing and finalization of the data collection instruments and coordinated study progress. Authors MAE, KTG and SHH performed the statistical analysis and wrote all versions of the manuscript. All authors critically revised and approved the final manuscript.

Article Information

DOI: 10.9734/AIR/2019/v20i230155
(1) Dr. Carlos Humberto Martins, Professor, Department of Civil Engineering, The State University of São Paulo, Brazil.
(1) Kamil Alakuş, Ondokuz Mayıs University, Turkey.
(2) N. S. Kannan, Pondicherry University, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/52009

Published 16 October 2019

ABSTRACT

Background: Hypertension is a worldwide public-health challenge and one of a leading modifiable risk factor for cardiovascular disease and death.

Aims: The aim of this study was compare parameter estimations using both Bayesian and classical approaches and to detect out potential factors that affects survival probability of hypertension patient’s under follow up.

Materials and Methods: A simple random sampling technique was used to select 430 patients among a total of 2126 hypertension patients who had been under follow up at Yekatit-12 Hospital in Ethiopia from January 2013 to January 2019. Parametric distributions: Exponential, Weibull, Lognormal and loglogistic are studied to analysis survival probabilities of the patients in both Bayesian and classical approaches. The model selection criteria are employed to identify the model with best fit to the data. Bayesian estimation approach was smaller deviance information criteria as compare to classical estimation approaches for the current data set.

Results and Conclusion: The analysis Bayesian Weibull results indicate that the baseline age of the patient, gender, family history of hypertension, tobacco use, alcohol use, khat intake, blood
cholesterol level of the patient, hypertension disease stage, adherence to the treatment and related disease were significantly associated with survival time of hypertension patients. Patients with raised blood Cholesterol level at baseline tend to have shorter survival time as compare to one with normal blood cholesterol level at baseline. Society and all stakeholders should be aware of the consequences of these factors which can influence the survival time of hypertension patients.

Keywords: Bayesian approach; classical approaches; estimation methods; hypertension; survival probability.

1. INTRODUCTION

Hypertension is a worldwide public-health challenge and a leading modifiable risk factor for cardiovascular disease and death. According to the World Health Organization (WHO) Global Health Observed Report, globally, the overall prevalence of Hypertension in adults aged 25 and over was around 40% in 2008 and was is estimated to cause 7.5 million deaths, about 12.8% of the total of all deaths worldwide. Blood pressure is summarized by two measurements, systolic and diastolic, which depend on whether the heart muscle is contracting or relaxed between beats. Normal blood pressure at rest is within the range of 100-140 mmHg systolic (top reading) and 60-90 mmHg diastolic (bottom reading). High blood pressure is said to be present if it is persistently at or above 140/90 mmHg [1]. Globally the number of people with uncontrolled hypertension rise by 70% between 1980 and 2008. The rising epidemic of hypertension is thought to be due to mechanization, population growth and ageing [2,3]. Hypertension doubles the risk of cardiovascular diseases such as coronary heart disease, congestive heart failure, stroke, renal failure and peripheral arterial disease [4]. According to Bygbjerg [5] an increasing prevalence of hypertension in developing countries, possibly caused by urbanization, ageing of population, changes to dietary habits, and social stress.

According to Opie LH and Seedat [6] Hypertension in Africa has now changed from a relative rarity to a major public health problem. Disease estimates for Sub-Saharan Africa are based on sparse data; however projections indicate increases in non-communicable diseases caused by demographic and epidemiologic transitions. According to Tesfaye, et al., [7] community-based cross-sectional study in urban Addis Ababa showed that the age adjusted prevalence of high blood pressure was 31.5% among males and 28.9% among females. Recently comprehensive assessment of the evidence concerning hypertension in Ethiopia does not exist. However, recent evidences indicate that hypertension and raised blood pressure are increasing partly because of the increase in risk factors. The goal of this study was to compare parameter estimations using both Bayesian and classical approaches and to detect out potential factors that affects survival probability of hypertension patient's under follow up.

2. DATA AND METHODOLOGY

2.1 The Data

The data for this study was collected from 430 randomly selected hypertension patients among a total of 2126 hypertension patients who had been under follow up over the time period from January 2013 to January 2019 at Yekatit-12 Hospital in Ethiopia. The data was extracted from the patient’s chart which contains epidemiological, laboratory and clinical information of all hypertension patients who start follow-up January 2013 to January 2019 at Yekatit-12 Hospital.

The explanatory (independent) variables of interest in this study include demographic factors disease, and medicine related factors, and characteristics of the disease. The response (dependent) variable is continuous; it is length of time of treatment for hypertension patients. It is the waiting time until the occurrence of an event (dead: 1, alive or censored: 0). Observations are censored, in the sense that, for some units, the event of interest has not occurred at the time the data are analyzed. It was calculated in months, taking into account the dates of starting follow up and the occurrence of the event (death) or censoring.

Predictors or explanatory variables which are called covariates are those whose effect on the waiting time we wish to assess. The predictor
3. METHOD OF DATA ANALYSIS

3.1 Parametric Models

In survival analysis, a parametric model that provides an alternative to the commonly used Proportional Hazards (PH) models for the analysis of survival time data. Under parametric models, we measure the direct effect of the explanatory variables on the survival time. The time variable is modeled in this article with several parametric distributions. It was used for multivariate analysis to identify factors associated with death from hypertension. We applied four parametric models (Exponential, Weibull, Lognormal and Loglogistic) and the models are given by Collett [8], Agresti and Hitchcock [9]:

3.2 Bayesian Inference

Bayesian inference expresses the uncertainty of parameters in terms of probability distribution and integrates them out of distribution of interest. In the Bayesian framework, there are three key components associated with parameter estimation: The likelihood function, the prior distribution and the posterior distribution.

3.2.1 Prior and posterior distributions

In a Bayesian approach, model parameters are treated as random variables and assigns probability to each, which is the major difference to the likelihood approach. The assumed distributions for the parameters are called prior distributions. Bayesian estimation and inference is based on the posterior distribution which is the conditional distribution of unobserved quantities given the observed data. The posterior distribution for all unknown parameters $\beta$ is then given by:

$$ f(\beta | T, X) = \frac{f(T | \beta) f(\beta)}{\int f(T | \beta) f(\beta) d\beta} $$

Where: $\beta$’s is the regression coefficients the predictors; $X$ is the predictors covariates; $f(\beta | T, X)$: is the posterior probability distribution; $f(\beta)$: is prior probability distribution and $f(T | \beta)$: is the likelihood distribution [10]. In the Bayesian framework, inference follows from the full posterior distribution. Bayesian inference is then based on samples drawn from the posterior distribution using a Markov Chain Monte Carlo (MCMC) algorithm such as the Gibbs sampler and Metropolis Hastings. The posterior means and variances of the parameters can be estimated based on these samples, and Bayesian inference can then be based on these estimated posterior means and variances. This sampling can be done using Open Bayesian inference Using Gibbs Sampling (BUGS) software. We used Non-informative prior distributions with large variance in our Open BUGS analysis. That is, the priors will have minimal impact relative to the data.

Table 1. Model of parametric distributions

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>$f(t)$</th>
<th>$h(t)$</th>
<th>$S(t)$</th>
<th>log ($T$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>$\lambda$</td>
<td>$\lambda \exp (-\lambda t)$</td>
<td>$\lambda$</td>
<td>$\exp (-\lambda t)$</td>
<td>Extreme value</td>
</tr>
<tr>
<td>Weibull</td>
<td>$\rho, \lambda = \eta^{-\rho}$</td>
<td>$\lambda \rho^{-1} \exp (-\lambda t^\rho)$</td>
<td>$\lambda \rho^{-1}$</td>
<td>$\exp (-\lambda t^\rho)$</td>
<td>Extreme value</td>
</tr>
<tr>
<td>Lognormal</td>
<td>$\mu, \tau$</td>
<td>$\frac{\sqrt{\tau}}{\sqrt{2\pi t}} \exp \left(-\frac{(\ln(t) - \frac{1}{2}(\ln(\mu)))^2}{\ln(\mu)}\right)$</td>
<td>$\frac{f(t)}{S(t)}$</td>
<td>$1 - \Phi \left(\frac{\ln(t) - \mu}{\sqrt{\tau}}\right)$</td>
<td>Normal</td>
</tr>
<tr>
<td>Loglogistic</td>
<td>$\rho, \lambda = \mu^{-\rho}$</td>
<td>$\frac{\lambda \rho t^{-\rho-1}}{(1 + \lambda t^\rho)^2}$</td>
<td>$\frac{\lambda \rho t^{-\rho-1}}{1 + \lambda t^\rho}$</td>
<td>$1$</td>
<td>Logistic</td>
</tr>
</tbody>
</table>
3.3 Diagnostics and Model Selection

Model comparison and selection are important to identify the best model that fit the data among different models. In this study, the model selection procedure was based on the deviance information criteria (DIC), Akaike information criteria (AIC), and Bayesian information criteria (BIC) [11].

For assessing convergence, we have used multiple chains. If parallel chains with varying starting values give the same solution that will increase our confidence for convergence. A simple method of assessing chain convergence is to look at the history of iterations using a time series plot, autocorrelation plot and Gelman-Rubin statistic. If the chains show a reasonable degree of randomness between iterations, it signifies that the Markov chain has found an area of high likelihood and is integrating over the target density [12] and hence indicating that it has converged. In this study, the Statistical Package for the Social Sciences (SPSS) version 20; R version 3.5.3 and Open BUGS were used to analyze the data.

4. RESULTS AND DISCUSSION

The study considered 430 hypertension patients under follow-up at Yekatit-12 Hospital in Ethiopia. Among those patients included in the study 77 (17.9%) experienced the event or died while the remaining 353 (82.1%) are censored. The death proportion of male was 45 (23.4%) which is greater than female patients 32 (13.4%).

Regarding the resident area approximately 88% of the patients live in rural areas and only 12% of the patients reside in urban areas with death proportion 19.6% and 17.7% respectively. Also 256 (59.5%) patients had negative family history of hypertension disease and the rest 144 (40.5%) patients have a positive family history of hypertension case and there death proportion is 19.6% and 14.6% respectively. In addition, the proportion of death was varied by the alcohol consumption of the patient. The highest proportion of death was observed from a patient who consume alcohol (21.5%) whereas the lowest proportion of death (15.8%) was recorded among a patient doesn't use alcohol.

To compare the event experiencing time of two or more groups the survival function used of the groups is good indication. To obtain a closer look at estimate of the survival time we used the Kaplan-Meier estimation technique. The pattern of survivorship function lying above another means the group defined by the upper curve had a better survival than the group defined by the lower curve. Fig. 1 exhibits that there were differences among survival curves of family history of hypertension, Blood cholesterol level and Diabetes Mellitus. Based on log-rank test result, they were significant in survival experience of the patients in khat intake ($\chi^2 = 9.7$ with 1 df, $p = 0.002$), blood cholesterol ($\chi^2 = 5.9$ with 1 df, $p = 0.02$), stage of hypertension ($\chi^2 = 59.7$ with 3 df, $p = 0.00$), adherence ($\chi^2 = 12$ with 1 df, $p = 0.00$) and related disease ($\chi^2 = 29.6$ with 2 df, $p = 0.00$).

4.1 Results of the Parametric Model Comparison

Most often the proportional hazards (PH) models are used for modeling survival data. However, when the PH assumption is violated, accelerated failure time models (parametric regression model) is an alternative approach [8]. The proportional hazard assumption fails as the covariates such as khat intake, age and stage of hypertension are found to be time dependent. Therefore cox proportional hazard model is not appropriate to fit the data so we extend to the parametric regression models.

We applied four parametric models namely exponential, Weibull, Lognormal and Log-logistic models as a parametric distribution model of survival time T. To select the appropriate parametric model for the hypertension patient data the common model comparison and selecting criterion AIC, BIC and DIC were used. Analysis of model comparison of four parametric distributions is given in Table 2. Estimates of total DIC for the four models are 1259.081 Exponential, 1181.56 for Weibull, 1198.91 for Log-normal and 1195.70 for Loglogistic model.

The Weibull model has the smallest DIC, AIC and BIC compare to other models. To estimate the parameter we extend the comparison between classical Weibull and Bayesian Weibull based on the analysis result of Table 2.

The parameter estimation using classical and Bayesian Weibull model are presented in the Table 3. The analysis indicate that baseline age of the patient, family history of hypertension, khat intake, blood cholesterol level of the patient, hypertension disease stage, adherence to the treatment and related disease were significantly associated with survival time of hypertension patients in both classical and Bayesian estimation methods. But the factors gender, Tobacco use and Alcohol use were significant only in Bayesian estimation methods.
The model comparison result in Table 4 shows that Bayesian Weibull models is smaller AIC, BIC and DIC than classical Weibull model. So, this recommends that Bayesian Weibull is the appropriate model to estimate the parameters for the current data set.

![Kaplan-Meier estimate](image1)

![Kaplan-Meier estimate](image2)

![Kaplan-Meier estimate](image3)

**Fig. 1.** Plot of Kaplan-Meier survival function curves of hypertension patients under Yekatit 12 hospital (a) Family History of Hypertension (b) Cholesterol level (c) diabetes Mellitus status

**Table 2.** Model Comparison among parametric models for hypertension data

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Exponential</th>
<th>Weibull</th>
<th>Log-normal</th>
<th>Log-logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>939.783</td>
<td>901.150</td>
<td>909.093</td>
<td>905.721</td>
</tr>
<tr>
<td>BIC</td>
<td>1012.932</td>
<td>978.362</td>
<td>986.305</td>
<td>982.933</td>
</tr>
<tr>
<td>DIC</td>
<td>1259.081</td>
<td>1181.56</td>
<td>1198.91</td>
<td>1195.70</td>
</tr>
</tbody>
</table>
Table 3. Parameter estimation results for Weibull model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classical</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Category</td>
<td>Estimate</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>0.301</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>&lt; 25</td>
<td>1.532</td>
</tr>
<tr>
<td></td>
<td>26-50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>-1.500</td>
</tr>
<tr>
<td>Place of Residence</td>
<td>Urban</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td></td>
</tr>
<tr>
<td>Family History of Hypertension</td>
<td>Positive</td>
<td>-0.665</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>Khat Intake</td>
<td>No</td>
<td>-0.854</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Tobacco use</td>
<td>No</td>
<td>0.426</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Alcohol use</td>
<td>No</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Cholesterol level</td>
<td>Normal</td>
<td>0.518</td>
</tr>
<tr>
<td></td>
<td>Raised</td>
<td></td>
</tr>
<tr>
<td>Stage of Hypertension</td>
<td>Stage 1</td>
<td>1.134</td>
</tr>
<tr>
<td></td>
<td>Stage 2</td>
<td>1.834</td>
</tr>
<tr>
<td></td>
<td>Stage 3</td>
<td>2.610</td>
</tr>
<tr>
<td></td>
<td>Stage 4</td>
<td></td>
</tr>
<tr>
<td>Adherence</td>
<td>Low</td>
<td>-0.703</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Related disease</td>
<td>None</td>
<td>0.798</td>
</tr>
<tr>
<td></td>
<td>Stroke</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heart case</td>
<td>1.564</td>
</tr>
</tbody>
</table>

Table 4. Comparison of classical and Bayesian Weibull Model

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical Weibull</td>
<td>901.150</td>
<td>978.362</td>
</tr>
<tr>
<td>Bayesian Weibull</td>
<td>870</td>
<td>964</td>
</tr>
</tbody>
</table>

Results presented in Table 3 above indicate the parameter estimates of coefficients for the covariates in Weibull model. Survival time of hypertension patients were significantly related with baseline age, family history of hypertension, khat intake, blood cholesterol level, stage of hypertension, adherence and related disease. The hazard rate of the patients who began treatment at ages between 26 up to 50 and above 50 years was 0.216 and 0.223 respectively when keeping other covariates constant. Similarly, patients who had positive family history of hypertension have shorter survival time than patient who had no family history of hypertension. When we see the covariate hypertension stage, mortality is monotonic and the hazard rate increased for patient in stage 4 followed by stage 3 and stage 2. Additionally, the patients in the stage 4, stage 3 and stage 2 of the disease exhibited an increased risk of dying compared to those who in the stage 1. Patient who had raised blood cholesterol level has significantly higher death hazard than patient who had normal cholesterol level. By letting other covariates constant the hazard rates of patient who took khat had been increase by 57.4%. Related disease is also another predictor variable related with risk of death of patients. Patient who had stroke and heart case had higher hazard rate than those who had none of related disease and the estimated hazard ratio is 2.222 and 4.780 respectively.

Comparisons of survival models under different distributions of the hazard function provide the best model for fitting the specific data with appropriate inference. In this study, the Weibull survival model has the smallest DIC, AIC and BIC in both classical and Bayesian methods, indicating its ability to fit the data. Previous
survival studies in southwestern Ethiopia have also recognized the Weibull regression model as the best model for fitting the time until event data on Human Immune-Deficiency Virus (HIV) [13].

In this study the finding regards the association between place of residence of patient and the survival time until hypertension-related death was similar with finding in Ashanti, West Africa reported that place of residence has significant effect on hypertension patients [14]. We have also found a Family history of hypertension was a statistically significant risk factor for death in hypertension patient; this finding similar with finding of other researchers [15,16]. Although this study did not find any association between gender and the survival time until hypertension-related death, other studies in Ethiopia and India have reported that the covariate gender doesn’t have any association with hypertension [17-22]. Christian et al. (2013) indicated Alcohol use is an independent risk factor of hypertension and they also found Hypertension was significantly higher in individuals who take alcohol than those who did not [22].

The finding in this study indicate that the medical factor diabetes mellitus there is association between diabetes mellitus and hypertension related death, possibly because all of the diabetic hypertension patients received opportune of diabetes treatment. Blood cholesterol level was identified to be the significant factor to hypertension related death. The findings of this study showed that a patient who had raised blood cholesterol exhibited a significantly higher death hazard than those who has normal blood cholesterol. Other studies conducted in Delhi, Nepali and Kingdom of Saudi Arabia also indicate blood cholesterol level has an effect on hypertension related mortality [17,21,22] Our study found that khat intake was associated with survival time of hypertension patients. These results are consistent with findings from a study of Ayana Am, et al., [23] they identified khat chewing is one of the main risk factor of hypertension.

4.2 Assessing Convergence

In Bayesian models, three parallel sampling chains of 50000 iterations with different three starting values are generated. Some plots are given in Fig. 2. Inferences are made based on samples of the posterior distributions that are taken with thinning of 10 after burn-in of 10000. Time series plots of the history of the simulations show a reasonable degree of randomness and they may convergence to same values. Auto-correlations and Gelman-Rubin statistics are also used to assess convergences. Finally independent samples are taken from the posterior distribution after convergence of the realization with specified burn-in and thinning values, and then all inferences are made using those samples. Sample of assessments plots are displayed in Fig. 2 from the analysis of Data using the Bayesian Weibull distribution are illustrated. Therefore, the Gibbs sampler has converged to the target density.
Fig. 2. Plots from analysis of Data using the Bayesian Weibull for simulations of parameter of age and khat intake (a) Time Series Plots of (b) Plots of Gelman-Rubin statistics (c) Autocorrelation plots

5. CONCLUSION

The objective of this study was to compare parameter estimations using both Bayesian and classical approaches and to detect out potential factors that affects survival probability of hypertension patient’s under follow up in Yekatit-12 Hospital. The proportional hazard assumption is violated. Because of this fitting cox proportional hazard model is not appropriate for current hypertension data set. Weibull model is the most appropriate model among the parametric models considered in this study for modeling hypertension data.

Model comparison criteria show that Bayesian Weibull models is smaller AIC, BIC and DIC than classical Weibull model and Bayesian Weibull is the appropriate model to estimate the parameters for the current data set.

The analysis both Bayesian and classical approach indicate that baseline age of the patient, family history of hypertension, khat intake, blood cholesterol level of the patient, hypertension disease stage, adherence to the treatment and related disease were significantly associated with survival time of hypertension patients. But the factors gender, Tobacco use and Alcohol use were significant only in Bayesian Weibull model.

In conclusion significant numbers of hypertension patients are found in study area and so greater attention and intervention needed on identified risk factor. We recommend Bayesian estimation methods for to screen out potential factor that affects the survival probabilities of hypertension patients.

6. LIMITATIONS

In this study some potential factors like weight, physical activity, and body mass index are not included. Because of cost constraints, we could not use the whole cohort. Also, patients who start follow-up before January 2013 and start after January 2019 excluded from the study.

CONSENT

For the purpose of confidentiality, there were no linkages with individual patients and all data had no personal identifier and were kept confidential and therefore did not require informed consent.

ETHICAL APPROVAL

The Ethical clearance was checked and approved by ethical clearance committee of Arba Minch University department of Statistics and...
Addis Ababa Administration Health bureau Yekatit 12 Hospital Medical College. The Addis Ababa Administration Health bureau Yekatit 12 Hospital Medical College medical director’s office granted permission to use the patients’ data for this study.

ACKNOWLEDGEMENTS

The authors are grateful to all health professionals, Addis Ababa Administration Health bureau Yekatit 12 Hospital Medical College medical director’s office, Addis Ababa Administration Health bureau Yekatit 12 Hospital Medical College for all support during data collection. We would like to thank hypertensive patients whom their data were used in the study.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. World Health Organization. A global brief on hypertension: world health day; 2013. [Internet].
17. Bonsa F, Gudina EK, Hajito KW. Prevalence of hypertension and associated factors in Bedele Town, Southwest


© 2019 Erango et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/52009