Survival Analysis of Factor Affects Survival Time of Hypertension Patients

Abtsega Samuel Ayalew, Markos Abiso Erango*, Kabtamu Tolosie Gergiso

College of Natural Science, Department of Statistics, Arba Minch University, Arba Minch, Ethiopia
Email: *markos.erango73@gmail.com

Abstract

Hypertension is a major long-term health condition and a leading modifiable risk factor for cardiovascular disease and death. The aim of this study was to examine major factors that affect survival time of hypertension patients under follow-up. We considered a total of 430 random samples of hypertension patients who had been under follow up at Yekatit-12 Hospital in Ethiopia from January 2013 to January 2019. Four parametric accelerated failure time distributions: Exponential, Weibull, Lognormal and loglogistic are used to analyse survival probabilities of the patients. The Kaplan-Meier estimation method and log-rank tests were used to compare the survival experience of patients with respect to different covariates. The Weibull model is selected to best fit to the data sets. The results indicate that the baseline age of the patient, place of residence, 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 factor like gender, tobacco use, alcohol use, diabetes mellitus status and fasting blood sugar were not significantly associated factors. Society and all stakeholders should be aware of the consequences of these factors which can influence the survival time of hypertension patients.

Keywords

Hypertension, Survival Analysis, Parametric Models

1. Introduction

Hypertension, sometimes called arterial hypertension, is a chronic medical condition in which the blood pressure in the arteries is elevated. This requires the heart to work harder than normal to circulate blood through the blood vessels. Blood pressure is summarized by two measurements, systolic and
diastolic, which depend on whether the heart muscle is contracting (systole) or relaxed between beats (diastole). 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].

Hypertension is a major long-term health condition and is the leading cause of premature death among adults throughout the world, including both developed, developing, and lesser developed countries [2].

Hypertension is a worldwide public-health challenge and a leading modifiable risk factor for cardiovascular disease (CVD) and death. According to the 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. 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 [3] [4].

Hypertension doubles the risk of cardiovascular diseases such as coronary heart disease, congestive heart failure, stroke, renal failure and peripheral arterial disease [5]. According to Global burden of cardiovascular disease over a fairly short period is attributable mainly due to changes in lifestyle such as diet and physical activity [6]. A meta-analysis also reported that lower values of blood pressure are associated with higher risk of cardiovascular disease and also with chronic kidney diseases [7].

Hypertension in Africa has now changed from a relative rarity to a major public health problem [8]. Current disease estimates for Sub-Saharan Africa are based on sparse data, but projections indicate increases in non-communicable diseases caused by demographic and epidemiologic transitions; however, hypertension control assumes a relatively low priority and little experience exists in implementing sustainable and successful programs. There is a wide disparity (0.4% to 43%) in the prevalence of hypertension and obesity in Sub-Saharan Africa. The detection rates in most high-income countries vary from 32% - 64% while in many low-income countries, the reported detection rates are substantially lower [9] [10].

Ethiopia is a country currently prioritizing prevention of communicable and nutritional deficiency diseases. However, it experiencing double mortality burden as evidenced among adult population in Addis Ababa and 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. However, only 35.2% of the hypertension subjects were aware of their high blood pressure and only 11% were on treatment with target blood pressure attained in 25.6% [11]. A study conducted in Ethiopia in the last decade showed that the prevalence of cardiovascular diseases risk factor increased rapidly [12]. In addition, a study conducted by Awoke, et al., [13] among 679 subjects aged above 35 years resulted 28% of prevalence of Hypertension whereas done in
Addis Ababa in 2010 on 1935 subjects of working age reported higher prevalence in men 22% than among women 14.9% [12]. A study which conducted in Tikur Anbessa Hospital in Addis Ababa, Ethiopia [14] [15] showed that hypertension is the leading cause of cardiovascular diseases such as stroke and ischemic heart diseases. 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 examine risk factors that affect survival time and comparing survival probabilities of hypertensive patients in Yekatit-12 hospital, Addis Ababa.

2. Data and Methodology

2.1. The Data

The data for this study was collected from hypertension patients under follow up at Yekatit-12 Hospital located in Addis Ababa, Ethiopia from January 2013 to January 2019. The data was extracted from the patient’s chart which contains epidemiological, laboratory and clinical information of all hypertension patients under follow-up. A simple random sampling technique was used to select 430 patients among a total of 2126 hypertension patients under follow up.

2.2. Variables in the Study

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 (covariate) variables which are assumed to influence the survival of hypertension patients included in the model are Gender (Female, male), Age (less than or equal 25, 26 - 50, greater than 50), Place of Residence (Urban, Rural), Family history of hypertension (Negative, positive), Tobacco use (No, Yes), Alcohol use (No, Yes), Khat intake (No, Yes), Related Disease (None, Stroke, Heart case), Blood Cholesterol Level (Normal, raised), Diabetes Mellitus Status (No, Yes), Fasting Blood Sugar (continuous), Stage of Hypertension (Stage 1, Stage 2, Stage 3, Stage 4) and Adherence (Low, High).

2.3. Ethical Consideration

The Ethical clearance was checked and approved by ethical clearance committee of Arba Minch University department of Statistics and Addis Ababa Adminis-
traction 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. 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.

3. Method of Data Analysis

Survival models are important statistical methods to describe and analyze the time-to-death events of hypertension patients. The study focused on time to event (time to death by hypertension), so the appropriate method of this particular study was survival analysis. We have used Kaplan-Meir estimator and parametric hazard model for the analysis and model building. We have also used log-rank tests for comparison of survival functions. Kaplan Meier analysis was used to study survival pattern; the KM plot, which is a step function, gives some indications about the shape of the survival distribution [16]. The figure in general shows if the pattern of one survivorship function lies above another which means the group defined by the upper curve lived longer, or had a more favourable survival experience than the group defined by the lower curve.

3.1. Parametric Regression Models

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 [17] [18] (Table 1).

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

3.2. 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) [19].

4. Results

The statistical packages SPSS version 20 and R version 3.5.3 have been used to analyze the data.
4.1. Summary Statistics

Among 430 samples of hypertension patients considered, 238 (55.3%) were female and 192 (44.7%) were male patients. 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. The majority of the patients were detected at the age of 25 - 50 years (53.9%) followed by age group above 50 years (42.3%) and a few cases at age below 25 years (3.8%). 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.

4.2. Descriptive Survival Analysis

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 survivor function lying above another means the group defined by the upper curve had a better survival than the group defined by the lower curve. Figure 1 exhibits that there were differences among survival curves of sex group, Family History of Hypertension, Age groups, Cholesterol level, Diabetes Mellitus, Stages of hypertension. However, there were not clear survivor probability differences among the categories of covariate Place of residence, Tobacco use and Alcohol use. It means there is no significant difference of survival experience in each group of covariates.

Based on log-rank test result, they were significant in survival experience of the patients in different categories of gender ($\chi^2 = 5$ with 1 df, $p = 0.03$), age group ($\chi^2 = 9.7$ with 2 df, $p = 0.02$), 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, $p = 0.00$). But, they are not significant in survival experience of the patients in different categories of place of residence, family history of hypertension, tobacco use, alcohol use and diabetes mellitus status ($\alpha = 5\%$).

4.3. Results of the Parametric Regression Model

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 [20] [21].
In Figure 2, the plots of scaled Schoenfeld residuals against transformed time for covariates khat intake, age, stage of hypertension and related disease show there are systematic departure from a horizontal line that indicates violation of the proportional hazard assumption [17] [20]. 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, Akaike information criterion (AIC) and Bayesian information criterion were used. The model comparison analysis in Table 2 indicated that Weibull model has smallest AIC and BIC compare to other models.

### 4.3.1. Weibull Regression Model

It is essential to include statistically important and clinically relevant covariates into the model in fitting parametric regression model. We included all available covariates into the model to rank their statistical importance. This is often the case that we have no prior knowledge on which variable should be included.
Figure 2. Plot of scaled-Shefnonfield residuals and their LOESS smooth curves for the covariates (a) khat intake, (b) Age group and (c) Stage of hypertension (d) Related disease.

Table 2. Model comparison among parametric models for hypertension data.

<table>
<thead>
<tr>
<th>Models</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>
</tbody>
</table>

Model development was done with backward elimination on covariates; the method starts with a full model that included all available covariates and then applies Wald test to examine the relative importance of each variable. Statistical significance level for a covariate to stay in a model can be specified.

The parameter estimation using finally Weibull regression model in Table 3 show that, the baseline age of the patient, place of residence, 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.

4.3.2. Interpretation of the Results

Results presented in Table 3 below indicate the parameter estimates of coefficients
Table 3. Parameter estimation for final Weibull regression model.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Category</th>
<th>Parameter</th>
<th>SE</th>
<th>Wald</th>
<th>Sig</th>
<th>Exp(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>&lt;25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26-50</td>
<td>1.532</td>
<td>0.510</td>
<td>8.58</td>
<td>0.003</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>-1.500</td>
<td>0.538</td>
<td>7.29</td>
<td>0.006</td>
<td>0.223</td>
</tr>
<tr>
<td>Place of residence</td>
<td>Urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>0.770</td>
<td>0.349</td>
<td>4.84</td>
<td>0.027</td>
<td>2.160</td>
</tr>
<tr>
<td>Family history of hypertension</td>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>-0.665</td>
<td>0.312</td>
<td>4.45</td>
<td>0.034</td>
<td>0.514</td>
</tr>
<tr>
<td>Khat intake</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>-0.854</td>
<td>0.311</td>
<td>7.23</td>
<td>0.007</td>
<td>0.426</td>
</tr>
<tr>
<td>Blood cholesterol level</td>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Raised</td>
<td>0.518</td>
<td>0.248</td>
<td>4.24</td>
<td>0.039</td>
<td>1.678</td>
</tr>
<tr>
<td></td>
<td>Stage 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stage 2</td>
<td>1.134</td>
<td>0.395</td>
<td>7.12</td>
<td>0.004</td>
<td>3.108</td>
</tr>
<tr>
<td></td>
<td>Stage 3</td>
<td>1.834</td>
<td>0.414</td>
<td>17.13</td>
<td>0.000</td>
<td>6.259</td>
</tr>
<tr>
<td></td>
<td>Stage 4</td>
<td>2.610</td>
<td>0.428</td>
<td>28.72</td>
<td>0.000</td>
<td>13.59</td>
</tr>
<tr>
<td>Stage of hypertension</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adherence</td>
<td>High</td>
<td>-0.703</td>
<td>0.239</td>
<td>6.35</td>
<td>0.003</td>
<td>0.495</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Related disease</td>
<td>Stroke</td>
<td>0.798</td>
<td>0.263</td>
<td>77.84</td>
<td>0.002</td>
<td>2.222</td>
</tr>
<tr>
<td></td>
<td>Heart case</td>
<td>1.564</td>
<td>0.361</td>
<td>14.97</td>
<td>0.000</td>
<td>4.780</td>
</tr>
</tbody>
</table>

for the covariates in the final Weibull regression model along with the associated standard error, significance level and hazard ratio. Survival time of hypertension patients was significantly related with baseline age, place of residence, 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. Keeping other covariates constant, the hazard rate of patient who live in rural area were 2.16 times more than patient who reside in urban area. Similarly, Patients who had positive family history of hypertension have shorter survival time than patient who had no family history of hypertension and the hazard ratio is 0.524.

Patient who had raised blood cholesterol level has significantly higher death hazard than patient who had normal cholesterol level (HR = 1.678). By letting other covariates constant the hazard rates of patient who took khat had been increase by 57.4%.

When we see the covariate hypertension stage, mortality is monotonic and the hazard rate increased for patient in stage 4 (HR = 13.59) followed by Stage 3 (HR = 6.259) and Stage 2 (HR = 3.108). Additionally, the patients in Stage 4,
Stage 3 and Stage 2 of the disease exhibited an increased risk of dying compared to those who in Stage 1.

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. Finally, by keeping other covariate constant the patient with high adherence has $0.495 - 1 = -0.505$ or 50.5% lower hazard than patient with low adherence.

4.3.3. Discussion

Comparisons of survival models under different distributions of the hazard function provide the best model for fitting the specific data with appropriate inference [22]. In this study, the Weibull survival model has the smallest AIC, 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 HIV [23]. This finding also agrees with a study that analyzes parametric models for diabetes mellitus survival data in Addis Ababa [24].

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 a significant effect on hypertension patients [25]. 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 [26] [27]. 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 [28]-[33]. 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 [33].

The current study found that an age was associated with survival time of hypertension patients. These results are consistent with findings from a study in Uganda by (Wamala, et al., [34] but this finding contradictory to a study conducted in Somalia regional state, eastern Ethiopia by [26] [28] [32] indicated that there exists a relationship between hypertension patients and their age.

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 had 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 [28] [32] [33]. Our study found
that khat intake was associated with survival time of hypertension patients. These results are consistent with findings from a study of Andualem, et al., [35] they identified khat chewing is one of the main risk factor of hypertension.

5. Conclusions

The objective of this study was to examine major factors that affect survival time of hypertension patients under follow-up in Yekatit 12 Hospital. A total of 430 patients were considered, 77 (17.9%) experienced the event whereas the remaining 353 (82.1%) are censored. The proportional hazard assumption is violated. Because of this fitting cox proportional hazard model is not appropriate for current hypertension data sets. Weibull regression model is the most appropriate model among the parametric models considered in this study for modeling hypertension data.

This study examined the demographic variables, behavioral and health-related determinants of survival time were related with baseline age, place of residence, family history of hypertension, khat intake, Blood Cholesterol level, Stage of hypertension, Adherence and related disease are significantly affect the survival time the hypertension patients.

The findings indicate that female hypertension patients had better survival probability than male hypertension patients and hypertension patients with age less than 25 year is the lowest survival probabilities when compare other age groups in the study areas. Hypertension patient with raised cholesterol level has lower survival probability as compared to that with a normal cholesterol level.

In conclusion, a significant number of patients found that lack of knowledge about behavioral risk factors of hypertension and so they need great attention. Therefore, teaching the population or patients about the effect of behavioral risk factor of hypertension like alcohol use, tobacco use and khat intake is highly recommended and all stakeholders in synergistic approach towards awareness creation regarding hypertension, promotion of healthy lifestyle and improving health checkups among the community and early screening of those who have a family history of hypertension should be suggested. Based on the identified risk factors, appropriate interventions and implementation of community-based screening programs for early detection of hypertension and treating related disease of hypertension are recommended.

Acknowledgements

The authors would like to sincerely thank the Yeketit12 Hospitals for providing us the data sets used in this study. We acknowledge the anonymous reviewers for their detailed comments and suggestions.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.
References


