International Journal of Public Health and Epidemiology
IJPHER
Vol. 5(1), pp. 071-082, January, 2019. © www.premierpublishers.org, ISSN: 1406-089X

Research Article

Modeling Time to Death of HIV Infected Patients on Antiretroviral Therapy in case of Hossana Queen Elleni Mohammad Memorial Hospital, South Ethiopia

***Adinew Arficho Handiso[1], Yasin Negash[2], Getachew Tekle Mekiso[3]**
[1]Department of Statistics, Wolaita Sodo University P.O.Box 138, Wolaita Sodo, Ethiopia
[2]Department of Statistics, Jimma University P.O.Box 378, Jimma, Ethiopia
[3]Department of Statistics, Wachemo University P.O.Box 667, Hossana, Ethiopia

The main aim of this study was modeling the factors that affect survival time of HIV infected patients by using Cox ph and parametric survival regression models. The analytical methodologies were used the Kaplan-Meier and Log Rank Test to estimate Descriptive analysis, Cox's regression model was employed to identify the covariates that have a statistical significant effect on the survival time of HIV infected patients and exponential, weibull, log logistic and log-normal survival regression models were applied to compare efficiency of the models. The overall mean estimated survival time of patients was 51.5 months. The Cox Proportional Hazards Regression Model result revealed that baseline weight, ART adherence, baseline CD4 count, WHO clinical stage, level of education, substance use and TB co-infection of patients are the major factors that affect significantly survival time of HIV infected patients. Among the parametric regression models, based on model Comparison methods, the Weibull regression model is better fit. The Weibull regression model results revealed that baseline weight<50 kg, low CD4 count at baseline, no education, WHO stages III and IV, poor ART adherence, co-infection with TB and substance abuse are the categories that reduce the survival probability of HIV infected patients.

Key Words: Survival analysis, Cox Proportional Hazard Regression model, Weibull Regression Model, Hazard ratio, HIV

INTRODUCTION

Background of the study

A pattern of highly unusual infection in otherwise healthy young adults emerged in the early 1980s in the United States of America. This pattern or clusters of diseases that appeared in those whose immune system being attacked came to be called Acquired Immune Deficiency Syndrome (AIDS). Between the 1983 and 1994 a new virus called Human Immunodeficiency Virus (HIV) has been identified as a cause of AIDS (UNAIDS, 2005). Human Immunodeficiency Virus (HIV) is the virus that causes Acquired Immune Deficiency Syndrome (AIDS). People are said to be HIV positive when the HIV antibody is detected in their blood. HIV attacks and destroys certain types of white blood cells that are essential to body's immune system, the biological ability of the human body to fight infections.HIV infects primarily vital cells in the human immune system such as helper T cells (to be specific, CD4+ T cells), macrophages, and dendritic cells that are necessary to activate B-lymphocytes and induce the production of antibodies. The infected person becomes susceptible to a wide range of opportunistic infections, such as tuberculosis and Pneumocistic Carinii Pnemonia, and rare cancer such as Caposis Sarcoma (WHO, 2007). From the total number of people who have died due to HIV/AIDS in 2006 alone was 88,997 and in 2007 it was estimated that 71,902 people would die (FMOH, 2007). In 2010, AIDS related death is expected to decline to 28,073

***Corresponding Author**: Adinew Arficho Handiso, Department of Statistics, Wolaita Sodo University P.O. Box 138, Wolaita Sodo, Ethiopia.
Email: adinew.handiso@yahoo.co.uk.

which might be as a result of ART. Currently an estimated 1,217,903 people are living with HIV/AIDS. It is estimated that 398,717 of the HIV positive cases are in need of ART out of which 26,053(6.5%) are children under 15 years of age. It is also estimated that the all ages HIV prevalence in SNNPR in 2013 is0.9% with 18,557 male and 27,221 female cases who live with the virus, currently an estimated 45,778 people are living with HIV/AIDS this may increase the number of HIV positive patients in the region (NAIDSR, 2014).

Hadiya zone is one of South Nations, Nationalities and Peoples Region (SNNPR), Ethiopia. SNNPR is one of the largest regions in Ethiopia, accounting for more than 10 percent of the country's land area and the current population is approximately 17 million with an average household size of 4.8 in 2007. More than 91 percent of the SNNPR population lives in rural areas. The mid-2012 population was estimated at nearly 17,745,000. The region is divided into 13 administrative zones including Hadiya zone. Hadiya Zone has 10 woradas and one town administration with an estimated total population of 1.5 millionin2013. It has one zonal hospital, 37 functional public health centers and 282 health posts among which ten health centers and one zonal hospital provide a total of 2899 HIV infected patients have visited ART clinic, 2039 ever started ART of which 258 have died *(*HZHD, 2014).

Statement of the Problem

Today, Ethiopia has made progress in reducing the number of HIV/AIDS death nationally, but the observed changes are not sufficient enough compared to the desired goals of the response against the epidemic. Investigating the existence of significant associations between the different factors and HIV/ADIS mortality can provide evidence for informed protection mechanisms. Most of the researches conducted previously in Ethiopia focused more on the prevention of people from infection by HIV/ADIS (NAIDSRC. 2010), but it seems that little attention has been given to study high risk factors that facilitate mortality of those people living with HIV/AIDS. Furthermore, modeling time to death of HIV infected patients on ART is helpful to identify covariates that facilitate mortality of those people living with HIV/AIDS (Leigh *et al.,* 2009).

In addition, a study conducted previously in HQEMMH used the Multilevel logistic regression model (Gemechu, 2013), but the Multilevel logistic regression model is not well suited to survival data for several reasons. According toCollett work, the survival times are not normally distributed and the censored data are the result of missing values on the dependent variable, but in this study the survival analysis method has been used to identify the risk factors as well as to compare the efficiency of Cox phand parametric survival regression models (Collett, 2003). Many covariates will collect to reduce possible modeling bias, when a large semi parametric/parametric model is built. An important and the first challenging task are to efficiently select a subset of significant variables upon which the hazard function depends (Hosmer and Lemeshow, 1999). In general, the motivation behind this study is intended to address the following two major research questions:

- Which factors significantly affect survival time of HIV infected patients over ART?
- Which type of survival model, Cox ph or parametric regression model, predicts well the covariate that are associated with high risk of mortality?

Objectives of the study

The main objective of this study is modeling the factors that affect survival time of HIV infected patients by using Cox ph and parametric survival regression models based in HQEMMH.

DATA SOURCE AND METHODOLOGY

Data Source

This study is a retrospective cohort study based on data from the ART clinical in Hossana Queen Elleni Mohamad Memorial Hospital (HQEMMH), Hadiya Zone, SNNP Region of Ethiopia. The survival data were extracted from the patient's chart which contains epidemiological, laboratory and clinical information of HIV patients under ART follow-up including a detailed antiretroviral therapy history.

The study was conducted in Hossana Queen Elleni Mohamad Memorial Hospital, SNNPR, *and* Ethiopia, from1st February 2011 to 1stJanuary 2016. Hadiya zone is one of 13 zones in SNNPR. There are 10 woreda as and one town administration in the zone and Hosanna town its administrative center which is 235 km away from Addis Ababa. In the town there is one hospital and three health center which gives preventive, curative and rehabilitative service for the population. The hospital has a separate ART clinic and the clinic has one doctor, one nurse, one pharmacist and two data clerks.

The population of the study was All HIV positive patients who were 15 years old and above placed under ART in between 1stFebruary 2011 to 1stJanuary 2016in Hossana Queen Elleni Mohamad Memorial Hospital. This study was based on a review of the patients' intake forms and follow-up cards of HIV patients. For uniformity use in the country so that those forms can be used to document almost all relevant clinical and laboratory variables. In this study were a total of 933 HIV infected patients were investigated who ever started ART.

The data were extracted from the available standard national medical registers which have been adopted by Federal Ministry of Health (FMOH) to be uniformly used by clinicians to simply identify and document clinical and laboratory variables. The registers include pre-ART register and follow up form, ART intake form, patients' card and death certificate complemented registration by home visitors. The data were collected by data clerks working in the clinic and coded and analyzed using the statistical packages STATA and R.

Methodology of the Study

Methods of Survival Analysis

Survival analysis is an important statistical technique used to describe and model time to event data. The purpose of survival analysis is to model the underlying distribution of the failure time variable and to assess the dependence of the failure time variable on covariates. The term survival analysis suggests that the event is death, but that is not necessarily so. Events could also denote success, such as recovery from therapy. Survival time then describes the time from a certain origin to the occurrence of an event.

Descriptive Methods for Survival Data

In any applied setting, a statistical analysis should begin with description of the data. In particular, an initial step in the analysis of a set of survival data is to present numerical or graphical summaries of the survival times in a particular group. Routine applications of standard measures of central tendency and variability will not yield estimates of the desired parameters when the data include censored observations. In summarizing survival data, the two common functions of applied are the survivor function and the hazard function (Hosmer and Lemeshow, 1999).

The survivor function, S (t), is defined for both discrete and continuous distribution as the probability that an individual survives beyond time t i.e., for continuous random variable T, the density function, f (t), is given by

$F(t) = \mathrm{P}\,(\mathrm{T} < t) \;= \int_0^t f(u)du\,, \mathrm{t} \geq 0$

Which represents the probability that a subject selected at random will have a survival time less than some stated value t. Then, the survival function $S(t)$ is defined as:

$$S(t) = \mathrm{P}\left(\mathrm{T} \geq \mathrm{t}\right) = 1 - F(t)$$

The survivor function can be used to represent the probability that an individual survives from the time origin to sometime beyond t and then relationship between the probability density function f(t) and S(t) will be:

$f(t) = \frac{d(1-S(t))}{dt} = \frac{-\mathrm{dS(t)}}{\mathrm{dt}}$(3.3)

The hazard function $\mathrm{h(t)}$is defined by:

$$\mathrm{h(t)} = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathrm{p\{an\ individual\ fails\ in\ the\ time\ interval(t,t+\Delta t)\backslash it\ survived\ until\ time\ t\}}}{\Delta \mathrm{t}}$$

$$\mathrm{h(t)} = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathrm{P}\,[\mathrm{t} < T < t + \Delta \mathrm{t} \backslash \mathrm{T} \geq \mathrm{t}]}{\Delta \mathrm{t}}$$

By applying the theory of conditional probability and the relationship in equation (3.4), the hazard function can be expressed in terms of the underlying probability density function and the survivor function as follows (Collett, 2003).

$$\mathrm{h(t)} = \frac{f(t)}{S(t)} = -\frac{d}{dt}\{\log S(t)\}$$

The corresponding cumulative hazard function $\mathrm{H}(t)$is defined by:

$$H(t) = \int_0^t h(u)du = -\log S(t)$$

Hence the survival function can be rewritten as

$$S(t) = \exp\{-H(t)\}$$

Kaplan-Meier Estimator of the Survival Function

The Kaplan-Meier (KM) estimator proposed by Kaplan and Meier (1958) is the standard non parametric estimator of the survival function$(\mathrm{Collett,\ 2003})$. Which is also called the Product-Limit estimator incorporates information from all observations available, both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times. The Kaplan-Meier estimator is used to estimate the survival time (time of censoring) of a patient and construct survival curves to compare the survival experience of a patient between different categorical variables. The first step in the analysis of ungrouped censored survival data is normally to obtain the Kaplan-Meier estimate of the survivor function.

Then the Kaplan-Meier estimator of the survival function at time t is given by:

$\hat{\mathrm{S}}(\mathrm{t}) = \prod_{j=1}^{r} \left\{\frac{n_j - d_j}{n_j}\right\}$(3.8) for $\mathrm{t_{(k)}} \leq \mathrm{t_{(k+1)}}$,

j=1,2,…,r, with $\hat{\mathrm{S}}(\mathrm{t})=1$ for $\mathrm{t} < \mathrm{t_{(1)}}$.

Where, n_j is the number of individuals who are at risk of dying at time $\mathrm{t_j}$and d_j is the number of individuals who failed (died) at timet_j. The variance of Kaplan-Meier survival estimator is estimated using Greenwood's formula (Collett, 2003) given as:

$\mathrm{var}\left(\hat{s}(t)\right) = (\hat{s}(\mathrm{t}))^2 \sum_{\mathrm{j}=1}^{\mathrm{r}} \frac{d_\mathrm{j}}{n_j(n_\mathrm{j} - d_j)}$(3.9)

Modeling Survival Data

Both the non-parametric methods defined earlier are examples of Univariate analysis; they describe the survival with respect to the factor under investigation, but necessarily ignore the impact of any others. In clinical investigations it is more common to have a situation where covariates potentially affect patient forecast. When investigating survival in relation to any one factor, it is often

desirable to adjust for the impact of others. Moreover, while the log-rank test provides a P-value for the differences between the groups, it offers no estimate of the actual effect size.

Fitting the Cox Proportional Hazard Regression Model

Fitting the Cox model to observed survival data requires estimating the unknown regression coefficients ($\boldsymbol{\beta}$). Also, the baseline hazard function must be estimated. It turns out that these two components of the model can be estimated separately. The coefficients should be estimated first and the estimates are then used to construct an estimate of the baseline hazard function. The regression coefficients in the proportional hazards Cox model, which are the unknown parameters in the model, can be estimated using the method of maximum likelihood (Collett, 2003).

In Cox proportional hazards model we can estimate the vector of parameters **β** without having any assumptions about the baseline hazard,$h_0(t)$. As a consequence, this model is more flexible and an estimate of the parameters can be obtained easily.

Maximum Likelihood Estimation

Suppose the survival data based on n independent observations are denoted by the triplet (t_i, δ_i, $\mathbf{X_i}$), i=1, 2...n.

Where
t_i- the survival time for the i[th] individual.
δ_i- an indicator of censoring for the i[th] individual. Given by i=0 for censored and i= 1 for event experience
$\mathbf{X_i}$ = (X_{i1}, X_{i2}...X_{im})'- column vector of m covariates for individual i.
The full likelihood function for right censored data can be constructed as:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} h(ti, \mathbf{X}i, \boldsymbol{\beta})^{\delta i} S(ti, \boldsymbol{X}i, \boldsymbol{\beta})$$

Where,$h(ti, \mathbf{X}i, \boldsymbol{\beta})$=$h_0(ti)e^{\boldsymbol{\beta}'\boldsymbol{X}i}$ is the hazard function for the i[th] individual.

$S(ti, \mathbf{X}i, \boldsymbol{\beta})$ = $[So(ti)]^{\exp(\boldsymbol{\beta}'\boldsymbol{X}_i)}$ is the survival function for the i[th] individual. It follows that, $L(\boldsymbol{\beta})$

$$= \prod_{i=1}^{n} \left[h_0(ti)e^{\boldsymbol{\beta}'\boldsymbol{X}i}\right]^{\delta i} [So(ti)]^{\exp(\boldsymbol{\beta}'\boldsymbol{X}_i)}$$

The full maximum likelihood estimator of **β** can be obtained by differentiating the right hand side of equation (3.20) with respect to the components of **β** and the base line hazard, $h_0(t)$.

This implies that unless we explicitly specify the base line hazard, $h_0(t)$, we cannot obtain the maximum likelihood estimators for the full likelihood. To avoid the specification of the base line hazard, Cox (1972) proposed a partial likelihood approach that treats the baseline hazard as a nuisance parameter remove it from the estimating equation.

Partial Likelihood Estimation

Instead of constructing a full likelihood, we consider the probability that an individual experiences an event at time $t_{(i)}$ given that an event occurred at that time. Suppose that data are available for n individuals, amongst them there are r distinct failure times and n - r right-censored survival times, and assume that only one individual was died at each ordered failure time, so that there are no ties. The r ordered failure times will be denoted by $t_{(1)}<t_{(2)}<\ldots.<t_{(r)}$, so that $t_{(i)}$ is the i[th] ordered failure time. The set of individuals who are at risk at time $t_{(i)}$ is the i[th] ordered failure (experiences an event) time, and denoted by R ($t_{(i)}$). And let $\mathbf{X}_{(i)}$ be the vector of explanatory variables for an individual who experiences an event at $t_{(i)}$.

The partial likelihood function is derived by taking the product of the conditional probability of a failure at time $t_{(i)}$, given the number of individuals who are at risk of experiencing the event at time $t_{(i)}$.Then, P(j[th] individual will experience an event at time $t_{(i)}$)= $\frac{\exp(\boldsymbol{\beta}'\boldsymbol{X}(i))}{\sum_{j\in \mathrm{R}\,(\mathrm{t(i)})}\exp(\boldsymbol{\beta}'\boldsymbol{X}j)}$

Where, the summation in the denominator is over all individuals in the risk set. Thus the partial likelihood is the product over all event time t(*i*) for *i*= 1,2,...,n of the conditional probability (3.21) to give the partial likelihood function and can be expressed in the form:-

$$L_p(\boldsymbol{\beta}, \boldsymbol{X}(i)) = \prod_{i=1}^{n}\left[\frac{\exp(\boldsymbol{\beta}'\boldsymbol{X}(i))}{\sum_{j\in \mathrm{R}\,(\mathrm{t(i)})}\exp(\boldsymbol{\beta}'\boldsymbol{X}j)}\right]^{\delta i}$$

The expression assumes that there are no tied times, and designed in such a way that it excluded terms when δ_t= 0, as a result the equation in (3.22) becomes. The product is over the r distinct ordered survival times. The corresponding log-partial likelihood function is given by:

$$logL_p(\boldsymbol{\beta}, \boldsymbol{X}(i)) = \sum_{i=1}^{n}\left\{\boldsymbol{\beta}'\boldsymbol{X}(i) - log\left[\sum_{j\in \mathrm{R}\,(\mathrm{t(i)})}\exp(\boldsymbol{\beta}'\boldsymbol{X}j)\right]\right\}$$

The maximum likelihood estimates of the regression parameters in the proportional hazards model can be found by maximizing the log-likelihood function in equation (3.23) using numerical methods. This maximization is accomplished using the Newton-Raphson procedure (Collett, 2003).The Newton-Raphson procedure is used to maximize the partial likelihood function based on the following iterative procedure. An estimate of the vector of **β**-parameters at the *(s+1)[th]* cycle of iterative procedure, $\widehat{\boldsymbol{\beta}}_{s+1}$, is given by:
$\widehat{\boldsymbol{\beta}}_{s+1}$=$\widehat{\boldsymbol{\beta}}_s$+$\mathbf{I}^{-1}(\widehat{\boldsymbol{\beta}}_s)\mathbf{U}(\widehat{\boldsymbol{\beta}}_s)$, for s = 0, 1, 2, ...

$$\mathrm{U}(\widehat{\boldsymbol{\beta}}_s) = \left(\frac{\partial logL_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_1}, \ldots\ldots\ldots, \frac{\partial logL_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_p}\right)$$

Where $\mathbf{U}(\widehat{\boldsymbol{\beta}}_s)$ is the $p \times 1$ vector of first derivatives of the log-likelihood function in equation (3.23) with respect to the $\boldsymbol{\beta}$-parameters and this quantity known as the vector of efficient scores evaluated at $\widehat{\boldsymbol{\beta}}_s$. $\mathbf{I}(\widehat{\boldsymbol{\beta}}_s) = -\frac{\partial^2 \log L_p(\boldsymbol{\beta})}{\partial \beta_j \partial \beta_k}$ is the $p \times p$ matrix and known as observed information matrix.

$$I(\widehat{\boldsymbol{\beta}}_s)_{pxp} = -\begin{bmatrix} \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_1^2}, \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}_2}, \dots, \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}_p} \\ \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_2 \partial \boldsymbol{\beta}_1}, \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_2^2}, \dots, \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_2 \partial \boldsymbol{\beta}_p} \\ \vdots \\ \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_p \partial \boldsymbol{\beta}_1}, \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_p \partial \boldsymbol{\beta}_2}, \dots \dots, \frac{\partial^2 log L_p(\boldsymbol{\beta}, \boldsymbol{X}(i))}{\partial \boldsymbol{\beta}_p^2} \end{bmatrix}$$

$\mathbf{I}^{-1}(\widehat{\boldsymbol{\beta}}_s)$ is the inverse of the observed information matrix evaluated at $\widehat{\boldsymbol{\beta}}_s$isthe variance-covariance matrix of $\widehat{\boldsymbol{\beta}}, var(\widehat{\boldsymbol{\beta}})$, can be approximated by the inverse of the information matrix evaluated at $\widehat{\boldsymbol{\beta}}_s$ i.e. $\mathbf{I}^{-1}(\widehat{\boldsymbol{\beta}}_s)$.

The partial likelihood derived above is valid when there are no ties in the data set. But in most real situations tied survival times are more likely to occur. In addition to the possibility of more than one experience an event at a time, there might also be more than one censored observations at a time of event. To handle this real-world fact, partial likelihood algorithms have been adopted to handle ties. There are three approaches commonly used to estimate regression parameters when there are ties. These are Breslow (1974), Efron (1977) and Cox (1972) approximations (Collett, 2003). The most popular and easy approach is Breslow's approximation. In many applied settings there will be little or no practical difference among the estimators obtained from the three approximations. Because of this, and since the Breslow approximation is more commonly available, otherwise, analysis presented in this study was based on it.

Parametric Survival Regression Models

In the analysis of survival data, survival models can also be used in addition to hazards model. One advantage of such models is that the proportionality assumption of the hazards is not required. The parametric survival regression models work analogous to the multiple linear regression of logarithm of survival time on explanatory variables. Such survival models are termed as parametric accelerated failure time models or simply AFT models. Because these models work on survival, the complementary concept of hazard, the sign of the regression coefficients in an AFT model will be opposite to those in PH models (Klein and Moeschberger 1997).

Most commonly used parametric Survival Regression models are Exponential, Weibull, Log-Logistic and Log-normal. Exponential and Weibull parametric models can work both in PH and in AFT models. These models are equally appropriate viewed in either model. And one can transform regression coefficients computed in PH model into the regression coefficient in AFT model or vice versa for Exponential and Weibull parametric survival models. That means:

- For exponential $\beta_i = -\alpha_i$, the exponential PH and AFT are in fact the same model, except that the parameterization is different, hence HR=exp $(-\alpha_i)$ is the hazard ratio of the i[th] group with the reference groups.
- For weibull, $\beta_i = -\rho\alpha_i$, where ρ is the shape parameter and hence, HR=exp $(-\rho\alpha_i)$ is the hazard ratio of the i[th] group with the reference groups.

Other parametric survival models such as Log-Logistic and Log-normal work only in AFT model as these models do not fit into the proportional hazards frame work.

Fitting parametric Survival Regression Models

The survival likelihood for Weibull distributed survival data with event times and right censored data is generally given by

$$L = \prod_{i=1}^{n} \left\{ (t_i \lambda \rho x_i{}^{\rho-1} \exp(-\lambda x_i{}^{\rho}))^{\delta_i} (\exp(-\lambda x_i{}^{\rho}))^{1-\delta_i} \right\}$$

Resulting in the log likelihood function

$$l = d \log(\lambda\rho) + (\rho - 1) \sum_{i=1}^{n} \delta_i log x_i - \lambda \sum_{i=1}^{n} x_i^{\rho}$$

with $d$the total number of events. Maximum likelihood estimators can be obtained by equating the first derivatives of l with respect to λ and ρ to zero and we get.

$$\hat{\lambda} = \frac{d}{\sum_{i=1}^{n} x_i{}^{\hat{\rho}}} \text{ and } \frac{d}{\hat{\rho}} + \sum_{i=1}^{n} \delta_i log x_i - \frac{d}{\sum_{i=1}^{n} x_i{}^{\hat{\rho}}} \sum_{i=1}^{n} x_i^{\rho} \, log x_i = 0$$

which is nonlinear in $\hat{\rho}$ and can only be solved by a numerical procedure such as the Newton Raphson algorithm.

The likelihood function is derived from the log-linear function of the model defined in equation (3.51).The likelihood function of n observed survival times, $t_1, t_2, \dots, t_n$ for the log-linear form of the parametric Survival Regression model is given by

$$L(\beta, \mu, \sigma) = \prod_{i=1}^{n} [f_i(t_i)]^{\delta_i} [S_i(t_i)]^{(1-\delta_i)}$$

Where $f_i(t_i)$ and $S_i(t_i)$ are the density and survival functions for the i^{th} individual at time t_i and δ_i is the event indicator for the observation and has value zero for censored and one for uncensored individuals. If $f_{\varepsilon_i}(z_i)$ and $S_{\varepsilon_i}(z_i)$ are probability density function and survival function respectively of the random variable $\boldsymbol{\varepsilon_i}$ in equation (3.54) in such a way that

$S_i(t_i) = S_{\varepsilon_i}(z_i)$and $f_i(t_i) = \frac{1}{\sigma t_i} f_{\varepsilon_i}(z_i)$

Where,$z_i = \left(\frac{\log t_i - (\mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_P x_{Pi})}{\sigma}\right)$

The resulting likelihood function using survival function and density function of assumed probability distribution represented by random variable ε_i is as follows:

$$L(\beta, \mu, \sigma) = \prod_{i=1}^{n} (\sigma t_i)^{-\delta_i} \left[f_{\varepsilon_i}(z_i)\right]^{\delta_i} \left[S_{\varepsilon_i}(z_i)\right]^{(1-\delta_i)}$$

The log-likelihood function is:

$$\log\left(L(\beta, \mu, \sigma)\right) = \sum_{i=1}^{n} \left\{-\delta_i \log(\sigma) + \delta_i \log f_{\varepsilon_i}(z_i) + (1 - \delta_i) \log S_{\varepsilon_i}(z_i)\right\} - \sum_{i=1}^{n} \delta_i \log t_i$$

The term $(-\sum_{i=1}^{n} \delta_i \log t_i)$ is omitted as it does not involve any unknown parameters. Hence the full log-likelihood function is given by

$$\log\left(L(\beta, \mu, \sigma)\right) = \sum_{i=1}^{n} \left\{-\delta_i \log(\sigma) + \delta_i \log f_{\varepsilon_i}(z_i) + (1 - \delta_i) \log S_{\varepsilon_i}(z_i)\right\}$$

The maximum likelihood estimates of the parameters are estimated by using iterative Newton-Raphson procedure.

RESULTS AND DISCUSSION

Descriptive Analysis of HIV patients

The study included 933 HIV patients, who started ART in Hossana Queen Elleni Mohamad Memorial Hospitalbetween1stFebruary 2011 to 1stJanuary 2016. Among those patients 15.3% were dead cases and the rest 84.7% were censored. The baseline socio-demographic variables of the cohort are summarized in Table 4.1. Out of these patients 501(53.7%) were females, death proportion were 14.2%. In case of age 530(57%) of the patients were less than 40 years old, death proportion were 13.2%. The residence of the patients were 481(51.5%) lived in rural out of Hossana town, death proportion were 15.8%. The mean survival time of patients based on different socio-demographic and clinical variables are summarized in Table 4.1below

Table 1: Summary of descriptive statistics for Socio-Demographic and clinical Variables

Covariates	categories	Status of patients				
		Total	Number of censored	Number Of death	**Mean**	95% CI
Base line weight	less than 50kg	363(39%)	293(80.7%)	70(19.3%)	48.99	(46.84 , 51.16)
	50kg or above	570(61%)	497(87.2%)	73(12.8%)	52.98	(51.55 , 54.42)
Marital status	Single	226(24.2%)	187(82.7%)	39(17.3%)	47.84	(44.58 , 51.10)
	Married	526(56.4%)	453(86%)	73(14%)	52.78	(51.32 , 54.24)
	Divorced	80(8.6%)	63(79%)	17(21%)	48.57	(44.23 , 52.93)
	Windowed	59(6.3%)	50(84.7%)	9(15.3%)	51.86	(47.68 , 56.04)
	Separated	42(4.5%)	37(88%)	5(12%)	54.86	(50.58 , 59.14)
Functional status	Working	608(65%)	529(87%)	79(13%)	52.71	(51.28 , 54.15)
	Ambulatory	242(26%)	195(80.6%)	47(19.4%)	49.28	(46.74 , 51.83)
	Bedridden	83(9%)	66(80%)	17(20%)	49.96	(45.96 , 53.96)
Drug regimen	D4T-3TC-NVP	255(27%)	207(81.2%)	48(18.8%)	49.80	(47.46 , 52.14)
	AZT-3TC-NVP	280(30%)	238(85%)	42(15%)	51.23	(48.91 , 53.57)
	TDF-3TC-EFV	398(43%)	345(86.7%)	53(13.3%)	52.68	(50.92 , 54.45)
Education level	no education	234(25%)	183(78%)	51(22%)	47.35	(44.68 , 50.06)
	Primary	383(41%)	324(84.6%)	59(15.4%)	52.09	(50.32 , 53.86)
	secondary and above	316(34%)	283(89.6%)	33(10.4%)	53.86	(51.93 , 55.80)
ART Adherence	Poor	174(18.6%)	132(75.8%)	42(24.2%)	49.38	(47.22 , 51.54)
	Good	759(81.4%)	658(86.7%)	101(13.3)	52.71	(51.27 , 54.15)
WHO clinical stage	stage I	263(28%)	233(88.6%)	30(11.4%)	53.20	(51.06, 55.36)
	stage II	279(30%)	246(88.2%)	33(11.8%)	53.70	(51.72, 55.68)
	stage III	295(31.6%)	236(80%)	59(20%)	47.76	(45.46, 50.07)
	stage IV	96(10.4%)	75(78%)	21(22%)	46.40	(42.59 , 48.22)
Base line CD4 cell counts	less than 200	426(45.7%)	346(81%)	80(19%)	50.32	(48.52 , 52.13)
	200 or above	507(54.4%)	444(87.6%)	63(12.4%)	52.59	(50.96 , 54.23)
Substance use (alcohol, soft drugs)	No	761(81.6%)	663(87%)	98(13%)	52.41	(51.06, 53.77)
	Yes	172(18.4%)	127(73.8%)	45(26.2%)	48.47	(45.76 , 51.18)
TB co-infection	No	743(79.6%)	642(86.4%)	101(13.6%)	52.86	(51.54 , 54.18)
	Yes	190(20.4%)	148(78%)	42(22%)	47.13	(44.32, 49.94)
Over All			84.7%	15.3%	51.50	(50.30 ,52.73)

Source: Hossana Queen Elleni Mohamad Memorial Hospital, SNNPR, Ethiopia; from 1st February 2011 to 1st January 2016; mean: mean survival time, CI: Confidence Interval for mean

Comparison of Survival Experience

The Kaplan-Meier survivor estimator is used to investigate the significance differences between the survival probabilities of different categories. In this study overall graph of the Kaplan-Meier survivor function showed that relatively small number of the deaths occurred in the earlier months of ART treatment which given in Figure 4.

Results of the Cox proportional hazards Regression Model

In order to study the relationship between survival time and covariates, a regression modeling approach to survival analysis using the Cox proportional hazards model can be employed for estimating the regression coefficients, making interpretation based on the hazard function, conducting statistical tests, constructing confidence intervals, checking the adequacy of model and its development precede interpretation of results obtained from the fitted model.

Consequently, the most important subset of these predictors to be included in the multivariable model will be selected by stepwise procedure, which based on their contribution to the maximized log partial likelihood of the model (-2*LL)*. The summary result indicates that the highest reduction in - 2*LL*(bˆ) is observed for drug regimen that reduced the value for the null/empty model, from 1707.449 to 1655.784, the difference is 51.66 and the next highest change is obtained for functional status of (48.789) followed by marital status (42.021). Therefore, all the covariates will be included in the multivariate study. The next step is to check the significance of the covariates in the multivariable model. The covariates which are not significant at 5% significance level, then those covariates eliminated from the model. Lastly, the final Cox ph regression model is fitted in Table 2 using the remaining significant covariates.

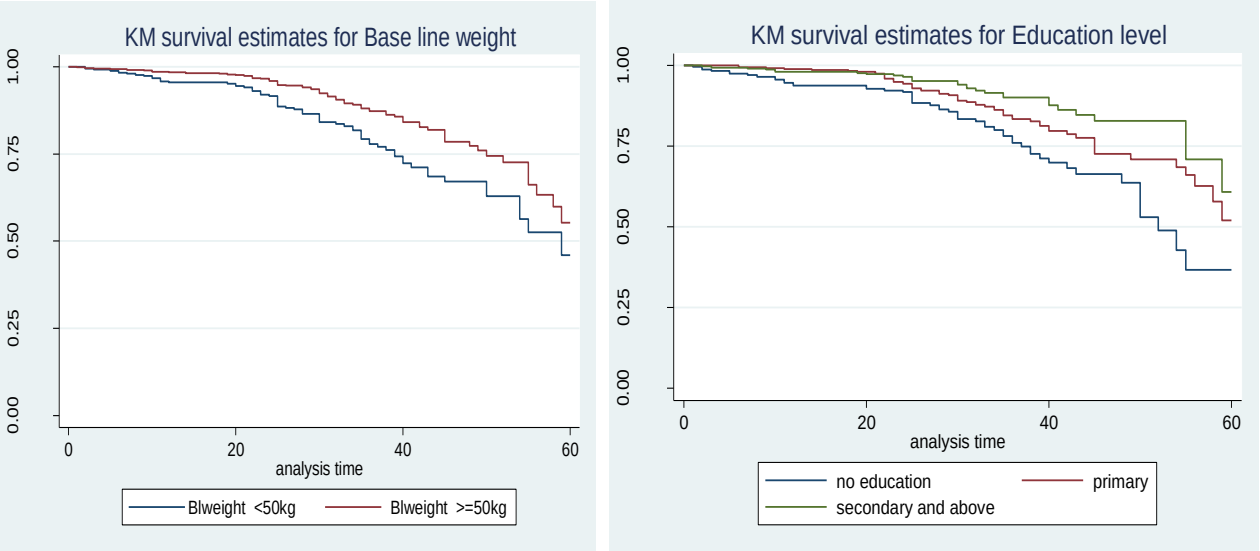


Figure 1: Plots of Kaplan-Meier survivor function estimates for the variable baseline weight and education level

Overall Goodness of Fit

The final step in the model assessment is to measure the overall goodness of fit. For this objective the study use the Cox-Snell residuals, R^2 and Likelihood Ratio, Score and Wald tests. Plot of the Cox-Snell residuals was applied to test the overall fit of the model. The plot of the Nelson-Aalen estimate of the cumulative hazard function against the Cox-Snell residuals is presented in Figure 4.4 below. It can be seen that the plot of the residuals in Figure is almost close to the 45^0 straight line through the origin. Thus, the plot is evidence that the model fitted to the data is satisfactory. However, there is little evidence of a systematic deviation from the straight line at the left, this can be expected even if we have a well-fitting Cox model because of the reduced effective sample size caused by prior failures and censoring (Khanal 2009).

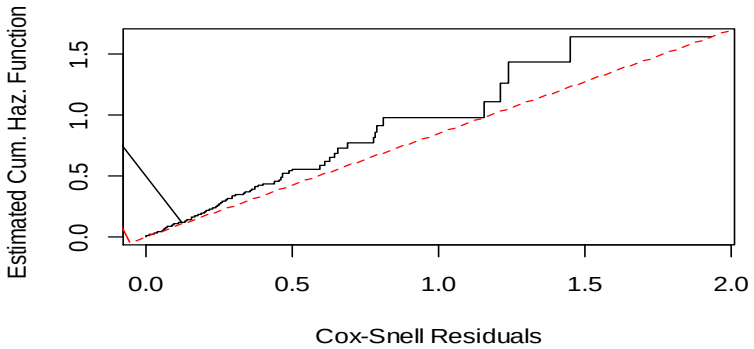


Figure 2: Cumulative hazard plot of the Cox-Snell residual for final Cox PH model

Table 2: The Parameter Estimates, Standard Errors and the Hazard Ratios of the Final Cox Proportional Hazard Regression Model

Covariates	DF	Parameter Estimate	SE	Wald	P-Value	HR	95.0% CI for the HR	
							Lower	Upper
Baseline weight <50kgms (Ref.) ≥50kgms	1	-0.438	0.173	6.332	0.0128*	0.6455	[0.459	,0.906]
Education level no education (Ref.)	2			18.518	0.000*			
Primary	1	-0.504	0.188	7.145	0.008*	0.604	[0.417	,0.875]
secondary and above	1	-0.972	0.231	17.639	0.001*	0.379	[0.242	,0.596]
ART Adherence Poor (Ref.) Good	1	-0.7881	0.172	5.068	0.024*	0.454	[0.284	,0.749]
WHO clinical stage	3			13.923	0.003*			
stage I (Ref.)								
stage II	1	0.325	0.354	0.845	0.358	1.384	[0.692	,2.769]
stage III	1	0.507	0.351	0.159	0.022*	1.650	[1.578	,2.290]
stage IV	1	0.823	0.327	6.340	0.012*	2.278	[1.700	,4.323]
Base line CD4 < 200 cells/μl(Ref.) ≥ 200 cells/μl	1	-0.4033	0.1734	5.379	0.020*	0.685	[0.495	,0.907]
Substance use No (Ref.) Yes	1	0.6034	0.184	10.739	0.001*	1.828	[1.275	,2.621]
TB co-infection Not infected (Ref.) Co-infected	1	0.3775	0.188	4.021	0.045*	1.458	[1.008	,2.109]

Source: Hossana Queen Elleni Mohamad Memorial Hospital, SNNPR, Ethiopia; from 1st February 2011 to 1st January 2016 indicates statistical significance at 0.05 level of significance. SE= Standard Error, HR= Hazard Ratio, CI = Confidence Interval, Ref. = Reference, DF= Degrees of Freedom, AIC value*= 1698.571

An adequate model is a model with lowR^2 due to high percent of censored data. The value of the -2Log-Likelihood of the model with covariates in table 4.6 which is equal to 1649.303 and the -2Log-Likelihood for the null or empty model equals 1707.449. The measure of goodness of fit R^2_p is calculated as: R^2_p=1- exp[$\frac{2}{n}$(L_O - L_p)] = 1-exp[$\frac{2}{933}$((-853.7245-(-824.65))] = 0.0604. which is small, indicating that the model fit the data well. Furthermore, the results of theLikelihood ratio, Score and Wald tests for model goodness of fit displayed in Table 4.6 which suggest that the model is good fit (i.e. significant at 5% level of significance). Therefore, the model with estimates as given in Table 2 is the final Cox PH Regression model.

Table 3: The Likelihood Ratio, Score and Wald tests for overall measures of goodness of fit of the final Cox PH model in table 2

Test	Chi-Square	DF	Pr>Chisq
Likelihood Ratio	72.34	10	<.0001*
Score	76.23	10	<.0001*
Wald	72.94	10	<.0001*

***Source**: Hossana Queen Elleni Mohamad Memorial Hospital, SNNPR, Ethiopia; from 1st February 2011 to 1stJanuary 2016 * indicates statistical significance at 0.05 level of significance.*

The coefficient of the categorical covariates is interpreted as the logarithm of the hazard ratio of death to the baseline (reference group) hazard. That is, they are interpreted by comparing the reference group with others. Similarly, the coefficient for a continuous explanatory variable indicates the estimated change in the logarithm of the hazard ratio for a unit increase in the value of the respective covariate when the remaining covariates in the model are under control. Accordingly, the interpretation of the covariates included in the final Cox proportional hazard model of HIV infected patients in the case of HQEMMH is as follows.

The estimated hazard ratio of death for patients whose baseline weight is ≥50 kgms is $\widehat{HR}$= 0.6455 [95% CI: 0.4595-0.9068, p=0.012]. This means that the hazard rate of death of patients whose baseline weight ≥50 kgms reduced by 35.45%compared to patients whose baseline weight<50 kg controlling for other variables in the model. Similarly, the covariate baseline CD4 count is statistically significant influence on the survival time of the patients. The estimated hazard rate of death of patients whose CD4 count ≥200cells/μl is 0.687 [95% CI: 0.4957-0.9071, p=0.02].This indicates that the estimated hazard rate of death of patients whose CD4 count ≥200cells/μl reduced by 31.3% compared to patients whose CD4 count < 200cells/μl controlling for other variables in the model.

The estimated hazard ratio of death for patients who were abuse substance (tobacco, alcohol, soft drugs) was1.828times higher than those who didn't uses substance [95% CI: 1.275-2.621, p=0.001]. This indicates patients who were abuse substance was82.8% higher risk of death than patients who did not use substance controlling for other variables in the model. Similarly, the estimated relative risk of death for patients who were TB co-infected was 1.458 times higher risk of death than patients not TB co-infected [95% CI: 1.008-2.109, p=0.045] controlling for other variables in the model.

Parametric Model Comparison for Time to death of HIV infected Patients

From this Time to death of HIV infected patients the parametric regression models were fitted in Table 4.15of the Appendix. This study considers model Comparison after adjusting for the effect of covariates and also compare models by using graphical method based on the Cox-Snell residual plots and Akaikie information criterion (AIC). In case of Cox-Snell residual plot, if the model is good, the plot of Cox-Snell residuals versus cumulative hazard estimates line should passes through the origin. Here this study presents the Cox-Snell residual plots for model comparison in Figures 4.5 to 4.8. From those figures Cox-Snell residuals plot for Weibull regression model shows deviation from the straight line passing through origin, it indicates that the Weibull regression model fit the data better, otherwise that the exponential, log normal and log logistic regression models fit the data poorly.

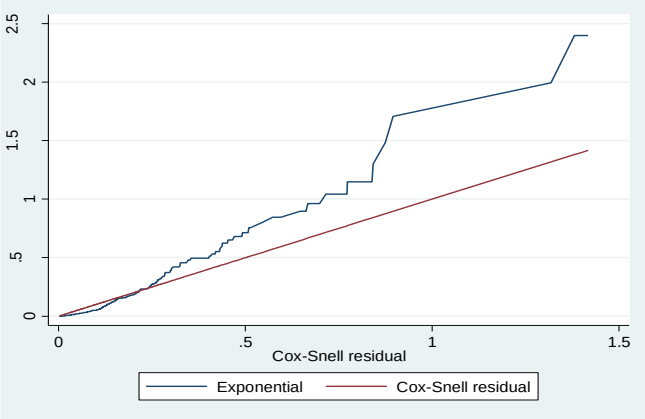


Figure 3: The Cox Snell plot after fitting Weibull regression model

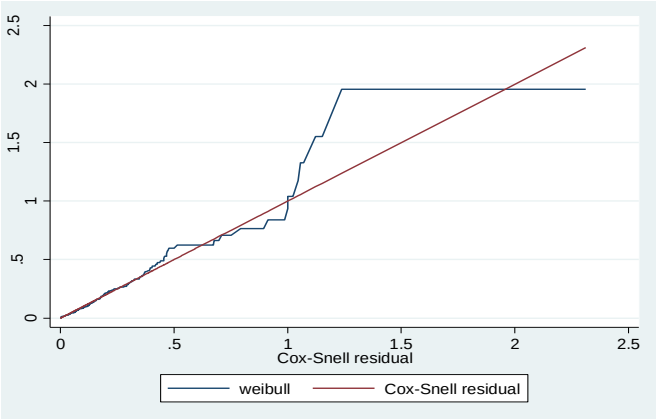


Figure 4: The Cox Snell plot after fitting Exponential regression model

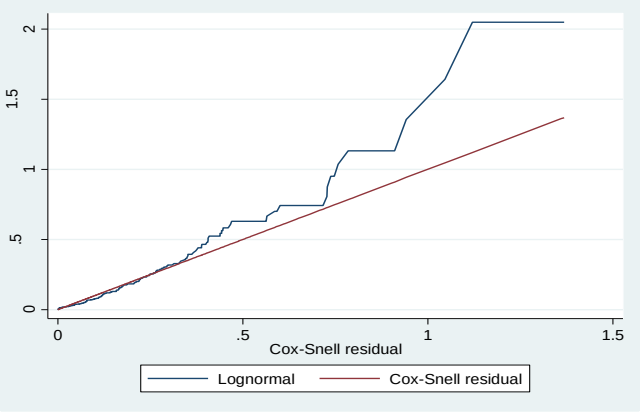


Figure 5: The Cox Snell plot after fitting log logistic regression model



Figure 6: The Cox Snell plot after fitting lognormal regression model

But graphical methods may not assure the result. In order to select the appropriate parametric survival regression model, the most common applicable criterion called Akaikie information criterion (AIC). Nevertheless, the results of cox-snell were consistent with the results based on Akaikie"s information criterion. Here, the models are not nested; it is not possible to compare the models using log-likelihood values. When the models were compared using AIC in Table 4.7, among the parametric models, the result of table reveal that the Weibull regression model has the smallest AIC, which shows that weibull model is the appropriate parametric survival regression model for HIV infected patients from Hossana Queen ElleniMohamad Memorial Hospital.

Table 4: Selection of parametric survival regression model by using Log likelihood and AIC

Model	log-likelihood	AIC
Exponential	-865.9	1759.879
Weibull	-827.1	1684.139
Log logistic	-828.6	1687.193
Lognormal	-840.8	1711.652

***Source**: Hossana Queen Elleni Mohamad Memorial Hospital, SNNPR, Ethiopia; from 1st February 2011 to 1stJanuary 2016*
AIC=Akaike's information criteria

Analysis of Weibull Regression Model

The result of relationship between covariates and survival probability of HIV infected patients modeled by Weibull regression model are presented in Table 4.10. It indicate the parameter estimates of coefficients for the covariates in the final Weibull regression model along with the associated significance level, hazard ratio with corresponding standard error and 95% confidence interval for the hazard ratio. Survival time of HIV infected patients were significantly associated with baseline weight, WHO clinical stage, education level, ART adherence, baseline CD4, substance use and TB co-infection as can be seen from the Table 4.10.

Table 5: Summary result of Parameter Estimates, Standard Errors and the 95% CI of the final multivariate Weibull regression model Analysis

Covariates	Parameter Estimate	SE	Z	P-Value	Hazard Ratio	95.0% CI for the H R	
Baseline weight							
≥50kgms (Ref.)							
< 50kgms	0.215	0.084	2.511	1.10e-02*	1.24	[1.048,	1.461]
Education level							
no education (Ref.)							
Primary	-0.230	0.084	-2.357	1.63e-02*	0.792	[0.619,	0.981]
secondary and above	-0.463	0.117	-3.948	8.62e-05*	0.629	[0.402,	0.845]
ART Adherence							
Poor (Ref.)							
Good	-0.589	0.084	-2.240	3.06e-02*	0.554	[0.389,	0.716]
WHO clinical stage							
Stage I (Ref.)							
stage II	0.086	0.122	0.712	5.04e-01	1.090	[0.844,	1.366]
stage III	0.237	0.107	2.207	2.90e-02*	1.267	[1.064,	1.476]
stage IV	0.648	0.170	0.866	3.86e-03*	1.711	[1.517,	2.044]
Base line CD4							
≥ 200 cells/μl(Ref.)							
<200 cells/μl	0.289	0.084	2.234	2.19e-02*	1.335	[1.170,	1.499]
Substance use							
No (Ref.)							
Yes	0.492	0.091	3.196	1.29e-03*	1.636	[1.427,	1.734]
TB co-infection							
Not infected (Ref.)							
Co-infected	0.388	0.092	2.049	3.46e-02*	1.473	[1.298,	1.654]

***Source**: Hossana Queen Elleni Mohamad Memorial Hospital, SNNPR, Ethiopia; from 1st February 2011 to 1st January 2016* indicates statistical significance at 0.05 level of significance. SE= Standard Error, HR= Hazard Ratio, CI = Confidence Interval, Ref. = Reference, AIC value*= 1684.139

In this study the baseline hazard for final weibull regression model obtained from equation (3.42) and with the parameters found in Table 4.10, the survival time of HIV patients with Weibull distribution can be expressed as t ~ Weibull(ρ,λ), with parameters $\lambda = \exp(\frac{-\mu}{\sigma})$= 2.98e^{-4} and $\rho = \frac{1}{\sigma}$ = 2.056 this shows hazard increases monotonically with time , time ~ Weibull (2.056, 2.98e^{-4}). By substituting the parameters in the final Weibull model with substitution of $\lambda = 2.98\mathrm{e}^{-4}$ and $\rho = 2.056$, the Weibull hazard regression model that predicts the hazard rate of patients with identical data settings is:

$h_i(t, X, B) = \mathrm{h_0(t)} \exp(\beta' x_i) = \lambda\rho t^{\rho-1} \exp(\beta' X) = 2.98\mathrm{e}^{-4} * 2.056 * \mathrm{t}^{1.056} \exp(\beta' X)$

Form the final Weibull regression model the baseline hazard vary with $\lambda\rho t^{\rho-1}$; so the base line hazard function of HIV infected patients for HQEMMH is given with formula of (4.2) in every increase in time

$h_o(\mathrm{t}) = \lambda\rho t^{\rho-1} = 2.98\mathrm{e}^{-4} * 2.056 * \mathrm{t}^{1.056}$ (4.2)

The importance of this interpretation is that for those data where it was considered reasonable to apply Cox regression to estimate the underlying hazard ratio, it

should also be reasonable to apply a Weibull analysis to estimate the hazard ratio and using the estimated scale parameter. In this study Weibull regression model was considered as better fit to the data, and also both hazard ratio and survival probabilities can be still interpreted as the hazard rate of death or survival probabilities increase/decrease in survival time on the reference group relative to others.

The abuse substance (alcohol, soft drugs) had also a significant effect on the survival probability of HIV patients. After adjusting other covariates, the hazard rate of patient who were abuse substance was 1.636 times higher than those patient who didn't use substance (adjusted HR= 1.636, CI=1.427-1.734), this pointed out that the survival probability of patients who use substance was reduced by 63.6%. Similarly, After adjusting other covariates, the hazard rate of patients who were co-infected with TB was 1.473times higher than patients who had not co-infected (adjusted HR=1.473, 95% CI: 1.298-1.654). This means that the survival probability of HIV patients who TB co-infected was declined by 47.3%.

CONCLUSION AND RECOMMENDATION

Conclusion

The results of Kaplan-Meier and log-rank test showed that patients who had: baseline weight 50kgms or above, working functional status, secondary and above education level, good ART adherence , ≥200 line CD4 count, not abuse substance and no TB co-infected had better survival time compared with reference groups. Univariate Cox Proportional Hazards regression models were developed to assess the relation between each covariate survival status and their selected variables. The result of multivariate Cox proportional hazards regression model showed that baseline weight, ART adherence, baseline CD4 count, WHO clinical stage, education level, substance and TB co-infection of patient were the major factors that affect the survival probability of HIV infected patients. In the other hand it was found that factors which had no significant impact on the survival of HIV patients were gender, age group, residence of patients, marital status, functional status and drug regimen of patients.

For modeling time to death of HIV patients Exponential, Weibull, lognormal and log logistic parametric regression models were applied. Among these using Cox-Snell residuals plot and AIC for model comparison, the Weibull survival regression model was better fitted model for time to death of HIV infected patients in case of Hossana Queen Elleni Mohamad Memorial Hospital than the other remaining parametric models. The Weibull regression model results revealed that baseline weight<50 kg, low CD4 count at baseline, no education, WHO stages III and IV, poor ART adherence, co-infection with TB and substance abuse are the categories that reduce the survival probability of HIV infected patients. Finally, The Weibull survival regression model provides better predictions to the survival probability of HIV patients.

Recommendation

Based on this study finding, the following recommendations can be forwarded for government program planners, decision makers, ART program implementers at different level and other stakeholder who work in the areas of giving care, support and treatment for HIV/AIDS patients. Health workers should be cautious when a patient has lower baseline CD4 and lower baseline weight. Health workers need to support those patients with no or little education by continuous awareness creation of taking care of themselves and knowing what factors facilitate death. Hence, education level of the patients has an important role in increasing their quality of life.

Prompt initiation of TB treatment in order to reduce patient mortality and Patients who drink alcohol need to be given advice to reduce excessive drinking. And also careful follow up for poorly adhered patients and giving them drug counseling is crucial to improve survival. It is imperative to integrating the HIV care with other developmental organizations like NGOs, Religious leaders and community supporters.

For Health workers, peer educators, data clerks and working with patients under ART should be given special training on support especially on how to recognize and manage patients with high risk, and to improve the quality of the data records of patients. Moreover, attempt should be made to investigate the causes of deaths that occurred out of hospitals, and mechanisms should be devised to trace patients lost to follow up. For future researchers on this area should apply Weibull survival regression model because Weibull distribution is unique that means only one that simultaneously both proportional and accelerated so that both relative event rates and relative extension in survival time can be estimated and it predict the survival probability of HIV patients well. Furthermore, Bayesian approached is recommended to compare the results obtained in the current technique.

Competing interests

The authors declare they have no competing interests.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification", IEEE Transactions on Automatic Control, **19** (6): 716–723, doi:10.1109/TAC.1974. 1100705, MR 0423716.

Bijaria, S. (2008). Dynamics of Naïve and Memory CD4+ T Lymphocytes in HIV-1 Disease Progression. Journal of Acquired

Immune Deficiency Syndrome, **30**:41-58.

Breslow, N. (1974). Covariance analysis of censored survival data. Biometric 30 (1974), 89.100

Cain, K. and Lange, N. (2003). Approximate case impudence for the proportional hazards Regression. Approximate case influence for the proportional hazards Regression model with censored data. Biometrika40 , 493-499.

Cox, C.(1998). Delta method in Encyclopedia of Biostatistics P. Armitage and T. Colton, Eds., vol.2, pp. 1125.1127.74.

Efron ,B. (1977). The efficiency of Cox"s likelihood function for censored data.; 72: 557-65.*Ethiop. J. Health Dev*; **22**(1).

Gehan, E. (1965). Generalized Wilcoxon test for comparing arbitrarily singly censored samples. Biometrika52 , 203.223.

HZHD,(2014*).* Hadiya zone Health Department: Health Sector Development program IV: Annual Health performance report*.*

Johansen, S.(1978). The product limit estimate as a maximum likelihood estimate. Scan- dinavian Journal of Statistics 5, 95.199.

Kaplan, E., and Meier, P.(1958).Nonparametric estimation from incomplete observations. Journal ofAmerican Statistical Association 53 , 457.481.

Lavori,l,D., Moorgar, G.,. (2005). Declining morbidity and mortality among patients with advanced human immune deficiency virus infection. The New England Journal of Medicine 338, 13, 853.860.

Lawless ,J.(1998). Parametric models in survival analysis. In encyclopedia of Biostatistics. Wiley: New York, 3254-64.

Liu,C., Johnson, L., Ostrow, D., Silvestre, A., Visscher, B., Jacobson, L. (2006).Predictors for Lower Quality of Life in the HAART Era Among HIV-Infected Men. J Acq; **42**:470- 477.

Schoenfeld, D. (1982). Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika; 7: 145-153.*

Stephen, D., Lawn, A., Harries, X., Landon, M. and Robin, W. (2011). Early mortality among Adults accessing antiretroviral treatment programmes in sub-Saharan Africa*. AIDS; 22:1897–1908.*

Sethi, E. (2009). Factors affecting CD4+T-lymphocyte count response to HAART in HIV patients. HIV medicine journal. Volume 8, Blackwell publishing, uk.pp135-141(7).

Sterensund, H. (1989).HIV medicine journal Volume 8, Blackwell publishing, uk.pp135-141(7).

Vernonet, T. (2013). Explaining urban-rural variations in health: A review of interactions between individual and environment. Social Science and Medicine42: 923-935.

Weibull, W. (1939). A Statistical Theory of the Strength of Materials. (1951), A Statistical Distribution of Wide Applicability. Journal of Applied Mathematics, 18, 29 3-297.

WHO, (2007). Report on the global AIDS epidemic, treatment and car

Accepted 13 December 2018

Citation: Handiso AA, Negash Y, Mekiso GT (2019). Modeling Time to Death of HIV Infected Patients on Antiretroviral Therapy in case of Hossana Queen Elleni Mohammad Memorial Hospital, South Ethiopia. International Journal of Public Health and Epidemiology Research, 5(1): 071--082.



Copyright: © 2019 Handiso *et al.* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.