College of Natural and Computational Sciences
Department of Statistics

Survival Analysis of HIV/AIDS Patients on Anti-retro Viral Therapy and Progression of HIV/AIDS Disease: The case of ALERT Hospital in Addis Ababa, Ethiopia

By: Fatuma Ebrahim
Adviser: Dejen Tesfaw (PhD)

A Thesis Submitted to the Department of Statistics in Partial Fulfillment of the Requirements for Degree of Master of Science in Statistics (Biostatistics)

June 2020
Addis Ababa, Ethiopia

Declaration

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of material used for the thesis have been duly acknowledged.

Fatuma Ebrahim ....

Name Signature Date

This thesis has been submitted for examination with my approval as an advisor.

Dejen Tesfaw (PhD) ....

Name Signature Date

Addis Ababa University
College of Natural and Computational Sciences
Department of Statistics

Survival Analysis of HIV/AIDS Patients on Anti-retro Viral Therapy and Progression of HIV/AIDS Disease: The case of ALERT Hospital in Addis Ababa, Ethiopia

By
Fatuma Ebrahim

As members of the Board of Examiners of M.Sc. thesis open defense examination of the above title, we have evaluated the thesis and examined the candidate.

Mekonnen Tadesse (Associated Prof.)	....	....
Examiner	Signature	Date
Bedilu Alamirie (PhD)	....	....
Examiner	Signature	Date
Merga Belina (PhD)	....	....
Chairman, Department graduate committee	Signature	Date

Acknowledgment

First and foremost, I would like to acknowledge the one and the true God "Allah" for blessing me with the ability and guidance to succeed my MSc study over the past years.

This study would never have been completed with out the contribution of many people to whom I would like to express my thinks. My special gratitude goes to my advisor Dr. Dejen Tesfaw for their kind support, assistance, valuable comments and advice through the initial to final level, which enable me to develop understanding of the subject.

Lastely but not the last, I would like to extend grateful acknowledgment to my family, friends and class followers for there blessed support and my deep appreciation goes to my beloved sister Toyba Ebrahim and heart-felt thanks to my friends Kemal Sherefa and Samson Gerawork, for their endless support and motivation.

Abstract

HIV is a virus that causes AIDS which 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. The main objective of the study was to identify the factors that affect the survival time of HIV/AIDS patients and progression of HIV/AIDS disease of an individual patient under ART follow-up. A retrospective cohort study was conducted on 729 HIV/AIDS patients who enrolled anti-retro viral therapy from September 2013 to August 2017 at ALERT hospital in Addis Ababa, Ethiopia. Kaplan-Meier plots and Log-Rank test were used to compare the survival experience of different categories and semi-parametric survival model and AFT models were employed to identify survival time of the patients. Semi-markov model also employed to identify the progression of the diseases. From 729 patients, 86(11.8%) died in the follow-up period. Log-logistic AFT model is better fit the data then other AFT models. The result of this model shows that the survival time of HIV patients under HAART significantly affected by functional status, clinical stage, adherence level, CD4 count, weight and BMI. The result of homogeneous semi-Markov model suggest that at the time increase, the probability to remain in the same state is decreasing while the probability of the patient transiting to a next worse state is increasing.

Contents

List of Tables

List of Figures

List of Acronyms

AIDS	Acquired Immunodeficiency Syndrome
HIV	Human Immune Deficiency Virus
ART	Anti-Retro viral Therapy
HAART	Highly Active Anti-Retro viral Therapy
PLHIV	Patient Live with Human Immune Deficiency Virus
ALERT	All Africa Leprosy and TB rehabilitation Training center
WHO	World Health Organization
BMI	Body Mass Index
AIC	Akaike's Information Criterion
PH	Proportional Hazard

1 Introduction

1.1 Background

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 identified in their blood. HIV attacks and destroys certain types of white blood cells that are essential to the body's immune system, the biological ability of the human body to fight infections. Although important progress has been achieved in preventing new HIV infections and in lowering the annual number of AIDS-related deaths, the number of people living with HIV continues to increase. The AIDS epidemic has put a spotlight on the many fault lines in society. There are inequalities, power imbalances, violence, marginalization, taboos and stigma, discrimination, and HIV takes hold (UNAIDS, 2019).

Standard Anti-Retro viral Therapy (ART) consists of the combination of anti-retro viral (ARV) drugs to maximally suppress the HIV and stop the progression of HIV disease. ART also prevents the onward transmission of HIV. Huge reductions have been seen in rates of death and infections when use is made of a potent ARV regimen, particularly in the early stages of the disease. World Health Organization (WHO) recommends ART for all people with HIV as soon as possible after diagnosis without any restrictions of CD4 counts. It also recommends an offer of pre-exposure prophylaxis to people at substantial risk of HIV infection as an additional prevention choice as part of comprehensive prevention. Countries are now following to adapt and implement these recommendations within their epidemiological settings (WHO, 2019).

ART should be initiated in all adults living with HIV, regardless of the WHO clinical stage and at any CD4+ cell count. Initiating ART early in PLHIV is associated with reduced mortality and ill health. Untreated HIV infection may be associated with the development of serious co-morbidity such as cardiovascular, kidney and liver diseases, cancers, and mental illness. Early initiation of ART serves the useful purpose of preventing the occurrence of these diseases. An additional advantage of early initiation of ART is that it substantially reduces the risk of sexual transmission to HIV-negative partners (FMOH, 2016).

Due to its high prevalence, the HIV pandemic created an unprecedented burden on the economies and health care systems of affected countries, particularly in Sub-Saharan Africa. The total number of adults living with HIV/AIDS in 2018 worldwide 36.2 million including 1.6 million newly infected of which 710,000 new cases are in Eastern and Southern Africa, 220,000 new cases are in West and Central Africa, and 18,000 new cases are in Middle East Africa during 2018 (UNAIDS 2019). The 25 countries with the highest numbers of new HIV infections selected for the Global HIV Prevention Coalition. Of these, 17 are African countries, including Ethiopia (FHAPCO, 2018).

According to UNAIDS, around 37.9 million peoples were living with HIV throughout the world, and 36.2 million of them were adults. 62% of patients aged 15 years and older living with HIV had access to treatment. 68% and 55% of female and male patients aged 15 years and older had access to treatment, respectively (USAID, 2019).

In Ethiopia, observed remarkable progress over the past two decades in reducing HIV prevalence rate from 3.3 percent in 2000 to 0.9 percent in 2017, and AIDS-related deaths from 83,000 deaths in 2000 to 15,600 in 2017, thus being on the right track to deliver on its commitments. However, the gains made so far seem to be challenged by complacency regarding primary HIV prevention (FHAPCO, 2018). The number of HIV infections among adult Ethiopians was estimated at 722,248 in 2017, increasing by 3,748 numbers patients from 2016. The highest estimated prevalence among adults in 2017 was in Addis Ababa and Gambella. The two administrative regions, Gambella and Addis Ababa, persist with a high load of HIV cases (Kibret et al., 2019).

1.2 Statement of the problem

Ethiopia has made progress in reducing the number of HIV/AIDS death nationally. However, the observed changes are not sufficient enough compared to the desired goals of the response against the epidemic. Most of the researches conducted previously in Ethiopia focused more on the prevention of people from infection by HIV/ADIS and the prevalence of HIV/AIDS. It seems that little attention has given to study the risk factors that facilitate the mortality of those people living with HIV/AIDS.

HIV infected patients start ART to reduce AIDS-related morbidity and mortality. Currently, countries including Ethiopia are trying to provide ART drugs to people who are living with HIV/AIDS, to decrease opportunistic infectious diseases, HIV/AIDS-related death, and to improve quality of life to those infected with HIV.

1.3 Objective of the study

1.3.1 General objective

- To identify the main factors that affect the survival time of HIV/AIDS patients and the progression of HIV/AIDS disease of patient under ART follow-up at ALERT hospital in Addis Ababa, Ethiopia.

1.3.2 Specific objective

- To describe socio-economic and biological factors on the survival time of HIV/AIDS patients.

- To test whether there are significant differences in survival or cumulative incidence of event among different groups of patients.
- To predict individual clinical progression AIDS diseases.
- To estimate the duration a patient stay in each state.

1.4 Significance of the study

As we know, HIV/AIDS has a serious burden on the economic, social, and political issues of the world, particularly in developing countries. It is also well understood that the benefit of ART for people living with HIV/AIDS in terms of improving quality of life and reducing morbidity and mortality. The findings or results obtained from this research are useful in many ways;

- \- Create general awareness on the disease HIV/AIDS, on the factors that influence the survival status of HIV patients.
- \- Governmental and non-governmental organizations could take intervention measures and set appropriate plans to reduce morbidity and mortality by identifying and giving priority for significant predictors of mortality among HIV-infected patients on ART.

2 Literature Review

The risk of HIV infection increases by the number of sexual partners, intravenous drug use, any sex without condom, alcohol and other drug use, tattoos, and body piercing with contaminated needles or instruments. Since AIDS first identified, researchers and scientists have tried their best to find medicine and vaccine, but the results everywhere were not successful. Consequently, the lives of human beings have observed to threaten due to infection with the virus. On top of all, since the disease is affecting the most productive citizens, it is natural that it causes damage to the national economy (UNAIDS, 2019).

In the study of examined the changes in three year survival and life expectancy of patients starting combination antiretroviral therapy (ART) between 1996 and 2013 using a collaborative analysis of cohort studies, the result showed that from 88,504 patients, 2,106 died during the first year of ART and 2,302 during the second or third year of ART. Between 1996 and 2013, the survival of people living with HIV in the first three years since ART initiation improved substantially. Life expectancy in patients starting ART has increased by about 10 years during the ART era, but remains lower than in the general population. Survival during the first three years of ART continues to improve, which probably reflects a transition to less toxic antiretroviral drugs, improved adherence, prophylactic measures, and management of comorbidity (Trickey et al., 2017).

In concentrated epidemic settings, HIV has spread rapidly in one or more populations but is not well established in the general population. Adult HIV prevalence is high enough in one or more sub-populations, such as people who inject drugs (PWID) or sex workers and their clients who maintain the epidemic in this sub-population, but the virus has not spread in the general population. Adolescent girls and young women aged 15-24 years have up to eightfold higher rates of HIV infection compared to their male peers. In several countries, HIV prevalence is nearly 20 times higher amongst high-risk sub-populations such as men who have sex with men (MSM) and sex workers compared to adult HIV prevalence in the general population (Ayesha et al., 2016).

Infection with HIV causes immunologic deficiency that result in the depletion of CD4 cells and suppression of cell-mediated immune defenses. HIV infects CD4 cells by interacting with their CD4 receptors, which allows the virus to gain entry into the cells. The invading HIV replicates within the CD4 cells, destroys them, and spreads to other CD4 cells, depleting the CD4 cell population. Because CD4 cells direct and activate immune responses, many immune functions degrade as a result of HIV infection. Individuals with weakened immune defenses are susceptible to infections caused by opportunistic pathogens that do not normally cause disease for immune-competent individuals. These infections are known as opportunistic infections. Once an opportunistic infection has begun, it can rapidly spread throughout the body via the circulatory system, damages vital organs, and becomes fatal (Virco, 2008).

A study conducted on the treatment outcomes of an HIV-infected cohort on antiretroviral therapy (ART) and the predictors in terms of immunological recovery and virological response in Indonesia, the result indicated that patients overall showed improvements in both virological and immunological response whilst on ART (Limmade et al., 2019). The majority of patients showed improvements in CD4 count recovery over time, and immunologic failure was significantly associated with lower initial CD4 counts and age. They observed a sharp temporal decline in CD4 counts between 24 and 30 months amongst patients starting ART with higher counts at treatment start. This finding is similar to other longitudinal studies conducted in Ethiopia (Reda et al.,2013, and Asfaw et al., 2015).

Ikhelowa et al. (2019) studied the survival analysis of HIV/AIDS patients under Anti-retroviral Treatment at Central Hospital, Agbor, Nigeria. From a sample of 1000 HIV/AIDS patients, 64.2 percent were female, and 35.8 percent were male, and 8.6 percent of the patients reported dead, while 91.4 percent of patients censored. Using Cox-PH regression analysis the result indicated that the survival time of the HIV/AIDs patient is significantly affected by gender, ART, enrolment date, and current age.

A study conducted by Ayele et al. (2015) entitled to treatment outcomes and their determinants in HIV patients on anti-retro viral treatment. This retrospective cohort study conducted among 730 adult HIV/AIDS patients who enrolled in anti-retro viral therapy from 2007 to 2011 in four selected health facilities. Kaplan-Meier survival function used to estimate survival probability, and the Cox proportional hazards regression model was used to identify factors associated with time to death. The result of the study indicated that the various baseline factors like the WHO clinical stage, CD4 counts, functional status, regimen adherence, and hemoglobin level had a significant effect on five years of survival time. Substantial efforts need to make to move patients into care earlier in their disease progression to obtain the maximum benefit from ART. CD4 counts and weight of the patients significantly increased at different follow-up periods.

A retrospective cohort study conducted among 350 patients on ART from September 2010 to August 2015 using Kaplan Meier estimator and Cox PH regression models at Jinka Hospital, South Omo, Ethiopia, 315 (90.0%) censored, and 35 (10.0%) died. The overall survival of patients on HAART was 64.00% (95% confidence interval [CI], 61.85 to 66.21%) at 72 months of follow-up. Patients with a history of tuberculosis treatment, ambulatory or bedridden functional status, WHO clinical stage disease, and substance abusers, found to be a significant effect on the survival time of HIV/AIDS patients (Tachbele et al. 2016).

Bedilu (2017) conducted a study on the determinant of mortality in HIV infected people on anti-retro viral therapy among 2.655 people at Mizan Hospital from 7 January 2005 to 8 May 2013 in Southern Ethiopia; survival probability at the sixth month after initiation of the treatment was 96, 94, 96 and 96% for pediatrics, teenagers and adults, respectively. The likelihood of survival by the sixth month after the beginning of ART was highest for adults compared to children and adolescents. Clinical stage, functional status, low CD4 cell count, Isoniazid preventive therapy, and tuberculosis co-infection were variables that predispose for mortality. HIV infected individual identification and early initiation of ART in the early stage of the disease should be given priority in the treatment modality.

Another study conducted by Tadege (2018) on time to death predictors of HIV/AIDS infected patients on anti-retro viral therapy in Ethiopia, a 600 HIV patients were collected from two hospitals and six health centers record from January 2003 to December 2017. From patients, there were 91 deaths from 600 patients, yielding death prevalence density are around 16 out of 100 and time to death is related to baseline CD4 level, occupation, TB status, base line weight and WHO clinical stage at 5% of the significant level. The risk of death for patients who lived in TB was about 2.872 times higher than those patients who were negative . Most of the HIV/AIDS patients on anti-retro viral therapy were died in a short period due to tuberculosis comorbidity, began with lower amount of CD4, being underweight, merchant, and being on WHO clinical stage IV.

Another study conducted by Tadege (2018) on time to death predictors of HIV/AIDS infected patients on anti-retro viral therapy in Ethiopia, 600 HIV patients collected from two hospitals and six health centers record from January 2003 to December 2017. From patients, there were 91 deaths from 600 patients, yielding death prevalence density are around 16 out of 100, and time to death is related to baseline CD4 level, occupation, TB status, baseline weight, and WHO clinical stage at 5% of the significant level. The risk of death for patients who lived in TB was about 2.872 times higher than those patients who were negative. Most of the HIV/AIDS patients on anti-retro viral therapy were died in a short period due to tuberculosis comorbidity, began with a lower amount of CD4, being underweight, merchant, and being on WHO clinical stage IV.

A study conducted by Handiso et al. (2019) on modeling the factors that affect the survival time of HIV infected patients by using Cox ph and parametric survival regression models, the overall mean estimated survival time of patients was 51.5 months. The Cox Proportional Hazards regression model result showed 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 a 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.

In the study of investigated the predictors of treatment failure among adult ART clients in Bale Zone Hospitals, Southeast Ethiopia, from 4,809 adult ART clients, the incidence rate of treatment failure found 9.38 per 1000 person-years. The highest rate of treatment failure has occurred between 6 months and ten months of ART initiation. Prevention and control of TB and other opportunistic infections, promotion of ART initiation at higher CD4 level, and better functional status, improving drug adherence are important interventions to reduce treatment failure among ART clients in Southeast Ethiopia (Haile et al., 2016).

Dessie et al. (2014) did a study on Predicting AIDS disease progression using longitudinal CD4 count among adult HIV/AIDS patients in Amhara Region, Ethiopia: Application of semi-Markov process. The results indicated that the probability that an HIV/AIDS patient with any one of the living states will transition to the death state is greater with increasing age, irrespective of the current state and age of the patient. More generally, the probability of dying decreases with increasing CD4 counts over time.

Goshu and Asena (2017) applied semi-Markov models to the HIV/AIDS disease progression and compared two sojourn time distributions, namely, exponential and Weibull probability distributions. The data obtained from 370 HIV/AIDS patients who were under the ART follow-up at the Yirgalem General Hospital between 2006 and 2015 in Ethiopia. The results of the study showed that the transition probability from a given state to the next worse state increases within the good states as time gets optimum then decreases with increasing the time during a follow-up.

3 Data and Methodology

This chapter deals with a detailed description of the data source, study variables, and the methods used in the study. The chapter also addresses the method of parameter estimation for survival data analysis. However, the Semi-Markov model is used to explore the general average progression of HIV/AIDS disease over time.

3.1 Data source

The data for the present study were extracted from inpatients' medical cards stored at The All Africa Leprosy and TB Rehabilitation Training Center (ALERT) hospital, Addis Ababa, Ethiopia. In this study, the data uses of HIV/AIDS patients who initiated Anti-retro viral Therapy in the ART clinic of ALERT, Addis Ababa, Ethiopia. The survival data were extracted from the patient's chart in which contains an epidemiological, laboratory, and clinical information of HIV patients under ART follow-up including a detailed anti-retroviral therapy history.

A retrospective study is applied to obtain data on HIV/AIDS patients that was recorded in the Oncology Department of ALERT hospital, Addis Ababa, Ethiopia. The population of the study was all HIV positive patients who are 15 years old and above placed under ART follow-up between September 2013 to August 2017 in ALERT hospital. This study was based on a review of the patient's 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. Thus, the data has been collected from patient follow up records based on the variables in the study.

3.2 Study Area

ALERT is the highest Level of Referral Hospital for Leprosy Complications in and also a WHO recognized international Leprosy training Centers. Few years after it establishment ALERT was accepted as training center for Africa by head of state of AU. ALERT has been administrated by the Federal Democratic Republic of Ethiopia's Ministry of Health since 2002.

3.3 Sample size determination

In a survival study, the occurrence of censoring means that it is not usually possible to measure the actual survival times of all patients in the study. We calculate the number of patients needed for there to be a certain chance of declaring θ (log-hazard ratio) to be significantly different from zero. Here, θ_R is reference value of θ. In practice, θ_R might be chosen on the basis of the increase in the median survival time that is to be detected, or in terms of the probability of survival beyond some specific time (Collett, 2003).

Typically, individuals are recruited over an accrual period of length $a = 60$ month. After recruitment is complete, there is an additional follow-up period of length $f = 12$ month. The required total number of deaths, d, can be found using the value of significance level $\alpha = 0.05$, power $1 - \beta = 0.8$ and log-hazard ratio is $\theta_R = 0.392$ (Abebe et al., 2013; Tsegaye and Worku, 2011; Worku and San Sebastian, 2009).

$$d = \frac{4(Z_{\alpha/2} + Z_\beta)^2}{{\theta_R}^2} = \frac{4(7.85)}{0.154}) = 204,$$

where:

$Z_{\alpha/2}$ and Z_β are the upper $\alpha/2$ and upper β-points, respectively, of the standard normal distribution.

To calculate the actual number of individuals that are required in a survival study, we need to consider the probability of death over the duration of a study. That is

$$\begin{aligned} P(death) &= 1 - \frac{1}{6}\big(S(f) + 4S(0.5a + f) + S(a + f)\big) \\ &= 1 - \frac{1}{6}\big(S(12) + 4S(42) + S(72)\big) = 0.28 \end{aligned}$$

where: $S(t)$ is the estimated survival function at time t.

The probability of a patients dying in the study has been evaluated and the required total number of patients will be;

$$n = \frac{d}{P(death)} = \frac{204}{0.28} = 729,$$

Therefore, 729 patients is need for the study over the accrual period of 60 months. To select the patients to be included in the sample from the population we use systematic random sampling technique.

3.4 The Study Variables

3.4.1 Response variable

In this study, the response variable is the survival time of HIV patients after starting ART. This is measured from the time the patient began to receive treatment until the time to an event (death) or censored using month as a measurement of time unit. For HIV/AIDS disease progression, the response variable is the change in CD4 cell lymphocyte level of patients in an interval of 12 months.

3.4.2 The Exploratory Variables

The explanatory variables included in the study are factors that are assumed to affect the survival time of patients. The description of explanatory variables are given in Table 3.1.

Table 3.1: Description and codes of the explanatory variables

	Variables	Description	Value 'code' (if any)
1	Sex	Sex of the patients	0=Male,1=Female
2	Age	Age of the patients	Continuous variable
3	Adherence Level	A patient's ability to follow a treatment plan	0=Poor, 1= Fair, 2= Good
4	MarSt	Marital status	0=Married, 1=Single, 2=Others
5	TB Screen	TB Screen of the patients	0=Negative, 1=Positive
6	BMI	Body mass index(kg/m^2)	Continuous variable
7	CD4	Baseline CD4 count ($cells/mm^3$)	Continuous variable
8	WHOct	WHO clinical stage	0=Stage I, 1=Stage II, 2= Stage III, 3=Stage IV
9	RegCl	Regimen class of patients	0= AZT-3TC-NVP, 1=AZT-3TC-EFV, 2=TDF-3TC-EFV, 3= TDF+3TC+NVP, 4= ABC+3TC+EFV, 5=Second Line
10	FunSt	Functional status of patients	0=Working, 1=Ambulatory, 2=Bedridden
11	Weight	Baseline weight (kg)	Continuous variable

3.5 Methods of data Analysis

3.5.1 Survival Data Analysis

Survival analysis is a set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest. It is used in analyzing the time-to-event data arising in several applied fields like medicine, biology, public health, epidemiology, demography (Aalen, 2008). Censoring is an important issue in survival analysis, representing a particular type of missing data. There are three categories of censoring such as right censoring, left censoring and interval censoring (Klein, 1997). The presence of patients in the data set who have not yet experienced a failure by the end of the study period.

3.5.2 Descriptive Statistics

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. In summarizing survival data, the two common functions applied are the survivor function and the hazard function (Hosmer and Lemeshow, 1999).

3.4.2.1 Survival Function

The survivor function is the probability that the survival time of a randomly selected subject is greater than or equal to some specified time. Thus, it gives the probability that an individual surviving beyond a specified time. The distribution of survival time is characterized by survivorship, probability density and hazard function. Let T be a random variable associated with the survival times and t be the specified value of the random variable T and $f(t)$ be the underlying probability density function of the survival time T. The survivor function, $S(t)$, is

$$S(t) = P(T > t) = 1 - F(t), t \geq 0$$

where, $F(t)$ is cumulative distribution function, which represents the probability that a subject selected at random will have a survival time less than or equal to some stated value t, is

$$F(t) = P(T \leq t) = \int_0^t f(u)du, t \geq 0$$

The probability density function, $f(t)$, is given by:-

$$f(t) = \frac{d}{dt}F(t) = \frac{-d}{dt}S(t)$$

3.4.2.2 Hazard Function

The hazard function is a measure of the risk of the event happening at any point in time. It is the instantaneous probability of having an event at time t (per unit time) given that one has survived (i.e. not had an event) up to time t (Kleinbaum and Klein, 2011). It is given by:

$$\lambda(t) = \frac{f(t)}{S(t)} \quad F(t) = \frac{-d}{dt}lnS(t)$$

The cumulative hazard function is given by:

$$\Lambda(t) = \int_0^t \lambda(u)du = -lnS(t)$$

Thus;

$$S(t) = e^{-\Lambda(t)}$$

3.6 Non-Parametric Estimation of Survival Function

3.6.1 The Kaplan-Meier Estimator

The Kaplan-Meier estimator is a non-parametric method of estimator which is used to estimate the survival function with censoring, which is not based on the actual observed event and censoring times, but rather on the order in which events and censored observations occur. This estimator incorporates information from all the observations available, both uncensored (event times) and censored, by considering survival to any point in time as a series of steps defined at the observed survival and censored times (Kaplan and Meier, 1958). Therefore, the Kaplan-Meier estimate of the survival function at time t, $\hat{S}(t)$, is given by:

$$\hat{S}(t) = \prod_{j:\tau_j \leq t} [1 - \frac{d_j}{r_j}] \quad ; j = 1, 2, ..., r$$

where:

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

3.7 Comparison of Survival Function

The Kaplan-Meier plots are used to see whether there is a difference in survival time or not between groups of covariates under investigation. But, the KM plot cannot be used to decide whether the survival time of patients living with HIV in each covariate is significantly different or not. Instead, we use log-rank test.

3.7.1 Log-rank test

The log rank test, developed by Mantel and Haenszel, is a non-parametric test for comparing two or more independent survival curves. Since it is a non-parametric test, no assumption about the distributional form of the data is required. This test is most powerful in detecting a higher cured proportion in one group than the other group (Mantel and Haenszel, 1959). The log-rank test statistic for comparing two groups is given by:

$$\chi^2_{logrank} = \frac{[\sum_{j=1}^{r}(d_{1j} - n_{1j}\frac{d_j}{n_j})]^2}{\sum_{j=1}^{r}\frac{n_{1j}n_{2j}d_j(n_j-d_j)}{n_j^2(n_j-1)}} \sim \chi^2_{(1)}$$

where:

- r is the total number of rank ordered event (death) times.
- d_{1j} is the number of failure in j^{th} time of 1^{st} group
- d_{2j} is the number of failure in j^{th} time of 2^{nd} group
- d_j is the number of failure in j^{th} time $(d_{1j} + d_{2j})$
- n_{1j} is the number at risk at j^{th} time of 1^{st} group
- n_{2j} is the number at risk at j^{th} time of 2^{nd} group
- n_j is the number at risk at j^{th} time $(n_{1j} + n_{2j})$

Under the null hypothesis that two survival functions are equal, the log rank test statistic, $\chi^2_{logrank}$, has chi-square distribution with one degree of freedom, $\chi^2_{(1)}$, for large sample size. The null hypothesis of equality of survival functions will be rejected for large values of $\chi^2_{logrank}$. The log-rank test can also be extended for comparing three or more groups of survival experience.

3.8 Survival models

3.8.1 Semi-Parametric proportional-hazards model

The Cox proportional-hazard model is widely used for survival analysis because it is simple; it can easily accommodate right censoring. It is used to relate several risk factors or exposures, considered simultaneously, to survival time. In a Cox semi-parametric proportional hazard regression model, the measure of effect is the hazard rate, which is the risk of failure (i.e, the risk or probability of suffering the event of interest), given that the participant has survived up to a specific time (Cox, 1972). Thus, the relationship of predictors and the time-to-event in survival analysis is given through hazard function as follows:

$$\lambda(t|Z) = \lambda_0(t)e^{\beta' Z} = \lambda_0(t)e^{\beta_1 Z_1 +,...,+\beta_p Z_p}$$

where:

- $\lambda(t|Z)$ is the hazard at time t for patients with a set of predictors $Z_1, Z_2, ..., Z_p$
- $\lambda_0(t)$ is the baseline hazard function and
- $\beta_1, \beta_2, ..., \beta_p$ are the parameters describing the effect of the predictors on the overall hazard rate.

The corresponding survival function for Cox PH model is given by :

$$S(t|Z) = [S_0(t)]^{e^{\beta_1 Z_1 + \beta_2 Z_2,...,+\beta_p Z_p}}$$

where:
$S_0(t)$ is the baseline survival function and $S(t|Z)$ is the survival function at time t for patients with a set of predictors $Z_1, Z_2, ..., Z_p$.

Assumptions of Cox-proportional hazards model

1. Independence of survival times between distinct individuals in the sample.
2. A multiplicative relationship between the predictors and the hazard (as opposed to a linear one as was the case with multiple linear regression analysis), and
3. There is a baseline hazard function $\lambda_0(t)$ common to all individuals in all the study groups.

3.8.2 Accelerated Failure Time Models

The AFT models is an alternative to the PH model for the analysis of survival time data when the Proportional hazard assumptions doesn't hold. The key differences between the Cox-PH model and AFT models are the baseline hazard function and ways of estimating coefficients (Kleinbaum and Klein, 2011). The AFT is obtained by regressing the logarithm of the survival time over the covariates and the effect of the explanatory variables on the survival time is directly measured. Some of the standard parametric AFT models are exponential, weibull, log-normal and log-logistic (Datwyler and Stucki, 2011).

The survival function of an individual with covariate X at time t, in the accelerated failure time models, is the same as the baseline survival function at time $t * exp(\beta_1 X_{1i} + \beta_2 X_{2i}+, ..., +\beta_p X_{pi})$, where $\beta_1, \beta_2, ..., \beta_p$ are coefficients of the regression models. Thus, the survival function of time t, $S(t|X) = S_0[t * exp(\beta_1 X_{1i} + \beta_2 X_{2i}+, ..., +\beta_p X_{pi})]$ for all $t \geq 0$. The effect of the covariates on the survival function is that the time scale is changed by a factor $exp(\beta' X)$, called accelerated factor. The AFT model treats the logarithm of survival time as the response variable and includes an error term that is assumed to follow a particular distribution. The AFT model can be written as follows:-

$$logT_i = \mu + \beta_1 X_{1i} + \beta_2 X_{2i}+, ..., +\beta_p X_{pi} + \sigma \varepsilon_i$$

This model shows the log-linear representation of the AFT model for the i^{th} individual, where: μ is an intercept, $logT_i$ is the log-transformed survival time, $X_1, X_2, ..., X_p$ are explanatory variables with coefficients $\beta_1, \beta_2, ..., \beta_p$, ε_i represents residual or unexplained variation in the log-transformed survival times, μ and σ are the intercept and scale parameter, respectively.

Table 3.2: Commonly used distributions and parameters in AFT models

Distribution	f(t)	S(t)	λ(t)
Exponential	$\lambda e^{-\lambda t}$	$e^{-\lambda t}$	λ
Weibull	$\lambda \rho t^{\rho-1} e^{-\rho t}$	$e^{-\rho t}$	$\lambda \rho t^{\rho-1}$
Log-logistic	$\frac{\lambda \rho t^{\rho-1}}{[1+\lambda \rho t^{\rho}]^2}$	$\frac{1}{1+\lambda \rho t^{\rho}}$	$\frac{\lambda \rho t^{\rho-1}}{1+\lambda \rho t^{\rho}}$
Log-normal	$\frac{1}{\sqrt{2\pi}\sigma} exp[\frac{-[logt-\mu]^2}{2\sigma^2}]$	$1-\Phi[\frac{logt-\mu}{\sigma}]$	$\frac{\frac{1}{\sqrt{2\pi}\sigma} exp[\frac{-[logt-\mu]^2}{2\sigma^2}]}{1-\Phi[\frac{logt-\mu}{\sigma}]}$

where:

- λ and ρ denotes scale parameter and shape parameter, respectively, for Exponential, Weibull and Log-logistic distribution.
- σ and μ denote scale parameter and shape parameter, respectively, for Log-normal distribution.
- $\Phi(.)$ denotes the standard normal distribution function.

3.8.3 Methods of Parameter Estimation

The parameters of semi-parametric Cox PH model will be estimated by using partial likelihood estimation method. The partial likelihood estimation is a technique used to make an inference about the regression parameters, β, in the presence of nuisance parameters $\lambda(t|x)$ (Cox, 1972). Whereas, the fully likelihood estimation method is used to estimate the regression parameters and baseline hazard functions in AFT models (Collett, 2003).

3.9 Model Building

The methods of selecting a subset of covariates in Cox-PH and AFT models are essentially similar to those used in any other regression models. Hosmer and Lemeshow (1998) recommended the following steps in selecting the variables by:

1. The first step is to fit model that contain each of the variables one at a time.

2. We begin by fitting a multivariable model containing all variables significant in the univariable analysis at the 20-25 percent level, as well as any other variables not selected with this criterion but judged to be of clinical importance.

3. Use backward selection to eliminate non-significant variables and examine the effect of remaining variables.

4. Starting with step (3) model, consider each of the non-significant variables from step (2) using forward selection and do the analysis.

5. Fit the final model by omitting variables that are non-significant and adding variables that are significant.

3.10 Model Selection Criterion

Akaike's information criterion (AIC) was used to choose the best fit AFT model from exponential, weibull, log-logistic and log-normal models, that fit the data results. It is a method used to compare different models, models that are not nested, and/or models with different numbers of parameters (Akaikie, 1974). The model with smaller AIC better fit the data as compared to other models. AIC is obtained by:

$$AIC = -2log(L) + kp$$

where:

- p is the number of parameters in the model
- L is the likelihood function

- k is a constant and can be seen as a penalty for additional parameters between 2 and 6 (often 2). The recommendation is to use a larger k with small sample.

3.11 Model Diagnostics

For survival data the evaluation of model adequacy are often based on quantities known as residuals. Residuals for survival data are slightly different than for other types of regression models, due to censoring.

3.11.1 Assessing the proportional hazards assumption

The assumptions of Cox-PH model was checked by test of correlation (rho) and global test. The assumption is valid using test correlation (rho) and global test if the test result is insignificant. The graphical methods can be used to check if a parametric distribution fits the observed data. This would be important for checking the adequacy of AFT models.

3.11.2 Model checking in parametric models

Datwyler and Stucki (2011) represents to use the plots;

- $log[S(t)]$ versus t gives approximately a straight line that pass through the origin if exponential distribution is reasonable.
- $log[-log[S(t)]]$ versus $log[t]$ gives approximately a straight line (linear) if Weibull distribution is reasonable.
- $log[S(t)/1 - S(t)]$ versus $log(t)$ is linear if log-logistic distribution is reasonable.
- $\Phi^{-1}[1-S(t)]$ versus $log(t)$ is linear if log-normal distribution is reasonable.

where:

- $S(t)$ denotes the probability that a randomly selected individual survives from the time origin to sometime t or beyond time t.

- $\Phi(.)$ denotes the standard normal distribution function.

3.11.3 Residual Plot

The Cox-Snell residuals, provided for checking the overall fit of the model, is a residual plot used to check if a parametric distribution fits the observed data or not. The Cox-Snell residuals, r_j , is defined by:

$$r_j = \hat{\Lambda}(T_j|X_j)$$

where $\hat{\Lambda}$ is the cumulative hazard function of the fitted model. If the model fits the data, then the r_j's should have a standard ($\lambda = 1$) exponential distribution, so that a plot of r_j versus the Nelson-Aalen estimator of the cumulative hazard of the r_j's should be approximately straight line with slope 1 (Jiezhi, 2009).

3.12 Homogeneous Semi-Markov Process Model

Homogeneous semi-markov model was adopted for predicting the clinical progression of AIDS disease and the survival probability of an infected patient up to specified t month (Giuseppe et al., 2007). Immunological assessment using the CD4 count identifies the disease progression and determines initiation of treatment. Semi-Markov supposes that the probability of a patient going from one state to another does not depend on the past states where a patient has been, but depends on the current state. Moreover, probability that a patient jump from one state to another depends on the elapsed time in the current state.

3.12.1 The Transition-Specific Semi-Markov Model

A semi-Markov process is a stochastic process, X(t): $\geq$0, where an embedded Markov chain governs the state to state transitions of the process while

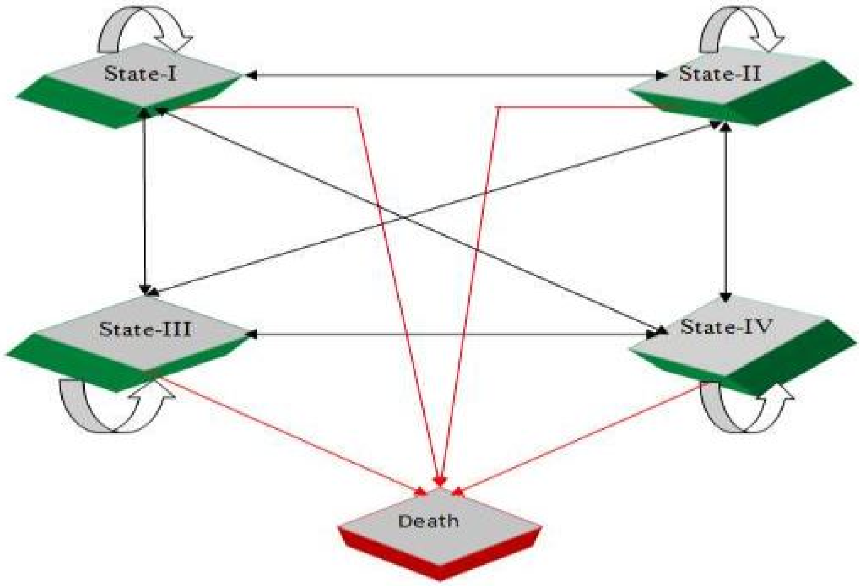

Figure 3.1: The Model of the Immunological Stages of HIV/AIDS Patients

a separate probabilistic mechanism determines the time spent in each state. It is assumed that the transition probabilities depend on the current state and the time spent in each state depends upon the current and next state.

The immunological classification of HIV/AIDS infected patients, have five states, where the first four states are the good states (transient states) and the last state is bad state or death state or absorbing state (Asena and Goshu, 2019). The states are defined as follows:
State I : CD4 count > 500 cells/microliter.
State II : $350 <$ CD4 count ≤ 500 cells/microliter.
State III : $200 <$ CD4 count ≤ 350 cells/microliter.
State IV : CD4 count ≤ 200 cells/microliter.
D : Death

Homogeneous semi-Markov process is defined based on two random variables running simultaneously.

$$X_n : \Omega \to S \quad T_n : \Omega \to \Re, \qquad n \in N,$$

X_n with state space $S = S_1, ..., S_m$ represents the state at the n^{th} transition. In the health care environment, the elements of S represent all the possible stages in which the disease may show level of seriousness. T_n, with state space equal to $\Re$, represents the time of the n^{th} transition. In this way, we cannot only consider the randomness of the states but also the randomness of the time elapsed in each state. The process (X_n, T_n) is assumed to be a homogeneous Markov process (Zelalem, 2014).

The kernel $Q = [Q_{ij}]$ associated with the process is defined as follows:

$$\begin{aligned} Q_{ij}(t) =& P(X_{n+1} = j, T_{n+1} - T_n \leq t | X_0, X_1, ..., X_{n-1}, X_n = i, T_0, T_1, ..., T_n) \\ =& P(X_{n+1} = i, T_{n+1} - T_n \leq t | X_n = i) \end{aligned} \tag{1}$$

$P_{ij} = \lim_{t \to \infty} Q_{ij}$ is the transition matrix of the embedded Markov chain in the process.

The probability that the process will leave a state i in a time t as

$$H_i(t) = P(T_{n+1} - T_n \leq t | X_n = i) = \sum_{j=1}^{m} Q_{ij}(t) \tag{2}$$

The distribution function of the waiting time in each state i given that the state subsequently occupied is known,

$$G_{ij}(t) = P(T_{n+1} - T_n \leq t | X_n = i, X_{n+1} = j) \tag{3}$$

The related probabilities can be:

$$G_{ij}(t) = \begin{cases} \frac{Q_{ij}(t)}{P_{ij}} & \text{if } P_{ij} \neq 0 \\ 1 & \text{if } P_{ij} = 0 \end{cases}$$

For any homogeneous semi-Markov process $[X(t), t \geq 0]$, the transition probabilities are given by:

$$\Phi_{ij}(t) = P[X(t) = j|X(0) = i] \tag{4}$$

They are obtained by solving the following evolution equations:

$$\Phi_{ij}(t) = \sigma_{ij}(1 - H_i(t)) + \sum_{l=1}^{m} \int_0^t Q_{il}(\gamma)\phi_{lj}(t - \gamma) \tag{5}$$

where

σ_{ij} represents the Kronecker delta

m = number of states of HSMP, which 5 in this case.

T = number of periods to be examined for the transient analysis of HSMP.

P = matrix of order m of the embedded Markov chain in HSMP.

G^T = square lower-triangular block matrix order T+1 whose blocks are of order m.

Q^T = kernel of SMP.

Φ^T = block vector of order T+1 the block which are square matrices of order m.

D^T = block vector of order T+1 the block which are the diagonal square matrix of order m.

V^T = square lower-triangular block matrix order T+1 whose blocks are of order m.

S^T = block vector of order T+1 the block which are the diagonal square matrix of order m. The diagonal element of each block t are $S_{ii} = \sum_{j=1}^{m} Q_{ij}(t)$

Sojourn Time

When the process enters state i, the time it spends there before moving to another state is called the holding time in state i or the sojourn time (Zare et al., 2014). We suppose that the sojourn or waiting time in a given state is random and has a distribution. The sojourn times of a continuous-time Markov process in a state j are independent, exponential random variables with mean $-\frac{1}{\lambda_{ii}}$ and the rate given by $-\lambda_{ii}$. Transition intensity or rate matrix where each element λ_{ii} represent rate at which transitions are made from state i to state i.

A homogeneous semi-Markov analysis was fitted in R software using 'msm' package. The package was developed by Jackson in 2011.

4 Results

4.1 Descriptive results

Table 4.1 and 4.2 shows that the descriptive result of survival time of HIV infected patients by socio-economic, demographic and biological factors. The total number of HIV/AIDS infected patients in the study was 729. Out of 729 patients, about 86(11.8%) died while 643(88.2%) censored during the follow-up period. The minimum and maximum follow-up times of patients was 1 and 59 months, respectively. The overall mean and median estimated survival times of patients was 49 and 52 months, respectively, during the study period.

As shown in Table 4.1, about 441(60.49%) of patients were female whereas 288(39.51%) of patients were male. With regards to patients status, about 45(15.62%) and 243(84.38%) of male patients and 41(9.3%) and 400(90.7%) of female patients died and censored, respectively. The table also shows about 279(38.27%) of patients were singel while 399(54.73%) and 51(7%) of patients married and others, respectively. And also 668(91.63%) patients were able to work whereas the remaining 25(3.43%) and 36(4.9%) were ambulatory and bedridden, respectively.

Table 4.1 shows the WHO clinical stage of HIV infected patients, out of 86(11.8%) died patient, about 35(4.8%) and 26(3.57%) were under stage II and III respectively, which indicates about 71% of died patients were under stage I and III. With regard to TB screen , about 683(93.69%) of patients had no TB. The adherence to treatment of HIV patients for poor, fair and good were 400(54.87%) , 272(37.3%) and 57(7.82%), respectively.

Table 4.1: Summary results of different socio-economic and biological factors

		Patients Status		
Covariates	Categories	Censored n(%)	Death n(%)	Total n
Sex	Male	243(84.38)	45(15.62)	288
	Female	400(90.7)	41(9.3)	441
	Total	643(88.2)	86(11.8)	729
Marital status	Married	346(86.72)	53(13.28)	399
	Single	247(88.53)	32(11.47)	279
	Others	50(98.04)	1(1.96)	51
	Total	643(88.2)	86(11.8)	729
Clinical Stage	Stage I	614(94.61)	35(5.39)	649
	Stage II	11(40.74)	16(59.26)	27
	Stage III	10(27.78)	26(72.22)	36
	stage IV	8(47.06)	9(52.96)	17
	Total	643(88.2)	86(11.8)	729
Regimen Class	1c	197(95.17)	10(4.83)	207
	1d	151(94.97)	8(5.03)	159
	1e	146(82.02)	32(17.98)	178
	1f	111(89.52)	13(10.48)	124
	1g	4(66.67)	2(33.33)	6
	Second Line	34(61.82)	21(38.18)	55
	Total	643(88.2)	86(11.8)	729
Functional Status	Working	627(93.86)	4(6.14)	668
	Ambulatory	10(40)	15(60)	25
	Bedridden	6(16.67)	30(83.33)	36
	Total	643(88.2)	86(11.8)	729
TB screen	Negative	626(91.65)	57(8.35)	683
	Positive	17(36.96)	29(63.04)	46
	Total	643(88.2)	86(11.8)	729

	Patients Status			
Covariates	Categories	Censored n(%)	Death n(%)	Total n
Adherence level	Poor	353(88.25)	47(11.75)	400
	Fair	241(88.6)	31(11.4)	272
	Good	49(85.96)	8(14.04)	57
	Total	643(88.2)	86(11.8)	729

As shown in Table 4.2, the median CD4 count of the patients was 468 $cells/mm^3$ [IQR: 296-665] and was lower among death compared to censored (death:178 $cells/mm^3$ [IQR:87-322], censored: 504 $cells/mm^3$ [IQR: 342-687]). The median ages of patients in this study was 42.6 [IQR:36-48] years and was lower among the death compared to censored.

Table 4.2: Summary statistics of baseline continuous variables

Variable	Status	Mean	St.dev	Min	Mix	Med	Q_1	Q_3
Time	Death	34.7	21.6	1	59	40	12	56
	Censored	51.4	3.26	26	56	52	49	54
	Overall	49	9.637	1	59	52	49	54
Age	Death	41	9.67	24	69	39	35	48
	Censored	43	9.25	19	76	42	36	49
	Overall	42.6	9.312	19	76	41	36	48
Weight	Death	50	9.52	32	74	48	43	55
	Censored	59.7	11.48	33	94	56	51	67
	Overall	58.6	9.312	32	94	41	36	48
CD4 count	Death	213.42	159.23	16	655	178	87	322
	Censored	529.67	255.91	27	1694	504	342	687
	Overall	492.36	266.719	16	1694	468	296	665
BMI	Death	20.29	3.28	11.75	31.2	20.3	18.14	22.03
	Censored	22.48	3.94	4.84	39.82	22.09	19.62	24.97
	Overall	22.23	3.934	4.84	39.82	21.87	19.48	12.89

4.1.1 The Kaplan-Meier Estimator

As shown in Figure 4.1, as the survival time of the patients increase the survival probability decrease. The number of death increased at the end of the follow-up time. The Kaplan- Meier plot of the survival probability by sex, TB screen, WHO clinical stage, adherence level and functional status are given in Appendix(see Figure 6.1).

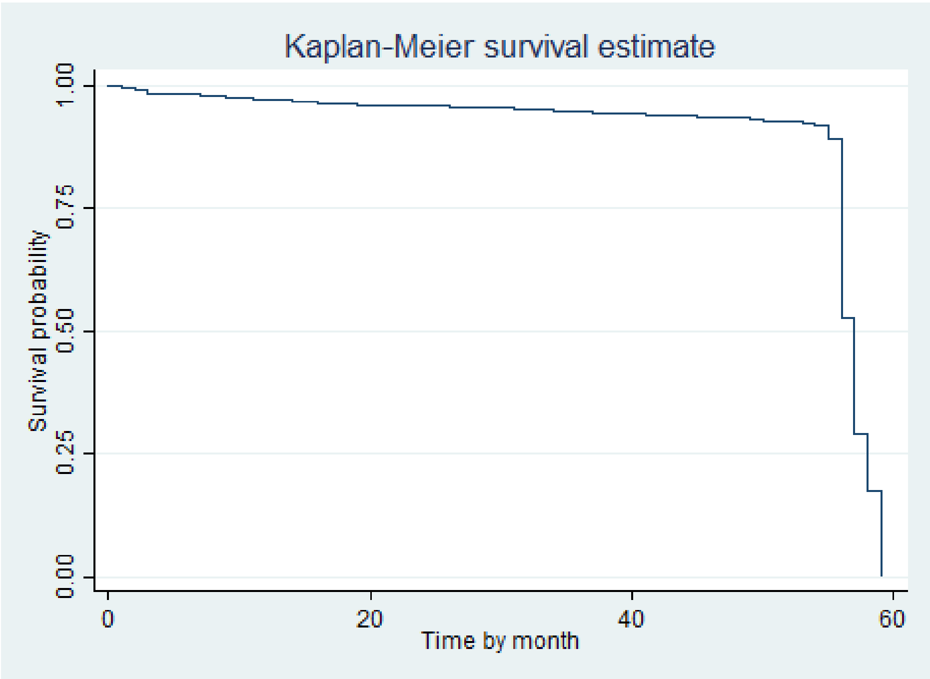

Figure 4.1: Estimated of survival function of HIV patients

4.1.2 Comparison of Survival function of HIV/AIDS patients

Table 4.3 shows comparison of survival function for each independent categorical variables. The table shows there is a significant difference in survival functions between the categories of sex, marital status, WHO clinical stage, functional status, TB screen and adherence level at 5% level of significance.

Table 4.3: Results of the log-rank test for each categorical variables of HIV/AIDS patients

Covariates	**Chi-square**	**df**	**P-value**
Sex	4.59	1	0.0321
Marital Status	6.71	2	0.0349
Clinical Stage	209.16	3	0.000
Regimen Class	54.58	5	0.840
Functional Status	156.84	2	0.000
TB Screen	65.41	1	0.000
Adherence level	0.34	2	0.0301

4.2 Cox Proportional Hazard Model

The fitted Cox-PH model, as shown from Table 6.1 in the appendices, shows that the survival time of HIV/AIDS patients significantly affected by sex, functional status, adherence level, WHO clinical stage, BMI, weight and CD4 count.

4.2.1 Checking the Assumption of Cox-PH Model

As it can be seen in Table 4.4, the p-value for functional status, age and adherence level are less than 5% level of significance and these shows the assumption of proportionality in Cox-PH model fails. In addition, using global test the assumption of proportionality in Cox-PH model also fails since the test result is significant (p-value=0.0009 $<$ 0.05).

Table 4.4: Shows test of assumption in Cox-PH model

Covariates	**rho**	**Chi-square**	**P-value**
Sex	-0.08218	0.77	0.3788
Age	-0.20919	3.91	0.0479
BMI	-0.1217	1.26	0.2626
CD4 Count	0.03981	0.18	0.6732
Clinical Stage	-0.05689	0.39	0.5306
Functional Status	-0.18817	4.99	0.0255
TB Screen	0.1023	1.35	0.2455
Adherence level	0.27267	8.12	0.0044
Weight	0.1922	3.56	0.0593
GLOBAL TEST	NA	29.90	0.0009

4.3 Accelerated Failure Time Models

We used univariable analysis before preceding to the multivariable analysis. The univariable analyses were fitted for each covariate with a p-value less than 0.25 by using different AFT models such as Exponential, Weibull, Log-logistic and Log-normal distribution. As shown Table 6.2, in the appendix, the AFT model show that the covariates like sex, age, BMI, CD4 count, WHO clinical stage, functional status, TB screen, adherence level and weight are found to be significant with survival times of of HIV/AIDS patients at 25% level.

The multivariable analysis of AFT models is done by using all significant covariates in univariable analysis at 5% level. We used backward elimination method to select the final significant covariates. The covariates such as CD4 count, WHO clinical stage, functional status and weight were significant in all AFT models at 5% level. Whereas adherence level for exponential and weibull distribution and BMI in log-normal and log-logistic distribution are also significantly associated with survival time of HIV patients. The model comparison was done using those significant covariates for each AFT models.

4.3.1 Model Selection

The value of AIC for all AFT models are displayed in Table 4.5. The AIC value for log-logistic AFT model is the smallest compared to other AFT models. This indicates the log-logistic AFT model better fits the HIV/AIDS patients data.

Table 4.5: Comparisons of AFT models using AIC

Distribution	AIC
Exponential	979.7633
Weibull	970.6942
Log-normal	969.1145
Log-Logistic	**965.9531**

4.3.2 Log-logistic Accelerated Failure Time Model

The estimated value of regression coefficients for log-logistic AFT model using multivariable analysis with backward elimination is shown in Table 4.6, the survival times of HIV patients significantly affected by functional status, clinical stage, adherence level, CD4 count, weight and BMI of patients.

Table 4.6 also shows, the estimated acceleration factor for those patients who had bedridden compared with working status is estimated to be 0.3167 with [95% CI: 0.1716, 0.5844]. This implies that the expected survival time of HIV patients decrease by 68.33% for bedridden group as compared to working group, holding other factors constant in the model. The estimated acceleration factor for those patients who had WHO clinical stage II, stage III and stage IV when compared to stage I were 0.2202 [95% CI:0.114, 0.4255], 0.4913 [95% CI: 0.2664, 0.9062] and 0.4669 [95% CI: 0.2156, 0.9894] respectively. Thus, the expected survival time of HIV infected patients decrease by 77.98%, 50.87% and 53.31% for WHO clinical stage II, stage III and stage IV ,respectively, as compared to patients who had WHO clinical stage I.

With regarding to adherence level of HIV patients, the estimated acceleration factor and 95% CI of acceleration factor for fair adherence level where compared to poor adherence level was 1.4179 and [1.0361, 2.0832], respectively. Thus, the expected survival time of HIV/AIDS patients increases by 41.79% for fair adherence level than patients having poor adherence level, keeping other factors constant in the model. Keeping the effect of other factors constant, for 10 kg change in the weight of patients the log of survival time is increased by 0.585. Similarly, for 100 $cells/mm^3$ change in CD4 cell count of HIV patients the log of survival time is increased by 0.37. Finally, for BMI of the HIV patients for 10 kg/m^2 change in the BMI of patients the log of survival time is decreased by 0.812. The value of the shape parameter $\hat{\rho} = 1.822$, hence this value is greater than unity the hazard function is unimodal.

Table 4.6: Result of maximum likelihood parameter estimates of the Log-Logistic AFT model

Covariates	Categories	$\hat{\beta}$	SE$[\hat{\beta}]$	ϕ	p-value	$[95\% CI \hat{\phi}]$
	Working					
Functional Status	Ambulatory	-0.3415	0.3484	0.7107	0.327	[0.359, 1.4068]
	Bedridden	-1.1499	0.3126	0.3167	0.000	[0.1716, 0.5844]
	Stage I					
Clinical Stage	StageII	-1.5131	0.3360	0.2202	0.000	[0.114, 0.4255]
	Stage III	-0.7107	0.3123	0.4913	0.023	[0.2664, 0.9062]
	Stage IV	-0.7617	0.3941	0.4669	0.049	[0.2156, 0.9894]
	Poor					
Adherence Level	Fair	0.3492	0.1963	1.4179	0.045	[1.0361, 2.0832]
	Good	0.5620	0.3585	1.7541	0.117	[0.8686, 3.542]
CD4 Count		0.0037	0.0006	1.0037	0.000	[1.0025, 1.0049]
Weight		0.0585	0.0149	1.0602	0.000	[1.0296, 1.0918]
BMI		-0.0812	0.0416	0.9220	0.050	[0.8498, 0.9997]
Intercept		2.6658	0.6058	14.3794	0.000	[4.3859, 47.1389]
shape parameter			$\hat{\rho} = 1.822$			

$\hat{\beta}$: coefficient estimate; ϕ: indicates Acceleration factor; 95%CI for ϕ: 95% confidence interval for acceleration factor; SE: standard error.

4.4 Model Diagnostics

4.4.1 The Cox Snell Residual Plots

The Nelson Aalen estimate of the cumulative hazard pf residuals against the Cox-Snell residuals is presented in Figure 4.3. It can be seen that the plot of the residual for the log-logistic AFT model is almost close to the 45^0 straight line through the origin. It indicates that the log-logistic AFT model fit the data very well.

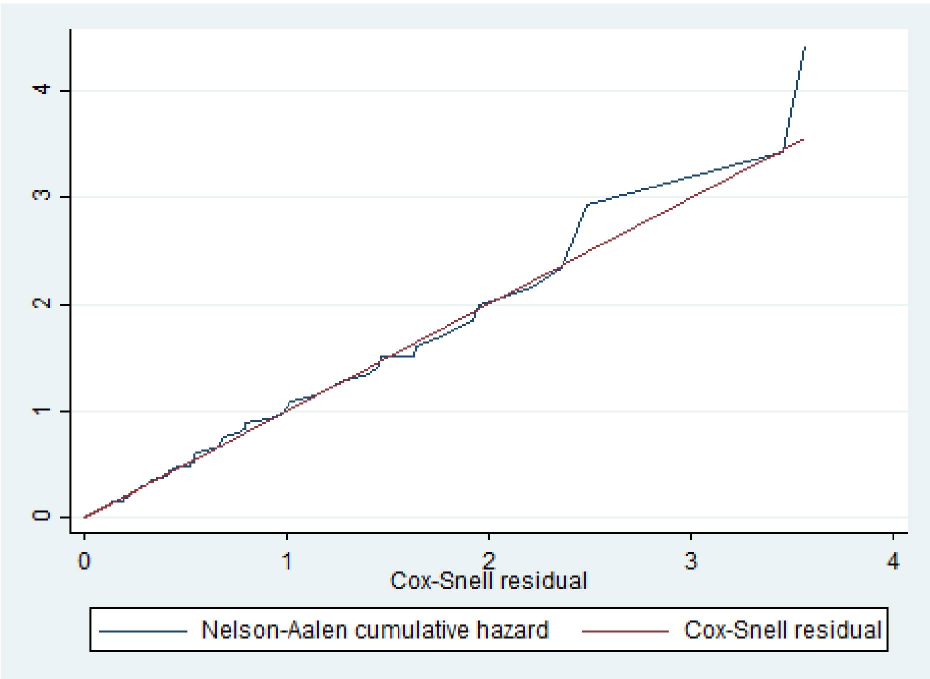

Figure 4.2: Cumulative hazard plot of the Cox-Snell residual for log-logistic AFT model

4.4.2 Likelihood Ratio Test for log-logistic AFT Model

As shown in Table 4.7, the likelihood ratio test for log-logistic AFT model is significant fit the HIV/AIDS patients data set since the likelihood value of

Table 4.7: The likelihood ratio test for the log-logistic AFT model

Loglik(model)	Loglik(intercept only)	χ^2	df	p-value
-203.1133	-337.3418	268.46	10	<0.0001

the full model shows an improvement after the covariates were included in the model than likelihood of null model.

4.5 Result of Clinical Progression of AIDS Disease

Figure 3.1 in chapter 3 shows all the immunological states of an HIV infected patients can go related to CD4 count. The states of the diseases progression adopted in the multi-state model were defined based on the following CD4 cell counts: (CD4 $>$ 500 $cells/mm^3$) as state I, (350 $cells/mm^3$ $<$ CD4 $\leq$ 500 $cells/mm^3$) as state II, (200 $cells/mm^3$$<$ CD4 $\leq$ 350$cells/mm^3$) as state III, (CD4 $<$ 200 $cells/mm^3$) as state IV, and the absorbing state death as state V (Giuseppe et al., 2007).

The frequencies and transition probabilities of the embedded Markov chain formed by the possible states of the process are shown in Table 4.8. The 729 patients contributed 4,065 transitions between the follow-up period of which 11.8% (n = 86) were death. The number of death observed from the successive working state I, II, III, and IV was 7, 10, 23, and 46, respectively. The number of death from state IV is high as compared to others. The probability of a patient dies with in state I, II, III, and IV are 0.0164, 0.0273, 0.2779, and 0.1253, respectively. Patients show improvement from state IV to III, from state III to II and state II to I with probability 0.2779, 0.2582, and 0.3383, respectively (Table 4.8).

Table 4.8: frequencies and probabilities of the transitions of the states of the process

State	State I	State II	State III	State IV	Death
State I	1082(0.8087)	205(0.1532)	34(0.0254)	10(0.0075)	7(0.0052)
State II	274(0.3383)	369(0.4556)	144(0.1778)	13(0.0160)	10(0.0123)
State III	46(0.0560)	212(0.2582)	448(0.5457)	92(0.1121)	23(0.0280)
State IV	6(0.0164)	10(0.0273)	102(0.2779)	203(0.5531)	46(0.1253)
Death	0	0	0	0	1

For specific month the solution of the evolution equation is presented in Table 4.9. The probability that an HIV positive patient being at time 0 in state i will be after t months, in the state j. Similarly, The conditional probability of a patient starting from state III at time zero, will enter after 2 years to state I, II and III is 0.0527, 0.1172 and 0.2908, respectively. The conditional probability of a patient starting from state IV at time zero, will enter after 2 years to I, II and IV is 0.1756, 0.2510 and 0.1338, respectively. A patient being at time zero in state IV, dies after 2, 4, and 6 years with probability 0.2026, 0.2924 and 0.3451, respectively. A patient being at time zero in state IV, stay in same state IV after 2 years with probability 0.3367.

The probabilities of direct transition from state I to state II, state II to state III and state III to state IV after 60 month are estimated to be 0.2268, 0.1667 and 0.0931 respectively. At the time increase, the probability to remain in the same state is decreasing while the probability of the patient transiting to a next worse state is increasing.

Table 4.9: Transition of the evolution equation from Month t

Transition	Month 24	Month 36	Month 48	Month 60	Month 72
1-1	0.7068	0.6430	0.5989	0.5664	0.5411
1-2	0.2054	0.2217	0.2265	0.2268	0.2252
1-3	0.6620	0.0953	0.1148	0.1274	0.1353
1-4	0.0091	0.0184	0.0271	0.0342	0.0395
1-death	0.0273	0.0217	0.0327	0.0453	0.0589
2-1	0.4249	0.458	0.4686	0.4692	0.4658
2-2	0.3097	0.2579	0.2373	0.2270	0.2205
2-3	0.1958	0.1861	0.1751	0.1667	0.1604
2-4	0.0421	0.0533	0.0573	0.0579	0.0570
2-death	0.0273	0.0442	0.0616	0.0791	0.0963
3-1	0.1756	0.2526	0.3043	0.3378	0.3588
3-2	0.2510	0.2385	0.2245	0.2137	0.2056
3-3	0.3781	0.2927	0.2427	0.2101	0.1873
3-4	0.13383	0.1236	0.1079	0.0931	0.0810
3-death	0.0613	0.0927	0.1206	0.1453	0.1678
4-1	0.0527	0.1059	0.1559	0.1967	0.2275
4-2	0.1172	0.1483	0.1596	0.1612	0.1588
4-3	0.2908	0.2685	0.2343	0.2024	0.1761
4-4	0.3367	0.2227	0.1577	0.1181	0.0926
4-death	0.2026	0.2546	0.2924	0.3216	0.3451

1 : State I; 2 : State II; 3 : State III; 4 : State IV

4.5.1 Prediction of Clinical AIDS Disease Progression in Individual Patient

The probability that an HIV/AIDS patient starting from state $i \in$ {I, II, III, IV} at time 0 enters state $j \in$ {I, II, III, IV and death} after month t is shown in Figure 4.4. The conditional probability of staying in the same state until a given number of month decreases with increasing the time. This shows patients change states with a non-zero probability after some time t given that patient was at some state at time t initially. This result indicates that the probability of being in the same state for an HIV/AIDS patient in a specific state of the disease decreases over time.

As shown in Figure 4.4A, the conditional probability that an HIV/AIDS patent starting from state I at time 0 enters state $j \in$ {I, II, III, IV} after month t. The probability of remaining in its stage is high as compared to other states after 120 months. The probability of a patient starting from state I at time 0, enters to state $j \in${I, II, III, IV and Death} after 108 months, estimated to be 0.4896, 0.2162, 0.1436, 0.0474 and 0.1031 respectively. The probability of a patient starting from state I at time zero enters to state $j \in$ {I, II, III, IV and Death} after 120 months, estimated to be 0.4769, 0.2127, 0.1436, 0.0484 and 0.1183, respectively.

Figure 4.4B shows the conditional probability that a patient starting from stage II at time zero enters after month t to state $j \in$ {I, II, III, IV}. As we can see the plot, the probability of remaining in state II is high as compared to others for the first 36 months. The probability of a patient starting from state II at time 0, enters to state $j \in${I, II, III, IV and Death} after 108 months, estimated to be 0.4472, 0.2068, 0.1476, 0.0528 and 0.1455 respectively. The probability of a patient starting from state II at time 0, enters to state $j \in$ {I, II, III, IV and Death} after 120 months, estimated to be 0.4399, 0.2029, 0.1444, 0.0515 and 0.1612 respectively.

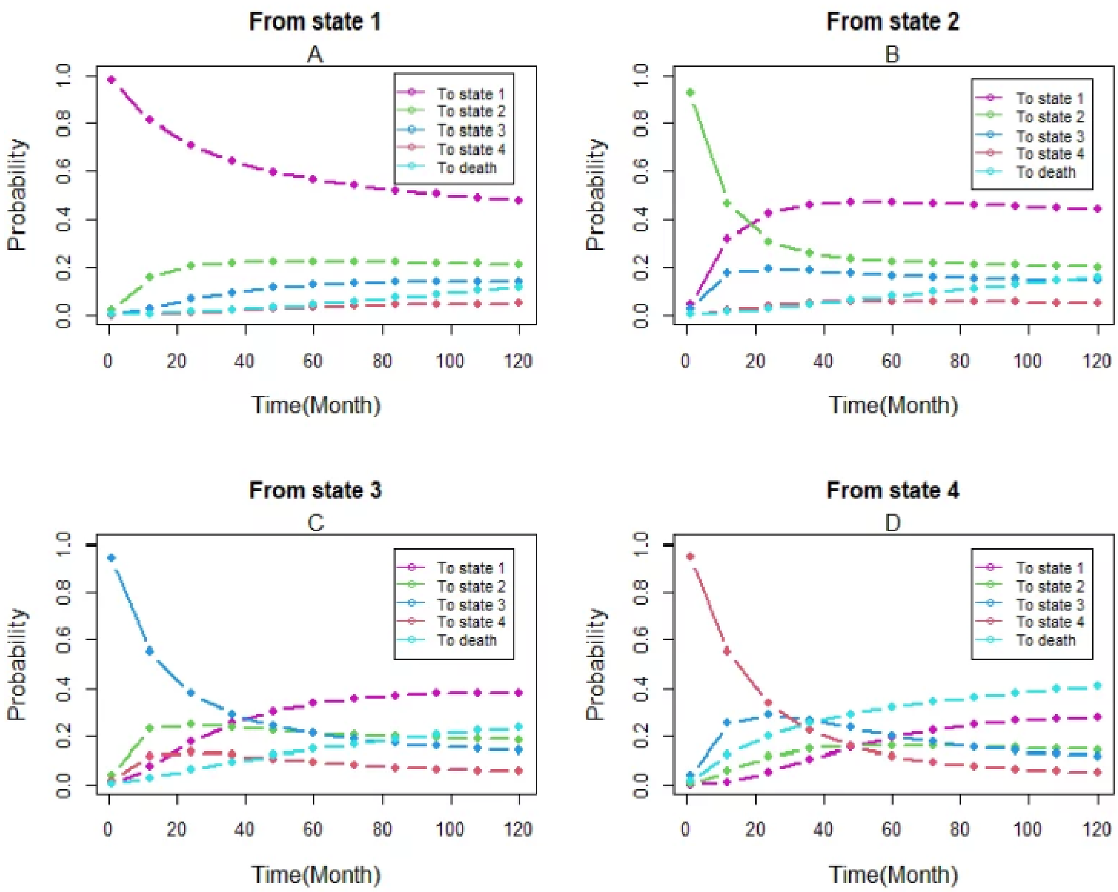

Figure 4.3: Conditional probabilities for each state

A) The probability that a patient at time 0 in state I will be in state $j \in$ {I, II, III, IV and death}, after month t;

B) The probability that a patient at time 0 in state II will be in state $j \in$ {I, II, III, IV and death}, after month t;

C) The probability that a patient at time 0 in state III will be in state $j \in$ {I, II, III, IV and death}, after month t;

D) The probability that a patient at time 0 in state IV will be in state $j \in$ {I, II, III, IV and death}, after month t.

Sojourn Time

The result in Table 4.10 show that when an individual enters state II ($350<$CD4 ≤ 500 $cell/mm^3$), the time he or she spends in this state for a single stay before moving to another state was estimated to be approximately 13 month on average. Since the waiting times for all states are relatively long, therefore, HIV disease progression in state I was slow as compared to other states. A patient in state I (corresponding to a CD4 $> 500 cell/mm^3$) have high waiting time and the CD4 counts will take time to decrease, because the patient is responding well to treatment.

Table 4.10: Sojourn time

i	Estimates	SE	L	U
State I	44.19	2.9859	38.72	50.46
State II	12.69	0.62	11.54	13.96
State III	16.47	0.9000	14.79	18.33
State IV	18.64	1.5485	15.84	21.94

L: 95% lower confidence interval; U: 95% upper confidence interval and SE: standard error.

5 Discussion

5.1 Survival Analysis Results

The aim of the study was to identify the main factors that affect the survival time of HIV/AIDS patients under ART follow-up at ALERT hospital in Ethiopia. The Cox-PH model was first used for this data. But since the assumption of proportionality in Cox-PH model was violated, the AFT models with baseline distribution: Exponential, Weibull, Log-logistic and Log-normal were considered. To compare different AFT models AIC, was used and log-logistic AFT model is found to be the best fit for the survival time of HIV/AIDS patients than others. The multivariable analysis using the log-logistic analysis showed that functional status, WHO clinical stage, adherence level, CD4 count, BMI and weight were significantly affect the survival time of HIV infected adult patients.

For the result of the study, WHO clinical stage is found to be significantly associated with time to death predictor variable. Thus, WHO clinical stage II, stage III and stage IV had decreased the survival rate of HIV infected adult patients as compared to stage I. This finding is in agreement with studies by Ayele et al. (2015), which indicates that Patients with WHO clinical stages III and IV had nearly twice higher risk of dying as compared with WHO clinical stage I.

A patient who had bedridden (inability to attain self-care in the daily living) had the shortest survival rate than working (able to perform routine activities). The result is agreed with various reports in Ethiopia (Bedilu G., 2017; Assefa and Wencheko, 2012). The study showed that bedridden patients the expected survival time of HIV infected adult patients decrease as compared to the group of working.

In this study, patients with fair adherence level were at low risk of death than those with poor adherence level. In line with this, the survival time of HIV infected adult patients increased by 41.79% for fair adherence level than HIV infected adult patients having poor adherence level. The study by Abebe et al. (2014) patients with fair and poor ART adherence were at high risk of death (2.16 and 1.88) times than those with good adherence.

A study conducted by Tachbele et al. (2016) suggested that using multivariable Cox regression analyses showed that patients of higher baseline CD4 cell count was associated with a lower risk of progression to AIDS. Our finding showed that HIV infected patients with higher baseline CD4 count had a better chance of survival. Another study done by Tadege (2018) shows that the survival time of HIV patients significantly affected by weight and CD4 count of patient. As the weight of patient increase cause to reduce mortality and the level of CD4 count increase cause to increase survival. This study was consistent with the current study.

Similarly, the survival time of HIV patients significantly affected by CD4 count of the patients and 100 $cells/mm^3$ change in CD4 count the log of survival time is increased by 0.37, keeping other factors held constant in the model. This result was consistent with the study done by Legesse and Sharmal (2016).

5.2 Disease Progression Results

Disease progression can be evaluated using time homogeneous semi Markov models using CD4 cell counts [Jackson, 2011]. Application of these models in the assessment of HIV progression has been used in the past decades [Longini et al. 1991], and the semi Markov model is a useful tool to predict the clinical progression of a disease. In this study, a continuous time homogeneous Markov process is used to model the progression of HIV/AIDS patients on anti retro viral therapy based on four CD4 cell counts transient states and death as the absorbing state. Disease progression of HIV/AIDS is studied for 729 patients under ART follow-up at ALERT hospital in Addis Ababa, Ethiopia.

The study indicated that the death of patients observed from state I, II, III, and IV was 7, 10, 23, and 46 ,respectively, during the study period. The probability of dying was increased from the worse transition states. This result is agreed with other studies [Seyoum et al., 2016; Goshu et al., 2017; and Zelalem, 2010]. Seyoum et al. (2016) conducted a study on predicting AIDS disease progression using longitudinal CD4 count among adult HIV/AIDS patients in Southwest Ethiopia: Application of semi-Markov process. The results shows that the death of patients observed from the state I, II, III, and IV was 3, 4, 15, and 40 ,respectively, during the study period. Another study conducted by Goshu et al (2017) in Yirgalem General Hospital, Ethiopia found that the transition probabilities from a given state to the next worse state increase with time and the probability of dying was increased from the worse transition states.

In this study, the conditional probability of staying in the given state decreases with increasing time, patients from the state I had high probability to remain in the ART starting state than other state. Other studies also had documented similar findings (Goshu et al. 2017 and Zelalem 2010).

Our result showed that the state II had shorter average waiting times than other states; this was also found in another study conducted by Matsena et al. (2019). The waiting time is important in disease modeling as it gives an indication of how rapidly the disease is progressing. Longer sojourn times in a disease state mean a slow progressing disease and shorter sojourn times mean a rapidly progressing disease.

6 Conclusion and Recommendations

6.1 Conclusion

The descriptive result shows that of the total 729 patients about 86(11.8%) died during the follow-up period. The estimated mean and median survival time was 49 and 52, respectively. The Kaplan-Meier plot also shows as the survival time increases the survival probability decreases. On the other hand, the homogeneous semi-Markov model was used to see clinical progression of HIV/AIDS disease.

The effect of time-to-death predictors of HIV patients under ART was functional status, WHO clinical stage, adherence level, weight, CD4 count and BMI. Patients with bedridden functional status, low BMI and being WHO stage II, III and IV had short survival. On the other hand, fair adherence level, high CD4 cell count and high weight of HIV patients have long survival.

A patient who is in the fourth state has the highest probability of dying after any given t months compared to the other states. The time increases the conditional probability of staying in the same state until a given number of month decreases.

6.2 Recommendations

Based on the findings of the study, we recommended as follows

- The ministry of health and policymakers should work on awareness about the risk factors of HIV/AIDS.
- Health workers should be careful follow up for low CD4 count, low weight and poorly adhered patients to giving them drug counseling is crucial to improve survival.
- Patients need to regularly check their CD4 count on the appropriate day in order to know their disease stage to improve survival probability and reduce mortality.

References

Aalen, O. (2008). *Survival and Event History Analysis.* Springer-Verlag, New York.

Aigboje Preston I.,Braimah Odunayo J., Umoru Oshioke M. and Aimufua G. Survival Analysis of HIV/AIDs Patients under Antiretroviral Treatment at Central Hospital, Agbor, Delta State, Nigeria. *International Journal of Scientific and Research Publications.* 2019; Volume 9, Issue 6, ISSN 2250-3153.

Akaike H. (1974). *A new look at the statistical model identification.* IEEE Trans. Automat. Contr. AC-19:716-23. [Institute of Statistical Mathematics, Minato-ku, Tokyo, Japan]

Asena, T.F. and Goshu, A.T.. Conditional Probabilities of AIDS Disease Transitions Using Semi-Markov Models. *Annual Research and Review in Biology*, 2019; pp.1-9.

Asfaw, A., Ali, D., Eticha, T., Alemayehu, A., Alemayehu, M. and Kindeya, F., 2015. CD4 cell count trends after commencement of antiretroviral therapy among HIV-infected patients in Tigray, Northern Ethiopia: a retrospective cross-sectional study. *PloS one*, 10(3), p.e0122583.

Assefa T, Wencheko E. . Survival analysis of patients under chronic HIV-care and antiretroviral treatment at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia. *Ethiop. J. Health Dev.* 2012; 26(1):22-29.

Ayele, W., Mulugeta, A., Desta, A. and Rabito, F.A., Treatment outcomes and their determinants in HIV patients on Anti-retro viral Treatment Program in selected health facilities of Kembata and Hadiya zones, Southern Nations, Nationalities and Peoples Region, Ethiopia. *BMC Public Health,* 2015; 15:826.

Ayesha B.M. Kharsany and Quarraisha A. Karim. HIV Infection and AIDS in Sub-Saharan Africa: Current Status,Challenges and Opportunities. *The Open AIDS Journal*, 2016; Volume 10.

Bedilu Girma W.. Determinant of mortality in HIV infected people on antiretroviral therapy in Southwest Ethiopia. *Journal of AIDS and HIV Research,* 2017; Vol. 9(9), pp. 187-193.

Central Statistical Agency (CSA) [Ethiopia] and ICF. 2017. 2016 Ethiopia Demographic and Health Survey Key Findings. Addis Ababa, Ethiopia, and Rockville, Maryland, USA. CSA and ICF.

Collett D. (2003). *Modeling survival data in medical research.* Chapman and Hall, London.

Cox,D.R. (1972). Regression models and life tables. *journals of royal statistical society*, Series B, ; 34: 187-202.

Custodio Martinez, R.L.M., 2007. Diagnostics for choosing between Log-rank and Wilcoxon tests.

Datwyler C., and Stucki T. (2011). *Parametric Survival Model* : https://stat.ethz.ch/education/semesters/ss2011/seminar handout-9.pdf.

Dessie, Z.G.. Multi-state models of HIV/AIDS by homogeneous semi-markov process. *American Journal of Biostatistics.* 2014; 4(2), p.21.

Federal HIV/AIDS Prevention and Control Office, Federal Ministry of Health 2018, HIV Prevention in Ethiopia National Road Map 2018 - 2020.

Federal Minister of Health. National Guidelines For HIV Prevention Treatment and Care 2016.

Gehan, E.A., 1965. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*, 52(1-2), pp.203-224.

Getiye Dejenu K., Aster F., Cheru Tesema L., Fasil W., Daniel Bekele K. and Animut A. Trends and spatial distributions of

HIV prevalence. *Infectious Diseases of Poverty*, 2019 8:90.

Giuseppe D, D'Amico G, Girolamo A, Janssen J, Iacobelli S, Tinari N, Manca RA. Stochastic model for the HIV/AIDS dynamic progression. *Mathematical Problems in Engineering.* 2007;20:656-674.

Haile, D., Takele, A., Gashaw, K., Demelash, H. and Nigatu, D.. Predictors of treatment failure among adult antiretroviral treatment (ART) clients in bale zone hospitals, south eastern Ethiopia. *PloS one* 11(10),2016 p.e0164299.

Handiso, A.A., Negash, Y. and Mekiso, G.T.. Modeling Time to Death of HIV Infected Patients on Antiretroviral Therapy in case of Hossana Queen Elleni Mohammad Memorial Hospital, South Ethiopia. textitInternational Journal of Public Health, 2019 5(1), pp.071-082.

Hosmer D., and Lemeshow S., (1999) *Applied Survival Analysis Regression Modeling of Time to Event Data.* John Wiley and Sons, New York.

Ikhelowa, A.P., Joseph, B.O. and Godstime, U.O.M.A., Survival Analysis of HIV/AIDs Patients under Antiretroviral Treatment at Central Hospital, Agbor, Delta State, Nigeria. *International Journal of Scientific and Research Publications,* Volume 9, Issue 6, 2019 ISSN 2250-3153.

Jackson CH. Multi-state models for panel data: *the msmpackage for R. J Statist Softw.* 2011; 38:1-28. doi: 10.18637/jss.v038.i08.

Jiezhi Qi. (2009). *Comparison of Proportional Hazards and Accelerated Failure Time Models.*

Kaplan EL. & Meier P. (1958). *Non-parametric estimation from incomplete observations.* J am Stat Assoc 1958(53):457?81.

Klein JP. & Moeschberger ML. (1997).*Survival analysis techniques for censored and truncated data.* Springer, New York.

Kleinbaum DG. & Klein M. (2011). *Survival analysis a self-*

learning text ($3^r d$ Ed.). Springer, New York.

Legesse, Haftu, and M. K. Sharma. "Parametric Modeling of Survival Data Based on Human Immune Virus (HIV) Infected Adult Patients under Highly Active Antiretroviral Therapy (HAART): A Case of Zewditu Referral Hospital, Addis Ababa (AA), Ethiopia." *International Journal of Statistics in Medical Research* 5, no. 4 (2016): 240-247.

Limmade Y., Fransisca L.,Rodriguez Fernandez R., BangsM J.& Rothe C. HIV treatment outcomes following antiretroviral therapy initiation and monitoring: A work place program in Papua,Indonesia. *PLoS ONE*, 2019; 14(2):e0212432.

Longini, I. M., Clark, W. S., Gardner, L. I., and Brundage, J. F. The Dynamics of CD4+ T-Lymphocyte decline in HIV-Infected Individuals: A Markov Modeling Approach. *Journal of Acquired Immune Deficiency Syndromes*, 1991; 4, 1141?1147.

Mantel N. & Haenszel W. (1959). *Statistical aspects of the analysis of data from retrospective studies of disease.* J Natl Cancer Inst 22:719?748.

Matsena Zingoni Z, Chirwa TF, Todd J and Musenge E. HIV Disease Progression Among Antiretroviral Therapy Patients in Zimbabwe: A Multistate Markov Model. *Front. Public Health*, 2019; 7:326. doi: 10.3389/fpubh.2019.00326.

Melaku T. Time to death predictors of HIV/AIDS infected patients on antiretroviral therapy in Ethiopia. *BMC Research Notes*, 2018; 11:761.

Reda, A.A., Biadgilign, S., Deribew, A., Gebre, B. and Deribe, K., 2013. Predictors of change in CD4 lymphocyte count and weight among HIV infected patients on anti-retroviral treatment in Ethiopia: a retrospective longitudinal study. *PLoS One*, 8(4), p.e58595.

Seid A., Getie M., Birlie B., Getachew Y. . Joint modeling of longitudinal CD4 cell counts and time-to-default from HAART treatment: a comparison of separate and joint models. *Electronic Journal of Applied Statistical Analysis*, 2014; 7(2): 292-314.

Unaids J. Global report: UNAIDS report on the global AIDS epidemic 2019. Geneva: UNAIDS; 2019.

Tachbele, E. and Ameni, G.. Survival and predictors of mortality among human immunodeficiency virus patients on anti-retroviral treatment at Jinka Hospital, South Omo, Ethiopia: a six years retrospective cohort study, *Epidemiology and Health*, 2016; 38:e2016049.

Trickey, A., May, M.T., Vehreschild, J.J., Obel, N., Gill, M.J., Crane, H.M., Boesecke, C., Patterson, S., Grabar, S., Cazanave, C. and Cavassini, M. . Survival of HIV-positive patients starting anti retro viral therapy between 1996 and 2013: a collaborative analysis of cohort studies,*The Lancet HIV*, 2017; 4(8), pp.e349-e356.

Virco (2008), *HIV resistance learning system, module: Overview of HIV/AIDS*, Virci Lab, New Jersey.

World Health Organization(WHO), Towards universal access by 2018: How WHO is working with countries to scale-up HIV prevention, treatment, care and support, Geneva: 201.

Zare A, Mahmoodi M, Mohammad K, Zeraati H, Hosseini M and Naieni KH. . Assessing Markov and Time Homogeneity Assumptions in Multi-state Models: Application in Patients with Gastric Cancer Undergoing Surgery in the Iran Cancer Institute. *Asian Pacific Journal of Cancer Prevention*, 2014; Vol 15, 441-447.

Appendices

Table 6.1: Result of Cox PH model

Covariates	Categories	Hazard ratio	SE[$\hat{\beta}$]	Z	p-value	[95%CI]
Sex	Male					
	Female	0.7214	0.1522	-2.10	0.036	[0.3396, 0.9638]
	Working					
Functional Status	Ambulatory	0.8770	0.4526	-0.25	0.799	[0.3190, 2.4113]
	Bedridden	2.6466	1.1028	2.34	0.020	[1.1695, 5.9894]
	Stage I					
Clinical Stage	StageII	10.3793	4.8071	5.05	0.000	[4.1873, 25.7275]
	Stage III	3.4621	1.5608	2.75	0.006	[1.4308, 8.3769]
	Stage IV	1.5364	0.7761	0.85	0.395	[0.5709, 4.1351]
TB	Negative					
	Positive	0.7865	0.2724	-0.69	0.488	[0.3989, 1.5505]
	Poor					
Adherence Level	Fair	0.4065	0.1067	-3.43	0.001	[0.2429, 0.680]
	Good	0.4721	0.2184	-1.62	0.105	[0.1907, 1.1688]
Age		0.9981	0.0122	-0.15	0.877	[0.9743, 1.0224]
CD4 Count		0.9958	0.0007	-5.37	0.000	[0.9943, 0.9973]
Weight		0.9287	0.0188	-3.64	0.000	[0.8920, 0.9664]
BMI		1.198	0.0735	2.95	0.003	[1.0625, 1.3516]

Table 6.2: Uni-variable AFT models

		Distributions			
		Exponential	Weibull	Log-normal	Log-logistic
Covariates	Categories	$\hat{\beta}$ [p-value]	$\hat{\beta}$ [p-value]	$\hat{\beta}$ [p-value]	$\hat{\beta}$ [p-value]
Sex	Male				
	Female	0.522[0.016]	0.4289 [0.019]	0.461[0.063]	0.427 [0.026]
Age		0.02[0.102]	0.016[0.106]	0.017 [0.198]	0.016[0.122]
BMI		0.1375[0.000]	.1126[0.000]	0.1490[0.000]	0.1210[0.000]
CD4 Count		0.0065[0.000]	0.0053[0.00]	0.0064 [0.000]	0.0054[0.000]
	Stage I				
Clinical Stage	Stage II	-2.83[0.000]	-2.228[0.000]	-2.899[0.00]	-2.46[0.000]
	Stage III	-2.943[0.000]	-2.304[0.000]	-2.932[0.000]	-2.414[0.000]
	Stage IV	-2.35[0.000]	-1.808[0.000]	-2.211[0.000]	-1.733[0.000]
Functional	Working				
Status	Ambulatory	-2.514[0.000]	-1.97[0.000]	-2.327[0.000]	-1.973[0.000]
	Bedridden	-3.098[0.000]	-2.431[0.000]	-3.26[0.000]	-2.734[0.000]
TB Screen	Negative				
	Positive	-2.256[0.000]	-1.834[0.000]	-2.672[0.000]	-2.009[0.000]
Adherence	Poor				
Level	Fair	0.0640 [0.178]	0.0558[0.768]	0.1548[0.0546]	0.082[0.68]
	Good	-1.438[0.782]	-0.117[0.709]	-0.092[0.839]	-0.073[0.827]
Weight		0.085[0.000]	0.069[0.000]	0.086[0.000]	0.071[0.000]

Figure 6.1: The survival time plots by sex, tb screen, WHO clinical stage, adherence level and functional status