Modelling HIV/AIDS Disease Progression: A Parametric Semi-Markov Model with Interval Censoring

Tilahun Ferede Asena[1] and Ayele Taye Goshu[2]
[1]Hawassa University, Hawassa, Ethiopia
[2]Kotebe Metropolitan University, Addis Ababa, Ethiopia
To be submitted for Publication

Abstract: *HIV/AIDS epidemic continues to be the main killer disease in sub-Saharan Africa. The main objective of this study is finding factors affecting HIV/AIDS Disease Progression. Patient follow-up data is obtained at Yirgalim General Hospital. Multivariate generalized hazard regression model were employed to investigate the disease progression using both time independent and time dependent covariates. Our study revealed that the risk of transition differs based on their Body Mass Index (BMI). The risk of transiting into the next higher states for patients whose BMI greater than 25 decreases as compared to patients BMI less than or equal to 25.This study revealed that the effect of sex, weight, age and BMI of patients are significantly associated with AIDS disease progression. In conclusion, transition probabilities are highly dependent on the choice of waiting times.*
Key words *:* AIDS disease progression, Choice of waiting time distributions, Covariate effect, Transition Probability

1. Introduction

HIV/AIDS epidemic continues to be the main challenge in the world. According to United Nations Programme on HIV/AIDS (UNAIDS) and the World Health Organization (WHO) reports of 2013, 35 million people were living with HIV worldwide, with 2.1 million new infections and with 1.5 million deaths occurred each year. Among these, 24.7 million lived in sub-Saharan Africa with 1.5 million new infections and 1.1 million AIDS deaths (UNAID,2013). The burden of HIV/AIDS in in sub-Saharan Africa and in Ethiopia in particular remains high.

Continuous-time multi-state models are widely used for categorical response data in the context of the natural history of chronic diseases (Titman and Linda, 2010). Inference is difficult when the process is only observed at discrete time points, with no information about the times or types of events between observation times. Studies show that mortality among HIV-infected individuals depend on their CD4 cell counts and other associated risk factors. CD4 cell counts and risk factors accelerate or decelerate the state transitions of the AIDS disease (Dinberu Seyoum et al., 2016).

Many studies are mainly concerned with estimation of transition and survival probabilities using Markov models without considering effects of covariates on the disease progression (Goshu and Dessie, 2013; Masala et al., 2014; Teni and Asena, 2015). The probability of dying increases in the worse transition states. The probability of being in healthy state after he/she started the treatment is higher as compared with any other working state (Dinberu et al., 2016). Some factors may increase/decrease the hazard rate of transition between disease states. When patients gets older and infected with TB, AIDS disease transitions increases (Dessie 2014a).

In Semi-Markov Models, we are able to choose a parametric distribution with more free-

dom than in traditional Markov Chain. Parametric semi-Markov models are powerful for studying chronic diseases and estimating factors associated with transitions between different stages of disease progression. Literature shows there exist a growing interest to apply parametric semi-Markov models for disease progression (Mathieu et al., 2013). The development of R packages such as the smm package (Jackson, 2016) for disease modeling is motivated by the interest to include time-varying characteristics of individuals to transition rates through a parametric approach. Epifani et al., (2014) used Bayesian estimation for a parametric Markov renewal model applied to seismic data. Their work focused on the development of methodology for Bayesian inference on a parametric Semi-Markov process, from the elicitation of the prior distribution to the computation of posterior summaries. The aim of this study is to investigate how some covariates affect the HIV/AIDS disease progression using parametric semi-Markov model with interval censoring.

2 Description of the Data

Data for this study were obtained from Yirgalem General Hospital. Yirgalem General Hospital is located 300 km south of Addis Ababa in Yirgalem town of Sidama zone, the Southern Nation's Nationalities Region. The Hospital was inaugurated in 1968. The ART clinic started follow-up on HIV patient in 2000. We adopted a simple random sampling procedure to select sample of HIV patients from the lists of patients who follows ART during the year 2008 to 2015. Accordingly, the following sample size determination formula Cochran (1997) is used:

$$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} = \frac{\left(Z_{\alpha/2}\right)^2 pq/d^2}{1 + \frac{1}{N}\left(\frac{\left(Z_{\alpha/2}\right)^2 pq}{d^2} - 1\right)} \tag{1}$$

where $z_\alpha/2$ is the upper $\frac{\alpha}{2}$ points of standard normal distribution with $\alpha = 0.05$ significance level, whose value is equal to $z_\alpha/2 = 1.96$. The term p represents proportion of death among HIV/AIDS patients. It was obtained from the previous comparable study conducted by Goshu and Dessie (2013) on data taken from Felege-Hiwot Referral Hospital where p = 0.134. The degree of precision d for this study was taken to be 0.03. With total number of N = 1570 HIV/AIDS patients at the Yirgalem General Hospital, the sample size for this study was estimated to be 375 patients. Referring the CDC (1992) immunological classification of HIV/AIDS infected patients, we have five states definde as follows. The first four states are the good states and the last state is bad state or death state. Thus, based on the seriousness of the cases we have the following states:

SI: CD4 count >500 x 10^6 cells/L

SII: 350 x 10^6 cells/L < CD4 count < 500 x 10^6 cells/L

SIII: 200 x 10^6 cells/L < CD4 count <350 x 10^6 cells/L

SIV: CD4 count < 200 x 10^6 cells/L.

D: Death.

The good states communicate with each other, and they alsocomunicatewith the absorbing state which is death state. The number of patients finaly analysed was 365. The covarites included are Age(Age), Sex(Sex), Place of residence (P_residence), Reli-

gion (Religion), Marital status (M_Status), Educational status (Edu_Status), Occupational status (Occupation), Drug use or not (Drug_Abuse), Knowledge of ART service (Know_Art), TB status (positive or negative) (TB_Status) and Post opportunistic infection (Post_oprtu).

3. The Semi-Markov Model

A Semi-Markov model is a statistical model with the same structure as a Markov model except that the sojourn time distribution is flexible in semi markov rather than an exponential distribution in Markov.

Suppose $S \in (1, 2, 3, 4, 5)$ is the state space. We define the following random variables.

$$X_n : \Omega \to ST_n : \Omega \to N^+ \tag{2}$$

J_n represents the state at the n^{th} transition and T_n represents the chronological time of the n^{th} transition. Let $N(t)$ be the counting process associated to the point process $(T_n, n \in N)$ defined for any time $t \geq 0$ by :

$$N(t) = Sup\{n, T_n \leq t\} \tag{3}$$

The random variable $N(t)$ represents the number of transitions occurred in the interval of time $(0, t]$. Let us define $(T_n)_{n \in N}$ as the duration process by :

$$T_0 = 0,$$

$$T_{n+1} = S_{n+1} - S_n \tag{4}$$

Here X_{n+1} represents the duration time spent in state J_n. The process $(J_n, S_n)_{n \in N}$ is called a Markov renewal process if

$$\begin{aligned} &P(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n = j, S_n = s, J_{n-1}, S_{n-1}, ..., J_0, S_0) \\ &= P(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n = i, S_n = s) \end{aligned} \tag{5}$$

and for $j \neq i$. The Semi-Markov kernel is given by

$$Q_{ij}(s, t) = P(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n = i, S_n = s) \tag{6}$$

The second component of Q_{ij}, namely t, represents a duration time whereas s represents a chronological time (Mathieu et al., 2013).

$$\begin{aligned} P_{ij}(s) &= lim_{t \to \infty} Q_{ij}(s, t), i, j \in E, i \neq j \\ &= P(J_{n+1} = j | J_n = i, S_n = s) \end{aligned} \tag{7}$$

Equation 33 represents the probability of a patient making its next transition to state j, given that he/she entered state i at time s and $\mathbf{P}(s) = [P_{ij}(s)_{i,j}]$ is the 5 by s matrix of transition probabilities of the embedded Markov chain $(J_n)_{n \in N}$, (Cox and Miller, 1965).

The probability density function of the waiting time in state i before passing to state j is given by:

$$f_{ij}(t) = \lim_{\Delta t \to 0} \frac{P(t < S_{n+1} - S_n < t + \Delta t | J_{n+1} = j, J_n = i)}{\Delta t} \tag{8}$$

From density function, we can derive the cumulative probability function, $F_{ij}(t)$ and the survival function $(S_{i.}(t))$ of waiting time in state i as defined by:

$$F_{ij}(t) = P(S_{n+1} - S_n \leq t | J_{n+1} = j, J_n = i) = \frac{Q_{ij}(s,t)}{P_{ij}(s,t)} \tag{9}$$

$$S_{i.}(t) = 1 - P(S_{n+1} - S_n \leq t | J_n = i) = \sum_{j \in E} P_{ij}(1 - F_{ij}(t)) \tag{10}$$

In semi Markov modeling one can choose any eligible waiting (sojourn) time distribution. Thus, we considered both exponential distribution and Weibull distribution as the waiting time distribution. For exponential distribution the hazard function is constant. The hazard function of the waiting time is given by:

$$\alpha_{ij}(t) = \frac{1}{\sigma_{i,j}}, \forall t \geq 0, \forall \sigma_{ij} > 0,$$

Further, the Weibull distribution generalizes the exponential distribution by using two parameters, which is more flexible and well adapted to various shapes. Its hazard function is defined by:

$$\alpha_{ij}(t) = \frac{\nu_{i,j}}{\sigma_{i,j}} (\frac{t}{\sigma_{i,j}})^{\nu_{i,j}-1}, \forall t \geq 0, \forall \sigma_{ij} > 0, \forall \nu_{ij} > 0$$

The hazard function of the Semi-Markov process, which represents the probability of transition towards state j between time t and $t + \Delta t$, given that the process is in state i for duration t can be derived as follows:

$$\begin{aligned} \lambda_{ij}(t) &= \lim_{\Delta t \to 0} \frac{P(J_{n+1} = j, t < S_{n+1} - S_n < t + \Delta t | J_n = i, S_{n+1} - S_n > t)}{\Delta t} \\ &= \frac{P_{ij} f_{ij}(t)}{S_{i.}(t)} = \frac{P_{ij} S_{ij}(t) \alpha_{ij}(t)}{S_{i.}(t)} \quad i, j \in E, i \neq j \\ \lambda_{ii}(t) &= -\sum_{i \neq j} \lambda_{ij}(t) \end{aligned} \tag{11}$$

Equation (37), can be interpreted as the subject's risk of progressing from state i to state j after having stayed in state i for aduration t (Listwon-Krol & Saint-Pierre, 2014).
Explanatory variables can be included at each level of the model through generalized regressions. A proportional hazards model can be used to relate the transition intensities $(q_{ij}(t))$ at time t to covariates z (t) at that time.

$$q_{ij}\{t, z(t)\} = q_{ij} exp\{\beta_{ij}^T z(t)\}$$

The parametric model allows to incorporate covariates in the distribution of sojourn times using a proportional-hazards regression model (Cox, 1972).

3. Results and Discussion

The aim of this study is to find factors affecting HIV/AIDS disease progression using a parametric Semi-Markov model with interval censored data. We use the R package

Factor		N	%	Factor		N	%
Age	>35 years	200	55.6%	EduStat	No Edu	151	41.9%
	$\leq 35 years$	160	44.4%		Primary	115	31.9%
Sex	Male	178	49.4%		Second	94	26.1%
	Female	182	50.6%	Occup.	Private	151	41.9%
					employ.	107	29.7%
P_resid	Rural	182	50.6%		Unempl.	83	23.1%
	Urban	150	41.7%		Student	19	5.3%
Religion	Orthodox	169	46.9%	Drug_Abu	Yes	248	68.9%
	Muslim	65	18.1%		No	112	31.1%
	Protest.	118	32.8%	ART	Aware	228	63.3%
	Others	8	2.2%		Not Aware	132	36.7%
Mar_St.	Single	106	29.4%	Past_oprt	No	288	80.0%
	Married	157	43.6%		Yes	72	20.0%
	Divorced	51	14.2%	TBstat	Negative	253	70.3%
	Widowed	46	12.8%		Positive	107	29.7%

Table 1: Summary of demographic, socio economic and health related factors.

semiMarkov to analyze the data and estimate the paramers of the Semi-Markov model. The results presented on Table 1 represents patient's demographic, socio economic and health related factors. The results show that about 55.6% of the respondents are of age above 35 years old. About 58.3% reside in urban areas and 41.7% in rural areas.
Table 2 describes time dependent continuous variables. The meanof body mass index of the patients is 22.37 with standard deviation 4.57. The mean and standard deviation of the patients' age are 35.14 and 9.64 years.

$Variable$	Minimum	Maximum	Mean	Std.dev
BMI	19	54.17	22.37	4.57
Age (in year)	2	70	35.14	9.641
Weight (in kg)	16	106	60.39	12.213
Height (in cm)	85	174.5	165.33	41.89

Table 2: Descriptive statistics for continuous variables for HIV/AIDS patients between 2008 and 2015).

4.1 Univariate Regression Analysis

Results in Table 3 are the univariate regression coefficients associated with the effect of covariates (fixed and time dependent)for HIV/AIDS disease progression analyzed using generalized hazard regression modeling and we test the effect by Wald test statistics ($H_0 : \theta_{ij} = 1$). The null hypotheses are the nullity of distribution parameters and the regression coefficients are equal to 1. In this model, each factor was added to the model one at a time. The regression coefficient β, the p-value of the Wald test when testing the absence of effect ($H_0 : \beta = 0$) is also provided. Moreover, results regarding each distribution parameter $\sigma($ or ν or $\theta)$ the p-value of the Wald test when testing $H_0 : \sigma = 1$ are provided. In this model, the regression coefficients are interpreted in

Variable	Transition	Coef	CI	P_Value
Sex: Male (Female, ref)	$1 \to 5$	14.0	$(-17.59 , 45.6)$	<0.0001
	$2 \to 5$	308.2	(308.1 , 308.2)	<0.0001
	$3 \to 5$	26.9	$(-0.16 , 69.7)$	0.4401
	$4 \to 5$	8.9	$(-2.55 , 20.44)$	0.091
Weight: >60kg ($\leq 60kg$, ref)	$1 \to 5$	129	(129 , 129.3)	<0.0001
	$2 \to 5$	15.7	(15.49 , 15.89)	<0.0001
	$3 \to 5$	24.4	(24.39 , 24.5)	<0.0001
	$4 \to 5$	57.5	(57.55 , 57.83)	0.231
Age: >35 years ($\leq$ 35 years, ref)	$1 \to 5$	54.7	(44 , 65.48)	<0.0001
	$2 \to 5$	22.5	(3.16 , 41.86)	0.1121
	$3 \to 5$	24.1	(9.36 , 38.74)	0.3451
	$4 \to 5$	18.1	(9.82 , 26.42)	0.7201
BMI : >25 ($\leq$ 25 , ref)	$1 \to 5$	86.0	(83 , 89.03)	<0.0001
	$2 \to 5$	19.1	(6.24 , 31.97)	0.331
	$3 \to 5$	24.4	(9.17 , 39.7)	0.0801
	$4 \to 5$	23.6	(10.96 , 36.33)	0.1561

Table 3: Estimates of the univariate generalized hazard regression coefficents. ref=Reference category, CI-confidence interval at $\alpha = 5\%$ likelihood ratio test is used for testing $H_0 : \beta_{ij} \neq 0$

terms of relative risk. The effect of covariates and the proportional hazard assumption is evaluated by representing the hazard rates. The proportional hazard regression model is considered to study the effect of covariates. The estimate of coefficient for sex of the patient associated to the transition from state one to state five is significantly different from 0 ($\beta = 14.0; CI : (-17.59, 45.6); p < 0.001$). This depicts that the risk of progression for male patients from the first state to enter the death state compared with female patient who is under follow up. Similarly, the estimated coefficient for sex of the patient associated to the transition from state two to state five is significantly different from 0 ($\beta = 308.2; CI : (308.1, 308.2); p =< 0.001$). The risk of progression for male patients from the second state to enter the death state has a lower risk of progression compared with female patient who is under follow up. However, the estimated coefficient for sex of the patient associated to the transition from state three to state five ($\beta = 26.9; CI : (-0.16, 69.7); p > 0.05$) and from state four to state five ($\beta = 8.9; CI : (-2.55, 20.44); p > 0.05$) are not significantly different from 0. This describes that for AIDS progression, the effect of gender difference is significant at the early stage (stage one and stage two) but does not have an effect at the latter stages (stage three and stage four) to progress to the death state. The risk of transiting from state one to five for patients whose body weight greater than 60kg has a lower risk as compared to patients having less than 60kg weight. Similarly, the risk of transition differs based on their Body Mass Index(BMI). The estimated coefficient of BMI associated to the transition from state one to state five is significantly different from 0 ($\beta = 86.0; CI : (83, 89.03); p < 0.001$). However, the effect of BMI for transitions from state two to state five ($\beta = 19.1; CI : (6.24, 31.97); p > 0.05$),from state three

to state five ($\beta = 24.4; CI : (9.17, 39.7); p > 0.05$), and from state four to state five ($\beta = 23.6; CI : (10.96, 36.33); p > 0.05$) are not significantly different from 0.

4.2. Multivariate Regression Analysis

In Table 4, we considered multivariate generalized hazard regression model to investigate disease progression by including both time independent and time dependent covariates such as sex, age, weight, and BMI of patient's simultaneously. In the multivarite regression model we consider factors only whose regression coefficients are significant in the univariate analysis.The estimated coefficient for sex of the patient associated to the transition from state one to state two ($\beta = 0.330; CI : (\text{-}0.02, 0.68); p < 0.001$) and from state two to state three ($\beta = \text{-}0.052; CI : (\text{-}0.45, 0.34); p < 0.001$) are significantly different from 0. The risk of transiting from state one to state two for patients whose age greater than 35 years old has a lower risk as compared to patients aged 35 and below. Whereas the effect of patients aged greater than 35 years has covariate effect on the transition of the disease from state three to state four. The risk of transiting

Variable	Transition	Coef	CI	P_Value
Sex: Male	$1 \to 2$	0.330	$(-0.02$, $0.68)$	<0.0001
(Female , ref)	$2 \to 3$	-0.052	$(-0.45$, $0.34)$	<0.001
	$3 \to 4$	0.581	$(-0.04$, $1.20)$	0.2146
Weight: >60kg	$1 \to 2$	0.013	$(0.00$, $0.02)$	0.0615
($\leq 60kg$,ref)	$2 \to 3$	-0.006	$(-0.03$, $0.01)$	<0.0001
	$3 \to 4$	-0.004	$(-0.02$, $0.01)$	<0.0001
Age: >35 years	$1 \to 2$	-0.003	$(-0.38$, $0.37)$	<0.0001
($\leq 35years$, ref)	$2 \to 3$	0.057	$(-0.34$, $0.46)$	<0.0001
	$3 \to 4$	0.022	$(-0.94$, $0.98)$	<0.001
BMI: >25	$1 \to 2$	-0.139	$(-0.59, 0.31)$	<0.0001
($\leq$ 25,ref)	$2 \to 3$	-0.275	$(-0.88, 0.33)$	<0.0001
	$3 \to 4$	0.432	$(-0.56, 1.42)$	<0.001

Table 4: Estimates of the multivariate generalized hazard regression coefficents. ,Coef-coefficent, CI-confidence interval at $\alpha = 5\%$ and likelihood ratio test is used for testing $H_0 : \beta_{ij} \neq 0$

from state two to state three ($\beta = \text{-}0.006; CI : (\text{-}0.03, 0.01); p < 0.001$) and state three to state four ($\beta = \text{-}0.004; CI : (\text{-}0.02, 0.01); p < 0.001$) for patients whose body weight is greater than 60 kg decreases compared to patients having less than 60kg weight. However, weight of patients is not important to understand the process of evolution of the disease from state one to state two. The estimation coefficient for BMI associated to the transition from state one to state two($\beta = \text{-}0.139; CI : (\text{-}0.59, 0.31); p < 0.001$),state two to state three($\beta = \text{-}0.275; CI : (\text{-}0.88, 0.33); p < 0.001$) and state three to state four ($\beta = 0.432; CI : (\text{-}0.56, 1.42); p < 0.001$) are significantly different from 0. It means that The risk of transiting into the next higher states for patients whose BMI greater than 25 decreases as compared to patients BMI less than or equal to 25.

Table 5 describes estimated parameters of mean waiting time distributions in months. Results shows that in comparison to weibull waiting time distributions (-2 * log-likelihood = 5056.362 and AIC=5064.362) and exponential waiting time distributions (-2 * log-

likelihood = 5159.791 and AIC=5167.791), the weibull waiting time distributions are preferred for HIV/AIDS disease progression. It is observed that there exist considerable differences in the estimated parameters of both distributions. Besides, differences in the transition estimates from state to state in both waiting time distributions, exponential distribution generates wider confidence intervals. This result shows that there is a significant difference in estimates when different waiting times are considered in the analysis. The Weibull waiting time distributions fits very well compared with the exponential waiting times.

4.3. Discussion

	Exponential	Weibuil	
Transition	Estimate (CI) for σ	Estimate(CI) for σ	Estimate(CI) for ν
$1 \rightarrow 2$	25.042(20.02,30.06)	28.462(24.84 ,32.09)	1.232(1.09,1.37)
$1 \rightarrow 3$	26.311(10.32,42.30)	19.311 (14.59 ,24.03)	1.791(1.29,2.29)
$1 \rightarrow 4$	15.916(1.58, 33.42)	14.441(4.35 ,24.53)	1.364(0.55 ,2.18)
$1 \rightarrow 5$	86.317(85.62 ,87.01)	26.075(18.58 ,33.57)	2.517(1.35 ,3.69)
$2 \rightarrow 1$	18.005(15.46 ,20.55)	18.136(16.09 ,19.93)	1.333(1.21 ,1.46)
$2 \rightarrow 3$	18.403(14.68 ,22.13)	14.235(15.78 ,20.49)	1.471(1.27 ,1.67)
$2 \rightarrow 4$	13.845(6.98 ,20.71)	14.235(10.13, 18.34)	1.614(1.15 ,2.08)
$2 \rightarrow 5$	19.543(6.05 ,33.04)	17.77 (16.14 ,84.35)	0.906(0.51 ,1.30)
$3 \rightarrow 1$	16.516(11.51 ,21.53)	50.244(16.14, 84.35)	1.17(0.93 ,1.41)
$3 \rightarrow 2$	14.371(12.04 ,21.53)	14.602(12.92 , 16.28)	1.436(1.27 ,1.60)
$3 \rightarrow 4$	16.611(10.49 ,22.73)	17.017(12.26 , 21.78)	1.372(1.02,1.72)
$3 \rightarrow 5$	25.244(11.53 ,38.95)	29.414(15.64 ,43.19)	1.326(0.82 ,1.83)
$4 \rightarrow 1$	9.716(3.87,15.57)	10.176(5.05 , 15.31)	1.205(0.65,1.76)
$4 \rightarrow 2$	8.871(5.54,12.20)	9.157(5.88 ,12.43)	1.086(0.81 ,1.37)
$4 \rightarrow 3$	10.367(7.9,12.83)	11.003(8.96 , 13.04)	1.291(1.07 ,1.51)
$4 \rightarrow 5$	20.526(12.02,29.03)	20.124(10.63 ,29.62)	0.87 (0.61 ,1.13)
-2 * log-likelihood	5202.282	5056.362	
AIC	5167.791	5064.362	

Table 5: Comparison of the estimated transition intensities for Semi-Markov model with Weibull and Exponential and Weibull distribution

In this paper, we studied the effects of covariates on the evolution of HIV/AIDS disease using parametric Semi-Markov model, Data were obtained from follow-up of patients at the Yirgalim General Hospital during 2008-2015. Both univariate and multivariate generalized hazard regression models are used to study the effects of covariates on the evolution of the disease. Based on the univariate generalized hazard regression model sex, age, weight and body mass index of the HIV/AIDS patient are found to be significantly associated with the AIDS disease progression. These results are consistant with those presented by Foucher et al (2006) and Masala (2014), studies based on large data sets and long follow-up time.

The risk of patients moving from state I to state II for patients aged greater than 35years of old were greater than those of patients aged 35 years of old and less. This might be

because of two reasons. First when patients become older their immunity for the disease decrease therefore this will increase the progression in latter stages. Second old patient are more exposed with different activities like drinking alcohols, smoking cigarettes than in the younger ages. Similarly, the risk of transition differs based on their Gender, weight and Body Mass Index(BMI).
The results are also consistant with reports by Goshu and Dessie,(2013), Dessie, (2014a), Dessie, (2014b) and Dinberu et al,(2016) from some country, Ethiopia for estimating the progression of disease without the effects of covariates.

Besides, the study confirmed that AIDS disease progression is affected by the choice of sojourn time distributions. Weibull waiting time distribution gives substantially narrower confidence intervals, revealing that it is more preferable in modelling the AIDS disease progression as compared with exponential waiting time distributions.

5. Conclusions

HIV/AIDS epidemic is the most killer disease in Sub-Saharan Africa and Ethiopia took the greatest share of it. In this study we aim to assess the effects of covariates on AIDS Disease progression. A follow-up data is obtained from Yirgalim General Hospital. Our study found that gender differences are very important during the early stage of the AIDS progression and it becomes unimportant at the latter stages (stage three and stage four). Results of multivariate analysis revealed that effects of sex, age, weight and body mass index are significantly associated with AIDS progression.

References

Asena, T. & Goshu, A. (2017). Comparison of Sojourn Time Distributions in Modeling HIV/AIDS Disease Progression. *Biometrical Letters*. 54(2),155-174. doi:10.1515/bile-2017-0009.

Cox DR., and Miller HD.(1965). The Theory of Stochastic Processes. Chapman and Hall, London.

Cox DR.(2011). Regression Models and Life-Tables. *Journal of the Royal Statistics Society B*, 34,187-220.

Dessie Z.G.(2014a). Modeling of HIV/AIDS Dynamic Evolution using Non- Homogeneous Semi-Markov Process.SpringerPlus. 3:2-11.

Dessie Z.G.(2014b). Multi-State Models of HIV/AIDS by Homogeneous Semi- Markov Process. *American Journal of Biostatistics*. 4: 21-28.

Dinberu Seyoum, Sisay Wondaya, Kibrealem Sisay, Mulualem Tadesse.(2016). Predicting AIDS disease progression using longitudinal CD4 count among adult HIV/AIDS Patients in Southwest Ethiopia: Application of Semi-Markov Process. *Int.J.Comput Bioinfo in Silico Model*. 5(2):808-814.

Epifani I., Lucia L., and Antonio P.(2014). Bayesian Estimation for a Parametric Markov Renewal Model Applied to Seismic Data. *Electronic Journal of Statistics*. 8:2264-2295.

Foucher Y, and Saint Pierre P., Pugliese P.,and Daures J.P. (2006).A Semi-Markov Frailty Model for Multi-State Survival Data: Illustration on Hiv Disease. *Far East Journal of Theoretical Statistics*. 47(19):185-201.

Goshu A.T., and Dessie Z.G.(2013). Modelling Progression of HIV/AIDS Disease Stages Using Semi-Markov Processes. *Journal of Data Science.* 11:269-280.

Jackson CH.(2016). Multi-State Models for Panel Data: The msm Package for R.*Journal of Statistical Software,* 38(8),1-29.
Listwon KA.,and Saint-Pierre, P. (2014). Package Semi-Markov. Reference Manual.

Listwon, A.,and Saint-Pierre P.(2013) Semi-Markov: An R Package for Parametric Estimation in Multi-State Semi-Markov Models.

Lutambi, A.M.(2015). Modelling the Impact of Stages of HIV Progression on Estimates. *Advances in Infectious Diseases.* 5,101-113.

Mandel M.(2010). Estimating Disease Progression Using Panel Data. *Biostatistics.* 10,0,1-13.Oxford University Press.

Masala G.,Cannas G., and Micocci M.(2014). Survival Probabilities for HIV Infected Patients through Semi-Markov Processes. *Biometrical Letters.* 51(1),13-36.

Mathieu E. Foucher Y. Dellamonica P., and Daures J.P. (2013). Parametric and Non Homogeneous semi-Markov Process for HIV Control. PNHSM Process for HIV Control.

Raquel.V.C.Oliveira,Silvia E.S ,Dayse.P.C.,Flaviana.P.V.S,Valdiléa G.V., Beatriz G., and Marilia C.(2013). Multi-State Models for Defining Degrees of Chronicity Related to HIV Infected Patient Therapy Adherence. *Cad. Saúde Pública, Rio de Janeiro,* 29(4):801-811, abr,.

Titman AC.,and Sharples LD.(2010). Semi-Markov Models with Phase Type Sojourn Distributions.*Biometrics.* 66,742-752.

UNAID. Global Report (2013). UNAIDS Report on Global AIDS Epidemic.

World Health Organization (WHO).(2008). The Global Burden of Disease: 2004 Update. Geneva: Wor1d Health Organization.