

Identification of factors affecting the survival lifetime of HIV+ terminal patients in Albert Luthuli municipality of South Africa

Pepukai Bengura (✉ 31659322@mylife.unisa.ac.za)

University of South Africa

Prince Ndlovu

University of South Africa

Annah Mulalo Managa

University of South Africa

Research Article

Keywords: Survival time, Follow-up time, HIV/AIDS disease, Antiretroviral therapy, ART adherence, CD4 cell count, Cox proportional hazards regression, Logistic regression, Quantile regression, Kaplan-Meier estimator, Log-rank test

Posted Date: March 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-328558/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Identification of factors affecting the survival lifetime of HIV+ terminal patients in Albert Luthuli municipality of South Africa

Pepukai Bengura^{1*}, Principal Ndlovu² and Mulalo Annah Managa²

Abstract

Background: South Africa has the biggest HIV epidemic in the world, with Mpumalanga province in which Albert Luthuli municipality is located having the second highest HIV prevalence rate after KwaZulu-Natal province. The objective of the study was to identify the factors that affect the survival lifetime of HIV+ terminal patients in rural district hospitals of Albert Luthuli municipality.

Methods: This is a typical retrospective cohort longitudinal design study whereby cohort of HIV+ terminal patients was retrospectively followed from 2010 to 2017 until a patient died, transferred to another hospital, lost to follow-up or was still alive at the end of the observation period. The follow-up time for each patient started at the time the patient got initiated to the ART programme at the hospital's wellness centre. Nonparametric survival analysis and semiparametric survival analysis methods were used to analyse the data.

Results: Through Cox proportional hazards regression modelling, it was found that ART adherence (poor, fair, good), Age, Follow-up mass, Baseline sodium, Baseline viral load, Follow CD4 count by Treatment (Regimen 1) interaction and Follow-up lymphocyte by TB history (yes, no) interaction had significant effects on survival lifetime of HIV+ terminal patients (p -values < 0.1). Furthermore, through quantile regression modelling, it was found that short, medium and long survival times of HIV+ patients, respectively represented by the 0.1, 0.5 and 0.9 quantiles, were not necessarily significantly affected by the same factors.

Discussion: The Cox PH modelling and the quantile regression analysis complemented each other in answering the research question. However, although the Cox PH modelling was the main approach in this study, the quantile regression analysis results are more informative than the Cox PH modelling results.

Conclusion: The study identified and modelled the factors affecting the survival of HIV+ terminal patients in Albert Luthuli Municipality by using Logistic regression, Cox PH regression, and Quantile regression modelling. Cox regression modelled the factors affecting the survival lifetime of HIV+ terminal patients as: ART adherence, Age, Follow-up mass, Baseline sodium, Baseline viral load and interactions of Follow-up lymphocyte by TB history and Follow-up CD4 by Treatment (Regimen 1).

Keywords: Survival time, Follow-up time, HIV/AIDS disease, Antiretroviral therapy, ART adherence, CD4 cell count, Cox proportional hazards regression, Logistic regression, Quantile regression, Kaplan-Meier estimator, Log-rank test.

Introduction

According to United Nations Agency for International Development (UNAIDS) Gap Report (2016), South Africa has the biggest HIV epidemic in the world, with estimated 7 million (12.7%) living with HIV in 2015. Statistics South Africa Report (2015) indicates a rise in AIDS-related deaths from 29.2% in 2014 to 30.5% in 2015. South Africa's Mpumalanga province has the second highest HIV prevalence rate after KwaZulu-Natal province, and the Gert Sibande district which is in Mpumalanga is leading all districts in the country with 46.1% prevalence rate (Masinga, 2014; Motsoaledi, 2013). The Gert Sibande district HIV prevalence stood at 40.5% in 2011 (Bezuidenhout et al., 2014) and at 38.6% in 2015 (National Department of Health, 2015). The Gert Sibande district has Albert Luthuli as one of its seven municipalities. In comparison to Gert Sibande, the HIV prevalence for Albert Luthuli stood at 43.2% in 2011 Nkosi (2017), and the prevalence stood at 44% in 2014 (Chief Albert Luthuli Municipality, 2018).

South Africa is known to be having the largest antiretroviral treatment (ART) programme globally. The number of patients in Gert Sibande receiving ART up to November 2011 was 32 979 (3.2%) (Bezuidenhout et al., 2014), and this number rose to 76 632 (6.75%) in 2015 (Mpumalanga Provincial AIDS Council, 2016). South Africa ART programme has saved millions of lives and infections averted. However, the dual burden of Tuberculosis (TB) and HIV disease continues unabated. South Africa continues to see an increase in HIV related TB incidences and has not yet felt the impact of HIV prevention programmes. There are some statistical methods such as survival analysis methods that can be used to analyse data for monitoring the survival time benefit of HIV/AIDS interventions, and for investigating potential factors that may affect the survival probability of HIV/AIDS patients.

Survival analysis methods are used to estimate the risk of death or progression of a disease and to provide predictions that help clinicians to estimate trends in their patient outcomes (Nakhae & Law, 2011). On one hand, the methods are used to estimate the time period during which an event (for example, death) can happen, and on the other hand, to estimate the impact of various independent factors on the time distribution to the occurrence of an event (Melnyk et al., 1995).

Despite an array of expensive treatments and preventive interventions used to combat HIV/AIDS in South Africa, the prevalence and death toll due to the disease, still remain too high. Consequently, the broad question is "Are there other factors, that include treatments and

preventive interventions, which negatively or positively significantly affect the survival lifetime of HIV/AIDS patients in South Africa despite the availability of ARTs?”. The main objective of this study is therefore to use survival analysis methods to identify factors that affect the survival lifetime of HIV/AIDS patients in South Africa (despite the availability of ARTs), using data from Albert Luthuli municipality hospitals.

Methods

Study design

The cohort of HIV+ terminal patients on ART was followed until a patient died or until a patient was censored as a result of loss to follow-up or transferred to another hospital or as a result of the conclusion of the study. The cohort of patients was divided into children (age ≤ 10 years), adolescents (age 11-19 years) and adults (age ≥ 20 years) as was done in previous studies by Bakanda et al. (2011) and by Moshago et al. (2014). Each stratum was further divided into males and females. The main event of interest in this research was death of an HIV+ terminal patient. The patients who transferred out of the hospital or got lost to follow-up were classified as drop-out patients.

Study variables

The variables which form part of the routine hospital records in Albert Luthuli municipality were used in this study and are described as follows. The length of the survival time of an HIV+ terminal patient was the dependent variable for the study. The categorical independent variables were gender, hospital (Carolina, Embhuleni), WHO stage (1, 2, 3, 4), marital status (single, married, staying together, widowed/separated/divorced), treatments (regimen 1)(NVP+D4T+3TC, EFV+D4T+3TC, EFV+AZT+3TC and EFV+3TC+TDF), ART adherence (poor, fair, good), transferred from hospital (yes, no) and lost to follow-up (yes, no). The continuous independent variables for the study which except age were classified into baseline and follow-up variables were mass, CD4 cell count, haemoglobin, lymphocyte, white blood cell count, viral load, creatinine, total protein, sodium and alanine transaminase. Diabetes and hypertension were excluded because their records were found in very few patient files.

Sampling procedures and demographics

An optimum sample size was determined using Monkey Survey online sample size calculator and also using Rao soft online sample size calculator (Rao & Rao, 2004), with a margin of error of 5% or 95% confidence level. Alternatively, considering Leulseged and Ayele (2019), the following statistical assumptions could have had been considered in determining the same sample size: survival probability of HIV patients of 0.85 (from Tadege, 2018), 5% precision or margin error, loss of 0.45 (from Damtew et al., 2015) and 95% level of confidence interval.

The application of sample size formula (Eneyew et al., 2016), $N = \frac{Z^2 p(1-p)}{\alpha^2}$, where N = sample size, $Z = 1.96$ (critical value at 95% level of confidence), p = proportion of deaths and α = type-1 error (0.05) would have yielded the same sample size as 357. The estimated total sample size was proportionally and randomly allocated to the two study sites (Embhuleni and Carolina hospitals with proportions of 79% and 21% respectively) and according to the age and gender proportions as per population proportions.

Statistical analysis

Basics in survival analysis

In the context of this study, suppose that T is a random variable representing the survival time of an HIV+ terminal patient during the observation period. Furthermore, suppose that $F(t)$ and $f(t)$ are the distribution and probability density functions of T , respectively. Then, the respective survivor and hazard functions of interest are (Box-Steffensmeir et al., 2004):

$$S(t) = P(T \geq t) = 1 - F(t) \text{ and } h(t) = \frac{f(t)}{S(t)} = -\frac{d[\ln S(t)]}{dt}. \quad (1)$$

Furthermore, the respective mean and the median survival times are given by:

$$E(T) = \int_0^{\infty} t f(t) dt = \int_0^{\infty} S(t) dt \text{ and } S^{-1}(0.5). \quad (2)$$

Also, of interest, can be the cumulative hazard function given by (Klein & Moeschberger, 2003; Kleinbaum & Klein, 2012):

$$H(t) = \int_0^t h(u) du = -\ln [S(t)]. \quad (3)$$

The Kaplan-Meier estimator of the survivor function

In the context of this study, let: $0 < t_1 < t_2 < t_3 < \dots < t_k < \infty$ be k observed times of death of patients in the cohort of HIV+ terminal patient during the observation period; $d_1, d_2, d_3, \dots, d_k$ be the respective number of deaths at each of these times; and $n_1, n_2, n_3, \dots, n_k$ be the corresponding number of remaining patients in the cohort at the respective times. The Kaplan-Meier (KM) estimator or the Product-Limit estimator of the survivor function $S(t)$ when death times are tied is given by (Etikan et al., 2017; Cleves et al., 2010):

$$\hat{S}(t) = \prod_{i|t_i \leq t} 1 - \left(\frac{d_i}{n_i}\right). \quad (4)$$

The cohort of HIV+ terminal patients in this study is not homogeneous with respect to their characteristics that may affect their survival. Hence, it will be necessary to test the equality of survivor functions among groups of patients. That is, to test the null hypothesis of the form

$$H_0: S_1(t) = S_2(t) = \dots = S_r(t) \equiv h_1(t) = h_2(t) = \dots = h_r(t), \quad (5)$$

for $t \geq 0$, where $S_i(t)$ and $h_i(t)$ are the respective survivor and hazard functions of the i^{th} group patients, and r is the number of groups. The alternative hypothesis of most interest is that the survival time of one group is stochastically bigger or smaller than the survival time for the other group, and is given by:

$$H_1: \text{either } S_i(t) \geq S_j(t), \text{ or } S_i(t) \leq S_j(t) \text{ for some } i \neq j. \quad (6)$$

A vertical gap between two graphs of the survival functions shows that at a given period of time, one group had a higher probability of surviving while a horizontal gap shows that one group took longer to experience an event (Etikan et al., 2017).

The log-rank test is the most commonly used statistical test for comparing survival functions of two or more groups (Clark et al., 2003; Etikan et al., 2017). A very important assumption for the appropriate use of the log rank test (and the Cox PH regression model) is that the hazards are proportional over time and this implies that the effect of a risk factor is constant over time (Sullivan, 2016).

The Cox proportional hazards (PH) model

The objective of the study is to identify factors that affect the survival of HIV+ terminal patients. The Cox proportional hazards (PH) model expresses the patient hazard rates as functions of potential such factors (covariates) as follows. Let $\mathbf{X}_i^T = (X_{i1}, X_{i2}, \dots, X_{ip})$ be a p -dimensional vector of the values of the covariates associated with the i^{th} patient. Then, the Cox proportional hazards regression model is as follows (Klein & Moeschberger, 2003; Etikan & Babatope, 2018):

$$h_i(t) = h_0(t) \exp\{\mathbf{X}_i^T \boldsymbol{\beta} = \sum_{j=1}^p \beta_j X_{ij}\}, \quad (7)$$

where $\boldsymbol{\beta}^T = (\beta_1, \beta_2, \dots, \beta_p)$ is a p -dimensional vector of regression coefficients to be estimated from the data, and $h_0(t)$ is the unspecified baseline hazard function that does not have to be

estimated. The hazard model in model (7) makes no assumptions about the shape of the hazard function over time. A hazard function could be constant, increasing, decreasing, or it could be a combination of two or three of these graph trends.

Model (7) can be written in terms of the survivor function (Kleinbaum & Klein, 2012) :

$$S_i(t) = S_0(t) \exp\{X_i^T \beta\}. \quad (8)$$

The model (7) assumptions which may be violated by the data, are: (i) the covariates X_i^T do not vary with time and hence the hazard rate ratios of pairs of patients do not vary with time; (ii) censoring and survival are independent; and (ii) the log hazard rate is indeed a linear function of the covariates.

The parameters are estimated as values of β which maximize the Cox likelihood (also called partial likelihood) function for censored data. The partial likelihood function is as follows (Hosmer et al., 2008; Cleves et al., 2010):

$$L(\beta) = \prod_{j=1}^k \left(\frac{\exp(X_j^T \beta)}{\sum_{i \in R_j} \exp(X_i^T \beta)} \right). \quad (9)$$

where R_j is the group of patients at the risk of death at time t_j . In particular, the maximum likelihood estimate of β , which is $\hat{\beta}$, is found by iteratively solving the equations $\frac{\partial \ln L(\beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}} = 0$. The popular iteration algorithms are the Fisher's scoring (Storvik, 2011) and Newton-Raphson algorithms (Zhou, 2016). The estimate of the covariance of $\hat{\beta}$ is a function of the inverse of the matrix $\frac{\partial^2 \ln L(\beta)}{\partial \beta^2} \Big|_{\beta=\hat{\beta}}$, and is of the form:

$$\widehat{Cov}(\hat{\beta}) = (X^T \widehat{W} X)^{-1}, \quad (10)$$

where \widehat{W} is a diagonal weight matrix, and X is the design/incidence matrix whose i^{th} row is X_i^T . The standard error of $\hat{\beta}_j$ is the square root of the j^{th} diagonal element of model (10).

A graph of the cumulative hazard of the Cox-Snell residuals versus the residuals was used to check the goodness of fit of the model. The Cox-Snell residual for the i^{th} observation from which the cumulative hazard of the residuals is (Ansin, 2015; Cleves et al., 2010):

$$CSr_i = \widehat{H}_0(t_i) \exp\{X_i^T \widehat{\beta}\}, \quad (11)$$

where $\widehat{H}_0(t_i)$ is an estimate of the cumulative baseline hazard rate which is obtained from the Cox PH model fit when all covariates are set to equal to zero ($\mathbf{X}_i^T = \mathbf{0}$ for all i). The graph of the cumulative hazard of the Cox-Snell residuals versus the Cox-Snell residuals is expected to approximately follow a 45° continuous straight line passing through (0,0) with a slope approximately equal to 1 if the conditional distribution of the cumulative hazard function given the covariate vector is reasonably well approximated by the exponential distribution.

Quantile regression in survival analysis

Recall that $F(t) = P(T \leq t) = 1 - S(t)$ (equation (1)) is the distribution function of the random variable T representing the survival time of a HIV+ terminal patient during the observation period. Then for $\alpha \in (0,1)$, $Q(\alpha) = F^{-1}(\alpha)$ is the quantile function of the random variable T . For example, when $\alpha = 0.5$, then $Q(0.5) = F^{-1}(0.5)$ is the median of the distribution of T . $Q(\alpha)$ for all $\alpha \in (0,1)$ provides a complete picture of the distribution of T . Suppose that the potential factors which affect the survival of HIV+ patients are represented as in the Kaplan-Meier estimator of the survivor function above, and that they are related to the survival time by the regression model:

$$T_i = \exp\{\mathbf{X}_i^T \boldsymbol{\beta} + \sigma \varepsilon_i\}, \quad (12)$$

where T_i is the survival time of the i^{th} patient, σ is an unknown scale parameter and the ε_i are independent random errors with an assumed distribution. Model (40) is called an accelerated failure time (AFT) model, from which the quantile regression model corresponding to the Cox regression model in is (Chaudhuri et al., 1997; Peng et al., 2008; Flom & Peter Flom Consulting, 2011):

$$Q_i(\alpha, \boldsymbol{\beta}^\alpha) = \exp\{\mathbf{X}_i^T \boldsymbol{\beta}^\alpha\} = \inf\{t: P(T_i \leq t | \mathbf{X}_i^T) \geq \alpha\} = F_i^{-1}(\alpha), \quad (13)$$

where $F_i(t)$ is the distribution function of T_i which is derived from the distribution of the ε_i , and $\boldsymbol{\beta}^\alpha$ is the unknown vector of effects of the covariates on the α^{th} quantile of the distribution of T_i . $\boldsymbol{\beta}^\alpha$ may vary with α which means $\boldsymbol{\beta}^\alpha$ measures the effect of covariates not only in the centre but also in the upper and lower tails of the survival time distribution.

Results

The Cox PH model

Table 1 shows that the main effects which are associated with poor survival probability at 0.1 significance level are: Age (Hazard Ratio=1.037, $p - value < 0.0411$), Baseline sodium (Hazard Ratio=1.048, $p - value < 0.0472$), \ln (Baseline viral load) (Hazard Ratio=1.853, $p - value < 0.0009$) and ART adherence (poor relative to fair ART adherence) (Hazard Ratio=5.334, $p - value < 0.0004$). This means an increase in Age by one year resulted in a statistically significant increase in the HIV hazard (by about 3.7%) among the HIV+ terminal patients, while holding all other covariates constant. An increase in Baseline sodium by one-unit resulted in a statistically significant increase in the HIV hazard (by about 4.8%) among the HIV+ terminal patients, while holding all other covariates constant. Similarly; 2.718 times increase in Baseline viral load holding all other covariates constant, is associated with (about 85.3%) increase in the hazard of HIV among the HIV+ terminal patients.

The estimated 95% confidence interval of the hazard ratio for Age is (1.001, 1.073), which implies that for every one-year increase in the Age of an HIV/AIDS patient, the hazard rate is as much as 7.3% low to 0.1% lower. With regards ART adherence, patients with a poor ART adherence faced an HIV hazard that was about 18.75% greater than of those with a fair ART adherence, while holding other covariates constant. The estimated 95% confidence interval of the hazard ratio for poor ART adherence is (2.104, 13.521), which implies that the hazard ratio for poor adherence to ART relative to fair adherence to ART is as high as 13.521 and as low as 2.104. The estimated 95% confidence interval of the hazard ratio for Baseline sodium is (1.001, 1.097), which implies that for every one-unit increase in Baseline sodium of an HIV/AIDS patient, the hazard rate is as much as 9.7% low to 0.1% lower. The estimated 95% confidence interval of the hazard ratio for \ln (Baseline viral load) is (1.289, 2.662), which implies that for every 2.718 times increase in Baseline viral load, the hazard rate is as much as 2.662 times low to 1.289 times lower. Table 3.2.7 also shows that the interaction effect associated with low survival probability at 0.1 significance level is Follow-up lymphocyte by TB history (yes relative to no TB history) (Hazard Ratio = 1.116, $p - value < 0.0248$). This means an increase in Follow-up lymphocyte by one-unit for patients having TB history relative to those without TB history resulted in a statistically significant increase in the HIV hazard (by about 11.6%). The estimated 95% confidence interval for the interaction Follow-up lymphocyte by TB history (yes relative to no TB history), is (1.043, 1.193), which implies that

the hazard ratio for this interaction is as much as 19.3% low to 4.3% lower.

The main effect which is associated with good survival probability at 0.1 significance level as in Table 1, is Follow-up mass (Hazard Ratio = 0.971, p – value < 0.0482). A one-unit increase in Follow-up mass resulted in a statistically significant decrease in the HIV hazard (by about 0.29%) among the HIV+ terminal patients while holding all other covariates constant. The estimated 95% confidence interval of the hazard ratio for Follow-up mass is (0.943, 1.000), which implies that for every one-unit increase in Follow-up mass of an HIV/AIDS patient, the hazard rate is as much as 5.7% low to 0% lower. It is noted that Follow-up mass has no significant effect on the survival of patients at the 0.05 level of significance as evidenced by the inclusion of 1 in the 95% confidence interval of its hazard ratio (0.943, 1.000). Table 1 also shows the interaction effect which is associated with high survival probability at 0.1 significance level as Follow-up CD4 by Treatment (Regimen 1) (Hazard Ratio = 0.936, p – value < 0.0024). This means an increase in Follow-up CD4 by one-unit for patients treated with (EFV+AZT+3TC) resulted in a statistically significant decrease in HIV hazard (by about 6.4%). The estimated 95% confidence interval for the interaction Follow-up CD4 by Treatment (Regimen 1) is (0.897, 0.977), which implies that the hazard ratio for this interaction is as much as 10.3% low to 2.3% lower.

Table 1: Maximum likelihood estimates of the final Cox PH model

Parameter	Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence	
Treatment (Regimen 1)=1 (NVP+D4T+3TC)	4.2922	2.1974	3.8156	0.0508	.	.	.
Treatment (Regimen 1)=5 (EFV+D4T+3TC)	3.5266	2.1456	2.7017	0.1002	.	.	.
Treatment (Regimen 1)=6 (EFV+AZT+3TC)	15.4000	3.3366	21.3030	<.0001	.	.	.
Treatment (Regimen 1)=19 (NVP+TDF+3TC)	-12.3093	1916	0.0000	0.9949	.	.	.
ART adherence=1 (poor)	1.6741	0.4746	12.4422	0.0004	5.334	2.104	13.521
ART adherence=3 (good)	-1.0505	0.8088	1.6869	0.1940	0.350	0.072	1.707
TB history=1 (yes)	-1.2713	0.7127	3.1820	0.0745	.	.	.
Age	0.0360	0.0176	4.1727	0.0411	1.037	1.001	1.073
Follow-up mass	-0.0293	0.0148	3.9037	0.0482	0.971	0.943	1.000
Follow-up lymphocyte	0.0080	0.0319	0.0634	0.8012	.	.	.
Baseline sodium	0.0467	0.0235	3.9391	0.0472	1.048	1.001	1.097
Follow-up CD4	0.0010	0.0043	0.0489	0.8250	.	.	.
<i>ln</i> (Baseline viral load)	0.6166	0.1850	11.1100	0.0009	1.853	1.289	2.662
Follow-up CD4*TreatmentR1=1	-0.0046	0.0050	0.8221	0.3646	0.996	0.991	1.002
Follow-up CD4*TreatmentR1=5	-0.0056	0.0047	1.4015	0.2365	0.995	0.992	0.999
Follow-up CD4*TreatmentR1=6	-0.0667	0.0220	9.2277	0.0024	0.936	0.897	0.977
Follow-up CD4*TreatmentR1=19	-0.0015	4.1730	0.0000	0.9997	0.999	0.000	3563.1
Follow-up lymphocyte*TB history=1	0.1013	0.0451	5.0405	0.0248	1.116	1.043	1.193

Key: TreatmentR1=Treatment (Regimen 1); Reference levels: (18=EFV+TDF+3TC) for Treatment (Regimen 1); 'fair' for ART adherence and 'no' for TB history.

Comment 1 on Table 1: The hazard ratio and the 95% hazard ratio confidence limits columns in Table 1, for the main effects of interaction are left empty. SAS omits hazard ratio entries for terms involved in interactions as a reminder that the hazard ratios corresponding to these effects depend on other variables in the model (University of California, Statistical Consulting Group, 2019).

Comment 2 on Table 2: The 95% hazard ratio confidence limits for covariates involving Treatment (Regimen 1) (NVP+TDF+3TC) are extremely wide supposedly because of small sample size and heavy censoring (low death rate) (Fay et al, 2013).

Nonparametric inferences about the survivor functions

The cohort of HIV+ terminal patients in this study is not homogeneous with respect to their characteristics that may affect their survival. Hence, it will be necessary to test the equality of survivor functions among groups (strata) of patients. Log-rank tests and the Kaplan-Meier functions presented in this section are for independent factors which are presumed by the researcher to be statistically and clinically significant. Hazard ratios for pairwise comparisons of strata are presented at the end of this section in Table 7.

Kaplan-Meier survival functions for ART adherence groups

Figure 1 shows that HIV/AIDS patients with poor ART adherence have the lowest survival probability (as low as around 40%) while patients with good ART adherence have the highest survival probability (at least around 95%). The hazard ratios in Table 7 show that patients with poor ART adherence and those with fair ART adherence experience the event (death) earlier than those with good ART adherence. The test statistics in Table 2 show that ART adherence strata in Figure 1 are statistically different ($p - value < 0.0001$). The analysis of hazard ratios shows that patients with poor adherence are about 6 times more likely to die from HIV/AIDS related illness relative to patients with fair ART adherence while patients with poor adherence are about 18 times more likely to die from HIV/AIDS related illness relative to patients with good adherence.

Table 2: Log-rank and other tests for equality of ART adherence survivor functions

Test of Equality over Strata (ART adherence)			
Test	Chi-Square	DF	Pr>Chi-Square
Log-Rank	52.9375	2	<.0001
Wilcoxon	41.3917	2	<.0001
-2Log(LR)*	45.0349	2	<.0001

Figure 1: Kaplan-Meier survival function estimates for ART adherence

Kaplan-Meier survival functions for Baseline viral load

Figure 2 shows that HIV/AIDS patients with Baseline viral load above 5000 HIV RNA copies/mm³ have the lowest survival probability (as low as around 65%) while patients with Baseline viral load below 50 HIV RNA copies/mm³ (undetectable viral load) have the highest survival probability (100%) throughout the follow-up period. The test statistics in Table 3 indicate that the survivor functions of the Baseline viral load strata are statistically different (p – value < 0.0001). The analysis of hazard ratios in Table 7 show that patients with Baseline viral load between 50 HIV RNA copies/mm³ and 5000 HIV RNA copies/mm³ have HIV hazard which is around 87% lower relative to patients with Baseline viral load above 5000 HIV RNA copies/mm³.

Table 3: Log-rank and other tests for equality of Baseline viral load survivor functions

Test of Equality over Strata of Baseline viral load			
Test	Chi-Square	DF	Pr>Chi-Square
Log-Rank	30.9448	2	<.0001
Wilcoxon	27.1307	2	<.0001
-2Log(LR)*	43.6135	2	<.0001

Figure 2: Kaplan-Meier survival function estimates for Baseline viral load strata

Kaplan-Meier survival functions for Follow-up lymphocyte by TB history groups

Patients with Follow-up lymphocyte below normal level and with TB history have the lowest

survival experience, and they experienced the event (death) earlier than other groups (see Figure 4). In addition, patients belonging to this stratum did not reach the end of the study alive. Patients with normal Follow-up lymphocyte levels and without TB history have the highest survival experience. The survivor functions of the strata in Figure 4 are significantly different as evidenced by the p-values in Table 4 ($p - \text{value} < 0.0001$). The hazard ratios in Table 7 show that patients with TB history relative to those without TB history (at Follow-up lymphocyte below normal) are about 2.8 times likely to experience HIV related hazard (death), while patients with TB history relative to those without TB history (at normal Follow-up lymphocyte) are about 52 times likely to experience HIV related hazard (death).

Table 4: Log-rank and other tests for equality of Follow-up lymphocyte by TB history survivor functions

Test of Equality over Strata (Follow-up lymphocyte by TB)			
Test	Chi-Square	DF	Pr>Chi-Square
Log-Rank	29.5848	3	<.0001
Wilcoxon	30.0687	3	<.0001
-2Log(LR)*	26.2431	3	<.0001

Figure 3: Kaplan-Meier survival function estimates for Follow-up lymphocyte by TB history strata

Kaplan-Meier survival function estimates for Follow-up CD4 by Treatment (Regimen 1)

In Figure 4, patients with Follow-up CD4 above 200 cells/mm³ and taking (EFV+3TC+TDF) or (EFV+D4T+3TC) have the highest survival experience if other covariates are kept constant. On the other hand, patients with Follow-up CD4 below 200 cells/mm³ and taking (EFV+AZT+3TC) have the lowest survival experience, if other covariates are kept constant. The difference across the Follow-up CD4 by Treatment strata are statistically significant (Table 5, $p - \text{values} < 0.0001$).

Table 5: Log-rank and other tests for equality of Follow-up CD4 count by Treatment (Regimen 1) survivor functions

Test of Equality over Strata (Follow-up CD4 by Treatment)			
Test	Chi-Square	DF	Pr>Chi-Square
Log-Rank	69.6268	9	<.0001
Wilcoxon	76.1848	9	<.0001
-2Log(LR)*	74.4060	9	<.0001

Figure 4: Kaplan-Meier survival function estimates for Follow-up CD4 by Treatment (Regimen 1) strata

Kaplan-Meier survival functions for Follow-up CD4 and related results

CD4 cell count is a key independent factor in this study because it remains the best measurement of a patient’s immune and clinical status, and supports diagnostic decision-making, particularly for patients with advanced HIV disease. Figure 5 shows some relevant strata in CD4 count while Figure 6 shows the bar graphs for Baseline CD4 and Follow-up CD4.

Table 6: Log-rank and other tests for equality of Follow-up CD4 strata survivor functions

Test of Equality over Strata (Follow-up CD4)			
Test	Chi-Square	DF	Pr>Chi-Square
Log-Rank	45.4274	3	<.0001
Wilcoxon	44.8056	3	<.0001
-2Log(LR)	51.3597	3	<.0001

Figure 5: Kaplan-Meier survival function estimates for Follow-up CD4 Count

The Kaplan-Meier survival functions for Follow-up CD4 count (figure 5) show that patients with: Follow-up CD4 count above 500 cells/mm³, Follow-up CD4 between 200 cells/mm³ and 500 cells/mm³ and Follow-up CD4 count below 200 cells/mm³ have comparatively high, moderate and low survival experiences respectively, assuming that other covariates are held constant. The difference across the Follow-up CD4 strata are statistically significant (Log-Rank, p – value < 0.0001). The big vertical gaps separating graphs confirm the difference

across the Follow-up CD4 strata. As shown in Figure 6, following ART initiation; out of the 357 patients, the Baseline CD4 counts were as follows: 254 (71.14%) had CD4 less than 200 cells/mm³, 80 (22.41%) had CD4 between 200 cells/mm³ and 500 cells/mm³ while 23 (6.44%) had CD4 greater than 500 cells/mm³. On the other hand, the follow-up CD4 counts were as follows: 139 (38.94%) had CD4 less than 200 cells/mm³, 165 (46.22%) had CD4 between 200 cells/mm³ and 500 cells/mm³ while 53 (14.85%) had CD4 greater than 500 cells/mm³. The condition of HIV disease when CD4 is below 200 cells/mm³ is called AIDS.

Figure 6: Bar graphs for the distribution of Baseline and Follow-up CD4 in HIV/AIDS patients

Table 7: Hazard ratios for the main and interaction effects strata

Factor	Point Estimate	95% Wald Confidence Limits	
TB history (yes vs no) at Follow-up lymphocyte =Below normal	2.847	1.445	5.608
TB history (yes vs no) at Follow-up	51.832	4.615	582.112
Treatment(Regimen1) (NVP+D4T+3TC vs EFV+3TC+TDF)	14.110	1.780	111.862
Treatment(Regimen1) (EFV+D4T+3TC vs EFV+3TC+TDF)	15.498	2.067	116.200
Treatment(Regimen1) (NVP+D4T+3TC vs EFV+D4T+3TC)	0.910	0.451	1.838
Treatment(Regime1) (NVP+D4T+3TC vs EFV+AZT+3TC)	2.941	0.636	13.595
Treatment(Regimen1) (EFV+D4T+3TC vs EFV+AZT+3TC)	3.230	0.753	13.851
Treatment(Regimen1) (EFV+AZT+3TC vs EFV+3TC+TDF)	4.798	0.435	52.932
TB history (yes vs no)	3.523	1.848	6.716
ART adherence (poor vs fair)	6.039	3.198	11.405
ART adherence (poor vs good)	18.441	4.233	80.332
Baseline viral load (viral load 50-5000 vs viral load	0.130	0.051	0.333
Follow-up CD4 (CD4 count <200 vs CD4 count >200)	5.493	2.731	11.048
Follow-up CD4 (CD4 count <200 vs CD4 count >500)	25.826	3.496	190.766
Follow-up lymphocyte (Below normal vs Normal)	1.810	0.557	5.878
Follow-up CD4 <200 vs >200 At Treatment (Regimen1) =(NVP+D4T+3TC)	8.381	1.805	38.906
Follow-up CD4 <200 vs >200 At Treatment (Regimen1) =(EFV+D4T+3TC)	4.092	1.834	9.128

Quantile regression modelling at quantile levels 0.1, 0.5 and 0.9

The covariates for the quantile models were selected using an automated ‘Proc QUANTSELECT’ procedure in SAS version 9.4. The results from quantile regression modelling the data produced the significant factors with negative significant effects on the 0.1st quantile of the log (survival time) at the 0.1 level of significance as: poor and fair ART adherence relative to good ART adherence, \ln (Follow-up viral load) and Follow-up white blood cell count while Follow-up CD4 and the interaction effect of Follow-up white blood cell count by Treatment (Regimen 1) have positive effects. Significant factors with negative significant effects on the 0.5th quantile of the log (survival time) at the 0.1 level of significance were WHO stage 3 relative to WHO stage 4, poor and fair ART adherence relative to good ART adherence and Baseline haemoglobin while the males relative females, Carolina hospital relative to Embhuleni hospital, WHO stage 1 relative to WHO stage 4 and Follow-up CD4 have positive effects. Lastly, the significant factors with negative significant effects on the 0.9th quantile of the log (survival time) at the 0.1 level of significance were poor and fair ART adherence relative to good ART adherence and \ln (Baseline viral load) while factors with positive effect are Carolina hospital relative to Embhuleni hospital, Follow-up CD4, Baseline white blood cell count and Follow-up sodium.

Some quantile process plots to study heterogeneity in the data

A visual comparison of the magnitudes of the effects of each covariate as shown in Figure 6 at quantile levels 0.1, 0.5 and 0.9 suggests the existence of heterogeneity in the data. In this section, heterogeneity is investigated using quantile process plots for (effects versus quantile levels) only selected highly significant covariates Follow-up CD4, \ln (Baseline viral load) and Treatment (Regimen 1). These covariates were significant either as main effects or as interaction effects in all of Cox regression, logistic regression and quantile regression modelling.

Figure 7: Estimated parameter versus quantile for log (survival time) with 95% confidence limits - Follow-up CD4, \ln (Baseline viral load) and Treatment (regimen 1)

Figure 7 shows that Follow-up CD4 has a positive effect on log (survival time) which increases with increase in the quantile levels. \ln (Baseline viral load) has a negative effect on log (survival time) which increases in magnitude with increase in the quantile levels up to the 0.8th quantile. Then the effect diminishes in magnitude. Thus, the effects of Follow-up CD4 and \ln (Baseline viral load) on log (survival time) are quantile specific. This is also true for the

intercept whose positive effect on log (survival time) increases with increase in the quantile levels. A closer look at the relationship between log (survival time) and \ln (Baseline viral load) (expressed by the slope) shows lack of significance for quantiles below 0.35th since zero lies within the 95% confidence limits. The Treatment (EFV+3TC+TDF), just as other treatments (EFV+D4T+3TC), (EFV+AZT+3TC) and (NVP+D4T+3TC), each had a sudden, negative, and worsening effect on log (survival time) up to around 0.3rd quantile level. As from around 0.3rd quantile upwards, the negative effect diminishes in magnitude for (EFV+3TC+TDF) and (EFV+AZT+3TC) while the diminishing effect for (EFV+D4T+3TC) and (NVP+D4T+3TC) takes place as from around the 0.8th quantile upwards. Thus, the effects of the covariates on \ln (survival time) are quantile specific.

Survival analysis of the effect of covariates on quantiles of the survival time of patients

According to Lin and Rodriguez (2013), the applications of quantile regression to survival analysis include studying the effect of a specific covariate on the survival time of an individual. The previous section has given us a picture of covariates which affect most patients with short, median and long follow-up time. The quantile regression analysis presented herein are for Follow-up CD4, \ln (Baseline viral load) and Treatment (Regimen 1). A given covariate may have a different effect on the survival of patients with low, medium and long follow-up time as shown in Figure 8. Figure 8 shows that an increase in Follow-up CD4 count results in high improvement in survival of patients in the 0.5th quantile while a gradual improvement is seen in 0.1st and 0.9th quantiles.

Figure 8: Survival analysis of covariate effect on quantiles of Survival time

In terms of survival, HIV/AIDS patients on ART for around 3.75 years respond better to increase in Follow-up CD4 count than patients on ART for any other duration within the 7.5 years. An increase in Baseline viral load is accompanied by a decrease in survival for all the quantiles of Survival time. In terms of survival, HIV/AIDS patients on ART for around 3.75 years are the most negatively affected by an increase in Baseline viral load than patients on ART for any other duration within the 7.5 years. In terms of survival, HIV/AIDS patients on ART for around 3.75 years respond better to Treatment (Regimen 1) than patients on ART for any other duration within 7.5 years. Treatments (EFV+3TC+TDF) and (NVP+3TC+TDF) are associated with higher survival at all quantiles than Treatments (NVP+D4T+3TC), (EFV+D4T+3TC) and (EFV+AZT+3TC).

Discussions

The findings in this study might not adequately compare with some related studies because of the following reasons:

- i) In the case of Damtew et al. (2015) and Moshago et al. (2014), the focus was mainly on death as the survival threat. This study, in addition to considering death as a threat to HIV/AIDS patients, it also considered patient transfer out of hospital and loss of patients to follow-up as phenomena which affect the long-term success of antiretroviral therapy programme and ultimately leading to increased death cases among HIV/AIDS patients especially in the case of lost to follow-up.
- ii) The national strategic plan of 2012-2016 by the Department of Health of South Africa stipulates that the retention rate of patients on ART should be at least 80% (Mlangeni & Senkubuge, 2016). The long-term retention of HIV/AIDS patients in Albert Luthuli was around 13% for the 2010-2017 period; this sharply deviates from the national strategic plan of 2012-2016. This sharp deviation in retention rates could be explained by challenges experienced in Albert Luthuli in form of high patient transfers (at 49.3%) and high patient loss to follow-up (at 26%).
- iii) Related studies done by Damtew et al. (2015) and Sieleunou et al. (2008) focused on adult patients and on patients with ages greater than 15 years. This study is related to the study by Moshago et al. (2014) in that the sample was exhaustively handled by dividing it into children, adolescents and adults. This was done with the view to find a solution to all groups at once using the limited resources available.
- iv) Most similar studies which were carried out were limited mostly to a follow-up period of 4-5 years as in the case of Sieleunou et al. (2008), Damtew et al. (2015) and Mlangeni & Senkubuge (2016). This research study spanned over a longer follow-up period of 7.5 years with the view to improve on the quality of the findings.
- v) The sample size for this research closely matches the sample sizes for studies by Tadege (2018), Teka et al. (2015), and Mlangeni & Senkubuge (2016), however, it is comparatively smaller than the sample sizes for studies by Shebeshi (2011), Damtew et al. (2015) and Sieleunou et al. (2008) who retrospectively followed greater numbers of HIV/AIDS patients which were 456; 784 and 1187 respectively.
- vi) This study reports a comparatively high percentage in terms of patients lost to follow-up and this sharply contrasts to the report by Sieleunou et al. (2008) on HIV/AIDS

patients in Cameroon where some intensive follow-up to those lost to follow-up was carried out. The follow-up in Albert Luthuli Municipality involves phone call, whereas in Cameroon follow-up involves phone calls and messengers.

Furthermore, one of the assumptions used in survival analysis is that patients who drop out or are censored have the same survival prospects as those who continue to be followed. This assumption may theoretically hold, but from the point of view of the practicalities of this study, patients lost to follow-up were defaulters to ART adherence and hence were presumed to be at higher risk to HIV/AIDS than those who were followed continuously.

- vii) Unlike most similar studies carried in South Africa for example by Mlangeni & Senkubuge (2016), and other countries which concentrated most on urban institutions; this study just like the study by Sieleunou et al. (2008), analysed the outcomes of ART in a marginalised rural set-up which presents its own peculiar challenges.
- viii) This research focused on complete hospital records only as independent factors, that means any other factors with the potential to affect the survival of HIV+ terminal patients which were not part of the hospital records or were incomplete hospital records (for example records on hypertension and diabetes) were not considered in this research.
- ix) Unlike in similar studies by Damtew (2015), Moshego (2012) and by Shebeshi (2011) where WHO stage emerged as a significant covariate, in this study it was not. The difference, the discrepancy or lack of statistical significance could be due to difference in calibration of the research tools, could be attributed to sample size or could probably be as a result of difference in research designs.

The application of quantile regression in identification of how covariates effects vary across patient quantiles, warrants the need for differential treatment based on lower tail (shorter length of follow-up time of a patient), median (median length of follow-up time of patients) and on upper tail (long length of follow-up time of patients). Thus, the study of quantile process plots may be applied in designing and timing of treatments which suit patients with short, medium or long follow-up times.

This study has shown that most patients (61.9%) died within 1 year from ART initiation and this concurs with similar studies by Shebeshi (2011) in which most of the HIV/AIDS patients also died during the first 12 months. In addition, this study has found out that the greatest number of patients died within the interval from 6 months to 1 year. These findings on the distribution of mortality among HIV/AIDS patients are important in the timing of the treatment interventions.

The Kaplan-Meier and hazard ratios in this study have shown that HIV/AIDS male patients have lower survival experience than HIV/AIDS female patients. Men are far more likely to die from an HIV-related illness than women, although women are becoming infected with HIV at a much faster rate (Merab, 2018). In addition, Ochieng-Ooko et al. (2010), pointed out that late treatment initiation predisposes men to a poor clinical outcome and to being lost to follow-up.

Although this study considered participants from Albert Luthuli municipality, the nature of the covariates used for modelling, makes the findings applicable to most settings where ART is administered among the HIV+ terminal patients. All the independent factors identified as significant factors affecting the survival time of HIV+ terminal patients in this study should be accurately recorded, tested, analysed and should be monitored closely in order to lengthen the survival time and to improve the wellness of the HIV+ terminal patients.

Some noticeable improvements in terms of immunologic, virologic and weight recoveries among the HIV/AIDS patients as reflected in quantile plots, suggest concerted effort given by the health care staff to ensure a long and healthy life for all South Africans as enshrined in the mission statement of the Department of Health. However, most Antiretroviral drugs (ARVs) have neuropsychiatric side-effects, for example efavirenz (EFV) is commonly associated with insomnia and headache (Reid et al., 2012). In addition, ART may lead to the manifestation of TDF nephrotoxicity as has been mentioned above.

Finally, Cox regression, logistic regression and quantile regression are compared in answering the research question of this study. Logistic regression analysis was performed not to directly answer the question but as an exploratory analysis to determine whether Patient status (censored, uncensored) was associated with the covariates. The Cox PH modelling should be preferred (versus the logistic modelling) as it directly answered the research question. In addition, Cox model should be preferred over the logistic model when survival time

information is available and there is censoring. As pointed out by Kleinbaum and Klein (2012), the Cox regression uses more information in form of ‘survival times’, whereas the logistic regression considers a (0, 1) outcome and ignores survival times and censoring.

The Cox PH modelling and the quantile regression analysis complemented each other in answering the research question. However, although the Cox PH modelling was the main approach in this study, the quantile regression analysis results are more informative than the Cox PH modelling results.

Conclusion

Cox regression modelled the factors affecting the survival lifetime of HIV+ terminal patients as: ART adherence, Age, Follow-up mass, Baseline sodium, Baseline viral load and interactions of Follow-up lymphocyte by TB history and Follow-up CD4 by Treatment (Regimen 1). Quantile regression determined significant factors on the 0.1st quantile of the log (survival time) as: ART adherence, \ln (Follow-up viral load), Follow-up white blood cell count, Follow-up CD4 and the interaction effect of Follow-up white blood cell count by Treatment (Regimen 1). The significant factors with significant effects on the 0.5th quantile of the log (survival time) at the 0.1 level of significance were determined as: WHO stage, ART adherence, Baseline haemoglobin, Gender, Hospital and Follow-up CD4. Lastly, the significant factors with significant effects on the 0.9th quantile of the log (survival time) at the 0.1 level of significance were determined as: ART adherence, \ln (Baseline viral load), Hospital, Follow-up CD4, Baseline white blood cell count and Follow-up sodium.

Acknowledgement

We are grateful to HIV/AIDS clinics staff for Carolina and Embhuleni hospitals. The Data capturers assisted with data collection while the administrative staff swiftly performed data collection permissions protocols.

Author information

Author notes

Pepukai Bengura, Prof Principal Ndlovu and Mulalo Annah Managa contributed equally to this work hence the corresponding author is transferable if the need arises.

Authors' contributions

PB got involved in obtaining ethics approval, data collection/analysis and drafting of the manuscript. PD and MAM contributed to data interpretation, project design and editing the manuscript. All authors read and approved the final manuscript.

Corresponding author

Correspondence to Pepukai Bengura.

Funding

Not applicable.

Ethics declarations

Ethics approval and consent to participate

This study was conducted in accordance with the South African local and national research guidelines. Ethics approval for this project was obtained from UNISA Ethics Review Committee (ERC) with the approval number being (2017/SSR ERC/005). The permission to conduct the study at Carolina and Embhuleni hospitals was obtained from Mpumalanga Department of Health with the permission number being (MP_201708_013). The consent to participate does not apply to a patient since no reference to an individual respondent was made, all results were handled in aggregate format.

Availability of data and materials

The data that support the findings of this study are available from the Department of health, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Department of health.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Statistics, Unisa, Pretoria, South Africa

²Department of Statistics, Unisa, Pretoria, South Africa

References

- UNAIDS. (2016). HIV and AIDS in South Africa. `Gap Report 2016'. Geneva, Switzerland: UNAIDS. Retrieved May 23, 2017, from www.unaids.org/en/media/unaids/.../unaidspublication/.../UNAIDS_Gap_report_en.pdf
- Statistics South Africa. (2011). *Census 2011 Municipal report*. Pretoria, South africa: Statistics South Africa. Retrieved June 13, 2017, from www.statssa.gov.za/census/census_2011/census_products/MP_Municipal_Report.pdf
- Masinga, S. (2014, April 4), Inside South Africa's HIV and AIDS capital Gert Sibande District. *Ziwaphi*. Retrieved from <https://www.ziwaphi.com/2014/04/04/inside-south-africas-hiv-and-aids-capital-gert-sibande-district/> on 14 May 2020
- Motsoaledi, A. (2013, November 22). Mpumalanga's Gert Sibande district has highest HIV rate. *City Press*. Retrieved April 12, 2017, from <http://www.news24.com/Archives/City-Press/Mpumalangas-Gert-Sibande>
- Bezuidenhout, S., Ogunsanwo, D.A., & Helberg, E.A. (2014). Patient satisfaction at accredited antiretroviral treatment sites in the Gert Sibande District. *Afr J Prm Health Care Fam Med*, 6(1). Retrieved January 22, 2018, from doi:<http://dx.doi.org/10.4102/phcfm.v6i1.627>
- National Department of Health. (2015). *The 2015 National Antenatal Sentinel HIV & Syphilis Survey, South Africa, National Department of Health*. Retrieved May 20, 2019, from <http://www.health.gov.za/index.php/shortcodes/download=2584:2015-national-antenatal-hiv-prevalence-survey-final-23oct17>
- Nkosi, J. (2017). *Chief Albert Luthuli local municipality Integrated Development Plan (IDP) 2017-2022*. Caroline. Mpumalanga. Chief Albert Luthuli Local Municipality
- Chief Albert Luthuli Municipality. (2018). *Integrated Development Plan (IDP) 2018/2019*. Retrieved January 12, 2018, from <https://cogta.mpg.gov.za/IDP/2018-19IDPs/Gert%20Sibande/ChiefAlbertLuthuli2018-19.pdf>
- Mpumalanga Provincial AIDS Council. (2016). *Annual progress report 2014/15*. Retrieved December 10, 2017, from http://sanac.org.za/download/563/resources/3385/mp_psp-annual-progress-report_final-report.pdf
- Nakhaee, F & Law, M. (2011). Parametric modelling of survival following HIV and AIDS in the era of highly active antiretroviral therapy: data from Australia. *Eastern Mediterranean Health Journal*, 17(3). Retrieved April 17, 2018, from <https://pdfs.semanticscholar.org/0046/d27a30da17665b5075263cd2b62b952206ae.pdf>
- Melnyk, A., Pagell, M and Jorae, G. (1995). Applying survival analysis to operations management: Analyzing the differences in donor classes in the blood donation process. *Journal of Operations Management* 13 (1995) 339-356. Retrieved May 15, 2020, from

<https://www.sciencedirect.com/science/article/pii/0272696395000313>

- Bakanda, C., Birungi, J., Mwesigwa, R., Nachega, J.B., Chan, K., Alexis Palmer, A., Ford, N & Mills, E.J. (2011). Survival of HIV infected adolescents on antiretroviral therapy in Uganda: Findings from a nationally representative cohort in Uganda. Uganda. *PLOS ONE*. Retrieved June 21, 2018, from <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0019261>
- Moshago, T., Haile, D & Enkusilasie, F. (2014). Survival Analysis of HIV Infected People on Antiretroviral Therapy at Mizan-Aman General Hospital, Southwest Ethiopia. *International Journal of Science and Research (IJSR)*, 3(5). Retrieved March 25, 2018, from <http://www.ijsr.net/archive/v3i5/MDIwMTMyMTAz.pdf>
- Rao, C.M., & Rao, S.R. (2004). Sample Size Calculator. Retrieved January 19, 2017, from Raosoft, Inc.: www.raosoft.com/samplesize.html
- Leulseged TW, Ayele BT (2019) Time to optimal glycaemic control and prognostic factors among type 2 diabetes mellitus patients in public teaching hospitals in Addis Ababa, Ethiopia. *PLoS ONE* 14(7): e0220309. <https://doi.org/10.1371/journal.pone.0220309>
- Tadege, M. (2018). Time to death predictors of HIV/AIDS infected patients on antiretroviral therapy in Ethiopia. Amhara, Ethiopia. *BMC Research Notes*. Retrieved April 12, 2019, from <https://doi.org/10.1186/s13104-018-3863-y>
- Damtew, B., Mengistie, B., & Alemayehu, T. (2015). Survival and determinants of mortality in adult HIV/Aids patients initiating antiretroviral therapy in Somali Region, Eastern Ethiopia. *Pan African Medical Journal*, 22: 138. Retrieved March 12, 2018, from doi: 10.11604/pamj.2015.22.138.4352
- Eneyew, K. , Seifu, D. , Amogne, W., & Menon, M. (2016). Assessment of Renal Function among HIV-Infected Patients on Combination Antiretroviral Therapy at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia. *Technology and Investment*, 7, 107-122. Retrieved November 5, 2018, from doi: 10.4236/ti.2016.73013
- Box-Steffensmeier, J.M., & Jones, B.S. (2004). *Event history modelling: a guide for social scientists*. Cambridge: University Press
- Kleinbaum, D.G., & Klein, M. (2012). Survival Analysis. A Self-Learning Text. *Springer*. Retrieved May 16, 2020, from DOI: 10.1007/978-1-4419-6646-9, pdf
- Etikan, I., Abubakar, S & Alkassim, R. (2017). The Kaplan Meier Estimate in Survival Analysis. *Biom Biostat Int J* 5(2): 00128. Retrieved January 21, 2018, from <http://medcraveonline.com/BBIJ/BBIJ-05-00128.php>
- Cleves, M., Gould, W.W., & Marchenko, Y.V., (2010). *An introduction to Survival Analysis Using Stata*. University of Arkansas: Stata press

- Clark, T.G., Bradburn, M.J., Love, S.B., & Altman, D.G. (2003). Survival Analysis Part I: Basic concepts and first analyses. *British Journal of Cancer*. Retrieved May 23, 2017, from www.slaop.org/pdf/814Journ7.pdf
- Sullivan, L. (2016). Survival Analysis. *Boston University School of Public Health*. Retrieved March 15, 2018, from https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/...Survival/BS704_Survival_print.html
- Klein, J.P & Moeschberger, M.L. (2003). Survival analysis Techniques for Censored and Truncated Data. Ohio. USA. *Springer*.
- Etikan, I., & Babatope, G. (2018). Survival Analysis: A Major Decision Technique in Healthcare Practices. *International Journal of Science and Research Methodology, Vol. 8 (4): 121-135*. Retrieved July 25, 2018, from <http://ijsrm.humanjournals.com/wp-content/uploads/2018/03/12.%C4%B0lker-Etikan-Ogunjesa-Babatope.pdf>
- Hosmer Jr, D.W., Lemeshow, S., & May, S. (2008). *Applied Survival Analysis: Regression Modeling of Time to Event Data: (Second Edition ed.)*. New York, NY. John Wiley and Sons Inc
- Storvik, G. (2011). Numerical optimization of likelihoods:Additional literature for STK2120. Oslo: *University of Oslo*. Retrieved May 16, 2018, from https://www.mn.uio.no/math/tjenester/kunnskap/.../num_opti_likelihoods.pdf
- Zhou, H. (2016). EM and MM Algorithms. *Los Angeles, USA: Department of Biostatistics University of California*. Retrieved June 10, 2018, from https://www.samsi.info/wp.../2016/08/Zhou_samsi-opt-summerschool-20160810-1.pdf
- Ansin, E. (2015). An evaluation of the Cox-Snell residuals. *Master's thesis*. Retrieved March 21, 2018, from <https://www.diva-portal.org/smash/get/diva2:826234/FULLTEXT01.pdf>
- Chaudhuri, P. K. (1997). On average derivative quantile regression. *The Annals of Statistics, 25(2): 715-744*. Retrieved March 12, 2019, from <https://www.jstor.org/stable/2242564>
- Peng, L., & Huang, Y. (2008). Survival Analysis With Quantile Regression Models. *Journal of the American Statistical Association*. Retrieved February 28, 2019, from doi:10.1198/016214508000000355
- Flom, P.L., & Peter Flom Consulting. (2011). Quantile regression with PROC QUANTREG. New York, USA: *NESUG 2011 Statistics & Analysis*. Retrieved March 15, 2019, from <https://www.lexjansen.com/nesug/nesug11/sa/sa04.pdf>
- University of California, Los Angeles, Statistical Consulting Group. (2018). *Introduction to SAS*. Retrieved March 4, 2018, from <https://stats.idre.ucla.edu/sas/modules/sas-learning-moduleintroduction-to-the-features-of-sas/>
- Fay, M.P., Brittain, E.H., & Proschan, M.A. (2013). Pointwise confidence intervals for a survival distribution with small samples or heavy censoring. *Biostatistics (2013), 14 (4)723-736*. Retrieved June 19, 2020, from doi:10.1093/biostatistics/kxt016

- Rodriguez, R., & Yao, Y. (2017). Five Things you should know about quantile regression. *SAS Institute Inc, Paper SAS525-2017*. Retrieved January 15, 2018, from <https://support.sas.com/resources/papers/proceedings17/SAS0525-2017.pdf>
- Mlangeni, N & Senkubuge, F. (2016). Antiretroviral therapy programme outcomes in Tshwane district, South Africa: A 5-year retrospective study. *South African Medical Journal, 106(4):365-368*. Retrieved June 24, 2018, from DOI:10.7196/SAMJ.2016.v106i4.9375
- Sieleunou, I., Souleymanou, M., Schönerberger, A.M., Menten, J., & Boelaert, M. (2008). Determinants of Survival in AIDS patients on antiretroviral therapy in a rural centre in the Far-North Province, Cameroon. *Tropical Medicine and International Health, 14(1): 36–43*. Retrieved April 17, 2018, from <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-3156.2008.02183.x/full>
- Teka, Z., Gizaw, Z., & Workneh, G. (2015). Investigating Correlates of the Survival of HIV/AIDS Patients Treated Under ART Follow-up: The Case of University of Gondar Hospital, Northwest Ethiopia. *Princeton University*. Retrieved May 28, 2019, from uaps2015.princeton.edu/abstracts/150168
- Shebeshi, D.S. (2011). Survival Analysis of adult HIV/AIDS patients and stochastic modelling of AIDS disease progression: A case study of Jimma University specialised hospital, Ethiopia. (*H. University, Ed.*) *Hawassa, Ethiopia*. Retrieved May 12, 2018, from DOI: 10.13140/2.1.3756.8323
- Merab, E. (2018). Why men are more likely to die of HIV. *Nation Media Group. Nairobi: Nairobi News*. Retrieved April 30, 2019 from <https://nairobi.news.nation.co.ke/news/why-men-are-more-likely-to-die-of-hiv>
- Ochieng-Ooko, V., Ochieng D., Sidle, J.E., Holdsworth, M., Wools-Kaloustian, K., Siika, A.M., Yiannoutsos, C.T., Owiti, M., Kimaiyo, S., & Braitstein P. (2010). Influence of gender on loss to follow-up in a large HIV treatment programme in western Kenya. *Bull World Health Organ, 88(9): 681–688*. Retrieved April 7, 2019, from doi: 10.2471/BLT.09.064329
- Reid, E., Orrell, C., Stoloff, K., & Joska, J. (2012). Psychotropic prescribing in HIV. *Southern African Journal of HIV Medicine, 13(4)*. Retrieved September 28, 2018, from September 28, DOI: 10.4102/sajhivmed.v13i4.115

Figures

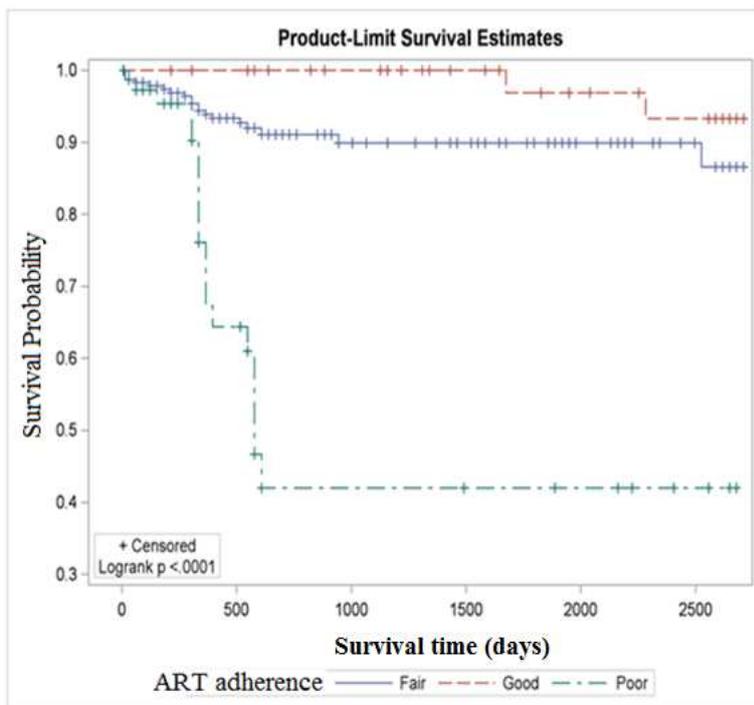


Figure 1: Kaplan-Meier survival function estimates for ART adherence strata

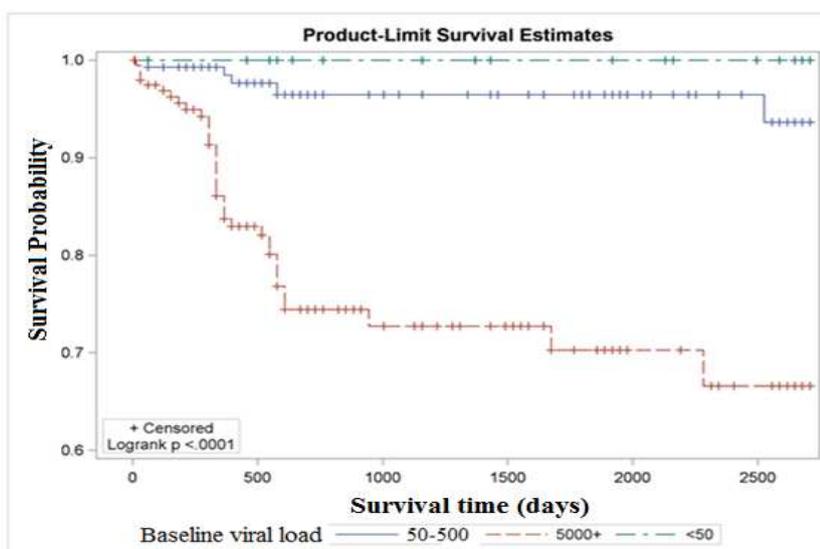


Figure 2: Kaplan-Meier survival function estimates for Baseline viral load strata

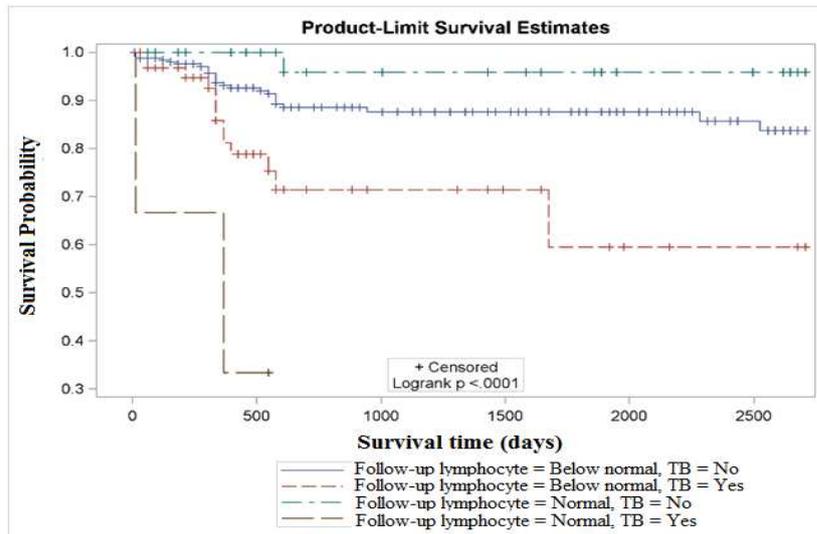


Figure 3: Kaplan-Meier survival function estimates for Follow-up lymphocyte by TB history strata

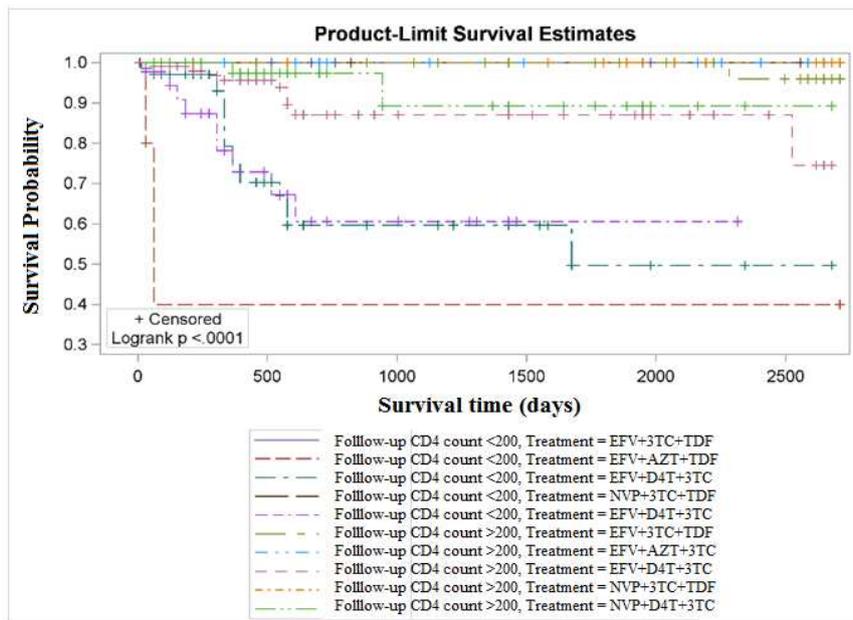


Figure 4: Kaplan-Meier survival function estimates for Follow-up CD4 by Treatment (Regimen 1) strata

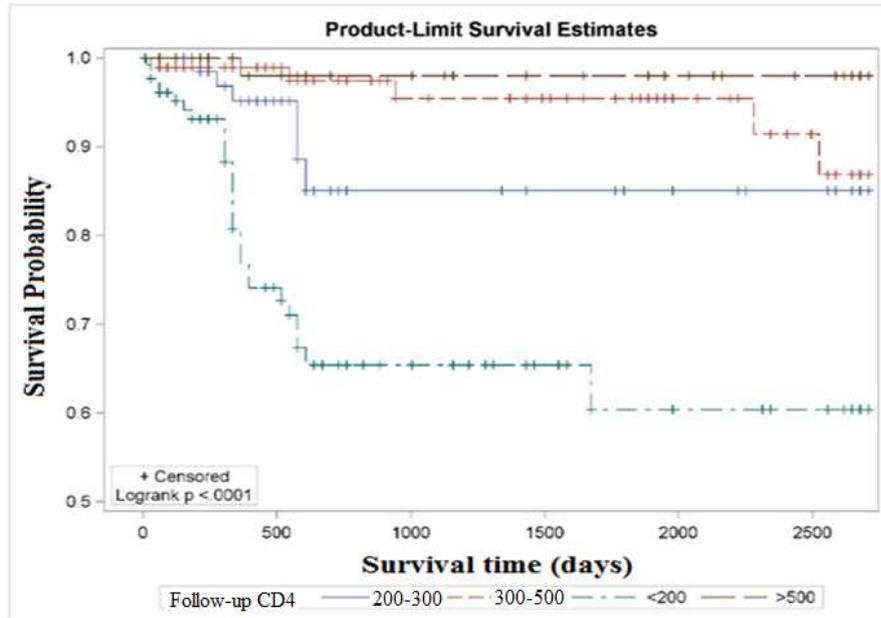


Figure 5: Kaplan-Meier survival function estimates for Follow-up CD4 Count

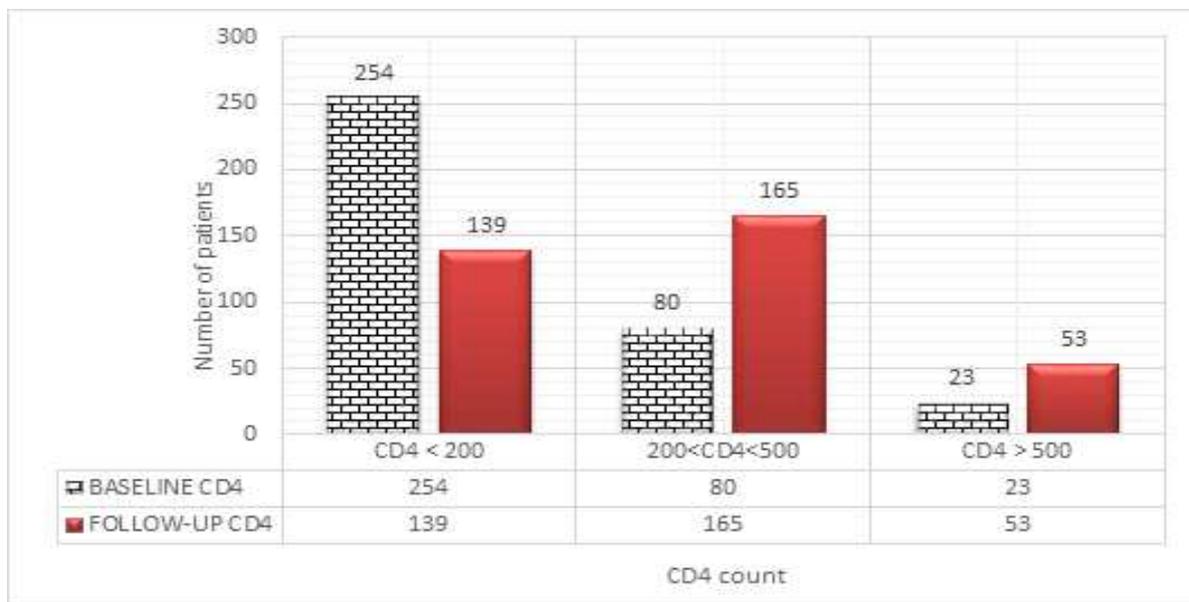


Figure 6: Bar graphs for the distribution of Baseline and Follow-up CD4 in HIV/AIDS patients

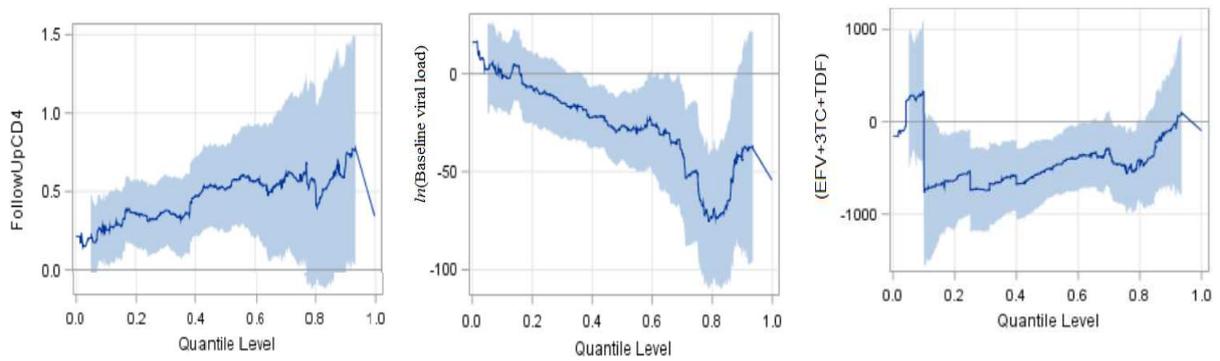


Figure 7: Estimated parameter versus quantile for log (survival time) with 95% confidence limits - Follow-up CD4, \ln (Baseline viral load) and Treatment (regimen)

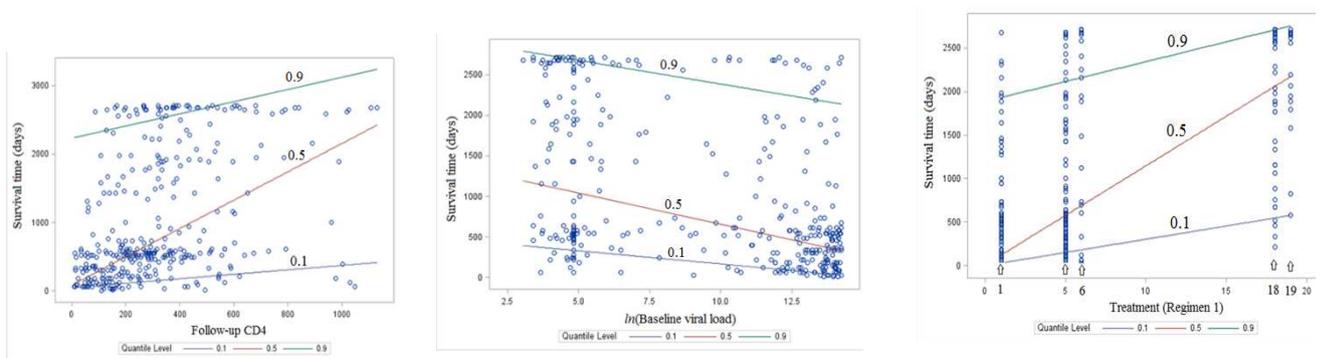


Figure 8: Survival analysis of covariate effect on quantiles of Survival time

Key: Treatment (Regimen 1): 1 = (NVP+D4T+3TC), 5 = (EFV+D4T+3TC), 6 = (EFV+AZT+3TC), 18 = (EFV+3TC+TDF) and 19 = (NVP+3TC+TDF).

Figures

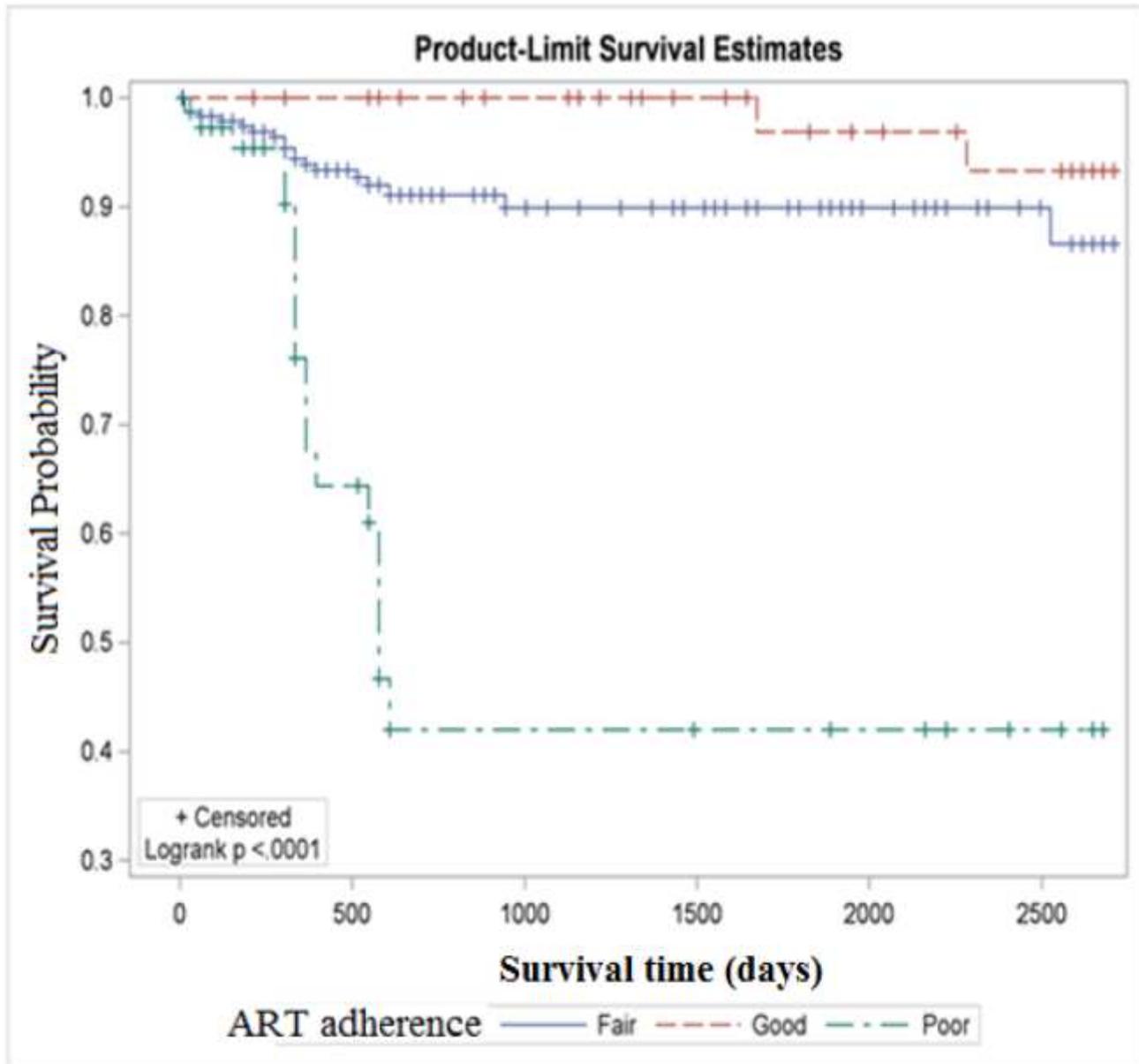


Figure 1

Kaplan Meier survival function estimates for ART adherence strata

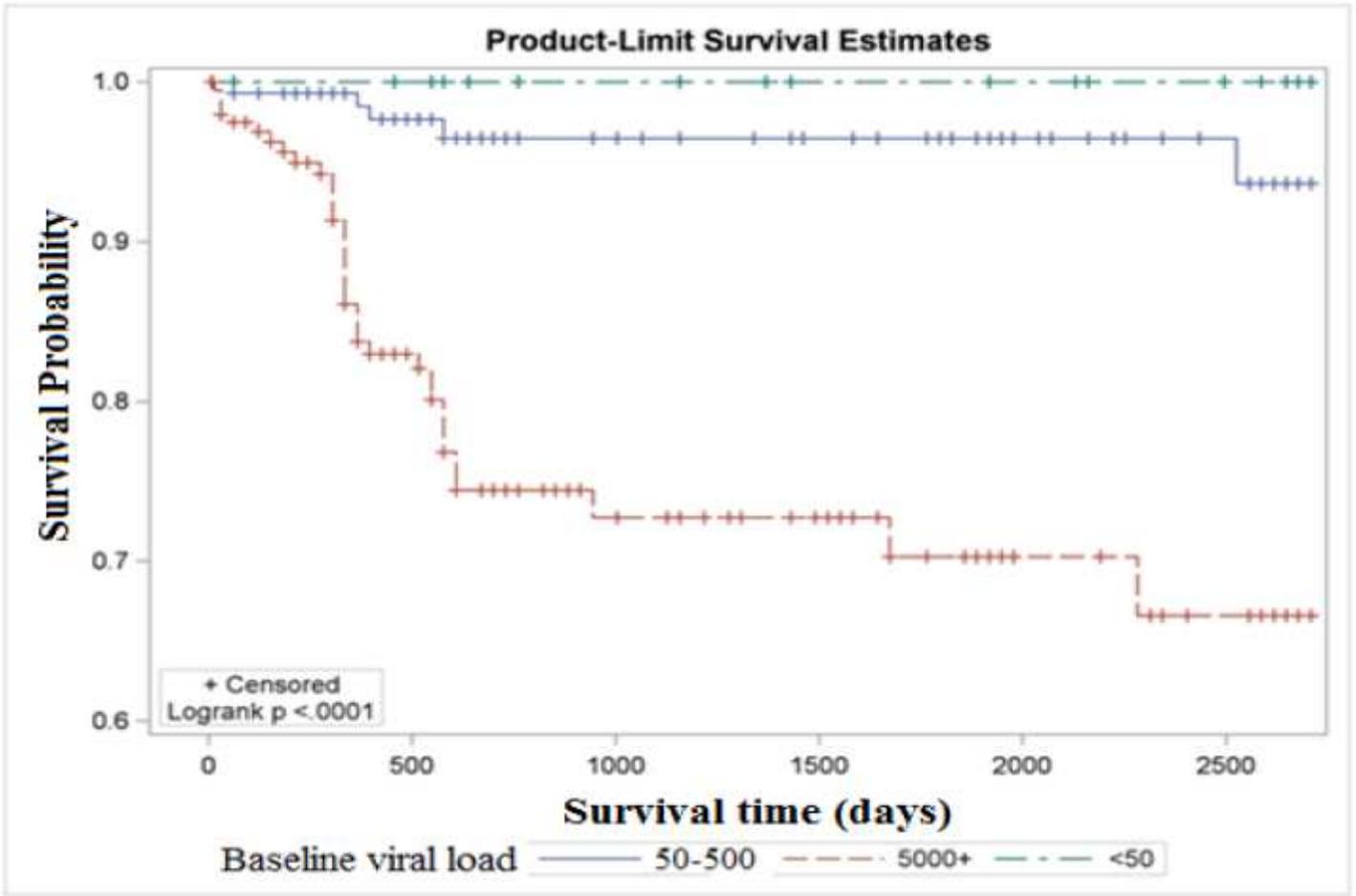


Figure 2

Kaplan Meier survival function estimates for Baseline viral load strata

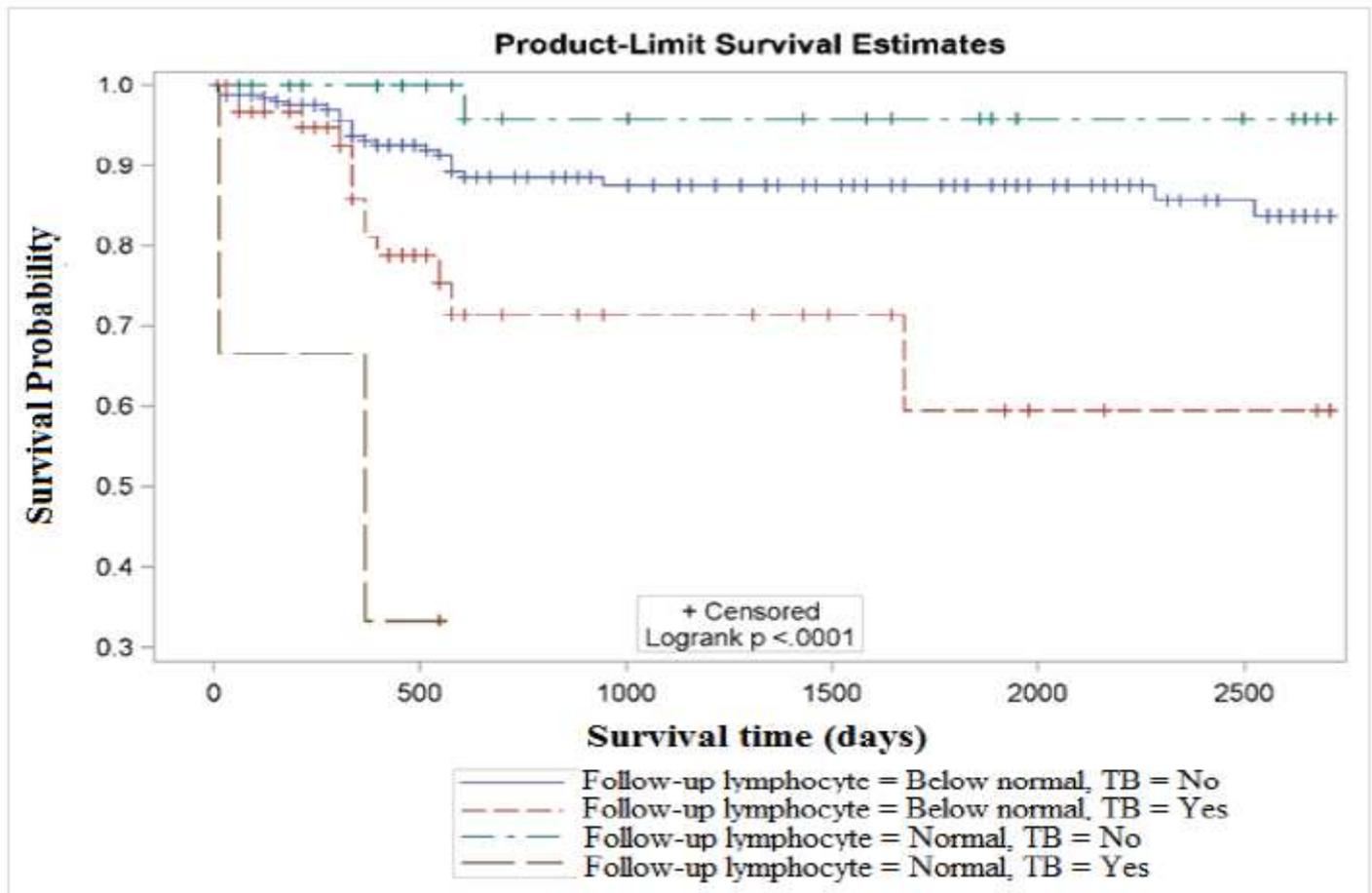


Figure 3

Kaplan Meier survival function estimates for Follow up lymphocyte by TB history strata

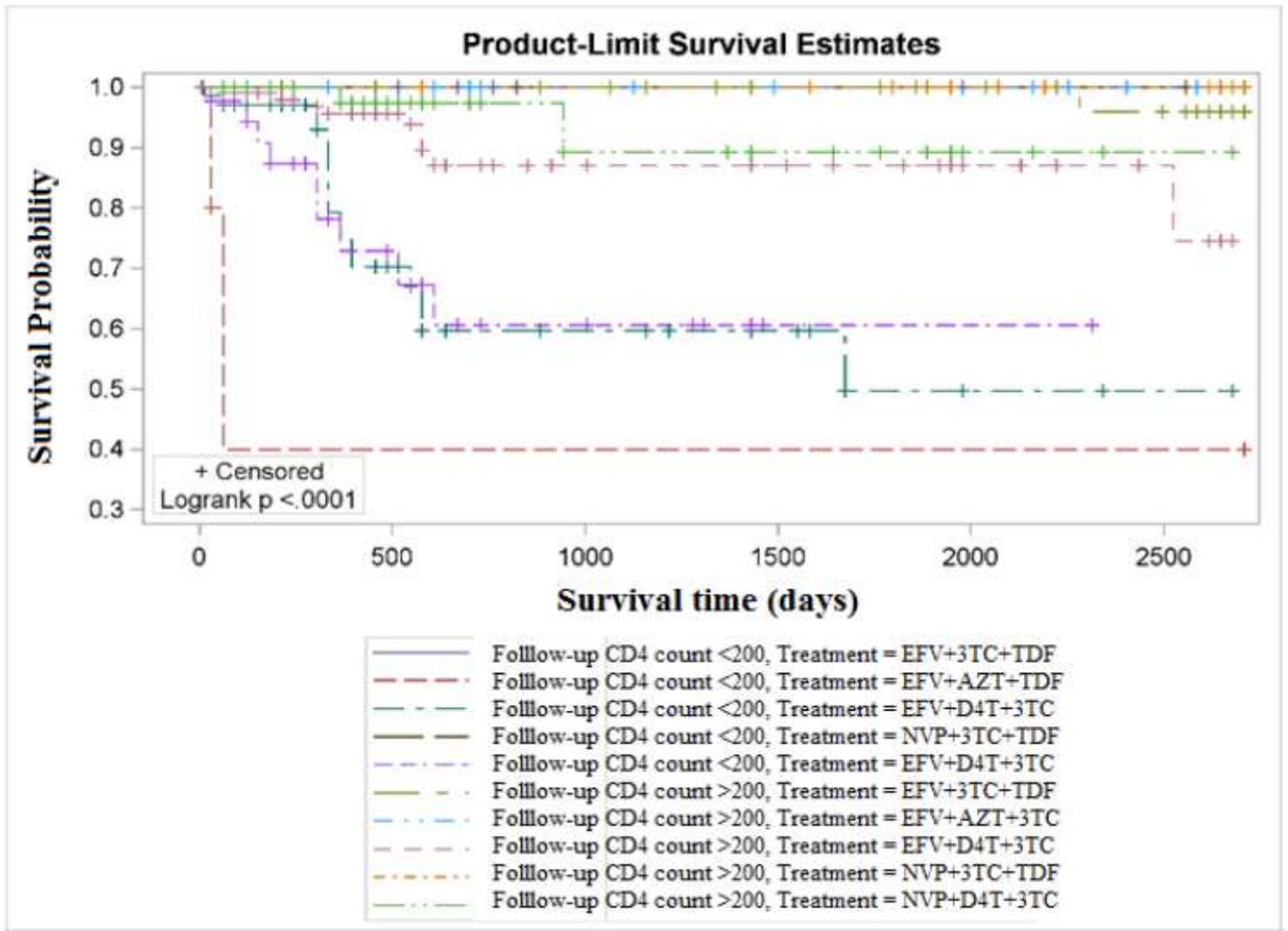


Figure 4

Kaplan Meier survival function estimates for Follow up CD4 by Treatment (Regimen 1) strata

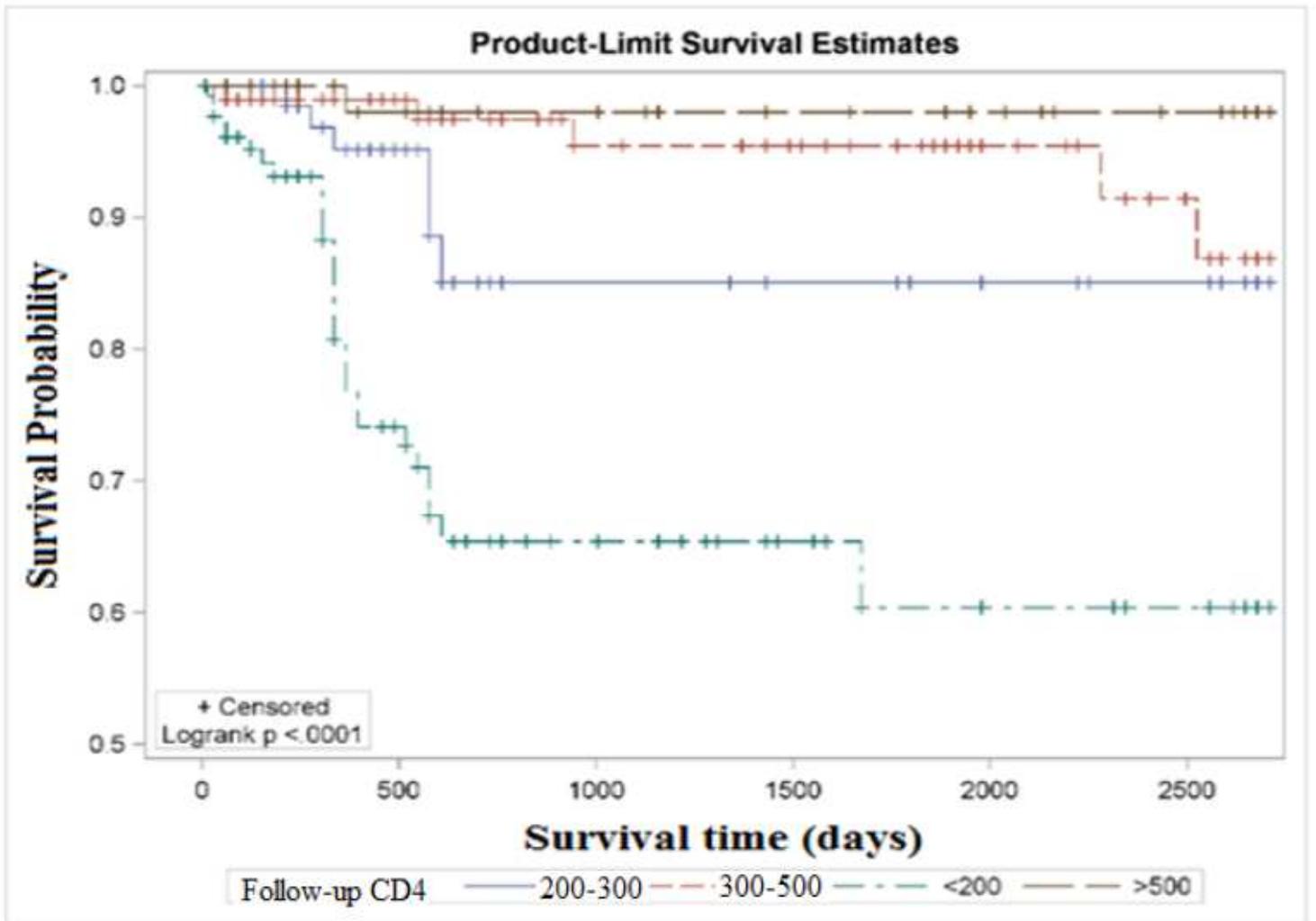


Figure 5

Kaplan Meier survival function estimates for Follow up CD4 Count

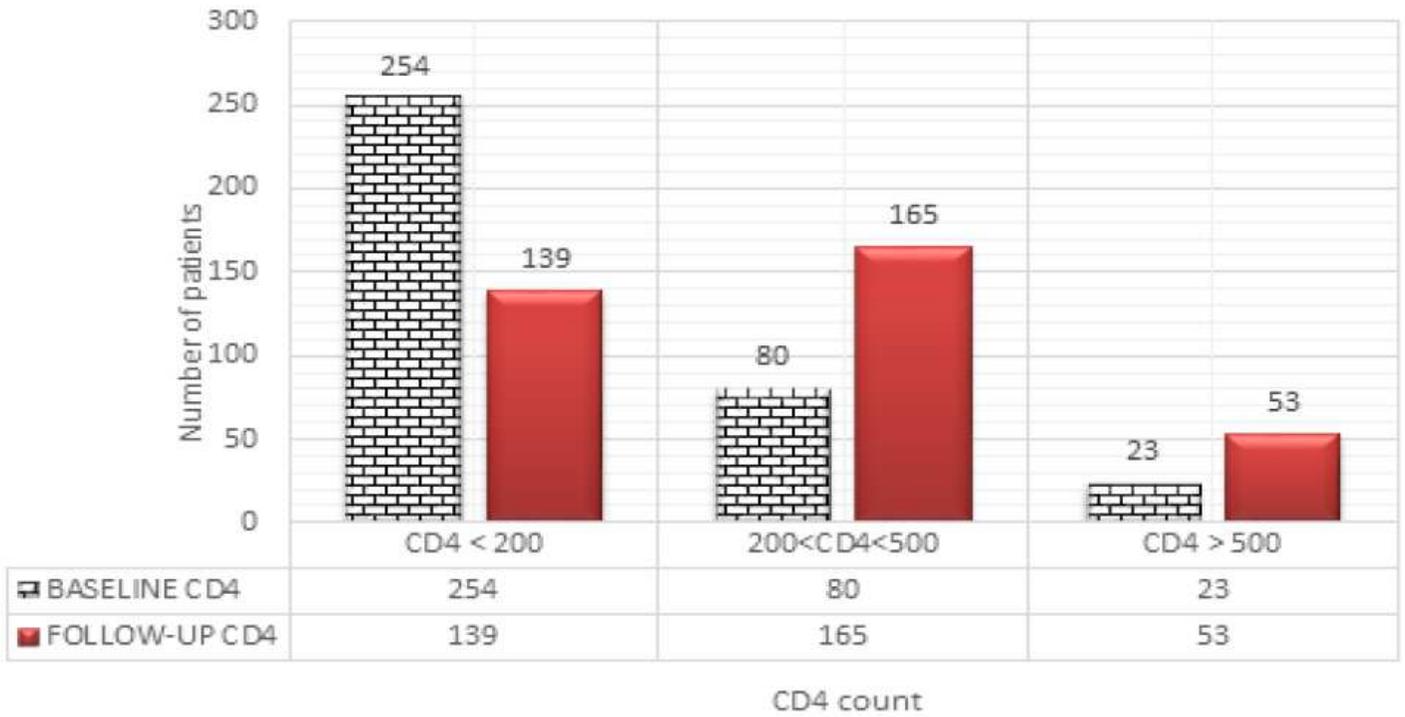


Figure 6

Bar graphs for the distribution of Baseline and Follow-up CD4 in HIV/AIDS patients

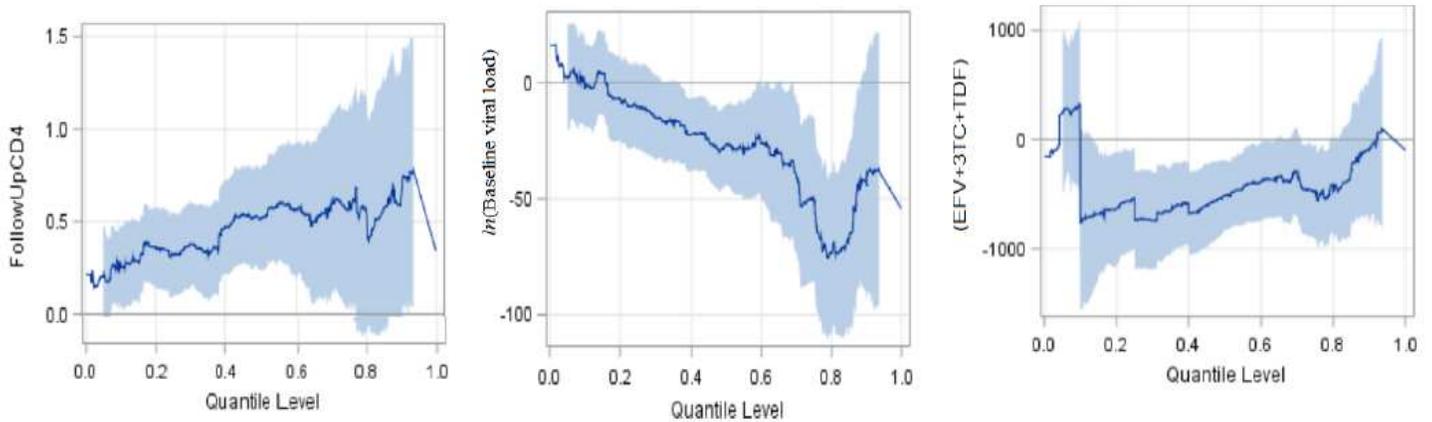


Figure 7

Estimated parameter versus quantile for log (survival time) with 95% confidence limits - Follow-up CD4, ln (Baseline viral load) and Treatment (regimen)

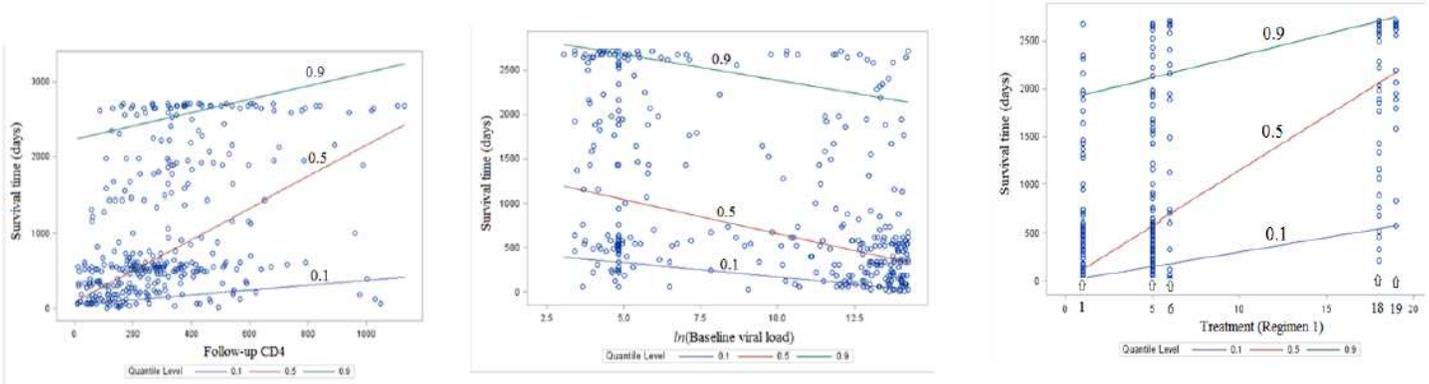


Figure 8

Survival analysis of covariate effect on quantiles of Survival time. Key: Treatment (Regimen 1): 1 = (NVP+D4T+3TC), 5 = (EFV+D4T+3TC), 6 =(EFV+AZT+3TC), 18 = (EFV+3TC+TDF) and 19 = (NVP+3TC+TDF).