Joint Analysis of CD4 Cell Counts and Time-to-Viral-Rebound of HIV-Infected Patients Initiating Antiretroviral Therapy at Jimma University Medical Center

Delelegn Lambamo, Abdisa Gurmessa, Kibrealem Sisay, Meskerem Getachew
Department of Statistics, College of Natural Sciences, Jimma University, Ethiopia

Abstract

Background: HIV infection leads to severe depletion of CD4 cells. The risk of viral rebound from undetectable levels increases when the immune system deteriorates due to severe depletion of CD4 cells. The main objective of this study was to compare separate and joint models for longitudinal and time-to-event outcomes as well as identify and analyze factors affecting changes in CD4 cell counts over time and time-to-viral-rebound of HIV-infected patients under antiretroviral therapy (ART).

Methods: A retrospective cohort study was conducted among 309 HIV-infected patients aged 18 and above under ART follow-up from February 2016 to May 2021 at Jimma University Medical Center. The number of CD4 cell counts and the time to viral rebound were first separately analyzed using linear mixed-effects and Cox PH models, respectively. Then, joint models with different random effects and shared parameters have been explored, and the best-fit model was selected based on information criteria.

Results: The estimated association parameter from the joint model was negative and statistically significant, indicating that higher CD4 cell count is associated with a significant reduction in the hazard of viral rebound. The results of the longitudinal sub-model revealed that patients' age, adherence to ART, functional status and WHO clinical stage had significant association with the square root of CD4 cell count. Furthermore, age, adherence to treatment, WHO clinical stage and peripheral neuropathy were found to be significant predictors of the risk of viral rebound from the survival sub-model.

Conclusion: Poor adherence to treatment and severe stage of HIV were significantly associated with lower CD4 cell count as well as higher risk of viral rebound. Moreover, patients with ambulatory and bedridden functional status were prone to significantly lower CD4 cell count, and those affected by peripheral neuropathy were more likely to be exposed to the risk of viral rebound. The significance of the association parameter, coupled with the relatively smaller standard errors of estimates, indicated that the joint model is more appropriate. Thus, joint modeling is a useful and efficient approach for evaluating the association between longitudinal and time-to-event outcomes.

Key words: longitudinal analysis, survival analysis, joint modeling, CD4 count, viral rebound

1. Introduction

The Human Immunodeficiency Virus (HIV) causes Acquired Immunodeficiency Syndrome (AIDS), a condition in which the immune system begins to decline, exposing infected individuals to life-threatening opportunistic infections. The first AIDS case emerged in the early 1980s and since then, the AIDS prevalence has been increasing. According to UNAIDS (2021), an estimated 37.7 million people were living with HIV globally in 2020. Among the total infected patients, 36.0 million were adults and 74% had access to antiretroviral treatment (ART). Among these 37.7 million infected patients, 25.6 million people were living in Africa and 16,384,000 (64%) had access to ART. There were 1.3 million newly infected people aged 15 years or older globally of which an estimated 765,000 (66%) were from Africa. Moreover, 680,000 people died due to AIDS and related illnesses globally in 2020, of which 467,900 (69%) were from Africa (UNAIDS, 2021).

In Ethiopia, an estimated 669,236 people were living with HIV, with 625,007 of them adults aged 15 years or older. There were 14,842 new infections of which 11,613 (78.2%) were adults aged +15 years. Furthermore, the number of deaths due to AIDS-related illnesses for the same period was estimated to be 11,546 in the country and 9491 (82%) were adults older than 15 years (EPHI, 2020).

The Global AIDS Update of UNAIDS (2021) highlighted that antiretroviral medicines were available to over 27.5 million people living with HIV in 2020. At least eight countries in a variety of geographic, epidemiological and socioeconomic settings have achieved the 90–90–90 testing and treatment targets of the global sustainable development goals (SDG). Globally in 2020, 84% (31.6 million) of people living with HIV knew their HIV status, 73% (27.4 million) were accessing ART and 66% (24.8 million) have achieved viral suppression. Despite the fact that this is a remarkable achievement, all of the global HIV targets for 2020 were missed.

The development of drug-resistant variants in HIV patients under ART makes it infeasible to completely eradicate the virus (Shoko and Chikobvu, 2018). But with proper adherence to treatment, ART has the potential to suppress viral replication, often below the level of detection by commercially available tests. ART can very effectively control the infection and hold the amount of circulating virus below the level detectable by a clinical assay, improving both the quality and length of life. Consequently, the standard of care for people living with HIV (PLWH) is to maintain life-long ART. However, there is heterogeneity in viral rebound times (Conway et al., 2019; Oguntibeju, 2012).

In recent years, the interest on longitudinal analysis has grown rapidly through the development of new methods and increased computational power to aid and further develop this field of research. When these processes are correlated, the use of independent models can cause biased estimates (Liu and Liu, 2015; Hsieh et al., 2006). In contrast, joint models result in a reduction in the standard errors, more accurate parameter estimates and valid inferences on the longitudinal and survival processes.

Joint modeling of longitudinal and time-to-event data is an area of increasing research (Tsiatis and Davidian, 2004). Joint modeling enables the simultaneous study of a longitudinal marker and a correlated time-to-event response. Among these, the shared random-effect models that are defined as a mixed model for the longitudinal marker and a survival model for the time-to-event including characteristics of the mixed-effects model as covariates received the main interest. Indeed, they extend naturally the survival model with a time-dependent covariate and offer a flexible framework to explore the link between a longitudinal biomarker and the risk of an event (Wang et al., 2014).

Following this line of thought, the primary aim of this study is to illustrate the virtues of joint modeling by comparing our results to those obtained from separate Cox PH and linear mixed-effects models. Another objective of this study is to identify and analyze factors affecting changes in the longitudinal CD4 cell count over time as well as time-to-viral-rebound of HIV-infected patients under ART.

2. Materials and Methods

2.1. Study area and design

This study was conducted at Jimma University Medical Center, Oromia Region, Ethiopia, and included all 18 years and above HIV/AIDS patients who initiated ART from February 1, 2016 until May 30, 2021. Data were obtained through a retrospective cohort study design where basically joint longitudinal and survival modeling were considered to determine potential predictors.

2.2. Source of data and quality assurance

The relevant data were extracted from charts of HIV/AIDS patients under antiretroviral therapy (ART) follow-up which contain epidemiological, laboratory and clinical information including detailed ART history. Patients who were 18 years old or older and who have attended a minimum of three visits for ART treatment in the stated period at the medical center were included in the study. The quality of data was controlled by data controllers from the ART section of the Jimma University Medical Center. The necessary amendments were made on the final data collection sheet and the filled formats were checked

daily by the supervisor and authors. The data extraction mechanism and variables included in this investigation ensured the reliability and completeness of the data.

2.3. Variables of the study

The response and predictor variables considered in this study are defined as follows. The survival outcome variable was the survival time-to-viral-rebound of the infected patients. Time-to-viral-rebound of the patients was the time from ART start date to viral rebound during the time period (in months). A patient who was lost to follow-up, transferred to another hospital before experiencing the event or did not experience rebound at 30 May 2021 was considered as censored (right censoring). The longitudinal continuous outcome which is a biomarker variable was the number of CD4 counts per cubic milliliter of blood which was measured in six months interval. The covariates considered in this study were those which are expected to potentially affect the CD4 cell progression and aggravate the viral rebound of HIV/AIDS patients. Specifically, the independent covariates considered for the longitudinal and survival modeling as well for the joint modeling were age, gender (female, male), marital status (single, married, divorced, widowed), place of residence (rural, urban), education level (not educated, primary, secondary, tertiary), treatment change (No, Yes), WHO clinical stage (I, II, III, IV), regimen type (ART regimens combinations), functional status (working, ambulatory, bed ridden), adherence (poor, fair, good) and peripheral neuropathy (Yes, No).

2.4. Model specification

To extract information and draw inference, the data were analyzed using different methods. Specifically, the study used the following three models:

- A linear mixed-effects model was used for the (continuous) longitudinal CD4 cell count response variable.
- Survival model (Cox proportional hazards model) was used for the (continuous) time-to-viral rebound response variable.
- Joint model of longitudinal CD4 cell count with survival time to viral rebound.

2.4.1 Linear mixed-effects model (LMM)

LMM is used to model longitudinal outcomes by accounting for within- and between-subject sources of variations. LMMs are statistical models for longitudinal or repeated-measures studies in which subjects are measured repeatedly over time and measurements in which the residuals are normally distributed but may not be independent. Moreover, LMMs are robust to missing data and irregularly spaced measurement occasions (Hedeker and Gibbons, 2006).

The linear mixed-effects model, proposed by Laird and Ware (1982) based on the work of Harville (1977), included a unified approach using growth models and repeated-measures models for the sequence of longitudinal measurements $y_{i1}, y_{i2}, \ldots, y_{in_i}$ for the i[th] subject at times $t_{i1}, t_{i2}, \ldots, t_{in_i}$:

$$\begin{aligned} y_i &= X_i'(t)\beta + Z_i'(t)b_i + \varepsilon_i \\ y_i &= \mu_i(t) + U_{1i}(t) + \varepsilon_i \end{aligned} \qquad \text{(1)}$$

where $y_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})'$ is the $(n_i \times 1)$ vector of observed response values, $X_i(t)$ is the $(p \times n_i)$ design matrix corresponding to the fixed-effects, β is the $(p \times 1)$ vector of fixed-effects parameters, $Z_i(t)$ is the $(q \times n_i)$ observed design matrix corresponding to the random-effects, b_i is the $(q \times 1)$ vector of random-effects with $b_i \sim N(0, D)$, and ε_i is the $(n_i \times 1)$ vector of within-group errors such that $\varepsilon_i \sim N(0, \sigma_\varepsilon^2 I)$.

In this study, $\mu_i(t) = X_i'(t)\beta$ represents the mean square root of CD4 measurement and $U_{1i}(t) = Z_i'(t)b_i$ incorporates the part of the random effects, that is, the true individual level CD4 measurement trajectories after they have been adjusted for the overall mean. The random effects b_i are introduced for each subject to account for the correlation between the repeated measurements within a subject.

Here, we can make a distinction between the random intercept model and random intercept-random slope model. The random-effects model or subject-specific model assumes that extra correlation arises among longitudinal response (Diggle et al., 2002). The random intercepts model allows intercepts to vary across groups. An intuitive extension that allows a random shift in the subject-specific slopes is known as random intercept and slope model.

2.4.2 Survival data modeling: Cox PH model

To determine if the variation in subjects' survival experience is partially explained by covariates or to find any possible relationship between survival times and important covariates, a popular approach in survival analysis is to model the hazard function rather than the mean of the survival times as in the classical regression models. In other words, survival models are most often defined in terms of the hazard function. The most commonly used semi-parametric survival model which does not require the distributional assumption of the survival time is the Cox proportional hazard (PH) model proposed by Cox (1972). This model expresses the hazard of an event at time t as:

$$\lambda(t) = \lambda_0(t)\exp\{W'\Upsilon\} \quad \ldots\ldots (2)$$

where W is a vector of baseline covariates, Υ is a vector of fixed-effects parameters and the term $\lambda_0(t)$ is the baseline hazard where the effects of covariates are all set to zero. The elements of W (covariates in survival model) may or may not be the same as those used to model the longitudinal response variable.

Suppose we increase the value of a continuous covariate W_k by one unit from w_k to ($w_k + 1$), while holding other covariate values fixed. Then the ratio of the corresponding hazards is equal to:

$$HR = \frac{\lambda(t, w_k + 1)}{\lambda(t, w_k)} = \frac{\exp(\gamma_1 w_1 + \ldots + \gamma_k (w_k + 1) + \ldots)}{\exp(\gamma_1 w_1 + \ldots + \gamma_k (w_k) + \ldots)} = \exp(\gamma_k) \quad \ldots\ldots (3)$$

Thus, $\exp(\gamma_k)$ can be interpreted as the hazard ratio corresponding to a one unit increase in the covariate W_k, while other covariate values being held constant.

For a categorical covariate W_j with ℓ levels, the model contains $(\ell - 1)$ dummy variables defined as $Z_i = 1$ if $W_j = i$, and 0 otherwise, $i = 1, 2, \ldots, (\ell - 1)$. Let $\gamma_{j1}, \gamma_{j2}, \ldots, \gamma_{j,\ell-1}$ denote the coefficients of the appropriate dummy variables. Then the ratio of the hazard of two subjects, one with W_j at level i and the other with W_j at level k, provided the values of all other covariates for these subjects are the same, is given by:

$$HR = \frac{\lambda(t, Z_i)}{\lambda(t, Z_k)} = \frac{\exp(\gamma_{ji})}{\exp(\gamma_{jk})} = \exp\{\gamma_{ji} - \gamma_{jk}\} \quad , \ i \neq k, \ i,k = 1, 2, \ldots, (\ell - 1) \quad \ldots\ldots (4)$$

2.5. The joint modeling structure

Joint modeling has expanded very rapidly in Biostatistics and medical research since it enables the simultaneous study of a longitudinal marker and a correlated time-to-event response. Among joint models, the shared random-effects model for the longitudinal marker and a survival model for the time-to-event outcome including characteristics of the mixed model as covariates have received the main interest. Such models extend the survival model with time-dependent covariates and offer a flexible framework to explore the link between a longitudinal biomarker and the risk of an event (McCrink et al., 2011).

The main aim of this study was also to relate the longitudinally measured CD4 biomarker with time to viral rebound for HIV patients to understand the association between the two processes. Therefore, after having appropriate separate models, the longitudinal sub-model has the same specification as the LMM in Equation (1). The survival sub-model includes a shared parameter association function to the specified Cox PH model in Equation (2).

For a single failure-time per subject, the hazard function for subject i at time t is given by:

$$\lambda_i(t \mid W) = \lambda_0(t)\exp\{W_i(t)'\Upsilon + U_{2i}(t)\} \quad \text{……} \quad (5)$$

where $\lambda_0(t)$ denotes the baseline hazard function, $W_i(t)$ is a vector of (possible time-varying) external covariates, Υ is a vector of fixed effects parameters, and $U_{2i}(t)$ is a latent process that captures the association structure between the measurement (longitudinally measured CD4 count) and event (time-to-viral rebound) processes.

The functional forms often considered for the association structure include: $U_{2i}(t) = \alpha' m_i(t)$, $U_{2i}(t) = \alpha' b_i$ and $U_{2i}(t) = \alpha'(\beta_b + b_i)$. Here, $m_i(t)$ denotes the true value of the longitudinal (CD4 cell measurement) marker at time t, the vector α quantifies the strength of association between the two processes, b_i is a vector of subject-specific random effects of the longitudinal part, and β_b is a vector of fixed effect parameters corresponding to the random effects.

2.6. Joint model estimation methods

The maximum likelihood estimation to jointly model the survival time and its longitudinal variables has been successful to model both processes in longitudinal data. Random effects in the longitudinal process are oftentimes used to model the survival times through a proportional hazards model, and this invokes an expectation-maximization (EM) algorithm used for the maximum likelihood (ML) estimates (Hsieh et al., 2006).

The standard ML method involves maximizing the log-likelihood corresponding to the joint distribution of the time-to-event and longitudinal data processes. Strictly, both processes share the same unobserved random effects and are conditionally independent given these random effects (Rizopoulos, 2012). The log-likelihood for the joint model is approximated using the EM algorithm since both the integral with respect to the random effects and the survival function typically do not have an analytical solution except in some special cases.

2.7. Model selection techniques

To select the model which best fits the given data, it is important to compare different models by using various techniques and methods. To come up with appropriate separate longitudinal and survival models, Akaike information criterion (AIC) and Bayesian information criterion (BIC) may be used. The model with smaller values of AIC and BIC is considered the preferred model.

3. Results and Discussion

3.1. Descriptive results

Descriptive summary statistics of baseline characteristics of HIV-infected patients is presented in Table 1. Among 309 HIV-infected patients considered in this study, 235 (76.1%) of them experienced viral rebound, while the remaining 74 (23.9%) were right censored. Gender-wise, about three-fifth of the patients were females and the proportion of those who experienced the event (viral rebound) was lower for males (70.2%) than females (80.0%). Moreover, the viral rebound proportion was highest for those patients who had no education (85.0%), while the lowest was for patients who had tertiary education (68.1%) and secondary education (71.0%).

About half of all HIV-infected patients included in the study were married. The percentage of those who experienced the event was almost the same among married, divorced and widowed patients. This figure was relatively lower for those who were single. The majority of the sample patients (218 or 70.6%) were from urban areas, and there was no marked urban-rural differential in the proportion of patients who experienced viral rebound.

Among the sample patients, there were 221(71.5%) who were able to work, 54 (17.5%) were ambulatory and 34 (11.0%) were bedridden. Almost all patients with ambulatory functional status (98.1%) had experienced the event. Regarding WHO clinical stages, 27.2%, 15.2%, 43.4% and 14.2% of patients under study were at stages I to IV, respectively. The viral rebound proportion was the highest for those at Stage-III (94.0%), followed by those at Stage-IV (88.6%). The reverse was true for patients at Stage-I (40.5%). Of the total of 309 patients, 195 (63.1%) had poor adherence to treatment and the highest percentage of viral rebounds occurred in this group (96.9%). Only one-fourth of those with good adherence experienced the event. Close to 100% of patients with regimen types TDF+3TC+EFV and AZT+3TC+DTG+DRV/r had experienced the event. Moreover, viral rebounds have occurred for over 90% of HIV-infected patients who developed peripheral neuropathy.

Table 1: Descriptive summary of baseline characteristics of sample HIV-infected patients under ART at Jimma University Medical Center

Covariate	Category	Total (%)	Censored		Event	
			n	%	n	%
Gender	Female	185 (59.9)	37	20.0	148	80.0
	Male	124 (40.1)	37	29.8	87	70.2
Marital status	Single	43 (13.9)	15	34.9	28	65.1
	Married	154 (49.8)	34	22.1	120	77.9
	Divorced	80 (25.9)	18	22.5	62	77.5
	Widowed	32 (10.5)	7	21.9	25	78.1
Functional status	Working	221 (71.5)	66	29.9	155	70.1
	Ambulatory	54 (17.5)	1	1.9	53	98.1
	Bed ridden	34 (11.0)	7	20.6	27	79.4
Regimen type	AZT+3TC+ATV/r	33 (10.7)	32	97.0	1	3.0
	AZT+3TC+LPV/r	35 (11.3)	34	97.1	1	2.9
	TDF+3TC+DTG	5 (1.6)	3	60.0	2	40.0
	TDF+3TC+EFV	76 (24.6)	3	3.9	73	96.1
	AZT+3TC+DTG+DRV/r	160 (51.8)	2	1.3	158	98.8
WHO Clinical stage	Stage-I	84 (27.2)	50	59.5	34	40.5
	Stage-II	47 (15.2)	11	23.4	36	76.6
	Stage-III	134 (43.4)	8	6.0	126	94.0
	Stage-IV	44 (14.2)	5	11.4	39	88.6
Place of residence	Rural	91 (29.4)	20	22.0	71	78.0
	Urban	218 (70.6)	54	24.8	164	75.2
Adherence	Poor	195 (63.1)	6	3.1	189	96.9
	Fair	28 (9.1)	8	28.6	20	71.4
	Good	80 (25.9)	60	75.0	20	25.0
Peripheral neuropathy	No	135 (43.7)	60	44.4	75	55.6
	Yes	174 (56.3)	14	8.0	160	92.0
Education level	Not educated	60 (19.4)	9	15.0	51	85.0
	Primary	109 (35.3)	23	21.1	86	78.9
	Secondary	93 (30.1)	27	29.0	66	71.0
	Tertiary	47 (15.2)	15	31.9	32	68.1
Treatment change	No	229 (67.0)	2	0.9	227	99.1
	Yes	80 (33.0)	72	90.0	8	10.0

3.2. Separate analysis of the longitudinal response

3.2.1 Exploring individual profiles and mean structure

In order to gain some insights on the data, the individual profile plots and mean structure plots were obtained (Verbeke and Molenberghs, 2000). Figure 1 depicts the variability (within and between patients) in the square root of CD4 measurements of HIV patients. Since the data were not balanced, loess

smoothing technique was used. The mean structure plot suggests a linear relationship between the mean of the square root of CD4 counts and time-to-viral rebound. Moreover, the individual profile plots exhibit a linearly increasing pattern, which rationalizes the use of linear mixed-effects model to analyze the trajectory of CD4.

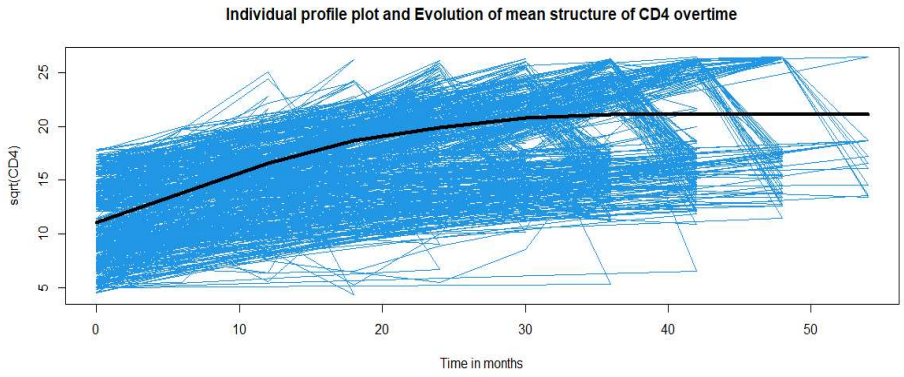

Figure 1: Individual profile plot with a loess smoothing technique

3.2.2 Comparison of linear mixed-effects models

In this study, four LMMs were considered. The values of the information criteria as well as the log-likelihood for each of these models are presented in Table 2. Since the random intercept - random slope model has minimum AIC and BIC values, it is the best fitting and preferred model. Thus, it is used for subsequent analysis.

Table 2: Information criteria for various random-effects models

Random effects model	AIC	BIC	Log-Likelihood
Random intercept	9272.29	9379.93	-4617.15
Random slope	9369.00	9506.63	-4670.50
Random intercept-Random slope	9237.49	9356.45	-4597.75
Random intercept and quadratic slope	9396.75	9633.13	-4682.37

3.2.3 Results of the random intercept-random slope model

Table 3 pertains to the results of the fitted random intercept-(linear time) random slope model. On average, a one year increase in the age of HIV-infected patients decreases their square root of CD4 cells by 0.42 (per cubic millimeter of blood), holding other covariates constant. Level of adherence to ART was also significant. The square root of CD4 cell count was 0.96 and 1.06 higher for patients with fair and good adherence to treatment as compared to those with poor adherence, respectively.

Regarding functional status, the square root of CD4 cell count was 3.09 and 2.19 lower for patients with bedridden and ambulatory functional status as compared to those with working functional status, respectively, controlling for the other covariates. The results also revealed that CD4 count goes down as HIV disease progresses. Patients at WHO clinical stage II tend to have 2.48 lower square root of CD4 cell count as compared to those at stage I. Moreover, ceteris paribus, the square root of CD4 count per cubic millimeter of blood was 5.47 and 8.55 lower for patients at WHO clinical stages III and IV as compared to those at stage I, respectively. The random effects portion of Table 3 also tells us that there is subject-specific variation in the repeated square root of CD4 cell measurements, and hence, subject-specific square root of CD4 count variances must be considered in the analysis.

Table 3: Fitted linear mixed-effects model with random intercept-linear time random slope

Fixed effect				**95% CI**	
Variable	**Coef.**	**Std. Err.**	**p-value**	**Lower**	**Upper**
(Intercept)	21.40	0.39	<0.001*	20.63	22.17
Age	-0.42	0.08	<0.001*	-0.58	-0.26
Adherence (Ref.=Poor)					
Fair	0.94	0.28	0.001*	0.39	1.49
Good	1.06	0.21	<0.001*	0.65	1.46
Functional status (Ref.= Working)					
Bedridden	-3.09	0.32	<0.001*	-3.71	-2.46
Ambulatory	-2.19	0.25	<0.001*	-2.67	-1.71
WHO clinical stage (Ref.=Stage-I)					
Stage-II	-2.48	0.29	<0.001*	-3.05	-1.91
Stage-III	-5.47	0.30	<0.001*	-6.06	-4.88
Stage-IV	-8.55	0.37	<0.001*	-9.28	-7.83
Time	-0.01	0.01	0.313	-0.04	0.01
Time: Adherence (Ref.=Poor)					
Time: Fair	0.03	0.01	0.009*	0.01	0.05
Time: Good	0.11	0.01	<0.001*	0.08	0.13
Time: Functional status (Ref.= Working)					
Time: Bedridden	-0.04	0.01	0.009*	-0.06	-0.01
Time: Ambulatory	-0.21	0.10	0.030*	-0.41	-0.01
Time: WHO clinical stage (Ref.=Stage-I)					
Time: Stage-II	0.04	0.01	<0.001*	0.01	0.07
Time: Stage-III	0.03	0.01	0.031*	0.03	0.05
Time: Stage-IV	0.06	0.02	<0.001*	0.03	0.09
Random Effect	**Std. Dev.**	**Corr.**			
Intercept	1.49	(Intr)			
Time	0.04	-0.81			
Residual	1.81				
AIC			**9151.22**		

*Significant at 5% level

3.3. Separate analysis of the risk of viral rebound

The results of the fitted Cox PH model for time to viral rebound are presented in Table 4. Age was found to be significantly associated with the risk of viral rebound of HIV-infected patients. The hazard rate increases by 1.78 with a one unit increase in age, keeping the other covariates constant. The estimated hazard ratio for a patient with no peripheral neuropathy was 0.56 (95% CI: 0.41-0.76). The implication is that the risk of viral rebound for a patient who didn't develop peripheral neuropathy was about 44% lower than that who suffered from the same, controlling for the other covariates in the model.

The other significant covariate was WHO clinical stage. The risk of viral rebound was 2.86, 2.94 and 3.23 times higher for HIV patients at WHO clinical stages II, III and IV compared to those in stage I, respectively. The results also indicated that the risk of viral rebound was 45% lower for patients with good adherence to ART compared to those with poor adherence (HR = 0.55; 95% CI: 0.30 - 0.99). However, there is no significant difference in the risk of viral rebound among patients with poor and fair adherence to treatment.

Table 4: Fitted multivariable Cox proportional hazards model

Variable	Coef.	HR	Std. Err.	p-value	95% CI	
					Lower	Upper
Age	0.56	1.78	0.21	<0.001*	1.18	2.67
Adherence (Ref.=Poor)						
Fair	-0.01	0.99	0.26	0.097	0.59	1.64
Good	-0.60	0.55	0.30	0.004*	0.30	0.99
WHO clinical stage (Ref.=Stage-I)						
Stage-II	1.05	2.86	0.35	0.002*	1.45	5.62
Stage-III	1.08	2.94	0.34	0.001*	1.49	5.78
Stage-IV	1.17	3.23	0.32	<0.001*	1.71	6.08
Peripheral neuropathy (Ref.=Yes)						
No	-0.58	0.56	0.16	<0.001*	0.41	0.76

*Significant at 5% level

3.4. Results using joint models

Table 5 pertains to the results of the joint models (shared random-effects model for the longitudinal marker and Cox PH survival model for the time-to-event outcome). In general, we observe that the estimated coefficients in the joint models have narrower confidence intervals as compared those in separate models. This indicates that the joint model is more precise than the separate models. Moreover, the association parameter was found to be significant. The estimated value of the association parameter (-0.102) indicates that the higher the square root of CD4 cell count, the lower the hazard of viral rebound.

Table 5: Results of the joint model for longitudinal and survival responses

Longitudinal sub-model	**Fixed effect**				**95% CI**	
		Coef.	**Std. Err.**	**p-value**	**Lower**	**Upper**
	(Intercept)	21.38	0.39	<0.001*	20.59	22.16
	Age	-0.23	0.01	<0.001*	-0.25	-0.21
	Adherence (Ref.=Poor)					
	Fair	0.95	0.28	0.001*	0.39	1.50
	Good	1.07	0.2	<0.001*	0.66	1.47
	Functional status (Ref.= Working)					
	Bedridden	-3.09	0.32	<0.001*	-3.72	-2.45
	Ambulatory	-2.19	0.24	<0.001*	-2.67	-1.71
	WHO clinical stage (Ref.=Stage-I)					
	Stage-II	-2.46	0.29	<0.001*	-3.04	-1.88
	Stage-III	-5.46	0.30	<0.001*	-6.06	-4.86
	Stage-IV	-8.54	0.37	<0.001*	-9.27	-7.80
	Time	-0.01	0.01	0.367	-0.03	0.01
	Time: Adherence (Ref.=Poor)					
	Time: Fair	0.03	0.01	0.010*	0.01	0.05
	Time: Good	0.11	0.01	<0.001*	0.08	0.13
	Time: Functional status (Ref.= Working)					
	Time: Bedridden	-0.04	0.01	0.010*	-0.06	-0.01
	Time: Ambulatory	-0.19	0.01	0.030*	-0.21	-0.17
	Time: WHO clinical stage (Ref.=Stage-I)					
	Time: Stage-II	0.04	0.01	<0.001*	0.01	0.06
	Time: Stage-III	0.04	0.01	0.035*	0.03	0.06
	Time: Stage-IV	0.06	0.02	<0.001*	0.03	0.09
	Random Effect	**Std. Dev.**	**Corr.**			
	Intercept	1.48				
	Time	0.03				
	Residual	1.80				
	(Intercept, Time)		-0.80			

Survival sub-model	**Fixed effect**	**Coef.**	**Std. Err.**	**p-value**	**HR**	**95% CI**	
						Lower	**Upper**
	Age	0.59	0.20	<0.001*	1.80	1.21	2.67
	Adherence (Ref.=Poor)						
	Fair	-0.74	0.25	0.008*	0.48	0.28	0.78
	Good	-0.65	0.29	0.002*	0.52	0.29	0.93
	WHO clinical stage (Ref.=Stage-I)						
	Stage-II	0.95	0.33	0.004*	2.59	1.34	5.01
	Stage-III	1.03	0.31	0.001*	2.80	1.51	5.21
	Stage-IV	1.00	0.33	0.002*	2.72	1.41	5.25
	Peripheral neuropathy (Ref.=Yes)						
	No	-0.53	0.15	<0.001*	0.59	0.43	0.79
	Assoc.	-0.102	0.02	<0.001*	0.90	0.86	0.94
	AIC			**6439.26**			

We can see from the results that, on average, when the age of HIV-infected patients increases by one (year), the square root of their CD4 cell count (per cubic millimeter of blood) decreases by 0.23, holding other covariates constant. Regarding disease severity, the square root of CD4 cell count was 2.46, 5.46 and 8.54 lower for patients at WHO clinical stages II, III and IV as compared to those at stage I,

respectively. Patients with fair and good adherence to treatment tend to have 0.95 and 1.07 higher square root of CD4 cell count as compared to those with poor adherence, respectively. Moreover, the square root of CD4 cell count was 3.09 and 2.19 lower for patients with bedridden and ambulatory functional statuses as compared to those with working functional status, respectively, controlling for the other covariates.

The interaction between time-to-follow-up and functional status was significant. This suggests that, as the follow-up time increases, the average square root of CD4 cell counts (per cubic millimeter of blood) of bedridden and ambulatory HIV-patients tends to be lower than those of working patients by 0.04 and 0.19, respectively. We can make similar interpretations for the interactions of WHO clinical stages as well as adherence to ART with follow-up times.

When we come to the fitted Cox PH sub-model, the hazard rate for age is 1.80. Thus, the hazard rate increases by 1.80 for a one unit increase in the age of HIV-infected patients. The risk of viral rebound was 52% (HR = 0.48; 95% CI: 0.28-0.78) and 48% (HR = 0.52; 95% CI: 0.29-0.93) lower for HIV patients with fair and good adherence to treatment, respectively, compared to those with poor adherence. Regarding disease severity, patients at WHO clinical stages II, III and IV were 2.59, 2.80 and 2.72 times more likely to experience viral rebound compared to those at stage I, respectively. Moreover, the risk of viral rebound was 41% lower for HIV patients who did not develop peripheral neuropathy compared to those who developed the same, keeping the other covariates constant.

The plots of marginal survival and marginal cumulative hazards against follow-up time are shown in Figure 2. We can see that the survival probability (probability of not developing viral rebound) appears to be linearly decreasing, while the cumulative probability of developing viral rebound (cumulative hazard) appears to be increasing as the follow-up time increases.

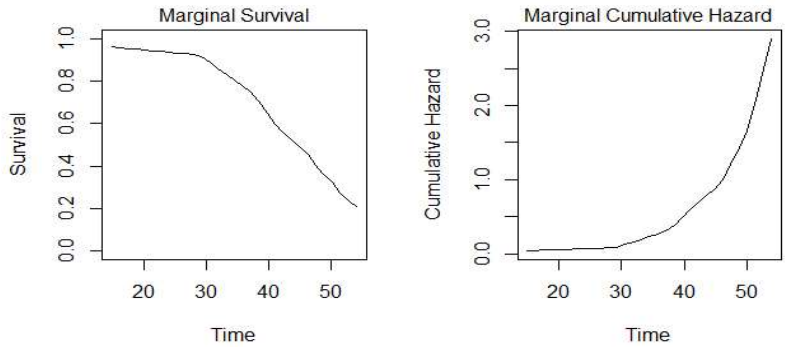

Figure 2: Marginal survival and marginal cumulative hazard plots

3.5. Joint model diagnostics

To validate the assumptions behind mixed models and relative risk models, standard types of residual plots can be used. The plot of the fitted values versus residuals indicates that the "cloud" of points is evenly distributed about zero with no systematic pattern, indicating no evidence of non-constant variance and that the square root of CD4 measure is linear in the parameters. The Q-Q plot also shows that the response variable in the longitudinal sub-model is approximately normally distributed (Figure 3).

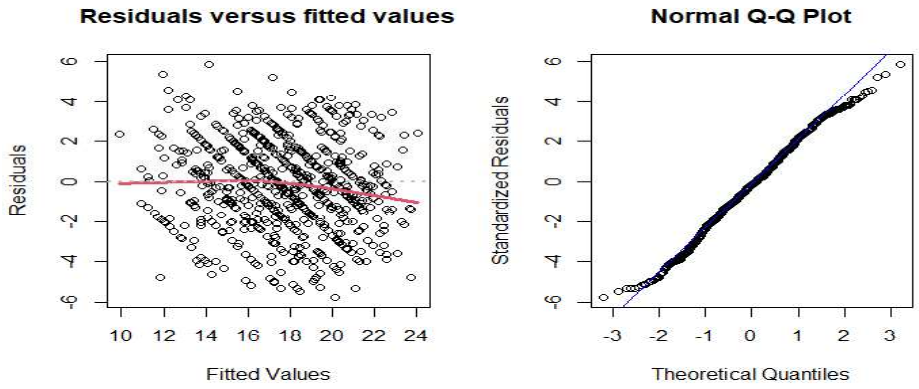

Figure 3: Residual plot against fitted values and Q-Q plot

Figure 4 is a plot of marginal residuals versus the marginal fitted values for the longitudinal process. Similar to our observation discussed above, the residuals roughly form a horizontal band around the y-axis (the zero line), suggesting that the variances of the error terms are equal (homoscedasticity). Moreover, the assumption of a linear relationship in the longitudinal sub-model seems reasonable since the residuals bounce randomly around the zero line.

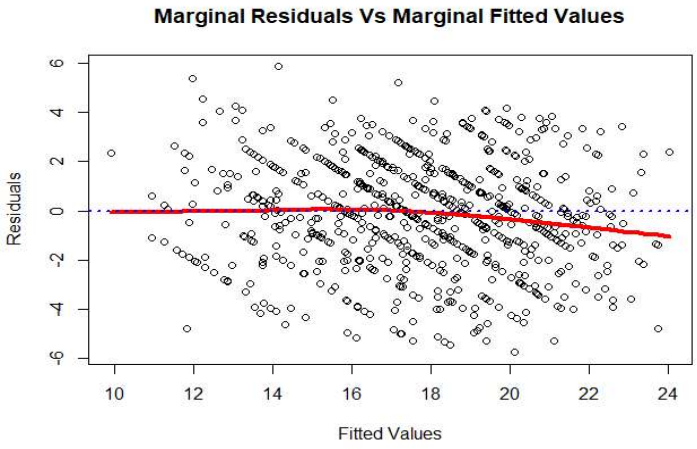

Figure 4: Plot of marginal residuals versus fitted values

3.6. Discussion

This study considered separate models for the time-to-event outcome (viral rebound) and the longitudinal response (CD4 cell counts) as well as a joint model for the two outcomes. We found that the fitted joint model was relatively more precise than the separate models since the estimated coefficients in the joint model have narrower confidence intervals. Moreover, the significance of the association parameter is perhaps an indication that the joint model is preferred to the separate models. Consistent with the findings of the study by Abdulbasit et al. (2018), the association parameter in the joint model was negative and significant, indicating that CD4 cell counts and time to viral rebound are negatively associated.

The study revealed that age, adherence to treatment and WHO clinical stage were significantly associated with both outcomes in the joint model. Additionally, peripheral neuropathy was a significant predictor of time to viral rebound (patient survival time), while functional status and interactions of adherence, functional status and WHO clinical stage with time were significantly associated with CD4 cell count.

Baseline age had a significant negative association with CD4 cell count, that is, the mean square root of CD4 cell count decreases as the age of a patient increases. This finding is in line with that reported by Tiruneh et al. (2021). The study revealed that adherence to ART has a significant effect on the CD4 cell count of HIV-infected patients. Patients with good and fair adherence to treatment have significantly higher CD4 cell count as compared to those with poor adherence, controlling for the other predictor variables in the model. This result is consistent with studies conducted by Maina et al. (2020) and Desta et al. (2020).

Considering functional status of patients on ART, the square root of CD4 count (per cubic millimeter of blood) was significantly lower for ambulatory and bedridden patients compared to that of working patients. This indirectly shows that patients with ambulatory and bedridden functional status have higher risk of viral rebound than patients with working functional status due to the significant and negative estimate of the association parameter between CD4 cell counts and time to viral rebound. These results are similar to the findings of the study conducted by Temesgen et al. (2018). WHO clinical stage was also found to have a significant effect on the CD4 cell counts of HIV-infected patients. The square root of CD4 cell count was significantly lower for patients at WHO clinical stages II, III and IV as compared to those at stage I.

Baseline age had a significant positive association with the hazard of viral rebound, that is, the risk of viral rebound for HIV-infected patients increases with age. However, the vast majority of studies reported

that younger adults experienced higher prevalence of viral rebound compared with older adults (e.g., Le Moing et al., 2002; Shoko and Chikobvu, 2018). Palmer et al. (2018) also found that for 1-year increase in age, there was a 1% decrease in the likelihood of viral rebound. They argued that older patients have a better understanding of the treatment therapy resulting in a better adherence to the treatment. The risk of viral rebound for HIV-patients with fair and good adherence to ART was lower than those with poor adherence to treatment. Our finding is similar to those of Shoko and Chikobvu (2018) and Le Moing et al., 2002). Non-adherence results in antiretroviral agents not being able to maintain sufficient concentration to suppress HIV replication in infected cells and to lower the plasma viral load (Chesney, 2000). Enhanced health education and close follow-up of PLHIV on antiretroviral therapy are crucial to reinforce adherence and maintain an undetectable viral load.

Another important variable significantly associated with the risk of viral rebound of patients was WHO clinical stage. The risk of viral rebound was higher for HIV-infected patients at stages II, III and IV compared to those at stage I, controlling for other predictor variables. Consistent with our finding, Maina et al. (2020) reported that the likelihood of viral rebound increased with a long duration on ART, poor adherence, and advanced disease progression. Jobanputra et al. (2015) also found that patients with advanced disease were more likely to have higher viral loads and least likely to achieve viral suppression when compared to the less advanced cohort. Consistent with the findings of the study by Shoko & Chikobvu (2018) and Simpson (2002), peripheral neuropathy was also significantly associated with the risk of viral rebound, with HIV patients not experiencing peripheral neuropathy having lower risk of viral rebound compared to those who developed peripheral neuropathy.

4. Conclusion

The results of our study indicated that higher CD4 cell count is associated with a significant reduction in the hazard of viral rebound. According to our finding, the risk of viral rebound for HIV-infected patients under ART increases with age. Moreover, the odds of viral rebound were significantly lower for patients with fair and good adherence to ART, those at WHO clinical stage I, and those who did not develop peripheral neuropathy. The study also revealed that patients with poor adherence to treatment, those with ambulatory and bedridden functional status, and those at WHO clinical stages II, III and IV had significantly lower CD4 cell count (per cubic millimeter of blood) compared to those with better adherence to treatment, working patients and those at lower disease severity level (stage I), respectively.

The results of both separate and joint analyses were found to be consistent. However, the significance of the association parameter that relates the hazard of viral rebound with the longitudinal response (CD4 cell

counts) indicates that the joint model is more appropriate. In addition, compared to the separate Cox PH and mixed-effects linear models, our joint model produced smaller standard errors, indicating an increased efficiency in the joint model. Thus, joint modeling should be preferred for simultaneous analyses of repeated measurements and survival data.

Limitations of the study

We were unable to include various important clinical, socio-demographic and socioeconomic variables that might have contributed to the viral rebound of HIV/AIDS patients. These include lactic acidosis, consumption of alcohol, smoking, income level, liver abnormality and diet style Moreover, the present study was restricted to the age group 18 and over due to the different measures of CD4 cell count for children and adult HIV/AIDS patients.

Acknowledgements

The authors would like to express their profound gratitude to Jimma University Medical Center for unlimited access to the pertinent medical registers from which the data used for the current study were extracted. Moreover, the authors are highly thankful to the Research and Postgraduate Coordinating Office, College of Natural Sciences, Jimma University, for the financial support to carry out this study.

References

1. Abdulbasit, A., Luguterah, A., Nasiru, S. and Abdul Rahaman, S. (2018). Joint Longitudinal and Survival Modeling Of HIV in the Upper West Region of Ghana. *Int J Health Sci*, 6(1):56-63.
2. Chesney, M. (2000). Factors Affecting Adherence to Antiretroviral Therapy. *Clin Infect Dis*, 30(Suppl 2): S171-S176.
3. Conway, J.M., Perelson, A.S. and Li, J.Z. (2019). Predictions of time to HIV viral rebound following ART suspension that incorporate personal biomarkers. *PLoS Comput Biol*, 15(7): e1007229.
4. Cox, D.R. (1972). Regression Models and Life-Tables, *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187-220.
5. Desta, A.A., Kidane K.M., Woldegebriel, A.G., Ajemu, K.F., Berhe, A.A. et al. (2020). Level of Adherence and Associated Factors among HIV-Infected Patients on Antiretroviral Therapy in Northern Ethiopia: Retrospective Analysis. *Patient Preference and Adherence*, 14:1585-1594.
6. Diggle, P., Heagerty, P., Liang, K. and Zeger, S. (2002). Analysis of Longitudinal Data. Oxford Statistical Science Series. Second Edition.

7. EPHI. (2020). HIV Related Estimates and Projections in Ethiopia for the Year-2019. https://ephi.gov.et/wp-content/uploads/2014/09/HIV_estimation_and_projection_for_Ethiopiawith-uncertainity_2019-PR.pdf
8. Harville, D.A. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *Journal of the American Statistical Association*, 72(358), 320-338.
9. Hedeker, D. and Gibbons, R.D. (2006). Longitudinal Data Analysis. Wiley-Interscience, New York.
10. Hsieh, F., Tseng, Y.K. and Wang, J.L. (2006). Joint Modeling of Survival and Longitudinal Data: Likelihood Approach Revisited. *Biometrics*, 62(4):1037-1043.
11. Jobanputra, K., Parker, L.A., Azih, C., Okello, V. et al. (2015). Factors Associated with Virological Failure and Suppression after Enhanced Adherence Counselling in Children, Adolescents and Adults on Antiretroviral Therapy for HIV in Swaziland. *PLoS ONE*, 10(2): e0116144.
12. Laird, N.M. and Ware, J.H. (1982). Random-effects Models for Longitudinal Data. *Biometrics*, 38(4):963-74.
13. Le Moing, V., Chêne, G., Carrieri, M.P. et al. (2002). Predictors of Virological Rebound in HIV-1-Infected Patients Initiating a Protease Inhibitor-Containing Regimen. *AIDS*, 16:21-9.
14. Liu, Y. and Liu, L. (2015). Joint Models for Longitudinal Data and Time-to-Event Occurrence. Routledge International Handbook of Advanced Quantitative Methods in Nursing Research. 1st Edition. Routledge.
15. Maina, E.K., Mureithi, H., Adan, A.A., Muriuki, J., Lwembe, R.M. and Bukusi, E.A. (2020). Incidences and Factors Associated with Viral Suppression or Rebound Among HIV Patients on Combination Antiretroviral Therapy from Three Counties in Kenya. *International Journal of Infectious Diseases*, 97:151-158
16. McCrink, L., Marshall, A.H. and Cairns, K. (2011). Joint Modelling of Longitudinal and Survival Data: A comparison of Joint and Independent Models. In Bulletin of the International Statistical Institute Proceedings of the 58th World Statistics Congress, Dublin, 4971-4976.
17. Oguntibeju, O.O. (2012). Quality of Life of People Living With HIV and AIDS and Antiretroviral Therapy. *HIV AIDS (Auckl)*, 4:117-24.
18. Palmer, A., Gabler, K., Rachlis, B., Ding, E. et al. (2018). Viral Suppression and Viral Rebound among Young Adults Living With HIV in Canada. *Medicine (Baltimore)*, 97(22):e10562.
19. Rizopoulos, D. (2012). Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. 1st Edition. Chapman and Hall/CRC, New York.
20. Shoko, C. and Chikobvu, D. (2018). Determinants of Viral Load Rebound on HIV/AIDS Patients Receiving Antiretroviral Therapy: Results from South Africa. *Theor Biol Med Model*, 15, 10.

21. Simpson, D.M. (2002). Selected Peripheral Neuropathies Associated with Human Immunodeficiency Virus Infection and Antiretroviral Therapy. *J NeuroVirol*, 8(2):33-41.
22. Temesgen, A., Gurmesa, A. and Getchew, Y. (2018). Joint Modeling of Longitudinal CD4 and Time-to-Death of HIV/TB Co-infected Patients at Jimma University Specialized Hospital. *Annals of Data Science*, 5(4):659-678.
23. Tiruneh, F., Chewaka, L. and Abdissa, D. (2021). Statistical Joint Modeling for Predicting the Association of CD4 Measurement and Time to Death of People Living with HIV who Enrolled in ART, Southwest Ethiopia. *HIV/AIDS-Research and Palliative Care*, 13:73-79.
24. Tsiatis, A.A. and Davidian, M. (2004). Joint Modeling of Longitudinal and Time-To-Event Data: An Overview. *Statistica Sinica*, 14:809-834.
25. UNAIDS. (2021). Confronting Inequalities: Lessons for Pandemic Responses from 40 Years of AIDS. https://www.unaids.org/sites/default/files/media_asset/2021-global-aids-update_en.pdf
26. Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. Springer, New York.
27. Wang, C.Y., De Dieu Tapsoba, J., Anderson, M.L., Vernon, S.W., Chubak, J., Fuller, S. and Green, B.B. (2014). Time to Screening in the Systems of Support to Increase Colorectal Cancer Screening Trial. *Cancer Epidemiology Biomarkers and Prevention*, 23(8):1683-1688.