

# Repeated Measure Analysis for the CD4<sup>+</sup> Cell Counts of HIV-Positive Patients Initiated to ART: A Case Study at Ambo Hospital

Endale Alemayehu\*, Tsigereda Tilahun

Departments of Statistics, Ambo University, Ambo, Ethiopia

## Email address:

endalestat@gmail.com (E. Alemayehu), tisgeti@gmail.com (T. Tilahun)

\*Corresponding author

## To cite this article:

Endale Alemayehu, Tsigereda Tilahun. Repeated Measure Analysis for the CD4<sup>+</sup> Cell Counts of HIV-Positive Patients Initiated to ART: A Case Study at Ambo Hospital. *International Journal of Data Science and Analysis*. Vol. 7, No. 2, 2021, pp. 32-38.

doi: 10.11648/j.ijdsa.20210702.12

**Received:** February 8, 2021; **Accepted:** March 24, 2021; **Published:** April 20, 2021

---

**Abstract:** Introduction: HIV is a virus that causes Acquired Immunodeficiency Syndrome (AIDS) by reducing a person's ability to fight the infection. It attacks an immune cell called the CD4 cell which is responsible for the body's immune response to infectious agents. Now a days anti retro viral therapy treatment is avail to elongate the life of patients. The treatment is given for patients to increase the CD4 counts of patients to keep the ability of body preventing the disease. Objectives: This study was aimed to identify the potential associated risk factors with CD4 counts of patients under ART treatment at public hospital in Ethiopia. The other was to fit linear mixed model by handling missing value of the data during follow up time. Method: To see the structure of the data, exploratory data analysis was conducted. Of the familiar variance structures, unstructured variance covariance is selected to be best and to fit the data under study, step-by-step procedure was passed to obtain best model. Results: The descriptive statistics directed that the progressive change in CD4 counts of females seems better than that of males. On the other hand, the output of the fitted model indicated that covariates significant with 5% level of significance is that baseline CD4, time, weight and interaction of Sex, baseline CD4 with time. Allowing the significance level to increase to 25% increases most covariates to be significant that help patients in a better awareness. Conclusion: With this result, full linear mixed with random intercept and slop is found to best model. There was high variability within patients over time and between patients and the interaction of time with covariates was also significant. Generally, the data was fitted by handling the missing value using multiple imputation technique.

**Keywords:** Longitudinal Data Analysis, CD4<sup>+</sup>, ART, Linear Mixed Model

---

## 1. Introduction

Human Immunodeficiency Virus (HIV) is a virus that causes Acquired Immunodeficiency Syndrome (AIDS) by reducing a person's ability to fight the infection. HIV attacks an immune cell called the CD4 cell which is responsible for the body's immune response to infectious agents [2]. HIV/AIDS is one of the major public health problems in world especially Sub-Saharan Africa, and Ethiopia, as one of these countries has been the most affected by the epidemic. In 2016, UNAIDS reported that the trend in new HIV infections among adults (aged 15 years and older) had remained static between 2010 and 2015 at 1.9 million new infections per year [2015 range of 1.7 million–2.2 million].

In 2017, however, UNAIDS estimates show a slightly different trend: new adult infections are estimated to have declined by 8% between 2010 and 2015 and by 11% between 2010 and 2016 [2].

Theoretically, the progressive change of HIV infection from an asymptotic stage to acquired immunodeficiency syndrome, AIDS, have a direct associated with a gradual decline in the total number of CD4<sup>+</sup> T cells in the blood. The biologist interpret this fact as, the decrease in the total number of CD4<sup>+</sup> T cells also correlates to an increase in the number of infected T cells and an increase in the amount of free virus in the blood [14].

However, HIV is not the only responsible for the reduction of CD4 from the body; rather other factors are also sharing

this infection. For instance, one with older age is under the risk of higher CD4 reduction. In regardless of sex, different suggestions were given by different studies. Some study indicated that females experienced better CD4 count response to ART compared to males as reported by a study done in North Ethiopia [3]; but others gave weight for male [7, 8]. Available literature indicated that there is significant positive relationship between the baseline CD4 counts and CD4 counts over time [8].

HIV/AIDS is one of the major public health problems in the world and Ethiopia, as one of these countries has been affected by the epidemic with a higher burden of prevalence rate. Now a day, even if many directional HIV/AIDS prevention methods were introducing, the number of infected people were increasing thoroughly worldwide. CD4 count is used to measures the number CD4 cells in the blood mostly during ART treatment to know the risk progression of HIV in the HIV infected patients. To attain on common idea, number of researchers have tried to model the CD4 counts inconsideration of longitudinal data and other correlated dataset. As there is no medicine to cure of this virus yet, the drug that used to reimburse the damaged CD4 counts which is key to elongate the life of infected people is the issue that investigators deals very currently [4, 16].

However, statisticians have no common view in modeling this natural count form of CD4 and they all are with impressive reason for their approach to use; especially it seems that they are in two groups; modeling CD4 counts in the form of Gaussian linear mixed model and non-Gaussian generalized linear mixed model. Since the nature of the CD4 count is naturally in the count form, investigators preferred to look for non-Gaussian especially Poisson distribution under Generalize linear mixed model and its extension [1, 8]. However, Poisson distribution has one drawback of having the same mean and variance which is very impossible in practice and even using its extension Negative-Binomial for CD4 count is not satisfactory; since assuming very disperse distribution of the virus is not expected in real world [5, 9, 15].

The general objective of this study was to model the progression of CD4<sup>+</sup> counts of HIV-positive patients on ART

for a period of four years

It is known that the burden of HIV/AIDS is a critical issue professional stress for long times to come up with drug which cure the virus. But, yet no one is lucky to do so, the number of infected person increase through time. As immediate solution for a being, CD4 compensator Anti-Retroviral treatment is avail. Since HIV is not the only responsible for the reduction of CD4 immune body, researchers conducted repetitive study by considering other genital and environmental factors [11].

Considering the sub-objective this study has addressed, different sectors and individuals will benefit from the relevant scientific result of the study. For instance, it will enhance the infected individuals to have careful attendance of ART in order to get regular increment of CD4 counts. In other way, this study will also use as input for the ART trainer so that they can give attention and follow those with less CD4 count based on identified factors in this study.

## 2. Data and Methodology

### 2.1. Source of Data

The data for this study has been obtained from Ambo Hospital in Ethiopia from 191 patients and 1528 measurements. As usual, all the relevant information was taken from patients' cards those starts ART treatment for the specified 4 year period of time.

### 2.2. Variables in the Study

In the study any research, selecting the outcome variable and the corresponding factors is very crucial to get potential results that agreed with real world facts.

#### Dependent Variable

The response variable of this study was the CD4 counts of HIV-infected patients during the follow up time of 4 year from the date of ART treatment initiation. The transformed square root CD4 was considered to assure normality assumption.

*Table 1. Independent Variables.*

Variables	Explanation	Code for categorical Variables
ID	Patients identification card	
Sex	Sex of Patients	0=Male and 1=Female
Age	Age of Patients at the baseline	
Marital status	Marital Status of the patient at the baseline	0=married, 1=Single and 2=Others
BaseCD4	Baseline CD4 counts	
Weight	Weight in Kg at the baseline	
Time	Time in months	0 to 7 (checked per six month)
WHO stage	WHO staging at ART initiation	0=Stage I, 1=Stage II, 2=Stage III and 3=Stage IV
Functional Status	Functional Status of patients	0=working, 1=ambulatory and 2=bed ridden)
Regimen Class	ART regimens taken	0=AZT-3TC-EFV, 1=AZT-3TC-NVP, 2=TDF-3TC-NVP, 3=d4t-3TC-NVP and 4=TDF-3TC-EFV

### 2.3. Methodology

To answer the basic research questions and realize the objectives of the study, different statistical methods have used.

For the analysis of the data R 4.0 and SAS 9.4 statistical software was used. To explore data techniques like individual plot profile, average plot profile, variance structure plot and correlation structure plot, and some summery numeric

statistics were used. The model used to fit the data was linear mixed model and other different techniques of model selection were stated step-wise here below.

#### *Longitudinal Data Analysis*

Longitudinal data are the repeated measurement of the same cases over time. It is commonly applicable especially in medicine and social science. Longitudinal data may come in many different forms that its measurement can be taken at different and irregular time points. The main advantage of longitudinal data analysis is that it relaxes the independence assumption of classical linear regression and handles the interdependency of observation [13].

#### *Linear Mixed Model (LMM)*

**Model Building:** When independence assumption is failed in general linear model because of dependence among observation or repetition of subjects under study, linear mixed model was used to produce proper inference by allowing residuals to be correlated.

**Stage 1:** The development of a linear mixed-effects model can be performed in two stages [10, 12]. In the first step, an ordinary linear regression model is fitted separately for each subject separately. The linear regression model is used to identify the within variability effect and denoted by:

$$Y_i = Z_i\beta_i + \varepsilon_i \quad (1)$$

Where:

1.  $Z_i$  is (nxq) matrix of known covariates
2.  $\beta$  is a dimensional vector of subject-specific regression coefficients
3.  $\varepsilon_i \sim N(0, \Sigma_i)$  is the vector of measurement errors (residuals). In many situations, a simple structure for the covariance matrix for measurement error is assumed.

**Stage 2:** In the second stage, between subjects variability is modeled. This is achieved by relating the estimated coefficients  $\beta_i$  with known covariates. Thus,

$$\beta_i = K_i\beta + b_i \quad (2)$$

Where:

1.  $K_i$  is (qxp) matrix of known covariates. Additional covariates other than time can be adjusted for in this step.
  2.  $\beta$  is a  $P$  dimensional vector of regression coefficients
- $b_i \sim N(0, D)$  and is a matrix of random effects indicating the deviation in individual subjects' measurements from the population average.  $D$  is the covariance matrix capturing the between-subject variability.

Generally, the linear mixed model from the above two stages were formulated as:

$$Y_i = Z_iK_i\beta + Z_i b_i + \varepsilon_i = X_i\beta + Z_i b_i + \varepsilon_i \quad (3)$$

**Missing Data:** The impact of missing data is a subject that most of us want to avoid. In most cases, you have to eliminate missing data before you address the substantive questions that led you to collect the data. You can eliminate missing data by (1) removing cases with missing data or (2) replacing missing data with reasonable substitute values. In

either case, the end result is a dataset without missing values. With this study, we dealt with missing values without deleting the row with missing data. To do so, first we identified the nature of missing values

## **3. Results and Discussion**

### **3.1. The Overall Descriptive Statistics**

With this sub section we presented some rough overall view of exploratory data analysis of CD4<sup>+</sup> counts for those patients registered under ART follow up treatments at some public hospital in Ethiopia for four consecutive years. The table 2 indicates the mean distribution of CD4 counts over follow up time. The mean average CD4 count of patients during the baseline is 264.4293 with standard deviation of 218.431 which is an indicator of large difference between patients when they registered. The total follow up time for individuals were 8 times which taken per six months and with this study patients were failed to follow up all the checkup time which totally there were 411 missing with CD4 outcome variable. When the number of patients under ART treatment is decreased over time, the mean counts of CD4 is increase except it slightly decreased at fifth follow up time.

### **3.2. Testing the Distribution of Outcome Variable (CD4<sup>+</sup>)**

The distribution of the dependent variable is the significant factor in identifying the model to use and is the base work before starting analysis. It gives potential clues on the nature and characteristics of the data and what to do next. To agree with this theory, we first tested the assumption of normality for the CD4 and sqrtCD4 with Shapiro test (for univariate at each follow up time), and q-q plot (discussed after the model was fitted)

#### **3.2.1. Mean Profile Plot**

The primary question under this method is how responses vary with time. The immediate impulse is to plot  $Y_{ij}$  against  $t_i$ , with values for each individual connected by line segments. These plots can be very messy, and practically useless [6]. To assess the trend with time requires some estimate of the trend to be plotted, along with enough information about the distribution of values to assess the strength of the trend.

The figure 1 indicated the mean profile plot of CD4 count over time using loess smoothing with missing value and by multiple imputation respectively. The red line on the left hand side figure revealed that there seems slight increase in average counts of CD4 and whereas similarly the next plot also assured that there was increase in the mean counts of CD4 expecting not constant increase at follow up time.

#### **3.2.2. Variance Structure Plot**

The variability structure plays a great role in identifying the pattern of response variable and gives clues what variance-covariance structure to be expected.

The figure 2 is the variance structure of CD4 counts. It indicated that there is CD4 count variability at baseline stage than at time 1. The CD4 variability increase very fast from time 1 to 2, whereas slightly increase to time 3. Then the

pattern shows that it is fast decrease except increase for the last follow up time.

### 3.3. Statistical Model Developments

In any data analysis, model inference is the crucial part of the research. In doing so, identifying variables selection techniques and considering the covariance between subjects over time is the basic issue to be undertaken. The general procedure included to develop the model was included identifying the appropriate covariance structure, correlation structure, forward variable selection method started from linear model to linear mixed model including interaction and finally the last model was selected by considering estimation techniques.

#### 3.3.1. Covariance-Variance and Correlation Structure

The variance covariance structure is the important issues to be included in the longitudinal data analysis. The table 3 indicated that unstructured variance- covariance has smaller AIC (8022.394) which is to mean that it was better in explaining the model under study.

Table 2. Variance-Covariance Structure.

Model	df	AIC	BIC	logLik
CS	25	8273.605	8406.898	-4111.802
SY	52	8217.472	8494.721	-4056.736
AR	25	8364.957	8498.250	-4157.479
UN	59	8022.394	8312.932	-3952.197
ARH1	32	8142.317	8336.965	-4039.159

Model selection technique for full model and the effect of time interaction on full model was described under table 3 below. The AIC result indicated that the full model with interaction of time is the best model as compared to the full model without interaction.

Generally, with not considering the random effect, the final comparative best model used under marginal model was that the one including all covariates and the effect of interaction effect. Once the best marginal model was selected, then we have seen the influence of random effect on the change of transformed square root of CD4.

Table 5. Covariance estimates for the random effect including only random intercept.

Effects	Subject	Estimate	std. Error	Z. Value	Pr> Z ICC
Intercept ID	17.38179	0.5194	9.4153	<.0001	0.538
Residual	12.4571	0.2520	19.471	<.0001	

In table 5 above the intra class correlation calculated to be 0.538. This is to mean that the 53.8% of variability in progressive change of individuals (within patients over time) is due to random effects. Thus, ID has significant effect in changing the CD4 counts of patients and has not been excluded from the model.

#### (ii). Random Intercept and Slop Model

The introduction of random effects has important ramifications for the interpretation of the “fixed-effects”

Table 3. Comparison of full model and full model with interaction under marginal model.

Model	AIC	logLik
Full	8773.432	-4375.875
Full (interaction of time)	8727.685	-4332.842

#### 3.3.2. Hierarchical Linear Model

In longitudinal data analysis techniques the effect of random terms (subject) have to be seen to determine the between subject variability and the dispersion measure of individuals over time. As it was discussed under figure 1 the average measure of CD4 count was seems linear and since the distribution of actual CD4 counts seems not normal, the transformed square root CD4 was better normal which was used as response variable of this study. Therefore, linear mixed model by including the random effect to determine subject specific (hierarchical) influence, was developed and finally seen whether it has significant effect or not.

The table 4 below is the result of full marginal model and hierarchical intercept model. Hence, since the mixed model has smaller AIC (8369.065), the random effect has influence in explaining the subject-specific variability and it is the best relative model.

Table 4. Comparison of full model between without random effect and with random effect.

Models	AIC	logLik
1. full model without random effect	8727.685	-4152.532
2. full model with random intercept	8369.065	-4332.842

#### (i). Random Intercept only Model

For the data with repeated follow up time, the effect of differences among the subjects (patients in our case) is expected to have significant influence on the event under study. The random intercept here is considered to see the between subject variability among patients not considering the effect of time. In fitting the random effect, estimating random intercept is used to test the intra class correlation among patients. This is because determining coefficients for all random intercept (ID) of patients are quit tedious and we are limited to see the variation in explaining the changes of CD4 counts over time.

regression parameters. The random slop is used to test the between patients variability over time in that to study the difference among subject change in progressive of follow up time.

As indicated under table 6, the model including the random effect specifically random slop is better than the marginal model. Therefore, including the between patient variability is important in explaining the model; so that time has significant effect in progressive change of the transformed square root CD4 counts.

**Table 6.** Parameter estimates of full linear mixed model including random intercept and slop.

Effects	Estimate	St.error	tcal	df	Pr (> t )
(Intercept)	14.735496	2.1225684	6.942248	493.67723	1.220024e-11
Time	0.293962	0.4413610	0.666035	24.96554	5.1149e-03***
Sex2	-0.274535	0.5996433	-0.457830	767.92216	6.472038e-01
Age	-0.010407	0.0364965	-0.2851413	697.82306	7.756205e-01
m.status2	-0.414728	0.7839822	-0.529002	975.52984	5.969244e-01
baseCD4	0.020467	0.0014219	14.39456	619.97320	0.0000e+00***
Weight	-0.048588	0.0238510	-2.037162	1197.51845	4.1853e-02**
Regimenclass2	0.58124	0.6873402	0.845639	817.04149	3.980019e-01
Regimenclass3	0.064489	0.9939908	0.064878	395.32704	9.483036e-01
Regimenclass4	-1.768142	1.1312539	-1.562994	297.09079	1.191186e-01
Regimenclass5	-0.72195	0.8570207	-0.842401	544.84325	3.999330e-01
FunctionalStatus2	0.02225	1.0375097	0.021449	473.55908	9.828966e-01
FunctionalStatus3	1.06391	1.3056478	0.814849	978.15287	4.153569e-01
WHOStage2	0.380573	0.7398036	0.514425	1225.28455	6.070476e-01
WHOStage3	-0.76570	0.7607236	-1.006539	1166.96259	3.143647e-01
WHOStage4	-0.68072	0.9998538	-0.680823	1171.91140	4.961182e-01
Time*Sex2	0.15334	0.1059811	1.446864	53.03054	1.5382e-02**
Time*Age	0.0060	0.0069125	0.871358	36.82226	3.892049e-01
Time*m.status2	0.04535	0.1272212	0.356491	105.99674	7.221812e-01
Time*baseCD4	-0.001141	0.000223	-5.107296	112.60807	1.3439e-06**
Time*Weight	0.005156	0.003545	1.454566	459.48057	1.464723e-01
Time*Regimenclass2	-0.035556	0.113993	-0.311912	91.22862	7.558194e-01
Time*Regimenclass3	-0.059401	0.141022	-0.421221	543.94861	6.737602e-01
Time*Regimenclass4	0.462390	0.293012	1.578058	63.17791	1.1954e-02**
Time*Regimenclass5	0.129409	0.152023	0.851247	46.57025	3.9898e-02**
Time*FunctionalStatus2	0.071926	0.172492	0.416985	66.03113	6.780419e-01
Time*FunctionalStatus3	-0.063032	0.213424	-0.295339	111.42551	7.682847e-01
Time*WHOStage2	0.070116	0.112962	0.620710	260.07647	5.353338e-01
Time*WHOStage3	0.020697	0.119903	0.172618	157.16691	8.631738e-01
Time*WHOStage4	0.145099	0.164441	0.882375	102.63129	3.796359e-01

### 3.4. Discussions and Conclusion

The basic objective of this study was to identify the major significant factors which associated with the progressive change of CD4 counts over follow up time of patients under ART treatment at public hospital in Ethiopia. Before the inference analysis of the finding, the exploratory data analysis was viewed that the mean structure of CD4 counts indicated that the average measure was increase through time except decrease at the fifth time. Besides, the variance structure is clearly indicated that there is high variability in progressive change of CD4 counts between patients under study.

Since the primary objectives of the study is to develop linear mixed model, identify the significant factors for the case under study, step-by-step variable relevance have been tested by fitting different models so as to show the priority of full linear mixed model. In first step of model fitting, full marginal model was found to be best model of all marginal model fitted using forward variable selection method as stated in table 6. To check the interaction effect of follow up time with all other covariates, we compared it with full model without interaction an since its AIC value is found to be smaller, time is considered to have significant effect as interaction of the covariates. Under marginal model the interaction of time with Sex and baseline CD4 is significant. The other interesting portion of the model is hierarchical modeling with random intercept and slop. As the usual, we

first compared the best model from the marginal family (full linear model with time interaction) with linear mixed model of having only intercept and because of the later model have smaller AIC 8369.065 as stated under table 4 it is relatively best model. Thus, the random effect has influence on the progressive changes of transformed square root CD4 counts of patients. The other major key concern is the effect of random slop which compared with the model having only random intercept and under this model since the AIC value of the model including slop is smaller; it was selected as the best relative model with REML estimation techniques.

Finally, we were interested to compare the two familiar estimation methods maximum likelihood and restricted maximum likelihood under the last selected model and dramatically full linear mixed model including both intercept and slop with ML estimation has smaller AIC (8286.651) and found to be the final best model selected to fit this data.

The dropdown of CD4 counts in human body has hurried patients to fall under a risk by reducing disease prevention immunity of their body. This study was therefore, aimed to identify the risk factors associated with the progressive change of patients registered under ART treatment at public hospital in Ethiopia for four year per six month follow up time the finding is consistent with the study by [1, 3, 4].

To identify the potential covariates associated with progressive change of CD4 counts of patients, the variables included in the study were Age, baseCD4, Functional Status, ID, Marital status, Regimen class, Sex, Time of follow up,

Weight and WHOStage. Thus, to fit the best model we compared different model from marginal family using forward method and hierarchical family considering random intercept and slop. Finally, the full linear mixed model with both intercept and slop under maximum likelihood estimation method is found to best model.

The result of the fitted model indicated that baseline CD4, weight, time and interaction of Sex, baseline CD4 significant with 5% level of significance. But, allowing 25% significance level make most variables to be significant as it

is observed from table 6 and it is recommended to consider covariates those may not be fall under obvious level of significance 5% to increase the awareness of patients. The result of residuals is also indicated that the underlined assumption of normality is meeting. It is based up on the transformed square root CD4 counts since the distribution of actual CD4 counts indicated that the data is not normal. With the increasing sample size the significant factors can identified. Researchers should look for other models that handle the dependency to identify more factors.

Table 7. Descriptive Statistics of CD4 counts.

Time	baseline	1	2	3	4	5	6	7
N	191	159	168	140	149	125	101	84
Mean CD4 <sup>+</sup>	264.4293	347.1258	378.4048	394.7500	416.7047	413.7680	427.2970	434.7738
Sd.DevCD4 <sup>+</sup>	218.4310	216.3049	233.6539	234.4038	211.4112	201.3266	182.6537	194.0842

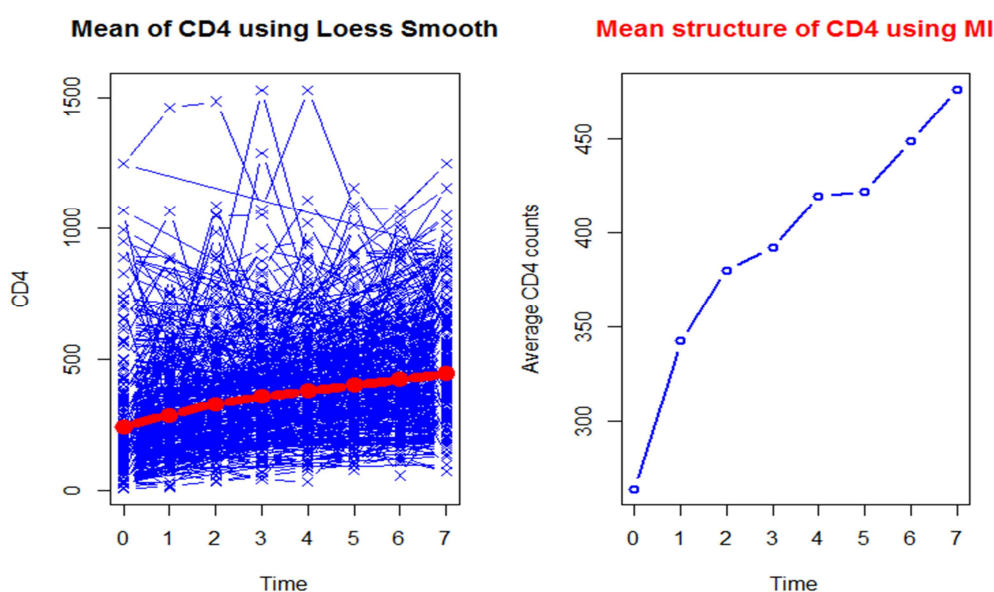


Figure 1. Mean profile plot of CD4 counts over time Using Loess smoothing and MI.

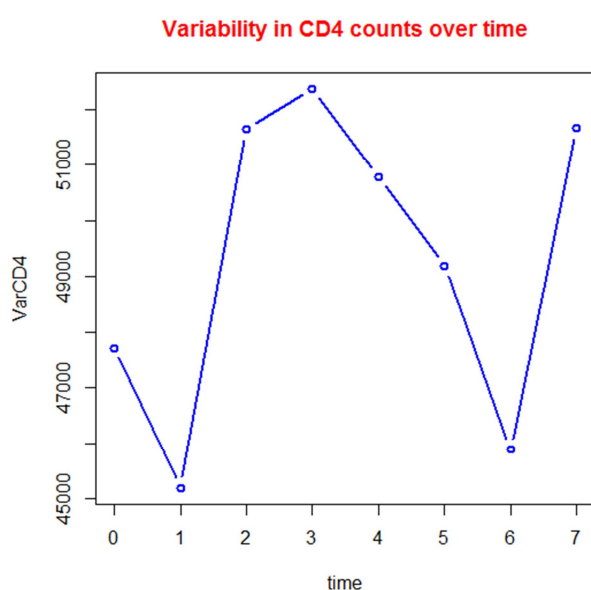


Figure 2. Variance structure of CD4 counts over time.

## Declarations

During conducting the study, the investigators have included the following declarations.

### Abbreviations

AIC: Akiaki information Criteria, AIDS: Acquired Immunodeficiency Syndrome, ART: Antiretroviral Therapy, HIV: Human Immunodeficiency Virus, HAART: Highly Active Antiretroviral Therapy, LMM: Linear Mixed Model, MI: Multiple Imputation.

### Availability of Data and Material

The data and materials will be provided at any time of request. There are no restrictions on the availability of data and R-code, and the investigators are willing to provide the code as well.

### Competing Interests

The authors declare that they have no competing interest.

### Authors' Contributions

This study was designed and compiled by EA as the principal investigator. The development of the basic research questions, identifying the problems and selecting appropriate statistical models have been done by him. Edition of the overall progress of the work and compiling the R-command was supported by TT.

### References

- [1] Aboma Temesgen (2017) Application of Poisson Mixed Combined Models for Identifying Correlations of CD4 Count Progression in HIV Infected TB Patients during ART Treatment Period. *International Journal of Statistics and Probability*; Vol. 6, No. 5; September 2017 ISSN 1927-7032 E-ISSN 1927-7040.
- [2] Adams, M. and Luguterah, A. (2013). Longitudinal analysis of change in CD4<sup>+</sup> cell counts of HIV-1 patients on antiretroviral therapy (ART) in the Builsa district hospital. *European Scientific Journal*, 9 (33), 1857-7881.
- [3] Addisu A, Dagim A, Tadele E, Adissu A, Mussie A, Filmon K. CD4 cell count trends after commencement of antiretroviral therapy among HIV infected patients in Tigray, northern Ethiopia: a retrospective cross-sectional study. *PLoS ONE*. 2015; 10 (3): e0122583. doi: 10.1371/journal.pone.
- [4] Awoke Seyoum, Principal Ndlovu and Zewotir Temesgen: Joint longitudinal data analysis in detecting determinants of CD4 cell count change and adherence to highly active antiretroviral therapy at Felege Hiwot Teaching and Specialized Hospital, North-west Ethiopia (Amhara Region). *AIDS Research and Therapy* DOI 10.1186/s12981-017-0141-3.
- [5] Burton, P., Gurrin, L., and Sly, P. (1998). Extending the Simple Linear Regression Model. *Statistics in Medicine*, 17, 1261-1291.
- [6] Diggle P. J., Heagerty P. J., Liang K. Y and Zeger S. L. (2002). *Analysis of Longitudinal Data*. (2nd Ed.). Oxford Science Publications. Oxford: Clarendon Press.
- [7] Fares Qeadan (2016). *Longitudinal Data Analysis by Example*. A seminar in biostatistics for the Mountain West Clinical Translational Research Infrastructure Network. University of New Mexico Health Sciences Center. Albuquerque, New Mexico.
- [8] Getachew Tekle, Wondwosen Kassahun and Abdisa Gurmessa Statistical Analysis of CD4<sup>+</sup> Cell Counts progression of HIV-1-positive Patients enrolled in Antiretroviral Therapy at Hossana District Queen Elleni Mohamad Memorial Hospital, South Ethiopia. *Biometrics & Biostatistics International Journal*.
- [9] Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, 72 (320-340).
- [10] Jiang, J., Rao, J. S., Gu, Z., and Nguyen, T. (2008). Fence methods for mixed model selection. *The Annals of Statistics* 36, 1669-1692.
- [11] Kaufmann RG, Perrin L, Pantaleo G, Opravil M, Furrer H, Telenti A, et al. CD4 T-lymphocyte recovery in individuals with advanced HIV-1 infection receiving potent antiretroviral therapy for 4 years: the Swiss HIV cohort study. *Arch Intern Med*. 2003; 163 (18): 2187-95.
- [12] Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38, 963-974.
- [13] Singer, J. D. & Willett, J. B. (2003), *Applied Longitudinal Data Analysis*, Oxford University Press, Oxford, UK.
- [14] Ve, V. Moing, L., Thiébaud, T., Chêne, G., Leport, C., Cailleton, V. Michelet, C., Fleury, H., Herson, S., Raffi, F.-Reviewed work (2002). Predictors of Long-Term Increase in CD4<sup>+</sup> Cell Counts in Human Immunodeficiency Virus-Infected Patients Receiving a Protease Inhibitor Containing Antiretroviral Regimen. *The Journal of Infectious Diseases*, Vol. 185, No. 4, pp. 471-480.
- [15] Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer-Verlag.
- [16] World health Organization (WHO), Regional office for South-East Asia New Delhi, (2007) Laboratory Guidelines for enumerating CD4 T Lymphocytes in the context of HIV/AIDS, ([www.who.int/hiv/amds/LaboratoryGuideEnumeratingCD4TLymphocytes.pdf](http://www.who.int/hiv/amds/LaboratoryGuideEnumeratingCD4TLymphocytes.pdf); accessed on December 11, 2016).