*International Journal of Current Research and Modern Education (IJCRME)*
*Impact Factor: 6.925, ISSN (Online): 2455 - 5428*
*(www.rdmodernresearch.com) Volume 5, Issue 1, 2020*

# AN APPLICATION OF LINEAR MIXED MODELS ON CD4+ COUNT PROGRESSION OF HIV+ PATIENTS UNDER ANTIRETROVIRAL THERAPY: A CASE OF DEBRE BERHAN REFERRAL HOSPITAL, DEBRE BERHAN, ETHIOPIA

**Abere Wondimu Kassie* & Damenech Syum****
Lecturer in Statistics, Department of Statistics, College of Natural and Computational Sciences, Debre Berhan University, Ethiopia

**Cite This Article:** Abere Wondimu Kassie & Damenech Syum, "An Application of Linear Mixed Models on CD4+ Count Progression of HIV+ Patients under Antiretroviral Therapy: A Case of Debre Berhan Referral Hospital, Debre Berhan, Ethiopia", International Journal of Current Research and Modern Education, Volume 5, Issue 1, Page Number 4-14, 2020.



**Abstract:**

In studying the progression of HIV-positive patients under ART, CD4+ count should be measured repeatedly per individual to decide about progression of patients to AIDS and monitoring the success of Therapy. The main aim of the study was modeling the progression of CD4+ counts of $HIV^+$ patients under ART in Debre Berhan Referral Hospital, Ethiopia. A retrospective cohort study was used and out of a population of HIV-patients who were taking ART in the hospital during September,2013toFebruary,2019,data on 322 patients was taken to the study; in which225 (69.88%) were females and 97(30.12%) were males. Linear Mixed Models used as data analysis which incorporate subject-specific random effects. The average baseline CD4+ count after patients initiated to ART was 335.7 and it changed to 408.6over specified period of time. Females'CD4+ count were better progression than males with an estimate, ($\beta$=0.7602, p=0.0215). Age had negative time interaction effect, ($\beta$=-0.01698, p=0.0092), implied older patients have lower mean change in CD4+ count than younger patients. Functional status, regimen class, age, marital status, WHO clinical stages had significant effect on patient's CD4+ count progression. The random intercept and random slope models have significant separate estimation with ($\beta$ =6.652, p=0.001) and ($\beta$=1.101, p=0.001) respectively. As conclusion, HIV-positive patients advised to start ART early with higher CD4+ counts as baseline and slope models have positive correlation estimate, 0.6621 implied patients who have higher baseline CD4+ count tend to have higher rate of increment than those with lower baseline CD4+ count.

**Key Words:** HIV, CD4+ count, Linear Mixed Model, HAART, Longitudinal Data Analysis

## 1. Introduction:

The CD4+ count is used as assessing the clinical status of HIV-infected patients, in making informed decisions regarding the initiation of antiretroviral therapy (ART). MeasuringCD4+ count is a strong predictor of progression to Human Immune deficiency Syndrome (AIDS) as well as a means of monitoring the success of such antiretroviral therapy. The possible increase or decrease in CD4+ counts are directly related to HIV replication. Low CD4+ counts are associated with a greater risk of patients living with HIV developing opportunistic infections, which may then progress to advanced disease and death (Langford et al., 2007 and Hoffman et al., 2010).

The CD4+ count is used as assessing the clinical status of HIV-infected patients', in making informed decisions regarding the initiation of antiretroviral therapy (ART).MeasuringCD4+ count is a strong predictor of progression to human immune deficiency syndrome (ADIS) as well as a means of monitoring the success such therapy. The possible increase or decrease in CD4+ counts are directly related to HIV replication. Low CD4+ counts are associated with a greater risk of patients living with HIV developing opportunistic infections, which may then progress to advanced disease and death (Langford et al., 2007 and Hoffman et al., 2010).

The use of combinations of antiretroviral drugs (ART) greatly results in the supervision of virus replication and hence increase levels of CD4+ count. The CD4+ count is also used to decide when to start antiretroviral therapy. Initiation on Highly Active Antiretroviral Therapy (HAART) at higher CD4+ counts has been demonstrated to "the risk of death, opportunistic infections and non HIV related co- morbidities". HAART consists of cocktails of at least two to three different classes of antiretroviral therapies and effectively lowers the concentration of the virus in the body by increasing the immune system which is called CD4+ count. In a healthy adult, a normal CD4+ count can vary enormously (by population, age group, etc.) but is typically around 500 to 1500 cells per cubic milliliter of blood (mm3). When it falls below 200, however, then the disease is technically classified as AIDS.

Limited studies were focused on modeling the CD4+ counts trend over time for patients on ART. The study by Ester and Ballard (1987) suggested that the progression of CD4+ T cells following a pattern of long period of slow declined just before one set of ADIS. Therefore this needs further investigation for the question

what happens about the progression after the start of HAART, which is examined in this study. A study showed that HAART brings a significance improvement to CD4+ count and provided further quantitative evidence about aspects of the therapy effect such as the changes in slope of CD4+ cell count shown by Berzuini and Allemani (2004). Moreover, patients who had already experienced an ADIS defining event of initiating HAART were also at higher risk of developing a new event, irrespective of their CD4+ count evolution during treatment.

Thus, to study the progression of HIV infection, the CD4+ count should be measured repeatedly per individual known as longitudinal study. Longitudinal studies are defined as studies in which the outcome variable is repeatedly measured; that is the outcome variable is measured in the same individual on several different occasions. In longitudinal studies the observations of one individual over time are not independent of each other, and therefore it is necessary to apply special statistical techniques, which take into account the fact that the repeated observations of each individual are correlated (Twisk J. 2003).

The generalized linear mixed models (GLMMs) represent one way to extend generalized models to longitudinal data; they extend bin a natural way the conceptual approach represented by the linear mixed effects model for continuous responses Larid and Ware (2004). Generalized linear mixed models are obt6ained from generalized linear models (McCullagh and Nelder, 1989) by incorporating random effects into the linear predictors, and include the well-known linear mixed models (LMMs) for normal responses as special case. These models are useful for modeling the dependence among response variable inherent in repeated measures, for accommodating over dispersion among binomial or poison responses, and for producing shrinkage estimators in multi-parameter problems.

**1.1 Statement of the Problem:**

The HIV infected patients have advised to start ART in order to reduce AIDS related morbidity and mortality by increasing their CD4+ count and to improve their quality of life. However, HAART improves the immune systems of patients, questions need to be answered about the numerical improvement of CD4+ counts over time after patients initiated to HAART and whether the evolution has different pattern depending on the patient's characteristics like gender, age, functional status, regimen class etc.

In line with this researchers interested to decide about the disease progress for a group of patients on their first occasions cross-section ally. But study the change in the number of cells over time is a good indicator of disease condition instead of studying the number of CD4+ count at one point in time; since it is not very instructive to tell about the disease status. In other words, modeling HIV biomarkers such as CD4+ count using standard generalized linear models without incorporating random effects to account between subject variability may lead to incorrect inferences. Because analyzing correlated data which measured repeatedly on same subject as it was independent; consequences inappropriate estimated standard errors and inefficient estimators. For this reason this study tried to examine the progression of the CD4+ count over time using linear mixed models by incorporating random and fixed effects to study within and between subjects variability and to assess associated determinant factors.

**1.2 Objectives of the Study:**

The aim of this study was to model the progression CD4+ counts over time among HIV+ patients under ART from September 11, 2013 to February 8, 2019 in Debre Berhan Referral Hospital.

The study has also the following specific objectives

- To examine the subject-specific evolution of CD4+ count over a specified period of time.
- To characterize the degree of heterogeneity across patients in the rate progression of CD4+ count among categories of ART regimen, gender, marital status, functional status and WHO clinical stages.
- ToidentifydeterminantfactorsforthechangeinCD4+ count of HIV-positive patients.

**2. Methodology:**

**2.1 Source Data:**

The data for this study obtained from Debre Berhan Referral Hospital, ART clinic by extracting from HIV-patient cards. This retrospective cohort study was used data of HIV/AIDS patients who initiated on Antiretroviral Therapy in the ART clinic of DBRH, Debre Berhan, Ethiopia during a period of September 11, 2013 to March 10, 2016 following up through the ART routine register records up to February 8, 2019. The study population was HIV- positive adults whose age was 16 years old and above initiated on ART treatment in the hospital. All patients who have initiated to ART and measured their CD4+ count at least two times including baseline and those whom started first line ART regimen class were included in the study. Patients whose age below 16 years and those who started ART before September 11, 2013 and after March 10, 2016 were excluded in the study and CD4+ count was taking approximately in every six months.

**2.2 The Study Variables:**

The response variable for this study is CD4+ count for each individual measured approximately in every six months interval. The CD4+ count is a continuous variable that measured on individual patients and expressed as cells per cubic millimeter (cells/mm3) of blood (Diggle et al, 2002).

The independent variables are baseline age of patients, baseline CD4+ count, observation time (in months), sex of the patients, marital status of patients at baseline, WHO clinical Stage, regimen class of patients, body mass

index (BMI) at baseline, functional status of the patients and educational level of patients.

Table 2.1: Description of the study variables and codes in the analysis

| No | Variables | Code |
|---|---|---|
| 1 | Sex of the patients | 0 = Male, 1 = Female |
| 2 | Age of the patients (years) | Continuous |
| 3 | Marital status | 0 = Married, 1 = never married, 2 = divorce, 3 = separated and widowed |
| 4 | Body mass index (kg/m2) | Continuous |
| 5 | Base line CD4+ count | Count |
| 6 | CD4+ count at each time | |
| 7 | WHO clinical stage | 1 = stage I, 2 = stage II, 3 = stage III, 4 = stage IV |
| 8 | CD4+recorded time (months) | Continuous |
| 9 | Regimen class of patients | 0 = AZT-3TC-EFV, 1 = AZT-3TC-NVP, 2 = TDF-3TC-NVP, 3 = TDF-3TC-EFV, 4 = others |
| 10 | Functional status of patients | 0 = Working, 1 = Ambulatory, 2 = Bedridden |
| 11 | level of education | 0 = No Education, 1 = Primary, 2 = Secondary, 3 = Tertiary |

**2.3 Statistical data Analysis:**

The study has been used exploratory data analysis to discover as much of the information regarding raw data as possible, plotting individual curves to carefully determine the data should be performed first before any formal model fitting is carried out. Thus, individual profiles plot were used to observe subject specific evolution over time and to decide on the random effects to be included in the model. To choose the fixed effects of the model mean profile plot is explored and to study the possible differences between groups, plotting the mean profiles for each subgroup separately is done. Finally, variance structure is plotted to explore variations of CD4+ count over time from baseline values.

**2.3.1 Linear Mixed Effects Model:**

The main feature of linear mixed effects model is that the mean response modeled as combination of population characteristics, that are assumed to be shared by all individuals, and subject-specific effects that are unique to a particular individual. The former is referred to as fixed effects, while the later referred as random effects Larid and Ware (2004).

Thus, the linear mixed effects model can be written as:

$$Y_i = X_i^{'}\beta + Z_i^{'}b_i + \epsilon_i \quad \text{_____} \quad (1)$$

where β is a $p \times 1$ vector of fixed effects, $b_i$ is a $q \times 1$ vector of random effects, $X_i$ is $n_i \times p$ matrix of covariates for fixed effect s, $Z_i$ is matrix of covariates for random effects with $q \leq p$ and $e_i's$ are regarded as measurement errors. The random effect s$b_i$, and the error term$\epsilon_i$, assumed to have multivariate normal distribution with$E\begin{bmatrix} b_i \\ \epsilon_i \end{bmatrix}$=$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $Cov\begin{bmatrix} b_i \\ \epsilon_i \end{bmatrix} = \begin{bmatrix} D \\ R \end{bmatrix}$, where, D is between-subject covariance matrix and R is within-subject covariance matrix.

The form of induced random effects covariance structure in linear mixed model, first distinguish the conditional mean of $Y_i$ given random effects$b_i$,

$$E(Y_i|b_i) = X_i^{'}\beta + \mathrm{Z_i^{'}b_i}$$

From the marginal or population averaged mean of$Y_i$, when averaged over the distribution of random effects $b_i$, is$E(Y_i) = \mu_i = X_i^{'}\beta$.

In a similar way, we can distinguish between conditional and marginal covariance. The conditional covariance of $Y_i$, given $b_i$, is $Cov(Y_i|b_i) = Cov(e_i) = R_i$, while the marginal covariance of $Y_i$, averaged over the distribution of $b_i$, is $Cov(Y_i) = Z_i D Z_i^{'} + R_i$.

**2.3.2 Parameter Estimation in Linear Mixed Models:**

Linear mixed models are likelihood based models then; maximum likelihood (ML) and/or restricted maximum likelihood (REML) were used for parameter estimation. The difference in two methods is in construction of the likelihood function. However, the two estimations are asymptotically equivalent and often give very similar results. The distinction between ML and REML becomes important only when the number of fixed effects is relatively large. After the appropriate covariance structure is selected, model building efforts should be directed at simplifying the mean structure of the model.

**2.3.3 Model Selection Criteria and Model Adequacy:**

In this study Akaike's information criterion (AIC), Bayesian information Criterion (BIC) and Likelihood ratio test were used to select the best fitted model. Model diagnostics are especially important in linear mixed models because likelihood based estimation methods are particularly sensitive to unusual observations. Thus, residual plots are used to check normality of these effects and to identify any outlying effect categories. And the normality for the within-group error is assessed with the normal probability plot of the

residuals by covariates. Statistical Analysis System (SAS) version 9.4 was used for statistical analysis. Five percent level of significance is considered for statistical tests.

**3. Results:**

Among the enrolled subjects, each of 322 patients was followed for six occasions with six month interval, and then finally there were 1932 observations involved.

**3.1 Descriptive Summary:**

Table 3.1: Socio-Demographic characteristics of patients on ART

| Variables | Categories | Frequency | Percentage |
|---|---|---|---|
| Sex | Males | 97 | 30.12 |
| | Females | 225 | 69.88 |
| Marital Status | Married | 155 | 48.14 |
| | Never married | 66 | 20.50 |
| | Divorce | 45 | 13.98 |
| | Separated and widowed | 56 | 17.39 |
| Educational Level | No education | 90 | 27.95 |
| | Primary | 132 | 40.99 |
| | Secondary | 70 | 21.74 |
| | Tertiary | 30 | 9.32 |

From the Table 3.1, of the total 322 patients who were included in the study, 225 (69.88%) were females and the remaining 97 (30.12%) were males. Regarding the marital status composition of patients, of the total of 322 HIV-infected patients in the study 155 (48.14%) were married, 66 (20.50%) were never married, 45 (13.98%) were divorced and 56(17.39%) were others respectively. When we look at the educational level categories of the HIV-positive patients, 90(27.95%) were no educated, 132 (40.99%) were primary school educated and the remaining 31.06 % were secondary and tertiary educated.

Table 3.2: Baseline Clinical characteristics of patients on ART

| Variables | | Frequency | Percent |
|---|---|---|---|
| Functional Status | Working | 296 | 91.93 |
| | Ambulatory | 23 | 7.14 |
| | Bedridden | 3 | 0.93 |
| WHO Clinical Stage | Stage I | 141 | 43.79 |
| | Stage II | 72 | 22.36 |
| | Stage III | 98 | 30.43 |
| | Stage IV | 11 | 3.42 |
| Regimen Class | AZT-3TC-EFV | 22 | 6.83 |
| | AZT-3TC-NVP | 36 | 11.18 |
| | TDF-3TC-NVP | 27 | 8.39 |
| | TDF-3TC-EFV | 215 | 66.77 |
| | Others | 22 | 6.83 |

From the Table 3.2among total HIV-positive patients who have included in the study 296 (91.93%) were in 'working' functional status, whereas the remaining 8.07%less than one-tenth were in ambulatory and bedridden status. Table 3.2 also depicts, 141 (43.79%) of the study patients were in WHO clinical stage I, 72(22.36%) were found in stage II, 98(30.43%) were in stage III and only 3.42% of patents were in stage IV. Regarding the composition of baseline ART regimen classes, 22(6.83%) of the patients were started by AZT-3TC-EFV class of ART regimen, 36(11.18%) were by AZT-3TC-NVP ART regimen type, 27(8.39%) were by TDF-3TC-NVP, 215(66.77%) were by TDF-3TC-EFV class of ART regimen and the remaining 22(6.83%) were initiated on other different ART regimen classes.

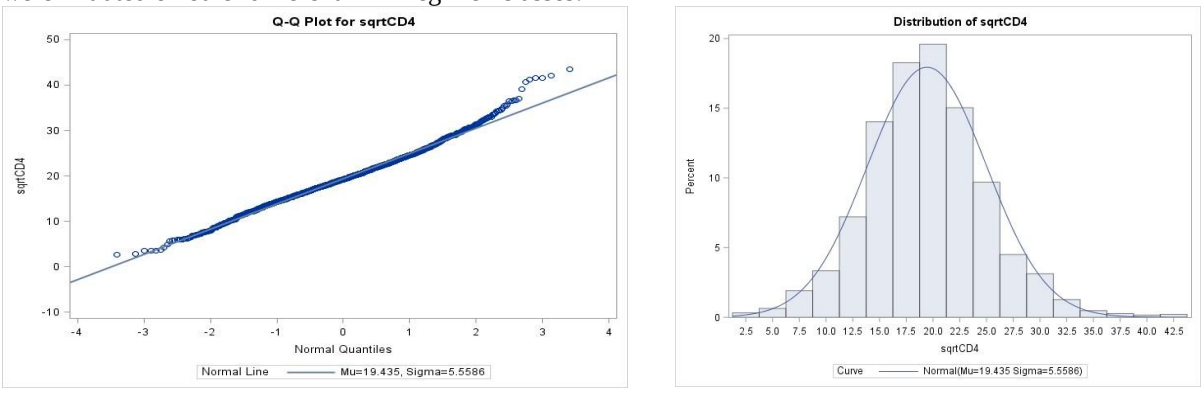



Figure 3.1: QQ-plot and Histogram for sqrtCD4+ count

The CD4+ count data was checked for normality assumption by using QQ-plot, histogram and Shapiro-wilk test. Since the actual CD4+ count was not normally distributed and then possible transformation were to be carried.

Then after taking log transformation and square root transformation methods, the data become approximately normal by square root transformation as shown in Figure 3.1 and Shapiro-Wilk test of normality in Table 3.3.

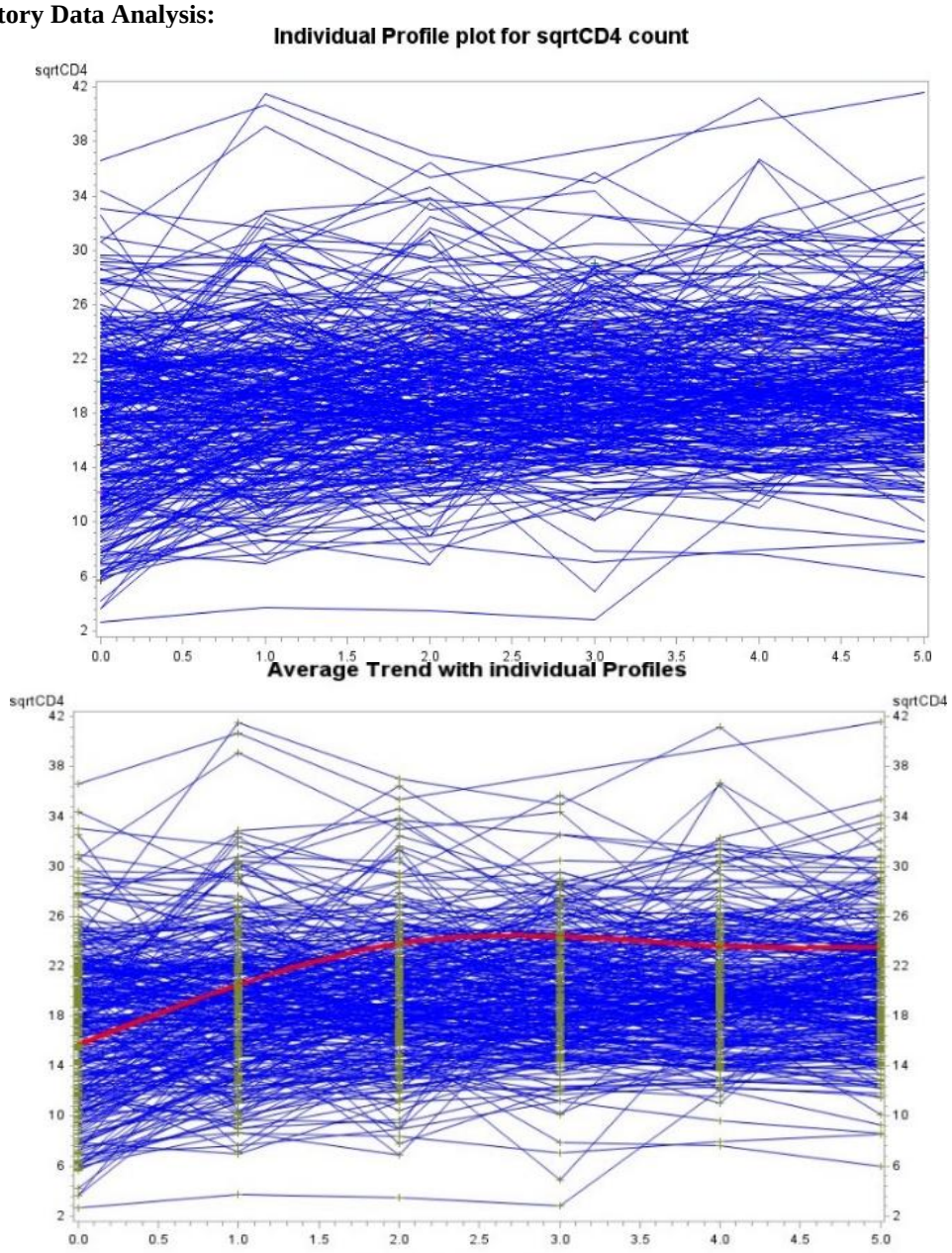
Table 3.3: Test of Normality

| | **Actual CD4+** | | **Log transformed CD4** | | **Sqrt CD4 transformed** | |
|---|---|---|---|---|---|---|
| Test | statistic | p-value | statistic | p-value | statistic | p-value |
| Shapiro-Wilk | 0.91258 | <0.0001** | 0.938595 | <0.0001** | 0.991197 | 0.432 |
| Kolmogorov-Smirnov | 0.084069 | <0.0100** | 0.074793 | <0.0100** | 0.030059 | 0.273 |

** P-value <0.0001

In Table 3.3 square root transformation test of normality has estimated with insignificant p-values in both Shaprio-Wilk and Kolmogorov-Smirnov tests. This implied that the actual and log transformed data were not normal while after square root transformation taken the data is almost normally distributed.

**3.2 Exploratory Data Analysis:**




3.2 (a) - Individual profile plot & 3.2 (b) - Mean profile plot

Figure 3.2: Individual and mean profile plot for sqrt CD4+ count

From Figure 3.2 (a), the individual profile plot showed that the subject-specific variations in CD4+ count over time. But the plot didn't essentially identify how the evolution of each subject seems over time

because of the number of observations' was large. In this case, a more meaningful plot is an overlay plot of the individual profiles and the average trend Thus, the smoothed cure in Figure 3.2(b) suggested that the average plot of square root CD4+ count has linear relationship with time. It indicates the square root CD4+ count increases up to twelfth months but the rate of increment become low after twelfth measurement occasions and then fairly constant around 24$^{th}$ month in square root CD4+ count value.

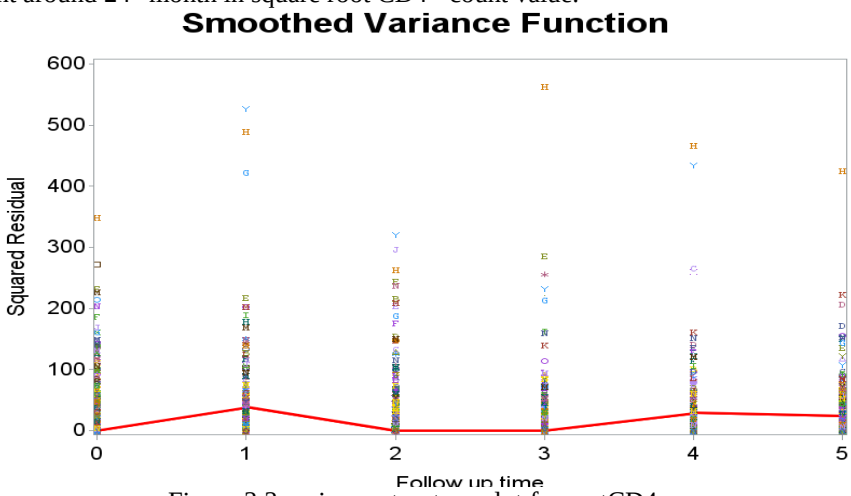
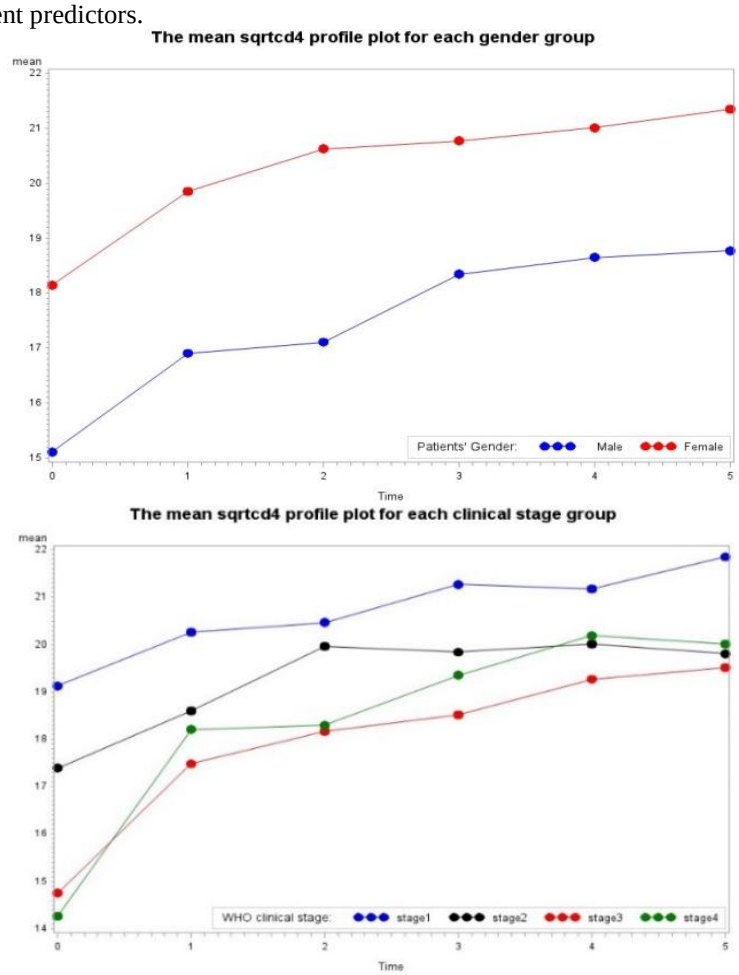



Figure 3.3 variance structure plot for sqrtCD4

Figure 3.3 explores the variability of patients' CD4+ count among measurement time-pints, and it shows the observed variance function changing across time. But the variability clearly occurred up to twelfth month from baseline time point and after eighteenth, while the variability between twelfth and eighteenth months is almost close to zero. Thus, other random effects in addition to the random intercepts would be needed in modeling patients' CD4+ count overtime. Generally, the variance of patients' CD4+ count fluctuates from time to time in different predictors.




3.4 (a) Mean profile plot for gender group & 3.4 (b) Mean profile plot for each WHO stage

Figure 3.4: Mean profile plot for gender and WHO clinical stage

Figure 3.4 indicates the degree of heterogeneity across patients in the rate of CD4+ count progression among categories of some predictors, here for example sex and WHO clinical stages. Figure 3.4(a) shows variations in average increment of CD4 count between males and females, thus female patients have higher CD4+ count change overtime. Here the effect of gender on evolution depends on the patients' baseline CD4+ counts, which implies female patients with high baseline CD4+ count would have higher rate of change over time. The plot for WHO clinical stages in Figure 3.4(b), indicates that the HIV-positive patients in stage one have higher increment rate in their CD4+ count as compared to patients in other stages; whereas patients in stages II, III and IV after eighteenth month measurement time points, the change in their CD4+ count is almost similar.

**3.3 Linear Mixed Models:**

Table 3.4 shows that comparison between -2res LL, AIC and BIC on three different covariance structures and then after unstructured covariance structure is selected due its smallest -2res LL, AIC and BIC values.

Table 3.4: comparing different covariance structures

| | **Exchangeable** | **Autoregressive (1)** | **Unstructured** |
|---|---|---|---|
| -2 Res Log Likelihood | 10319.8 | 10319.8 | 10318.7 |
| AIC (Smaller is Better) | 10323.8 | 10326.7 | 10323.8 |
| AICC (Smaller is Better) | 10323.8 | 10326.7 | 10323.8 |
| BIC (Smaller is Better) | 10331.4 | 10341.8 | 10331.4 |

Table 3.5: Variance component models with random intercept and random slope effects

| Covariance Parameter | | Subject | Estimate | Standard Error | Z Value | Pr. Z | 95% CI |
|---|---|---|---|---|---|---|---|
| Model 1 | Intercept | ID | 6.6519 | 0.7118 | 9.35 | <.0001** | 5.4508, 8.3011 |
| | Residual | | 11.8664 | 0.4269 | 27.80 | <.0001** | 11.0722, 12.7496 |
| Model 2 | Slope | ID | 1.1009 | 0.1136 | 9.69 | <.0001** | 0.9082,1.3624 |
| | Residual | | 10.1165 | 0.3688 | 27.43 | <.0001** | 9.4309, 10.8801 |

* P-value <0.05, ** P-value <0.0001

Table 3.5 indicates the separate estimated variance components for random intercept in Model 1andrandom slope in Model 2. The random intercept model partitions the total variation in the data into within-individual and between-individual components with its estimates 11.8664(p-value, 0.0001) and 6.6519(p-value, 0.0001) respectively. Similarly the variance components for random slope model are also shown in Model 2. From the two estimated models shown in the Table 3.5, the significant value for the within individual variation suggests that the data structure is best captured by using random effects model.

**3.3.1 Covariance Parameter Estimates**

Table 3.6: Covariance Parameter Estimates

| Covariance Parameter | Subject | Estimate | Standard Error | Z Value | Pr Z |
|---|---|---|---|---|---|
| UN(1,1) | ID | 1.2696 | 0.5683 | 2.23 | 0.0127* |
| UN(2,1) | ID | 0.6621 | 0.2083 | 3.18 | 0.0015* |
| UN(2,2) | ID | 0.6971 | 0.1065 | 6.55 | <.0001** |
| Residual | | 9.5860 | 0.3839 | 24.97 | <.0001** |

* P-value <0.05, ** P-value <0.0001

From Table 3.6, variance-covariance of the random effects and residual estimations are shown. The variance of random effects,$\boldsymbol{b_i}$,var($\boldsymbol{b_{1i}}$) = 1.2696, var($\boldsymbol{b_{2i}}$) = 0.6971, cov($\boldsymbol{b_{1i}}, \boldsymbol{b_{2i}}$) = 0.6621, and residual variance; var($\boldsymbol{\epsilon_{ij}}$) = 9.5860. Both the variances of intercepts and linear effects of time were significantly different from zero; which indicated that the CD4+ counts at baseline were vary across subjects and the change of CD4+ counts over time vary within subjects. The estimated total variability between subjects obtained from the summation of variances and covariance of random effects was 2.6288, whereas the total variability within individual was 9.5860. However; the total variation in square root CD4+ count was estimated to be 2.6288+9.5860= 12.2148.

Based on the intra-class correlation which is the proportion of total variability attributed to between individual variations in their square root CD4+ count was 0.2152. While the proportion of total variability that is attributed to within-patient variations obtained was 0.7848; this indicating that 78.48% of the variation in the data is accounted by within- subject variability; and the remaining 21.52% of the total variability was attributed by between subject variations.

$$\boldsymbol{D} = \begin{bmatrix} \mathbf{1.2696} & \mathbf{0.6621} \\ \mathbf{0.6621} & \mathbf{0.6971} \end{bmatrix}$$

An additional piece of information that can be captured from models that contain both random slopes and intercepts is the correlation between the random effects. From the matrix, the value 1.2696 indicates the variance of random intercepts; the value 0.6971 represents the variance of slopes for time effects and 0.6621 is the covariance of random intercept and slopes for time. The value 0.6621 indicates that a strong positive

correlation between the slope and the intercept. This means that individuals who have higher baseline CD4+ count tend to have higher rates of change over time, and individuals who have lower baseline CD4+ count tend to have lower rates of change over time.

**3.3.2 Random Intercept and Slope Model:**

Table 3.7: Fixed Effect Estimates in linear mixed model

| Effect | Estimate | Std. Error | DF | t Value | Pr> \|t\| |
|---|---|---|---|---|---|
| Intercept | 11.1556 | 2.3382 | 307 | 4.77 | <.0001** |
| Sex (female) | 0.7602 | 0.3303 | 1239 | 2.30 | 0.0215* |
| Sex (male) Reference | . | . | . | . | . |
| Age | -0.01852 | 0.01517 | 1239 | -1.22 | 0.0225* |
| BMI | 0.01026 | 0.05425 | 1239 | -0.19 | 0.0500* |
| Functional Status (working) | 0.6362 | 1.5111 | 1239 | 0.42 | 0.0138* |
| Functional Status (ambulatory) | 0.1817 | 1.5822 | 1239 | 0.11 | 0.9086 |
| Functional Status (bedridden) Reference | . | . | . | . | . |
| Marital Status (married) | 0.5163 | 0.4142 | 1239 | 1.25 | 0.0128* |
| Marital Status (never married) | 1.0629 | 0.5108 | 1239 | 2.08 | 0.0377* |
| Marital Status (divorced) | 1.0305 | 0.5237 | 1239 | 1.97 | 0.0693 |
| Marital Status (Separated and widowed) Reference | . | . | . | . | . |
| Regimen Class (AZT-3TC-EFV) | -0.6537 | 0.7987 | 1239 | -0.82 | 0.4132 |
| Regimen Class (AZT-3TC-NVP) | -0.04678 | 0.7348 | 1239 | -0.06 | 0.9492 |
| Regimen Class (TDF-3TC-NVP) | -1.2520 | 0.7705 | 1239 | -1.62 | 0.0044* |
| Regimen Class (TDF-3TC-EFV) | -0.8371 | 0.6121 | 1239 | -1.37 | 0.0217* |
| Regimen Class (others)Reference | . | . | . | . | . |
| Educational Level (no educated) | 0.4264 | 0.5557 | 1239 | 0.77 | 0.4430 |
| Educational Level (primary) | 0.3048 | 0.5255 | 1239 | 0.58 | 0.5621 |
| Educational Level (secondary) | 0.2923 | 0.5688 | 1239 | 0.51 | 0.6074 |
| Educational Level (tertiary)Reference | . | . | . | . | . |
| WHO Clinical Stage (stage I) | -0.1735 | 0.8330 | 1239 | -0.21 | 0.8350 |
| WHO Clinical Stage (stage II) | 0.05383 | 0.8687 | 1239 | 0.06 | 0.0306* |
| WHO Clinical Stage (stage III) | -0.5137 | 0.8418 | 1239 | -0.61 | 0.5418 |
| WHO Clinical Stage (stage IV)Reference | . | . | . | . | . |
| BaseCD4 | 0.01952 | 0.000665 | 1239 | 29.36 | <.0001* |
| Time | 2.2509 | 1.0388 | 305 | 2.17 | 0.0310* |
| Time*Time | -0.1493 | 0.02937 | 1541 | -5.08 | <.0001** |
| Age*Time | -0.01698 | 0.006513 | 1239 | -2.61 | 0.0092* |
| Time*Functional Status (working) | -1.1642 | 0.6720 | 1239 | -1.73 | 0.0434* |
| Time*Functional Status (ambulatory) | -0.9880 | 0.7043 | 1239 | -1.40 | 0.0309* |
| Time*Functional Status (bedridden) Reference | . | . | . | . | . |
| Time* Marital Status (married) | 0.4257 | 0.1843 | 1239 | -2.31 | 0.0211* |
| Time* Marital Status (never married) | 0.3034 | 0.2268 | 1239 | -1.34 | 0.1812 |
| Time* Marital Status (divorced) | 0.4366 | 0.2335 | 1239 | -1.87 | 0.0617 |
| Time* Marital Status (others) Reference | . | . | . | . | . |
| Time* WHO Clinical Stage (stage I) | 0.3958 | 0.3682 | 1239 | -1.08 | 0.0425* |
| Time* WHO Clinical Stage (stage II) | -0.4545 | 0.3861 | 1239 | -1.18 | 0.0594 |
| Time* WHO Clinical Stage (stage III) | 0.02136 | 0.3749 | 1239 | 0.06 | 0.0346* |
| Time* WHO Clinical Stage (stage IV) Reference | . | . | . | . | . |

* P-value <0.05, ** P-value <0.0001

From Table 3.7 the fixed effect parameters estimates in linear mixed effect model are shown and parameters have subject specific interpretations unlike marginal model. Thus, given the random effects$\boldsymbol{b_{1i}}$; the parameter estimate for intercept is, $\boldsymbol{b_{0i}}$ =11.1556. And the square of $\boldsymbol{b_{0i}}$ is 124.447 which indicates that the intercept of $\boldsymbol{i^{th}}$ patient deviated from population intercept$\boldsymbol{\beta_0}$.

The average CD4+ count for each female patients' was 0.5779 times higher than that of male individuals' given baseline random effects (β=0.7602, p= 0.0215). When a unit increase in BMI results an increase of mean change CD4+ count of patients by the square of 0.01026 (β=0.01026, p=0.0500). The estimate of married patients under ART is (β=0.5163, p-value=0.0128; which indicates married patients were 0.2665 times higher in average CD4+ count than those who were separated and widowed, keeping the effect of other variables. HIV-positive patients who were never married had square of 1.0692 times higher expected CD4+ count than patients who were separated and widowed fixing the effect of other variables (β=1.06292, p-value, 0.0377).

The coefficient of time is 2.2509 (p -value,0.0310), implying that the average CD4+ count increases

5.067 times higher for i[th] female individuals over time when the effect of other variables are hold constant.
In Table 3.7 age had negative time interaction effect(β=-0.01698, p= 0.0092); this implies that older patients had square of 0.01698 times lower mean CD4+ count than younger patients over time. Patient's working status also has significant interaction effect with time, its estimate -1.1642 indicates that the average CD4+ count of working patients was 1.3554 times lower compared to bedridden patients average CD4 count with time interaction effect.

$$\sqrt{\text{CD4ij}} = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 BMI_i + \beta_4 FuncSt_{0i} + \beta_5 MarSt_{0i} + \beta_6 MarSt_{1i} + \beta_7 RegCl_{2i} + \beta_8 RegCl_{3i} + \beta_9 WHOSt_{2i} + \beta_{10} BaseCD4 + \beta_{11} Time_i + (\beta_{12} Age_i + \beta_{13} FuncSt_{0i} + \beta_{14} FuncSt_{1i} + \beta_{15} MarSt_{0i} + \beta_{16} WHOSt_{0i} + \beta_{16} WHOSt_{3i}) t_{ij} + b_{0i} +, \boldsymbol{b_{1i}} t_{ij};$$

*where,* $FuncSt_{0i}, FuncSt_{1i}, MarSt_{0i}, MarSt_{1i}, RegCl_{2i}, RegCl_{3i}, WHOSt_{2i}$, are patients' functional status: working, ambulatory; marital status: married, never married; regimen class: TDF-3TC-NVP, TDF-3TC-EFV and second WHO clinical stage.

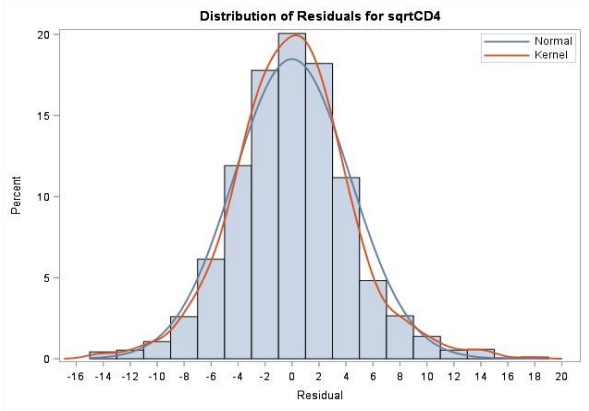
**3.4 Model Diagnosis:**

From Table 3.8, the Shapiro-Wilk test is not significant at 5% level of significance. Thus, there is insufficient evidence to conclude that the model residuals are not normally distributed.

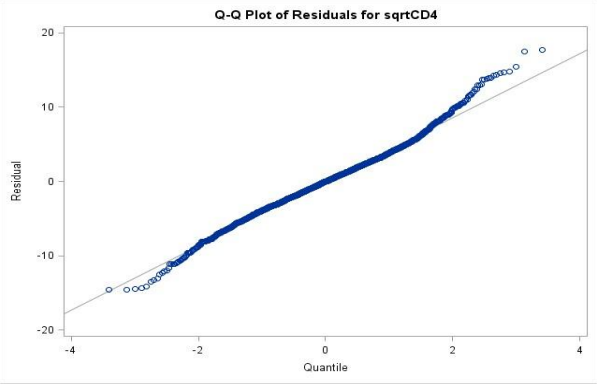
Table 3.8: Shapiro-Wilk normality test Tests for Normality

| Test | Statistic | p Value |
|---|---|---|
| Shapiro-Wilk | 0.989089 | <0.8431 |

Figure 3.5 displays histogram and QQ-plots for residual respectively suggested that the residuals for square root CD4 count of patients were normally distributed.




(a) Residual Histogram




(b) Residual QQ-plot

Figure 3.5 Normality of residual in linear mixed model

**4. Discussions:**

This study was aimed to investigate the progression of CD4+ count among patients under ART in Debre Berhan Referral Hospital. It was also aimed in determining whether the change in CD4+ count differs among selected patient characteristics. On the basis of those objectives, exploratory analyses which provide some initial basis followed by subsequent model-based results were conducted.

The normality of CD4+ count had been checked by using Q-Q plots, histogram and Shapiro-Wilk test before detail analyses. These plots indicated that CD4+ count needed to be transformed; and then after the square root transformation of the CD4+ count was found to be normal.

From exploratory analysis, the mean profile plot suggested that average CD4+ count increases up to twelfth month and then fairly shows constant increment rate over time. This supports the result of (Abrogoua, et al, 2012) which identified that after patients initiated to ART their CD4+ count increases up to third time points from baseline. We also observed that female HIV-positive patients under ART had higher mean change CD4+ count compared to males.

Unstructured covariance structure was selected after considering all possible covariance structures in their -2logLL, AIC and BIC. In conducting comparison of different variance components, the random intercept and random slope models were selected as final random effect model; this is based on the result of intra-class correlation that total variability attributed to each random model. The result of covariance parameter estimate suggested that subject-specific variability was significant indicating the validity of the heterogeneity of variances.

The result of this study showed that females have higher mean change rate of CD4+ count than males over time, with significant p-value 0.0215. This is in contrary with study by (Gezie et al, 2016), which found that when adjusted for other variables, sex was not found to be significant predictor of CD4+ count. However, this was in agreement with other study by (Asfaw et al, 2015), which showed that females had better response to ART as compared to males. Patients with higher baseline CD4 count tend to have higher rate of change over

time. This was in agreement with the result of study by Smith et al, 2004, which showed that patients with higher pre-ART CD4 counts were found to be associated with smaller long-term increases in CD4 counts during ART, which may reflect a greater scope for improvement among those with lower pre-ART CD4 cell counts.
Baseline age was negatively associated with CD4+ count. This is in line with the study by (Gezie et al, 2016). The finding of this study also showed that married and never married patients, respectively, were square of 0.5163 and 1.0629 times higher in mean CD4+ count compared to those with widowed and separated patients over time. In contrary to the study with (Assefa et al, 2015) which reported that the evolution CD4 count described by second order function of time, this study shows the change in CD4 count is negatively associated with quadratic function of time.

In co-variance parameter estimates, both the variances of intercepts and linear effects of time were significantly different from zero; which indicated that the CD4+ cell counts at baseline vary across subjects and the change of CD4+ cell counts over time vary within subjects. Age with time interaction was found to have significant negative effect implying that, the rate of progression in square root CD4+ count for older patients was lower than that of younger ones showing the same conclusion as study by (Picat et al, 2013).

The result of this study revealed that patients with working and ambulatory status were square of -1.1642 and -0.9880 respectively, times lower in average CD4+ count than patients with bedridden functional status. This is contrary to study by (Kebede et al, 2015) which suggested that patients who were bedridden when starting ART found to experience low CD4 count change compared to patients with working status.

## 5. Recommendations and Conclusion:

### 5.1 Recommendations:

- The HIV-positive patients might be advised to start ART treatment as early as possible, as the result of this study showed that patients who have higher baseline CD4 count tend to have higher rate of change in their CD4 count over time.
- The result indicates that male patients and patients who had lower baseline CD4 count had experienced less mean change CD4 count over time. Thus, patients with such characteristics need special guidance and due attention.
- This study suggested that the progression of patients' CD4 count using linear mixed effects model is advisable other than linear models the reason is to thatit can be incorporate within and between subject variations simultaneously.

### 5.2 Conclusions:

The study revealed that after initiation of the ART, the average baseline CD4+ count of HIV-positive was 335.7 and this changed to 408.6 overtime. The result of the study also identified variables which had significant effect on patients' CD4+ count progression those were included in the study. Thus, educational level did not have significant effect on CD4+ count progression of patients over time. Whereas, baseline CD4+ count, functional status, treatment regimen class, age, gender, marital status and WHO clinical stages had significant effect on the CD4+ count progression of patients.

From the result we observed that the female patients had higher mean increment on CD4+ count than male patients in the follow-up time. The evolution depends on baseline CD4+ count values and hence we concluded that patients with high baseline CD4+ would have higher rate of change in their average CD4+ count over time. Patients, who were bedridden, had higher mean change CD4+ count compared to ambulatory and working patients.

The result of this study also revealed the time interaction effects among categories of some predictor variables such as functional status, WHO clinical stage and marital status were significantly differ in mean change in CD4+ count of HIV-positive patients. And age had negative time interaction effect, this indicates that older patients were associated with lower average increment no their CD4+ count over a specified period of time.

## 6. References:

1. Abrogoua, D.P., Kablan, B.J., Kamenan, B.A.T., Aulagner, G., N'Guessan, K. and Zohoré, C.(2012). Assessment of the impact of adherence and other predictors during HAART on various CD4 cell responses in resource-limited settings. Patient preference and adherence, 6, p.227.
2. Asfaw, A., Ali, D., Tadele Eticha, A.A., Alemayehu, M. and Kindeya, F. (2015). CD4 cell count trends after commencement of antiretroviral therapy among HIV-infected patients in Tigray, Northern Ethiopia: a retrospective cross-sectional study. PloS one, 10(3).
3. Assefa, A.T. (2015). Modelling the Evolution of CD4+ Cell Counts and Hemoglobin Concentration Level for HIV-1 Patients on Antiretroviral Therapy (ART) in Mildmay Uganda (Master's thesis).
4. Berzuini, C. and Allemani, C. (2004). Effectiveness of potent antiretroviral therapy on progression of human immunodeficiency virus: Bayesian modeling and model checking via counter factual replicates. Journal of the Royal Statistical Society: Series C (Applied Statistics), 53(4), pp.633-650.
5. Diggle PJ, Heagerty P, Liang K-Y., (2002), Analysis of Longitudinal Data. 2nd ed. Oxford, United Kingdom: Oxford University Press.

6. Eyster, E. Gail, M. and Ballard, J. (1987). Natural History of Human Immunodeficiency Virus Infection in Hemophiliacs. Effects of T-Cell Subsets, Platelet Counts, and Age, Annals of Internal Medicine, 107, pp.1-6.
7. Gezie, L.D. (2016). Predictors of CD4 count over time among HIV patients initiated ART in Felege Hiwot Referral Hospital, northwest Ethiopia: multilevel analysis. BMC research notes, 9(1), p.377.
8. Hoffman, J., Van Griensven, J., Colebunders, R. and McKellar, M. (2010). Role of the CD4 count in HIV management. HIV therapy, 4(1), pp.27-39.
9. Kebede, M.M., Zegeye, D.T. and Zeleke, B.M. (2015). Predictors of CD4 Count Changes after Initiation of Antiretroviral Treatment in University of Gondar Hospital, Gondar in Ethiopia. Clin Res HIV/AIDS, *1*(2), pp.1-15.
10. Laird, N.M. and Ware, J.H. (2004). Random-effects models for longitudinal data. Biometrics, 38(4), pp.963-974.
11. Langford, S.E., Ananworanich, J. and Cooper, D.A. (2007). Predictors of disease progression in HIV infection: a review. AIDS research and therapy, 4(1), p.11.
12. McCullagh, P. and Nelder, J. (1989). Generalized Linear Models, (2nd ed.). London: 0Chapman and Hall. Pp.150-30.
13. Picat, M.Q., Lewis, J., Musiime, V., Prendergast, A., Nathoo, K., Kekitiinwa, A., Ntege, P.N., Gibb, D.M., Thiebaut, R., Walker, A.S. and Klein, N. (2013). Predicting patterns of long-term CD4 reconstitution in HIV-infected children starting antiretroviral therapy in sub-Saharan Africa: a cohort-based modelling study. PLoS medicine, 10(10), p.e1001542.
14. Smith, C.J., Sabin, C.A., Youle, M.S., Kinloch-de Loes, S., Lampe, F.C., Madge, S., Cropley, I., Johnson, M.A. and Phillips, A.N. (2004). Factors influencing increases in CD4 cell counts of HIV-positive persons receiving long-term highly active antiretroviral therapy. The Journal of infectious diseases, 190(10), pp.1860-1868.
15. Twisk, J.W. (2013). Applied longitudinal data analysis for epidemiology: a practical guide. cambridge university press, New York.