

Research Article

Linear Mixed Modeling of CD4 Cell Counts of HIV-Infected Children Treated with Antiretroviral Therapy

Belay Belete Anjullo  and **Derbachew Asfaw Teni** 

Department of Statistics, Arba Minch University, Arba Minch, Ethiopia

Correspondence should be addressed to Belay Belete Anjullo; belay.belete@amu.edu.et and Derbachew Asfaw Teni; derbachew.asfaw@amu.edu.et

Received 23 September 2020; Revised 7 January 2021; Accepted 15 January 2021; Published 29 January 2021

Academic Editor: Hamidreza Karimi-Sari

Copyright © 2021 Belay Belete Anjullo and Derbachew Asfaw Teni. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Human immunodeficiency virus (HIV) is a major health problem in the world, and failure to implement prevention programs results in an increased number of infections among newborns. The goal of this study was to investigate the evolution and determinants of cluster of differentiation four (CD4) cell count among HIV-infected children who were under antiretroviral therapy (ART). **Methods.** We follow up a cohort of 201 children aged under fifteen years from October 2013 to March 2017 at Adama Hospital in Ethiopia. To get insight into the data, exploratory data analysis was performed on the change in the longitudinal CD4 cell count. **Results.** At the baseline, the average number of CD4 cell counts was 468.5 cells/mm³ of blood with a standard deviation of 319.11 cells/mm³. Here, we employed the random intercept and the random slope linear mixed-effects model to analyze the data. Among predictor variables, observation time, baseline age, WHO clinical stage, the history of tuberculosis (TB), and functional status were determinant factors for the mean change in the square root of the CD4 cell count. **Conclusions.** The finding revealed that the change in the square root of the CD4 cell count increases with an increment of age at diagnosis. Regarding WHO clinical stages of patients, those who were in stage III and stage IV of the HIV/AIDs disease stages relatively had lower CD4 cell counts than stage I patients. This shows the change in the square root of CD4 cell counts of stage III and stage IV patients was 6.43 and 9.28 times lower than stage I patients, respectively. Similarly, we noticed that observation time, the history of TB, and functional status were significantly associated with the mean change in the square root of the CD4 cell count.

1. Introduction

HIV is a major health problem in the world, and failure to implement prevention programs results in an increased number of infections among newborns. HIV-infected children should start ART to reduce AIDs-related morbidity and mortality or to improve their survival time [1]. According to United States Agency for International Development (UNAIDS) [2], an estimated 1.8 million children globally were living with HIV of which 1.18 million are in sub-Saharan Africa. As per the same estimate, there are 180,000 new infected children globally with an estimate of 108,000 (60%) occurring in sub-Saharan Africa, and 111,000 children died due to AIDs and related illnesses globally followed by 72,000 (65%) children in sub-Saharan Africa. In

Ethiopia, an estimate of 729,089 people live with HIV including 80,923 children less than 15 years. As per the same estimate, there are 21,606 new infections of which 1,276 (5.9%) are children under 15 years. Furthermore, the number of deaths due to AIDs-related illnesses for the same period was estimated to be 10,960 in the country and 1,924 (2.4%) were children less than 15 years [3].

As studies reported, children experience more rapid HIV disease progression making them highly susceptible to opportunistic infections and death [4]. Antiretroviral therapy (ART) can restore immune function and has enormously reduced morbidity and mortality among HIV-infected children. Due to the advent of ART, many HIV-infected children can survive to adolescence and adulthood [5]. Current revised WHO guideline recommends that all

HIV-infected children should initiate ART irrespective of the clinical disease stage or degree of immune suppression [6]. Hence, one of the main interests in HIV clinical studies is the change in CD4 cell counts of patients who are receiving ART. The statistical modeling has greatly contributed to identifying the predictors related to the change in the CD4 cell count of patients initiating ART. The objective of this study was to investigate the evolution and determinants of the CD4 cell count among HIV-infected children initiating ART at Adama Hospital in Ethiopia.

2. Cohort-Based Data

The data used in this study came from a total of 591 children aged under 15 years old cohort-based retrospective study by reviewing patients' ART charts and electronic databases at Adama Hospital from 2013 to 2017 in Ethiopia [7]. A sample of 201 children who have full records or a complete history during the study period was considered in this study, and the data were retrieved by physicians who were working in the hospital. Those who had no full records or an incomplete history during the study period were excluded from the analysis. Here, we distinguished cause-effect specific related factors from various literature reviews. Thus, the change in the longitudinal CD4 cell count per mm^3 of HIV-infected children treated with ART was considered as a response variable of the study and measured every six months, at the baseline (first diagnosis), during the first visit (after 6 months), second visit (after 12 months), third visit (after 18 months), fourth visit (after 24 months), fifth visit (after 30 months), and finally at the sixth visit (after 36 months). The predictor variables considered in this study were those likely to affect the CD4 cell count of HIV-infected children including the age of children, hemoglobin, weight, gender (male/female), primary caregiver, caregiver HIV status (positive/negative), status of tuberculosis (positive/negative), functional status (ambulatory/bedridden/working), WHO clinical stage (stage I/stage II/stage III/stage IV), type of ART, and BMI.

3. Methodology

We conducted exploratory data analysis to investigate various structures and patterns exhibited in the dataset. This consists of obtaining the summary statistics such as mean and variance for the CD4 cell count. Besides, the individual profile plots, mean structure, and variance structure plots were used to gain some insights into the data. While, the individual profile plots and the variance structure were used to gain insight into the variability in the data and to determine which random effects to be considered in the linear mixed model. Also, the mean structure was used to gain intuition on the time function that can be used to model the data.

3.1. Linear Mixed-Effect Model (LMM). LMM is the most frequently used random effects model in the context of continuous repeated measurements from longitudinal responses when the measurements are taken on the same or

related subjects at different times; in both cases, the responses are likely to be correlated [8]. When modeling longitudinal data, our interest is to study the association between dependent variable and a set of explanatory variables [9]. In the LMM, we assume that the dependent variable is a linear function of independent variables with regression coefficients that vary randomly from one person to another. This variation among individuals arises because of unmeasured or hidden factors [10]. The term 'mixed' is used because the LMM includes both fixed and random effects [11]. The fixed part represents the mean response, while the random part is for the individual level responses. Hence, the LMM provide a general modeling framework for subject-specific random effects, assumed to follow a normal distribution and are included to account for the correlation [12, 13]. Here, the dependent variable was taken on the same subject at different times with different baseline characteristics. The assumed model captures for the correlation of the CD4 cell count taken on the same subject at different timepoints. To formalize, let β be a $p \times 1$ vector of unknown coefficients for the fixed effects part and \mathbf{X}_i be the $n_i \times p$ design matrix of fixed predictors linking β to the set of longitudinal measurements of CD4 cell counts labeled as \mathbf{Y}_i . Let \mathbf{u}_i be a $k \times 1$ vector of latent individual random effects and \mathbf{Z}_i denote a known $n_i \times k$ design matrix values of the random factors linking \mathbf{u}_i to \mathbf{Y}_i .

$$\mathbf{Y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\varepsilon}_i,$$

$$\begin{cases} u_i \sim N(0, D), \\ \boldsymbol{\varepsilon}_i \sim N(0, \Sigma), \\ u_1, \dots, u_n \text{ and } \boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n \text{ are independent,} \end{cases} \quad (1)$$

where \mathbf{Y}_i is the $n_i \times 1$ CD4 cell count for the i^{th} children, and $\boldsymbol{\varepsilon}_i$ distributed as $N(0, \Sigma_i)$ is a vector of residual components, combining measurement error and serial correlation. Then, u_i is distributed as $N(0, \Omega)$, independent of each other. That is, $\text{cov}(u_i, \boldsymbol{\varepsilon}_i) = 0$. Furthermore, $\Sigma_i = \delta^2 I_{n_i}$ is the $n_i \times n_i$ positive-definite variance-covariance matrix for the errors in subject i , where I_{n_i} denotes the $n_i \times n_i$ identity matrix. Among the commonly used variance covariance structure for random effects, compound symmetry, heterogeneous compound symmetry, first-order autoregressive, and unstructured were considered and compared. Furthermore, to select the model which appropriately fits the given data, Akaike's information criterion (AIC) and Bayesian information criterion (BIC) were used.

4. Results and Discussion

4.1. Exploratory Data Analysis. Exploratory analysis of longitudinal data seeks to discover patterns of systematic variation across groups of children, as well as aspects of random variation that distinguish individual children. Table 1 displays the summary statistics of the longitudinal CD4 cell count of HIV-positive children in different follow-up months. During the follow-up periods, the size of the cohort varied between 201, 185, 137, 125, 115, 92, and 54, respectively, for the baseline (first diagnosis), first visit, second

TABLE 1: Summary statistics of the longitudinal CD4 cell count of HIV-infected children in different follow-up months.

Time	Baseline	First visit	Second visit	Third visit	Fourth visit	Fifth visit	Sixth visit
<i>N</i> (%)	201 (100%)	185 (92.0%)	137 (68.2%)	125 (62.2%)	115 (57.2%)	92 (45.8%)	54 (26.9%)
Mean	468.50	615.05	868.68	874.65	912.56	911.21	893.52
SD	319.11	376.98	387.24	354.37	348.16	336.57	302.82

visit, third visit, fourth visit, fifth visit, and sixth visit. The number of the CD4 cell count of patients between follow-up periods was decreasing over time indicating that they leave the study due to several reasons including death, early withdrawals, lost to follow-up, and other reasons. It can be seen that the mean of the CD4 cell count of children increases with an increasing rate until 24th month (i.e., 4th visit time) and start decreasing afterwards. The same is true for their standard deviations, increasing until 12th month (i.e., 2nd visit time) and start decreasing slowly. The average number of baseline CD4 cells count was 468.50 cells per mm³ with a standard deviation of 319.11 per mm³ of blood, implying that children were at risk at baseline and the average CD4 cell count start increasing after initiation of ART.

4.2. Individual Profile Plot. Figure 1 shows the individual profile plot of the longitudinal CD4 cell count of all study subjects (left) and twenty randomly selected HIV-infected children's (right) by the follow-up time. Hence, it can be seen that some trajectories were steeper while others were almost horizontal, indicating the possible variability in the slope and intercept of CD4 cell counts. The plot also provides information on variability between CD4 cell counts and shows there is a change in the CD4 cell count over time. Some children' CD4 cell counts increase with an increasing rate and decrease over time, and the other has the erratic CD4 cell count. It appears that there is a fluctuation in the CD4 cell count over time after they initiated ART, and the variability of the CD4 cell count seemed larger at the beginning and lower at the end (Figure 1 (right)). Hence, there is variability between children in terms of their CD4 cell count. So, this is an indication of including random effects for each child to capture this variability and to allow the CD4 cell count for children within the same child to be correlated.

4.3. Mean Profiles Plot. The average evolution describes how the profile of many subpopulations (or the population as a whole) evolves. The results of this exploration will be useful to choose a fixed-effects structure for the linear mixed model. The average of the profiles increases over time by the sex group, the point at which the CD4 count was recorded. Therefore, the approximately straight-line trajectories linear time effect in the mixed effect model could be good in fitting the data (Figure 2).

4.4. Variance Profiles Plot. The variance of the CD4 cell count of children showed an irregular pattern over the follow-up period. It increases at some point and decreases at another point suggesting a nonconstant variation among children over the follow-up period. High variation was

observed among male until 12th month and higher for females afterwards. The variance of both genders increases at some point and decreases at another point and suggests there is no constant variation over time (Figure 3). Therefore, because of the variability in the intercept and slope of trajectories (Figure 1) and pattern observed in Figure 2, the mixed effect model could be the candidate model to fit the data.

4.5. Linear Mixed-Effect Model. Before modeling, normality of the CD4 cell count was checked, and the data on the CD4 cell count appeared to be right skewed, and then, the square root transformation of the CD4 cell count was considered to normalize the data (result not displayed here). In fitting the linear mixed effect model, a series of covariance structures of the longitudinal CD4 cell counts of HIV-infected children were considered. From the possible covariance structures, compound symmetry, unstructured, first order autoregressive, and heterogeneous compound symmetry considered, first order autoregressive having the smallest AIC and BIC values (result not displayed here) was considered in the study. The random effect to be included in the linear mixed-effect model was compared based on model selection criteria, and the model with random intercept and random slope (model 2) has relatively lower AIC and BIC values than the random intercept (model 1) model and random intercept and quadratic slope model (model 3) (Table 2).

Then, a random intercept and slope model was used in the linear mixed effect model to predict the mean change of the square root of the CD4 cell count over time in addition to potential predictor variables considered in the study as results are depicted in Table 3.

We found that among potential predictors considered in the study, the age of children, observation time, WHO clinical stage of the disease, the history of TB, functional status of children, and the interaction effect of the follow-up time with the age and WHO clinical stage were significantly associated with the mean change in the square root of the CD4 cell count at 5% level of significance (Table 3). Thus, being TB positive was associated with lower in the square root of the CD4 cell count. When all the other predictor variables were controlled, the change in the square root of the CD4 cell count was 2.88 times lower for TB-positive children compared to TB-negative children. This result is consistent with a study conducted by Marie-Quitterie et al. [14]. Regarding the age of children, the change in the square root of the CD4 cell count increases with an increment of age at diagnosis in agreement with a study by Marie-Quitterie et al. [14], and younger children have good potential for achieving high CD4 counts on ART [15]. Likewise, children at the stage of working functional

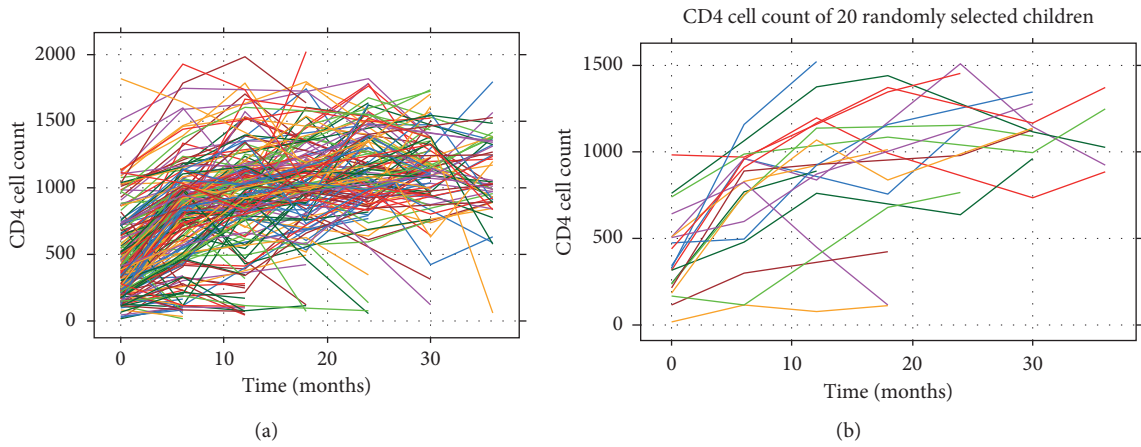


FIGURE 1: Individual profile plot of the CD4 cell count of all samples (left) and twenty randomly selected children's by the follow-up time (right).

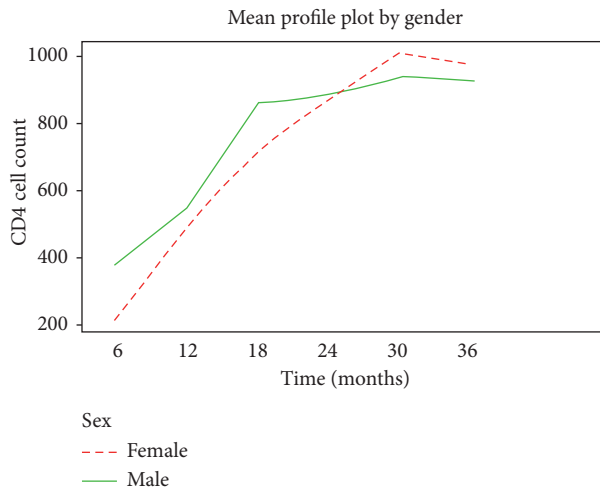


FIGURE 2: Mean profile plot of the longitudinal CD4 cell count by gender.

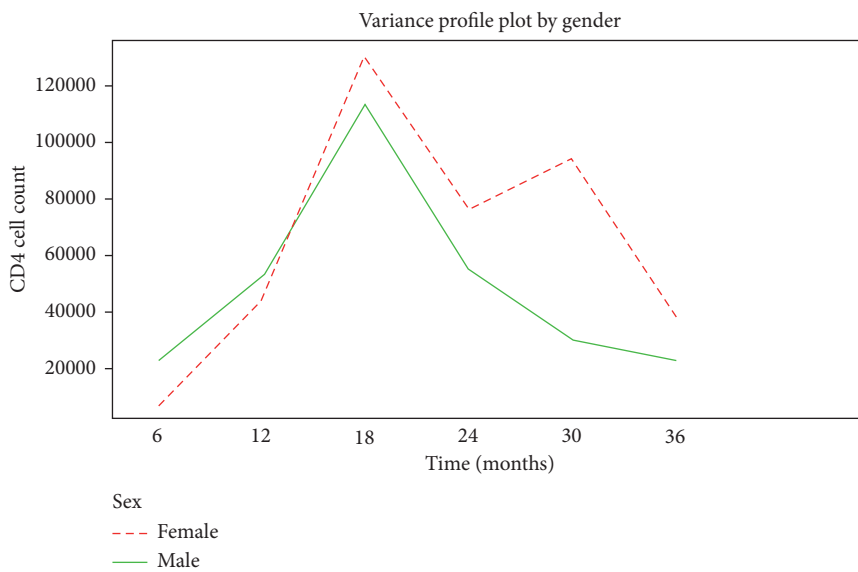


FIGURE 3: Variance profile plots of the longitudinal CD4 cell count by gender.

TABLE 2: Selection of random effect models to be included in the linear mixed-effect model.

Model	Random effect	AIC	BIC
1	Random intercept	5709.43	5800.55
2	Random intercept and random slopes	5705.43	5786.95
3	Random intercept and quadratic slope	5748.81	5839.93

TABLE 3: Parameter estimate of the linear mixed-effect model with random intercept and random slope.

Predictor variables	$\hat{\beta}$	SE ($\hat{\beta}$)	t value	p value	95% CI
Intercept	30.12	1.05	28.85	0.001 *	(28.05, 32.19)
Observation time	0.17	0.04	4.25	0.001 *	(0.08, 0.25)
Age	0.78	0.11	7.09	0.001 *	(-1.01, -0.54)
WHO stage; stage I (Ref)	—	—	-5.15	0.010 *	—
Stage II	-1.17	0.89	-1.31	0.190	(-2.93, 0.58)
Stage III	-6.43	1.19	-5.40	0.001 *	(-8.78, -4.08)
Stage IV	-9.28	1.54	-6.03	0.001 *	(-12.34, -6.23)
TB; negative (Ref)	—	—	1.38	0.034 *	—
Positive	-2.88	0.66	-4.36	0.001 *	(-4.20, -1.57)
Functional status; ambulatory (Ref)	—	—	1.78	0.045 *	—
Bedridden	-0.67	0.92	-0.72	0.163	(-2.50, 1.14)
Working	1.13	0.70	1.61	0.010 *	(-0.26, 2.52)
Time * WHO; time * stage I (Ref)	—	—	4.50	0.028 *	—
Time*stage II	0.05	0.04	1.25	0.184	(-0.02, 0.14)
Time * stage III	0.16	0.06	2.66	0.014 *	(0.03, 0.29)
Time * stage IV	-0.23	0.15	-1.53	0.113	(-0.53, 0.05)
Time * age	0.02	0.01	2.00	0.012 *	(0.003, 0.02)

Note: $\hat{\beta}$ represent the estimated parameters; SE ($\hat{\beta}$) is the standard error of estimated parameters; * represent the significant p value at 5% level of significance; and Ref denotes the reference category.

status have the higher square root of the CD4 cell count as compared to those who were bedridden and ambulatory. The change in the square root of the CD4 cell count was 1.13 times higher for children in working functional status compared to ambulatory and for the bedridden was 0.67 times lower compared to ambulatory. Regarding WHO clinical stages, the estimated coefficient for stage III and stage IV was negative and significantly different from zero indicating stage III and stage IV children had a lower CD4 cell count than stage I during the follow-up. Hence, for those children who are in stages III and IV, the change in the square root of the CD4 cell count is 6.43 and 9.28 times lower in the square root of the CD4 cell count compared to those who are in stage I, respectively. The findings are consistent with the study performed by Aboma et al. and Abdulbasit et al. [15, 16].

5. Conclusions

Among predictor variables, observation time, baseline age, WHO clinical stage, the history of TB, and functional status were determinant factors for the mean change in the square root of the CD4 cell count. Late WHO clinical stages, being TB positive, being ambulatory, and being bedridden are indicators of the disease progression. Therefore, children should need diagnosis and initiate ART early as per the recent WHO recommendation; HIV-infected children could better initiate ART treatment early in respective of the disease marker.

Data Availability

The dataset used to support the conclusions of this study is available from the corresponding author upon request.

Ethical Approval

The study was carried out after getting permission from the Statistics Department, Arba Minch University. In this regard, the official letter of cooperation referenced with stat/319/2011 was written to concerned bodies at Adama Hospital. Then, the Board of Ethical Approval Committees reviewed and approved the letter.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The first author designed the study, analyzed the data, drafted the manuscript, and critically reviewed the article. All authors read and approved the final manuscript.

Acknowledgments

This research was fully funded by Arba Minch University, community service, and research directorate. The authors thank Arba Minch University for providing necessary facilities, and their sincere gratitude goes to Mr. Tayu Nigusie

of Bule Hora University, Ethiopia, for his initial contribution to this work. They also extended their gratitude to staff in Adama Hospital who were working in the ART unit, for their kind cooperation in providing all the data.

References

- [1] UNAIDS, *Global Report: UNAIDS Report on the Global AIDS Epidemic 2012*, UNAIDS, Geneva, Switzerland, 2012.
- [2] UNAIDS, "Joint United nations programme on HIV/AIDS, global AIDS," 2018, <http://www.unaids.org/sites/default/files/media-asset/global-AIDS-update-2018-en.pdf>.
- [3] Ethiopian Public Health Institute, *HIV Related Estimates and Projections for Ethiopia*, Federal Ministry of Health, Addis Ababa, Ethiopia, 2017.
- [4] T. Fuu Jen, C. Chi Fung, L. Chih Ho et al., "Effect of anti-retroviral therapy use and adherence on the risk of hyper lipidemia among HIV infected patients, in the highly active antiretroviral therapy era," *Oncotarget*, vol. 8, no. 63, pp. 106369–106381, 2017.
- [5] A. L. Agwu and L. Fairlie, "Antiretroviral treatment, management challenges and outcomes in perinatally HIV-infected adolescents," *Journal of the International AIDS Society*, vol. 16, no. 1, p. 18579, 2013.
- [6] World Health Organization, *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach*, World Health Organization, Geneva, Switzerland, 2015.
- [7] T. Nigusie, D. Asfaw, and B. Belete, "Determinants of change in CD4 count and relationship with survival among children with HIV in Ethiopia," *AIDS Care*, vol. 33, pp. 1–5, 2020.
- [8] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, no. 4, pp. 963–974, 1982.
- [9] K.-Y. Liang and S. L. Zeger, "Longitudinal data analysis using generalized linear models," *Biometrika*, vol. 73, no. 1, pp. 13–22, 1986.
- [10] P. Diggle, *Analysis of Longitudinal Data*, Oxford University Press, Oxford, England, 2002.
- [11] G. Fitzmaurice, G. Molenberghs, M. Davidian, and G. Verbeke, "Generalized estimating equations for longitudinal data analysis," in *Longitudinal Data Analysis*, pp. 51–86, Chapman and Hall/CRC, Boca Raton, Florida, 2008.
- [12] G. Verbeke and G. Molenberghs, *Linear Mixed Models for Longitudinal Data*, Springer Science & Business Media, Heidelberg, Germany, 2009.
- [13] O. Adedeji, C. Thorne, R. Malyuta et al., "Modelling CD4 T cell recovery in hepatitis C and HIV co-infected children receiving antiretroviral therapy," *The Pediatric Infectious Disease Journal*, vol. 36, no. 5, pp. e123–e129, 2017.
- [14] P. Marie-Quitterie, L. Joanna, M. Victor et al., "Predicting patterns of long-term CD4 reconstitution in HIV-infected children starting antiretroviral therapy in sub-saharan Africa: a cohort-based modelling study," *PLoS Med*, vol. 10, no. 10, 2013.
- [15] T. Aboma, G. Abdisa, and G. Yehenew, "Joint modeling of longitudinal CD4 count and time to death of HIV/TB co-infected patients: a case of Jimma University Specialized Hospital," *Annals of Data Science*, vol. 5, no. 4, pp. 659–678, 2018.
- [16] A. Abdulbasit, A. Luguterah, N. Abdul-Rahaman, S. Nasiru, and S. Abdul Rahaman, "Joint longitudinal and survival modeling of HIV in the Upper West Region of Ghana," *International Journal of Health Sciences*, vol. 6, no. 1, pp. 56–63, 2018.