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

Belay Belete Anjullo (belay.belete@amu.edu.et)
Arba Minch University
Derbachew Asfaw Teni
Arba Minch University

Research

Keywords: Children aged under fifteen years, CD4 cell count, HIV/AIDs, Cohort data, Linear mixed model

Posted Date: October 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-85009/v1

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

Belay Belete[1*] and Derbachew Asfaw[2*]

[1*,2*] Department of Statistics, Arba Minch University, Arba Minch, Ethiopia

Email of corresponding author: belay.belete@amu.edu.et

Corresponding authors

1. Belay Belete Anjullo(Assistant Prof.)

MSc in Biostatistics and Applied Statistics

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

Email: belay.belete@amu.edu.et

Mobile: +25193773831

2. Derbachew Asfaw Teni (Assistant Prof.)

PhD in Statistics

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

Email: dasfaw469@gmail.com or derbachew.asfaw@amu.edu.et

Mobile: +251946555818

P.o.Box: 21

Abstract

Background: *HIV is a major health problem in the world and failure to implement prevention programs result in an increased number of infections among newborns.The goal of this study was to investigate the evolution and determinants of CD4 cell count among HIV-infected children who were under ART.*

Methods:*We follow-up a cohort of 201 children aged under fifteen years from Oct. 2013-to-Mar. 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/mm^3 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, history of TB and functional status were determinant factors for the mean change in the square root of CD4 cell count.*

Conclusions: *The finding revealed that; the change in the square root of 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 were 6.43 and 9.28 times lower than stage I patients respectively. Similarly, we noticed observation time, history of TB, and functional status had significantly associated with the mean change in the square root of CD4 cell count.*

Keywords: *Children aged under fifteen years, CD4 cell count, HIV/AIDs, Cohort data, Linear mixed model*

Introduction

HIV is a major health problem in the world and failure to implement prevention programs result 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 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%) those occurring in Sub-Saharan Africa and 111,000 children were died due to AIDs and related illnesses globally

followed by 72,000 (65%) children were in Sub-Saharan Africa. In Ethiopia an estimate of 729,089 people living 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%) of those are children under 15 years. Furthermore the numbers 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 revisedWHO guideline recommends 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 relatedtothe change in CD4 cell count of patients initiating ART. The objective of this study was to investigatethe evolution and determinants of CD4 cell countamong HIV infected children initiating ART at Adama Hospital in Ethiopia.

Cohort-Based Data

The data used in this study came from a cohort-based retrospective study on children aged under 15 years old byreviewing patients' ART charts and electronic databases at Adama Hospital from 2013-2017 in Ethiopia [7]. A total sample of 201 children who have full records or complete history duringthe study period was considered in this study. Here we distinguished cause-effect specific related factors from various literature reviews. Thus, the change in longitudinal CD4 cell count per $mm^3$of HIV-infected children treated with ARTwas 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 CD4 cell countof HIV infected children includingthe age of children,

hemoglobin, weight, gender, primary caregiver, caregiver HIV status, TB status, functional status, WHO clinical stage, type of ART and BMI.

Methodology

We conducted exploratory data analysis to investigate various structures and patterns exhibited in the data set. This consists of obtaining the summary statistics such as mean and variance for 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.

Linear Mixed Effects 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 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 LMMincludes both fixed and random effects [11]. The fixed part represents the mean response, while the random part is for the individual level responses. Hence, LMMprovide 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 characteristic. The assumed model captures for the correlation of CD4 cell count taken on the same subject at different time points. To formalize, let $\boldsymbol{\beta}$be a $p \times 1$ vector of unknown coefficients for the fixed effects part and $\boldsymbol{X}_i$be $n_i \times p$ design matrix of fixed predictorslinking $\boldsymbol{\beta}$ to the set of longitudinal measurements of

CD4 cell counts labeled as $\boldsymbol{Y}_i$. Let $\boldsymbol{u}_i$ be a $k \times 1$ vector of latent individual random effects and $\boldsymbol{Z}_i$ denote a known $n_i \times k$ design matrix values of the random factors linking $\boldsymbol{u}_i$ to $\boldsymbol{Y}_i$.

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{u}_i + \boldsymbol{\varepsilon}_i \text{ .... (1)}$$

$$\begin{cases} u_i \sim N(0,\ D) \\ \varepsilon_i \sim N(0,\ \Sigma) \\ u_1,\ldots,u_n \ \textit{and} \ \varepsilon_1,\ldots,\varepsilon_n \ \textit{are independent} \end{cases}$$

Where $\boldsymbol{Y}_i$ is the $n_i \times 1$ CD4 cell count for the ith children and $\boldsymbol{\varepsilon}_i$ is distributed as $N(0, \Sigma_i)$ is a vector of residual components, combining measurement error and serial correlation. The nu_i is distributed as $N(0, \Omega)$, independently of each other. That is, $cov(u_i, \varepsilon_i) = 0$. Furthermore $\Sigma_i = \delta^2 I_{ni}$ is the $n_i \times n_i$ positive-definite variance-covariance matrix for the errors in subject i, where I_{ni} 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.

Results and Discussion

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 1below display 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 visit, third visit, fourth visit, fifth visit and sixth visits. The number of CD4 cell count of patients between follow up periods was decreasing over times indicating that they leave the study due to several reason including death, early withdrawals, lost to follow up and due to other reasons. It can be seen that the mean of CD4 cell count of children were increases with an increasing rate until $24'th$ month (i.e. $4'th$ visit time) and

start decreasing afterwards. The same is true for their standard deviations; increasing until $12'th$ month (i.e. $2'nd$ visit time) and start decreasing slowly. The average number of baseline CD4 cells count was 468.50 cells per mm^3 with a standard deviation of 319.11 per mm^3 of blood, implying that children were at risk at baseline and the average CD4 cell count start increasing after initiation of ART.

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

Time	Baseline	Frist visit	Second visit	Third visit	Fourth visit	Fifth visit	Sixth visit
N(%)	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

Individual Profile Plot

Figure 1 shows individual profile plot of longitudinal CD4 cell count of all study subject (left) and twenty randomly selected HIV infected children's (right) by 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 change in CD4 cell count over time.Some children CD4 cell count increase with an increasing rate, decrease over time and the other have erratic CD4 cell count. It appears that, there is a fluctuation in CD4 cell count over time after they were initiated ART and the variability of 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.

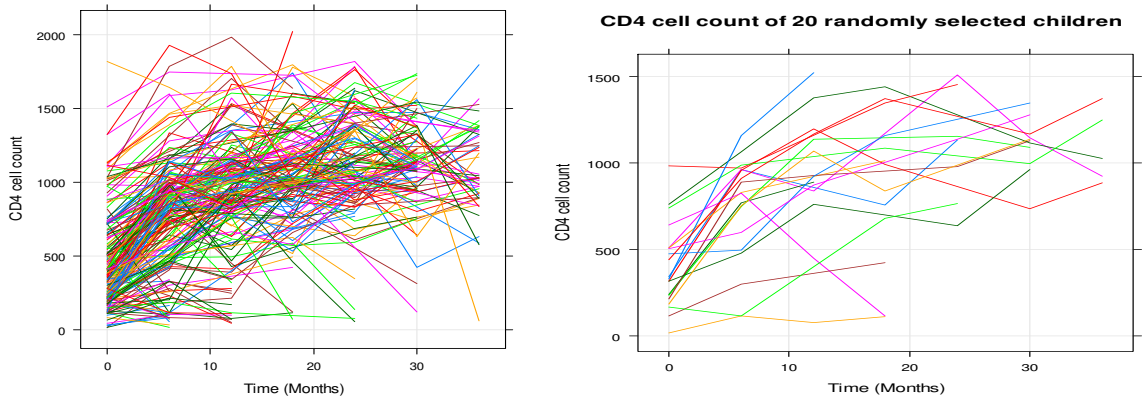

Figure 1: Individual profile plot of CD4 cell count of all samples (left) and twenty randomlyselected children's by follow up time (right)

Mean Profiles Plot

The average evolution describes how the profile of many subpopulations (or the population as a whole) evolves. From 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 sex group; point at which CD4 count was recorded. Therefore, because of the approximately straight-line trajectories linear time effect in mixed effect model could be good infitting the data.

Figure 2: Mean profile plot of longitudinal CD4 cell count by gender

Variance Profiles Plot

The variance of CD4 cell count of children showed an irregular pattern over follow up period. It increases at some point and decreases at another point suggesting a non-constant variation among children over follow up period. High Variation was observed among male until *12'th* month and higher for females afterwards. The variance of both genders increases at some point and decreases at another point 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, mixed effect model could be candidate model to fit the data.

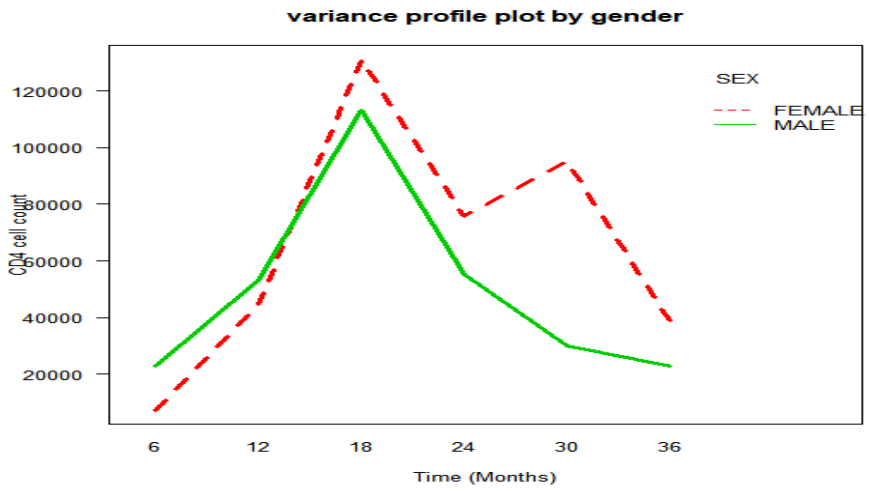

Figure 3: Variance profile plots of longitudinal CD4 cell count by gender

Linear Mixed Effect Model

Before modeling, normality of CD4 cell count was checked and the data on the CD4 cell count was appeared to be right skewed and then the square root transformation of 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 has the smallest AIC and BIC value(result not displayed here) was considered in the study. Random effect to be included in linear mixed effect model were 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 random intercept (model 1) and random intercept and quadratic slope model (model 3) (see Table 2).

Table **2**: Selection of random effect models to be included in 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

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

We found that among potential predictors considered in the study;the age of children, observation time, WHO clinical stage of the disease, history of TB, functional status of children and the interaction effect ofthe follow-up timewith the age and WHO clinical stage was significantly associated with mean change in the square root of CD4 cell count at 5% level of significance. Thus, being TB positive was associated with lower in the square root of CD4 cell count. When all the other predictor variables controlled, the change in the square root of 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 Adedeji et al. (2016)[14].Regarding the age of children, change in the square root of CD4 cell count increase with an increment of age at diagnosis in agreement with a study by Marie-Quitterie et al. (2013), younger children have good potential for achieving high CD4 counts on ART [15].Likewise,children at the stage of working functional status have the higher square root of CD4 cell count as compared to those who were bedridden and ambulatory. The change in the square root of 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 were negative and significantly different from zero indicating stage III and stage IV children had lower CD4 cell count than stage I during the follow-up. Hence, for those children who are in stage III and IVthe change in the square root of CD4 cell countis 6.43 and 9.28 times lower in the square root of CD4 cell count compared to those who are in stage I, respectively. The findings are consistent with the study done by Aboma et al and Abdulbasit et al [15,16].

Table **3**: Parameter estimate of linear mixed effect model with random intercept and randomslope.

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

Conclusions

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

Ethical Consideration

The study was carried out after getting permission from the Statistics Department, Arba Minch University. In this regard, the official letter of co-operation referenced with stat/319/2011 was written to concerned bodies.

Consent for publication

Not applicable.

Availability of Data and Materials

The dataset supporting conclusions of this article is available by contacting the authors.

Competing Interests

The authors declare that they have no competing interests.

Funding

This research was fully funded by Arba Minch University, community service, and research directorate.

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.

Acknowledgements

We thank Arba Minch University for providing necessary facilities and our sincere gratitude goes to Mr. Tayu Nigusie of Bule Hora University, Ethiopia for his initial contribution to this work. We also extended our gratitude to staff in Adama Hospital who was working in the ART unit, for their kind cooperation in providing all the data.

References

1. UNAIDS (2012). Global report: UNAIDS report on the global AIDS epidemic 2012.
2. UNAIDS (2018). Joint United Nations Programme on HIV/AIDS, Global AIDS. Available at URL: http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2018 en.pdf.
3. Ethiopian Public Health Institute (2017), Federal Ministry of Health. HIV Related Estimates and Projections for Ethiopia July , Addis Ababa, Ethiopia.
4. Fuu Jen, T, Chi Fung, C, Chih Ho, L, Yang Chang, W, Mao Wang H, Jen Hsien, W, Ni Tien, T, Xiang, L, Hsinyi, Ting Hsu, L, Chiu Chu, L, Shao Mei, H, Ju Pi, L, Jung Chun, L, ChihChien, L, Jin Hua, L & Ying-Ju L(2017). Effect of antiretroviral therapy use and adherence on the risk of hyper lipidemia among HIV infected patients, in the highly active antiretroviral therapy era: Oncotarget, 2017, Vol. 8, (No. 63), pp: 106369-106381.
5. Agwu, L, &Fairlie, L (2013). Antiretroviral treatment, management challenges and outcomes in perinatally HIV infected adolescents. Journal of the International AIDS Society, 16(1), 18579.
6. World Health Organization (2015). Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. World Health Organization.
7. TayuNigusie, DerbachewAsfaw& Belay Belete (2020) Determinants of change in CD4 count and relationship with survival among children with HIV in Ethiopia, AIDS Care, DOI: 10.1080/09540121.2020.1787941
8. Laird, NM & Ware, JH (1982). Random-effects models for longitudinal data. Biometrics, 38(4), 963-974.
9. Liang, KY, &Zeger, SL (1986). Longitudinal data analysis using generalized linear models. Biometrika, 73(1), 13-22.
10. Diggle, P, et al. (2002). Analysis of longitudinal data. Oxford University Press.

11. Fitzmaurice, G, Molenberghs, G, Davidian, M. &Verbeke, G (2008). Generalized estimating equations for longitudinal data analysis. In Longitudinal data analysis (pp. 51-86). Chapman and Hall/CRC.
12. Verbeke, G, &Molenberghs, G (2009). Linear mixed models for longitudinal data. Springer Science & Business Media.
13. Adedeji, O, Thorne, C, Malyuta, R, Volokha, A, Callard, RE., Klein, NJ, & Lewis, J (2016). Modelling CD4 T cell recovery in hepatitis C and HIV co-infected children receiving antiretroviral therapy. The Pediatric infectious disease journal, 36(5), e123.
14. Marie-Quitterie P, Joanna L, Victor M, Andrew P, Kusum N&Addy K (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 Med 10(10): e1001542.
15. Aboma, T, Abdisa, G, &Yehenew, G (2018). 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, 5(4), 659-678.
16. Abdulbasit, A, Luguterah, A, Nasiru, S, & Abdul Rahaman, S (2018). Joint Longitudinal and Survival Modeling of HIV in the Upper West Region of Ghana.International Journal of Health Sciences, 6(1), 56-63.

Figures

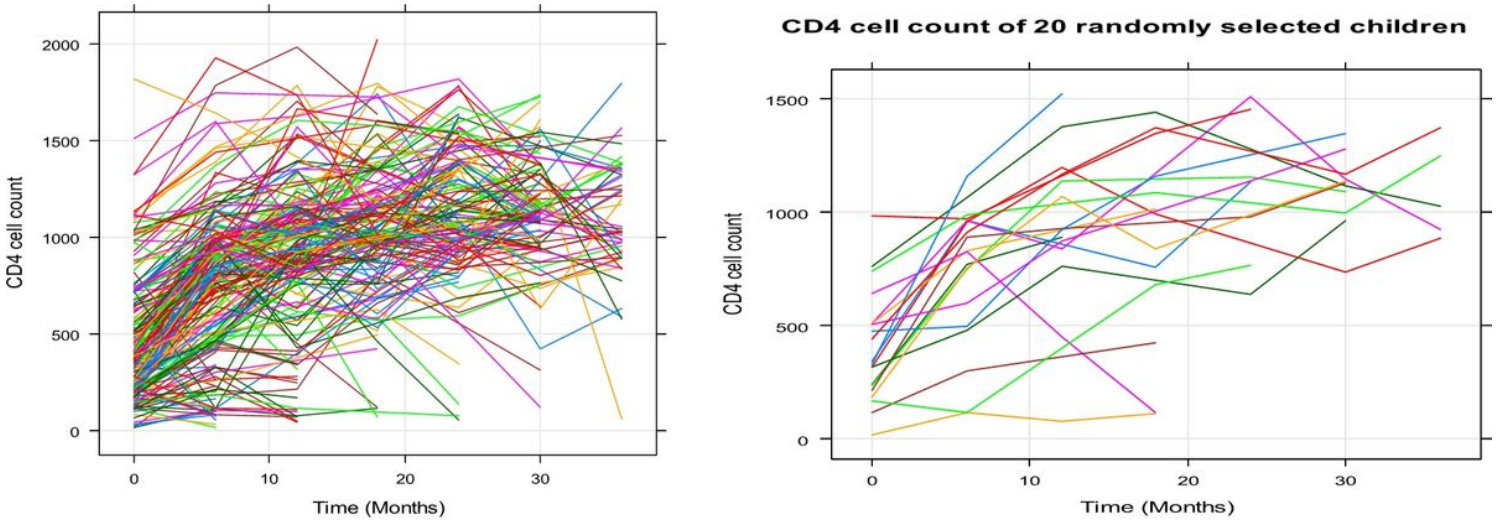

Figure 1

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

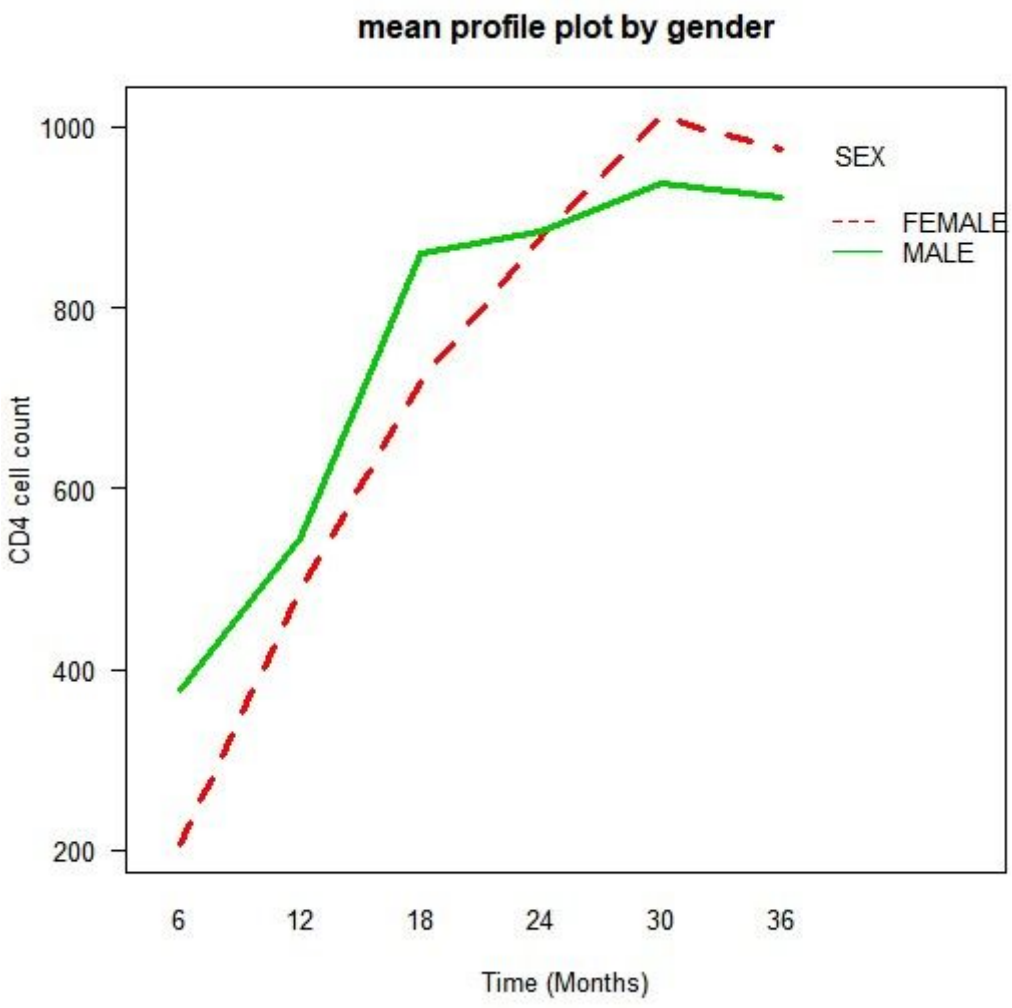

Figure 2

Mean profile plot of longitudinal CD4 cell count by gender

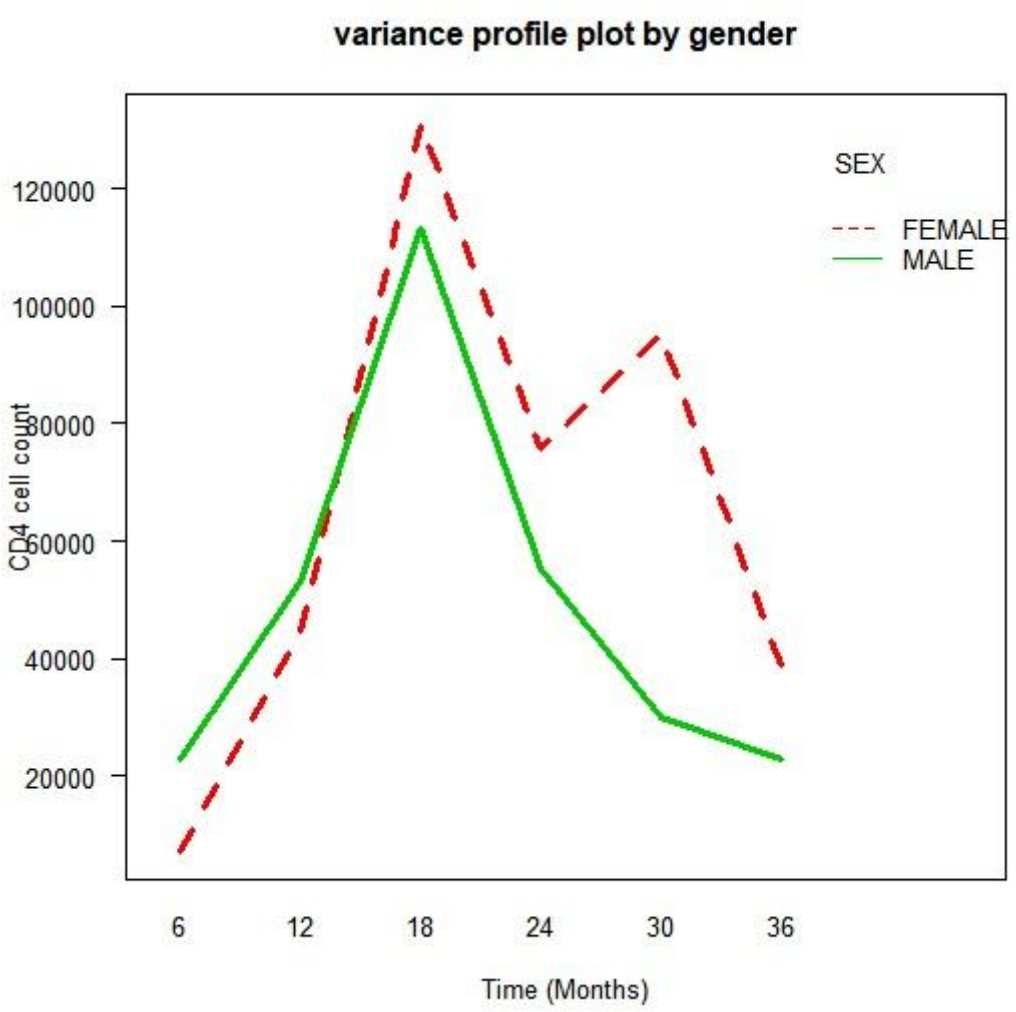

Figure 3

Variance profile plots of longitudinal CD4 cell count by gender