COLLEGE OF HEALTH SCIENCES

SCHOOL OF PUBLIC HEALTH

DEPARTMENT OF BIOSTATISTICS

JOINT MODELING OF TIME TO DEVELOP TUBERCULOSIS AND CHANGE IN CD4 COUNT AMONG HIV PATIENTS UNDER ART IN MEKELLE, ETHIOPIA, 2024

PRINCIPAL INVESTIGATOR: TEKLEBRHAN KINFE (MSc. FELLOW)

A THESIS REPORT SUBMITTED TO MEKELLE UNIVERSITY, COLLEGE OF HEALTH SCIENCES, SCHOOL OF PUBLIC HEALTH, DEPARTMENT OF BIOSTATISTICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE MASTER OF SCIENCE IN BIOSTATISTICS AND HEALTH INFORMATICS

ADVISOR: Mr. KIDANEMARIAM ALEM (ASSISTANT PROFESSOR)

CO-ADVISOR: Mr. GOITOM HALEFOM (MPH)

JUNE, 2025

MEKELLE, ETHIOPIA

MEKELLE UNIVERSITY

COLLEGE OF HEALTH SCIENCES

SCHOOL OF PUBLIC HEALTH

DEPARTMENT OF BIOSTATISTICS

A THESIS SUBMISSION FORM

INVESTIGATOR	TEKLEBRHAN KINFE (MSC. FELLOW)
NAME OF ADVISORS	MR. KIDANEMARIAM ALEM (ASSISTANT PROFESSOR) MR. GOITOM HALEFOM (MPH)
FULL TITLE	JOINT MODELING OF TIME TO DEVELOP TUBERCULOSIS AND CHANGE IN CD4 COUNT AMONG HIV PATIENTS UNDER ART IN MEKELLE, ETHIOPIA, 2024
DURATION OF THESIS	JANUARY-DECEMBER, 2024
STUDY AREA	ETHIOPIA
TOTAL COST OF THE THESIS	22,457.60 ETHIOPIAN BIRR
ADDRESS OF INVESTIGATOR	**EMAIL;** Teklebrhanki2@gmail.com **CELL PHONE**; +251911360758

Advisor Approval Sheet

This is to certify that the thesis entitled "Joint modeling of time to develop tuberculosis and change in CD4 count among HIV patients under ART in Mekelle, Ethiopia, 2024" is submitted in partial fulfillment of the requirements for the degree of Master of Science (MSc) in "Biostatistics and Health Informatics" to the Graduate Program of the College of Health Sciences of Mekelle University and has been carried out by **Teklebrhan Kinfe Gebru** ID No: **CHS/PR169486/12** under my supervision. Therefore, I recommend that the student has fulfilled the requirements and hence can submit this thesis to the department.

Mr. Kidanemariam Alem (Assistant Professor) ______________ ____________

Name of Main Advisor Signature Date

Mr. Goitom Halefom (MPH) _________________ _____________

Name of Co-advisor Signature Date

I hereby declare that this MSc. thesis is my original work and has not been presented for a degree at any other university, and all sources of material used for this thesis have been duly acknowledged.

Name: Teklebrhan Kinfe Gebru Signature: ______________ Date: ______________

This MSc research thesis has been submitted for examination with my approval as thesis advisor.

Mr. Kidanemariam Alem (Assistant Professor) ______________ ____________

Name of Main Advisor Signature Date

Examiner Approval Sheet

We, the undersigned members of the Board of Examiners of the final open thesis defense by "Teklebrhan Kinfe Gebru," have read and evaluated his thesis "Joint modeling of time to develop tuberculosis and change in CD4 count among HIV patients under ART in Mekelle, Ethiopia, 2024" and evaluated the candidate. This is therefore to certify that the thesis has been accepted in partial fulfillment of the requirements for the MSc. degree in Biostatistics and Health Informatics.

Name of chairperson	Signature	Date
Name of Major Advisor	Signature	Date
Name of Internal Examiner	Signature	Date
Name of External Examiner	Signature	Date

Final approval and acceptance of the thesis is contingent upon the submission of the final copy of the thesis to the candidate's Department through the office of the Department Graduate Program Coordinator.

Thesis approved by

Graduate Program Coordinator	Signature	Date

Certification of the final thesis

I hereby certify that all the corrections and recommendations suggested by the Board of Examiners are incorporated into the final thesis entitled "Joint modeling of time to develop tuberculosis and change in CD4 count among HIV patients under ART in Mekelle, Ethiopia, 2024", by Teklebrhan Kinfe Gebru.

Department Head	Signature and Stamp	Date

LIST OF ABBREVIATIONS AND ACRONYMS

ACSH	Ayder Comprehensive Specialized Hospital
AIC	Akaike Information Criterion
AIDS	Acquired Immuno-Deficiency Syndrome
CD4	Clusters of Differentiation 4
CI	Confidence Interval
CPT	Cotrimoxazole Preventive Therapy
HIV	Human immunodeficiency Virus
HR	Hazard Ratio
IPT	Isoniazid Preventive Therapy
IQR	Inter-Quantile Range
MGH	Mekelle General Hospital
MICE	Multiple Imputation by Chained Equations
PHA	Proportional Hazard Assumptions
PLHIV	People Living with HIV
Q–Q	Quantile–Quantile
SSA	Sub-Saharan Africa
STATA	Statistics and Data
TB	Tuberculosis
UNAIDS	Joint United Nations Programme on HIV/AIDS
VIF	Variance Inflation Factor
WHO	World Health Organization

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Background: In patients with HIV, tuberculosis remains the leading cause of mortality and morbidity. Little is known about the predictors and the median time to develop tuberculosis while considering for the effect of the variation of longitudinal CD4 cell count.

Objective: To investigate the time to develop tuberculosis accounting for longitudinal CD4 cell count change and its predictors among HIV patients who are under ART follow-up at Mekelle General Hospital and Ayder Comprehensive Specialized Hospital, Mekelle, Ethiopia, 2024.

Methodology: A facility-based retrospective follow-up study was conducted among 449 adult PLHIV under ART follow-up from March 2018 to May 2024. The study participant were selected via a simple random sampling. The secondary data were collected from the patients' medical records via Kobocollect version 2021.2.4 and exported to STATA version 17.0. The final model was a joint random intercept Cox-proportional hazard model. A model with the lowest Akaike information criterion and Bayesian information criterion was selected.

Results: The incidence density of TB disease was 6.77 cases/100 person-years with a restricted mean survival time of 60 months. The joint analysis provided an association parameter alpha with AHR=0.854; 95% CI(0.8-0.91), indicating that for a unit increase in the average $\sqrt{\text{CD4 cell count}}$, the hazard of TB infection decreased by 14.6%, keeping other variables constant. The study also revealed that advanced WHO clinical stage (AHR = 1.024, 95% CI: 1.017–1.033), sex (AHR= 1.62, 95% CI: 1.09,2.4), CPT intake (AHR= 0.55, 95% CI: 0.35,0.89), and adherence (AHR= 0.38, 95% CI: 0.28,0.52) were significantly associated with the time to develop tuberculosis. The random intercept model indicated that greater variation in CD4 counts at baseline contributed strongly to the hazard of tuberculosis.

Conclusion and Recommendation: This research highlights that PLHIV with a decreasing trajectory of CD4 count, advanced WHO clinical stage, female sex, history of CPT intake, and poor adherence have a higher risk of tuberculosis. On the basis of these findings, it is strongly recommended that the government and relevant health actors working on TB/HIV should intensify activities that improve patient adherence and a regular CD4 cell measurement.

Keywords: CD4 count, Time to TB, Joint modeling, PLHIV, ART, Ethiopia

1. INTRODUCTION

1.1. Background

Human immunodeficiency virus (HIV) is an infection that attacks the body's immune system, specifically white blood cells called clusters of differentiation 4 (CD4) T-cells. HIV destroys these CD4 cells, weakening a person's immunity against opportunistic infections, such as tuberculosis (TB) infections, fungal infections, severe bacterial infections, and some cancers. Acquired immunodeficiency syndrome (AIDS) is the most advanced stage of the disease (1).

To track the severity of individual HIV infection, CD4 cell count measurements have been used for many years. The CD4 cell count is the number of CD4 cells per cubic centimeter, with a normal range between 500 and 1500 cells/mm^3. For clinicians, the CD4 cell count has become a good indicator of disease progression, disease staging, and monitoring the effectiveness of antiretroviral treatment (ART) (1,2). The ART refers to the use of a combination of three or more antiretroviral drugs for treating HIV infection. The standard treatment consists of a combination of drugs (often called "highly active antiretroviral therapy" or HAART) that suppress HIV replication (3).

TB disease is a major source of morbidity and mortality among people living with HIV (PLHIV). TB is a bacterial infection caused by a bacillus called Mycobacterium tuberculosis. Approximately 5–10% of people infected with TB will eventually experience symptoms and develop the disease (4). Presumptive TB case in PLHIV refers to an individual who presents with persistent cough of any duration, fever, night sweats, unexplained weight loss or with chest x-ray abnormality suggestive of TB confirmed either through Xpert test, sputum microscopy test, or through emprical clinician decision (5).

TB and HIV/AIDS are the main burdens of infectious diseases in resource-limited countries. Two pathogens, HIV and Mycobacterium tuberculosis, exacerbate the deterioration of immunological function by potentiating the effects of one another. It hurts the immune response to HIV, accelerating the progression from HIV infection to AIDS. Up to 50% of PLHIV are expected to develop active TB (2). The appropriate use of ART, early screening for TB infection, and Isoniazid

Preventive Therapy (IPT) are the major intervention used to prevent TB disease among PLHIV (6).

1.2. Statement of the problem

TB disease is a major cause of morbidity and mortality among PLHIV. Currently, the public health burden of TB disease is on the rise globally by overtaking COVID-19 for the leading cause of mortality from a single infectious disease. PLHIV are 16 (14–18) times more likely to fall ill with TB disease than people without HIV. Without proper treatment, nearly all HIV-positive people with TB die. It also remained the leading cause of death among PLHIV (27% of PLHIV mortality). In 2022, approximately 167,000 people died of HIV-associated TB (4,7).

The World Health Organization (WHO) African Region has the highest burden of HIV-associated TB. According to a systematic review and meta-analysis performed in sub-Saharan Africa (SSA), the overall pooled incidence rate of TB in PLHIV was 3.49 per 100 person-years, and it was 3.79 per 100 person-years in adults (7,8). In Ethiopia, a study performed in Addis Ababa and northeastern Ethiopia revealed that among adult PLHIV the overall incidence density of TB was 6.82 and 8.6 per 100 person-years of observation, respectively (9,10).

The incidence of TB among PLHIV is affected by multiple factors. Several studies performed globally and nationally revealed that PLHIV with low CD4 cell count, underweight, male sex, no history of IPT, no history of Cotrimoxazole Preventive therapy (CPT), previous TB disease, bedridden, anemic, poor adherence, WHO clinical stages III and IV, first year of follow-up, and large family size are especially susceptible to TB illness (6,8–11).

Even though identifying the median time and associated factors of the time to develop TB while accounting for the variation of longitudinal CD4 cell count (surrogate biomarker) is quite essential in better prediction of TB development among PLHIV, however, little is known about the association between the longitudinal CD4 cell count and the time to develop TB.

In recent years, joint models for survival and longitudinal data have become quite popular in HIV/AIDS studies. Joint modeling reduces the biases of parameter estimates by accounting for the association and dependency between the longitudinal and time-to-event data (12,13). While multiple separate models of time to TB disease development and longitudinal CD4 cell count changes are widely presented in different studies, few studies have investigated the joint modeling of time to develop TB disease and longitudinal CD4 cell count changes among PLHIV who are receiving ART follow-up. Furthermore, to the best of our knowledge, no investigations have been conducted to identify the association parameters and key predictors that jointly influence these two outcomes: variation in CD4 cell count and the time to develop TB among PLHIV receiving ART at Mekelle General Hospital (MGH) and Ayder Comprehensive Specialized Hospital (ACSH). Therefore, this study was designed to investigate the time to develop TB disease accounting for longitudinal CD4 cell count change and its predictors among PLHIV Who are under ART follow-up at MGH and ACSH, Mekelle, Ethiopia.

1.3. Significance of the study

The findings of this study will provide an important input to the regional health bureau to reduce the risk of TB disease development among PLHIV under ART follow-up by focusing on the most significant factors. It will also provide a good input to the patients to decrease their risk of TB disease development. Furthermore, this study will help in the development of a better policy to prevent and reduce TB disease among PLHIV. Finally, the results could serve as a baseline for further studies.

2. LITERATURE REVIEW

2.1. Overview of TB among PLHIV

The global estimate of TB/HIV disease in 2022 was 710,000 (6.7% of all TB cases). TB has caused an estimated 167,000 deaths among people with HIV (7,14). In SSA, according to a systematic review and meta-analysis, the overall pooled incidence rate of TB in HIV-infected adults was 3.79 per 100 person-years (8). Similarly, several studies performed in Thailand and Tanzania reported that the incidence of TB disease among PLHIV under ART was 7.5 and 2.8 per 1000 person-years of follow-up, respectively (15,16). However, a study performed in the United Kingdom reported a relatively low (0.6/1000 person-years) TB incidence among PLHIV under ART (17).

In Ethiopia, multiple studies have shown a high incidence rate of TB among PLHIV under ART. According to a systematic review performed in Ethiopia, the estimated pooled incidence rate of TB among HIV-infected patients after initiation of ART was found to be 4.3 per 100 person-years of observation (18). In addition, according to studies performed in Addis-Ababa and Afar, the incidence density of TB among PLHIV under ART was found to be very high, at 6.82 and 8.6 per 100 person-years of observation, respectively (10,19). Similarly, a study performed in Mekelle, Tigray region of Ethiopia, revealed that the incidence rate of opportunistic infections after ART initiation was 7.5 cases/100 person-years, with TB being the highest. However, another study performed in Addis Ababa reported a 3.08/1000 person-year incidence rate of TB disease among PLHIV under ART (20). According to a study done in Iran an increasing trajectory of CD4 cell count was associated with a 68% decreased risk of TB co-infection (21).

The median survival time to TB development is an important milestone in TB disease prevention and control. According to a 5 years survival study done in Northeastern Ethiopia, the median survival time to develop TB was found to be 54 months (9). A 6 years followup study performed in Amhara, Ethiopia, revealed that the mean survival time to develop TB was found to be 60.8 months (22).

2.2. Predictors of the longitudinal CD4 cell count and time to develop TB

Many studies have shown that PLHIV under ART with low baseline CD4 cell counts are more likely to be TB coinfected. Several systematic reviews and meta-analyses performed in Africa have shown that PLHIV under ART with low baseline CD4 cell counts are more likely to be TB coinfected (8,23). Other studies performed in Ethiopia revealed that PLHIV who have low CD4 cell counts are more likely to develop TB than their counterparts are (10,24). This finding was similar to that of a study performed in Tanzania and Thailand, in which a high CD4 cell count was associated with a lower hazard of TB (15,16).

Age was found to be an important predictor of TB among PLHIV. For example, PLHIV under ART who are older (aged above 65 years) are more likely to be coinfected with TB (20). A joint analysis of longitudinal CD4 cell count and TB status performed in northwestern Ethiopia also revealed that as patients' age increase, the expected CD4 cell count decreased (25).

With respect to sex, studies have shown significant variation between males and females. A systematic review and meta-analysis in SSA revealed that male patients were approximately 1.43 times more likely to develop TB (8). Similarly, a study performed in Thailand revealed that male patients have a 1.4-fold greater risk of TB disease (15).

Low body weight was found to be positively associated with TB development among PLHIV under ART. A study of SSA revealed that PLHIV who are underweight (body mass index <18.5 kg/m^2) PLHIV have a 1.79-fold greater risk of acquiring TB than HIV patients who have a normal weight (8). Multiple studies performed in Ethiopia have shown that PLHIV under ART who are underweight (body mass index <18.5 kg/m^2) at baseline are more likely to be coinfected with TB (10,19,20,25). Another study performed in Oromia, Ethiopia, confirmed that PLHIV with moderate wasting have an approximately three (2.86) times greater risk of developing TB (26). In addition, a study performed in Thailand revealed that PLHIV with lower body weights (<50 kg) have a higher rate of TB (15).

Studies have shown that taking IPT has a protective effect against TB coinfection. A systematic review indicated that patients who do not take IPT have a triplet higher risk of acquiring TB in SSA (8). A study performed in Addis-Ababa and northeastern Ethiopia revealed that HIV patients who had taken IPT had 80% and 86% decreased risks of being TB coinfected, respectively (10,19).

Similarly, a study performed in Central Ethiopia revealed that PLHIV not receiving IPT were approximately eight times more likely to be TB coinfected (26).

PLHIV with an advanced WHO stage were more likely to have TB coinfection. A systematic review and meta-analysis performed in SSA revealed that PLHIV with advanced WHO clinical stages were twice as likely to be TB coinfected (8). In several studies performed in southwestern and northeastern Ethiopia, PLHIV under ART with an advanced WHO clinical stage (stages III and IV) were approximately seven and three times more likely to be coinfected with TB, respectively (19,24).

Studies have shown that functional status is also an important predictor of TB development among PLHIV. According to a study performed in SSA, bedridden or ambulatory patients have a 1.87-fold greater hazard of TB development than working patients do (8). According to a study performed in northeastern Ethiopia, PLHIV who were bedridden at baseline were 5.45 times more likely to develop TB than PLHIV who were not in bedridden (19). Those who were ambulatory had 3.95 lower CD4 cell counts than did those with working status (25).

PLHIV with low hemoglobin levels have a greater risk of developing TB coinfection. In SSA, anemic patients are 1.73 times more likely to be TB coinfected (13). Several studies performed in western and northeastern Ethiopia have shown that patients with low baseline hemoglobin levels have an approximately seven- and twofold-fold greater risk of developing TB than do those with normal hemoglobin levels (19,26).

ART medication adherence was found to be an important predictor of TB development among PLHIV receiving ART. A study performed in northwestern Ethiopia revealed that PLHIV with poor ART medication adherence had low CD4 cell counts and a relatively high risk of TB (25).

History of CPT utilization was found to be an important predictor of TB development among PLHIV under ART. A study performed in South Africa revealed that PLHIV who don’t utilize CPT was associated with a 42% decrease in TB co-infection (27).

Overall, many of the studies were cross-sectional and did not assess the influence of longitudinal CD4 cell count changes on the time to develop TB among PLHIV during ART follow-up. Therefore, this study aimed to investigate the influence of longitudinal CD4 cell count changes on the time to develop TB and predictor variables among PLHIV receiving ART follow-up in Mekelle, Ethiopia.

2.3. Conceptual Framework

This conceptual framework was developed after a review of different literature. The major predictors identified were classified as sociodemographic, clinical behavioral, and time-varying covariates. This conceptual framework considers repeated CD4 cell count measurements as an outcome variable and as a surrogate predictor of time to develop TB among PLHIV under ART.

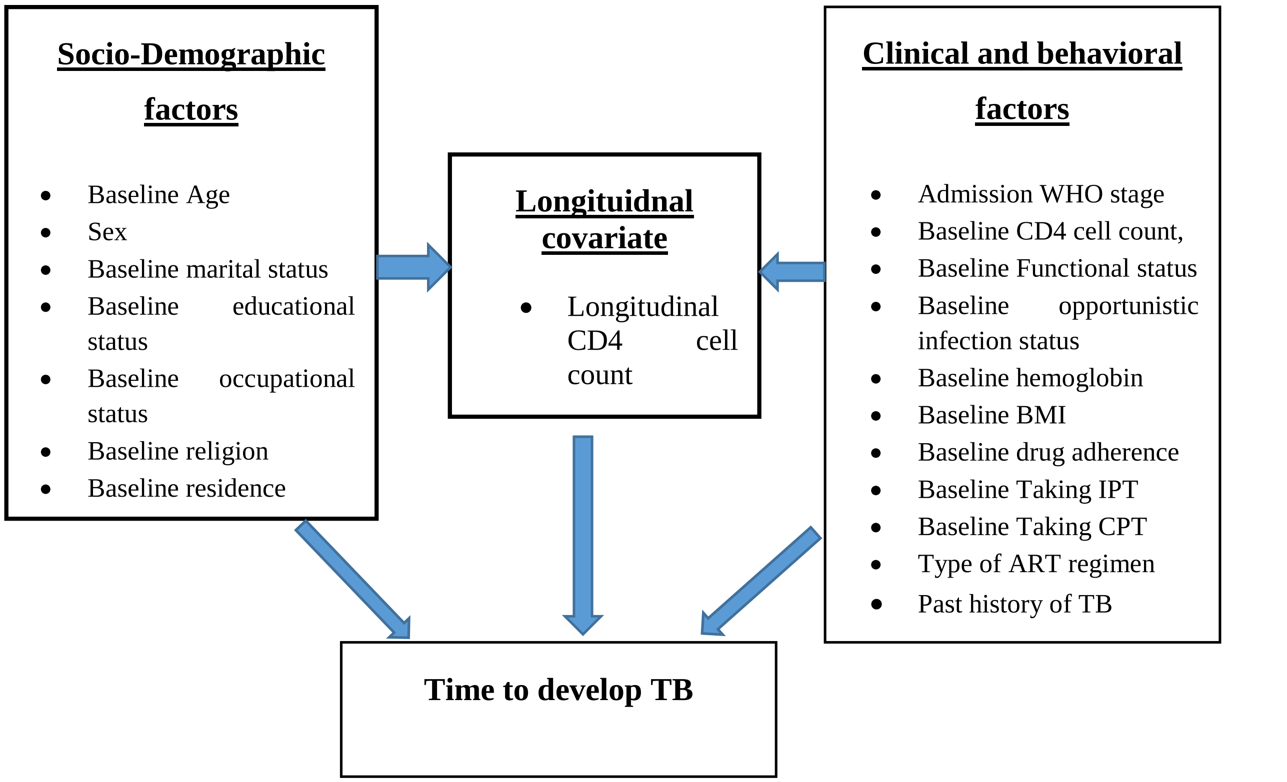

Figure 1: A schematic diagram showing the conceptual framework of predictors of developing TB disease among PLHIV under ART (8,10,16,19,23,26)

3. OBJECTIVES

3.1. General objective:

- To investigate the time to develop TB disease accounting for longitudinal CD4 cell count change and its predictors among PLHIV Who are under ART follow-up at MGH and ACSH, Mekelle, Tigray, Ethiopia, 2024

3.2. Specific objectives:

- To determine the TB incidence density among PLHIV who are receiving ART follow-up at MGH and ACSH, Mekelle, Ethiopia, 2024
- To determine the median time to develop TB disease among PLHIV who are receiving ART follow-up at MGH and ACSH, Mekelle, Ethiopia, 2024
- To determine longitudinal CD4 cell count changes among adult PLHIV who are receiving ART follow-up at MGH and ACSH, Mekelle, Ethiopia, 2024
- To predict the effects of longitudinal CD4 cell count changes and other associated factors on the time to develop TB disease among PLHIV who are receiving ART follow-up at MGH and ACSH, Mekelle, Ethiopia, 2024

4. METHODOLOGY

4.1. Study area and period

The study was conducted among people who initiated ART between March 2018 and February 2024 at MGH and ACSH. Those hospitals are located in Mekelle city, Tigray region, Northern Ethiopia. Mekelle city is situated approximately 783 kilometers north of Addis Ababa, the capital of Ethiopia.

In 1962, MGH officially opened to serve approximately 20,000 people living in and around the city catchment area. ACSH, a referral and teaching hospital that was opened in 2008, is the largest hospital in the Tigray Region. It serves a population of over 8 million from Tigray and parts of the Afar and Amhara regional states.

In 2003, Ethiopia introduced ART service in selected health facilities, and in early 2005 it started free ART service. By 2018, more than 431,939 adult PLHIV were on ART follow-up. In MGH, ART service was begun since 2004 and to date more than 10,000 PLHIV have received ART service (28). Both MGH and ACSH are the largest ART service-providing hospitals in the region (29).

4.2. Study design

A facility-based retrospective follow-up study design was used.

4.3. Population

Source population: All PLHIV aged 15 and above years who were under ART follow-up at MGH and ACSH.

Study population: All PLHIV aged 15 and above years who initiated ART follow-up from March 2018 to February 2024 at MGH and ACSH.

Study unit: A 15 and above years old PLHIV under ART follow-up with at least 2 CD4 measurements.

4.4. Eligibility criteria

4.4.1. Inclusion criteria

All 15-years and older ART enrolled PLHIV.

4.4.2. Exclusion criteria

All 15-years and older ART-enrolled PLHIV who had either TB disease at the time of admission and/or had fewer than 2 CD4 measurements.

4.5. Sample size and Sampling Procedures

The minimum sample size required for this study was calculated via a sample size determination formula for survival analysis by comparing the hemoglobin level, history of taking IPT, and history of opportunistic infections from a previous study performed in northeastern Ethiopia (19).

Table 1: Minimum sample size calculation using variables associated with time to develop TB among PLHIV

Variables	q_1	q_0	λ_1	λ_0	**HR**	**Sample Size**
History of taking IPT	0.0425	0.3221	0.0816	0.4873	0.14	448
Hemoglobin level (g/dl)	0.4673	0.2006	0.6369	0.3341	2.31	262
History of opportunistic infections	0.5588	0.2398	0.7170	0.3868	2.31	178

The log rank method of the sample size calculation formula is (30):

$$d = \frac{(z_{\alpha/2} + z_\beta)^2}{q_1(1-q_1)\,[log(\Phi)]^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

$$\boldsymbol{n} = \frac{\mathbf{2d}}{\mathbf{2 - exp}^{(-t\lambda_0)} - \boldsymbol{exp}^{(-t\lambda_1)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

where

d= Expected number of events

q_1= Proportion of exposed with event

q_0 =1-q_1= proportion of non-exposed with event

ϕ=hazard ratio

n= Required minimum sample size for the expected event

t= Follow-up period

λ_0= Hazard of event among exposed

λ_1= Hazard of event among non-exposed

From a previous study conducted in Afar, Ethiopia, using power =80%, α =0.05, β=0.2, and ϕ =0.14, q_1=0.0425, q_0=0.3221, λ_1=0.0816, and λ_0=0.4873.

Then, we calculate the sample size via Schoenfeld's formula:

$$d = \frac{(1.96+0.842\ )^2}{0.0425(0.3221)[\log(0.14)]^2} = 149$$

$$n = \frac{2*149}{2-\exp(-^{6*.0816})-exp^{(-6*.4873)}}$$

$$n = \frac{2*149}{2-.6129-0.0537} = 224 \text{ for one group}$$

$$n_{final} = 2n = 448$$

Therefore, on the basis of the above calculation, the minimum sample sizes for a history of IPT, hemoglobin level, and history of opportunistic infections were found to be 448, 262, and 178, respectively. Therefore, the minimum required sample size was 448.

From the output, 149 events (failures) were required to be observed in this study to ensure 80% power to detect a hazard ratio of 0.14 by using the log-rank test. The respective estimate of the total number of subjects required to observe 149 events in a 6-year study were 448 with 224 subjects per IPT usage group.

There were more than 4500 and 1500 adult PLHIVs on ART follow-up in both MGH and ACSH, respectively. To sample the study participants, 497 of those who ever started ART from 1st March 2018, up to 29th February 2024 were identified from the electronic medical record. Fourty-seven patients that had less than two CD4 measurments and/or develop TB during enrollment were excluded from the study. Subsequently, to fulfill the minimum sample size, 450 participants (300 from MGH and 150 from ACSH) who fulfilled the inclusion and exclusion criteria were enrolled in to the study. The study subjects were selected via a simple random sampling technique.

4.6. Data collection tools and procedures

The data were obtained from the MGH and ACSH ART clinics. Both the longitudinal and survival data were extracted via an adapted standardized extraction sheet from the electronic medical record databases and patient's chart under ART follow-up through Kobotoolbox version 2021.2.4. There were 4 data collectors who were supervised by the principal investigator. The data were collected from June 21 to July 30, 2024.

4.7. Study variables

4.7.1. Outcome variables

The outcome variables were the time to develop TB and repeated CD4 cell count measurements.

- **Survival outcome**—The survival outcome variable for this study was the time to develop TB disease among PLHIV under ART follow-up, which was measured in months. For this study, while the event was defined as the development of active TB during the follow-up period, censored TB was defined as failure to develop active TB until the end of the study period, death before developing active TB, and loss to follow-up before developing active TB.
- **Longitudinal outcome**: the longitudinal outcome variable for this study was a repeated measurement of the number of CD4 cells per milliliter of blood sample usually taken at a time interval of every six months.

4.7.2. Explanatory variables

The predictor variables include baseline sociodemographic characteristics (age, sex, occupation, religion, marital status, educational status, and place of residence), baseline clinical and behavioral characteristics (WHO clinical stage, CD4 count, past history of TB, functional status, history of opportunistic infection, hemoglobin, BMI, adherence, history of CPT intake, history of IPT intake, and type of ART regimen), and time-varying covariate: repeatedly recorded as the number of CD4 cell count (cells/mm^3).

4.8. Operational Definition

Time to TB development: This was the time from PLHIV enrollment dates ART sevice to ART clinic to TB diagnosis by laboratory investigation and/or by clinician decision during the follow-up period.

Adherence: Adherence is the extent to which a person's behavior involves taking medication, following a diet, and/or changing lifestyle corresponds with agreed-upon recommendations from a health worker.

Good adherence (≥95% adherence) was considered when only 1 dose out of 30 doses or 3 or lower doses from the 60 doses of the ART drug were missing. Good adherence is considered an appropriate level to achieve maximal viral suppression (31).

Fair adherence (85–94% adherence) was considered when 2–4 doses out of 30 doses or 4–9 doses from the 60 doses of the ART drug were missing.

Poor adherence (<85% adherence) was considered when ≥ 5 doses out of 30 doses or ≥ 10 doses from 60 doses were missing.

WHO clinical staging: The WHO has classified the clinical stage of HIV patients from asymptomatic (stage one) to advanced (stage IV) (32).

Lost to follow-up: A patient who had not received repeated ART for 3 months or longer and was not yet classified as "dead" or "transferred out" (28).

Adult PLHIV: PLHIV aged 15 years and above (6).

4.9. Data Quality Assurance

After the checklist was developed it was checked for its reproducability and vailidity on patient cards. Next, the four data collectors were assigned, and were trained on the objective of the study and how to review the documents as per the data extraction checklist. Data were extracted by reviewing follow-up charts and cards of patients for completeness and consistency and were supervised daily by the principal investigator. Moreover, data entry errors were minimized by the use of the Kobocollect toolbox.

4.10. Data management and analysis

4.10.1. Introduction

The data were first entered into Kobocollect version 2021.2.4 and then exported to STATA version 17.0 software for further analysis. In STATA, the data were first reshaped from wide to long format. Afterward, the rate of missing data in the longitudinal data was found to be 9.7%. The missingness pattern was diagnosed using logistic regression, and it was found to be missing at random since there was an association between the missing data and some of the observed predictor variables. Subsequently, the missing records in the longitudinal CD4 cell count were handled by multiple imputation by chained equations (MICE).

4.10.2. Descriptive statistics

Data were visualized via tables, graphs, and texts. Descriptive statistics such as frequency and percentage were used to summarize categorical variables. To summarize the continuous variables mean, standard deviation, median, minimum, maximum, and Inter-Quantile Range (IQR) were used.

4.10.3. Joint longitudinal survival models

Rationale for the use of joint modeling

Currently, studies that can assess longitudinal and time-to-event outcomes jointly are receiving much attention. Models such as the linear mixed model for longitudinal data and the Cox proportional hazards model for time-to-event data do not consider dependencies between these two different data types.

A powerful method that takes into account the dependency and association between longitudinal data and time-to-event data is joint modeling. Joint models for longitudinal and time-to-event data are models that bring these two data types together (simultaneously) into a single model so that one can infer the dependence and association between the CD4 count change and the time to develop TB. Joint modeling reduces bias in estimates of treatment effects and provides improvements in efficiency in the assessment of treatment effects and other prognostic factors. It accounts for informative drop-out by linking the two processes (12,13). Therefore, joint modeling of longitudinal data from repeated CD4 cell count measurements and time-to-event data on the time to develop TB among adult PLHIV under ART is employed.

Building the models

The joint model consists of two linked submodels, known as the longitudinal submodel, and the survival submodel, as subsequently given. The Cox proportional hazard model and the linear mixed-effects model were fitted for the longitudinal and survival submodels, respectively.

The association parameter (alpha value) from a fitted joint model was used to assess the association between the time to TB disease development and longitudinal CD4 cell count change. While, a significance level of 0.25 was used to select variables in the univariable analysis, a significance level of 0.05 was used to declar presence of statistically significant association in the multivariable analysis. The Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were used for model selection.

Longitudinal submodel

Initially, the normality of the longitudinal CD4 outcome data was visualized via a normal quntile Q–Q plot and histogram. Moreover, the Breusch–Pagan/Cook–Weisberg test for heteroskedasticity was employed to test constancy of variance. A linear mixed effect model was fitted to determine factors associated with a change in CD4 cell count on repeated measurements. Graphically, profile plots were used to visualize the patterns of individual and average changes in the CD4 cell count. Significant differences in CD4 cell counts were assessed via a repeated measure ANOVA. The formula for a random intercept linear mixed effect model is as follows:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + b_{0i} + \varepsilon_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

where:

y_{ij} is the CD4 cell count for the i^{th} individual at the j^{th} time point

β_0 and β_1 are the fixed effects representing the intercept and the slope of the longitudinal trajectory, respectively.

b_{0i} is the random intercept

X -is an independent variable

ε is an error term

The random intercept (b_{0i}) indicates the extent to which each subject intercept deviates from the mean (group) intercept. The random intercept assumes that the slopes of all the subjects are equal. All lines on the profile plot are parallel for all group means. Model comparison for the longitudinal submodel was performed via the AIC and BIC.

Survival submodel

The time to develop TB disease among different categories of the predictor variables was graphically visualized via Kaplan–Meier and statistically estimated via log-rank test. The median time to develop TB was estimated and the incidence density was computed as the number of new TB cases divided by patient-months at risk. Right-censored data resulted in a median survival time that was outside of the indicated follow-up time range. Therefore, for right-hand censored data, the restricted mean survival time (the area under the curve) was calculated. This means that changing the longest survival time to the event by providing a truncation time. Truncation time is the longest survival time in the shortest survival group.

Before fitting the survival submodel, proportional hazard assumption (PHA) was checked graphically using the cumulative log hazard and statistically using the Schoenfeld residual test as follows:

1, **Graphical technique:** the most popular method of PHA test is to compare the estimated -ln(-ln) survivor curve over different categorical variables. The parallel curve between the two comparisons indicates that the PHA is satisfactory.

The formula for the Cox proportional hazard model is as follows:

$$h_{(t,x)} = h_{0(t)} e^{(\sum_{i=1}^{p}) u \beta_i x_i)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

$$\phi = \frac{h_{1(t)}}{h_{0(t)}}$$

ϕ is the hazard ratio that indicates that over time, the hazard is constant and does not vary with time.

2**. Goodness-of-fit methods:** This approach provides a large Z or chi-square value calculated for variables in the model adjusting for the other variables. The P-value is used to check the PHA for that variable. A nonsignificantly large P-value indicates (>0.05) that PHA is fulfilled. A P-value less than 0.05 indicates that the PHA is not satisfactory.

3. **Time-dependent covariates:** When time-dependent covariates are used to assess the PHA for time-independent variables, Cox is extended to include an interaction term. If the coefficient of the interaction term is not significant, the PHA is violated.

The log-rank test was used to estimate the effects of categorical variables on the time to develop TB among PLHIV under ART. Those with a P-value of < 0.25 were included in the final Cox proportional hazard model. In addition, a univariable analysis of the Cox proportional hazard model was conducted to identify significant continuous predictor variables to include in the final model.

Joint modeling

The model that best fits and accounts for the informative dropouts was formulated by considering the joint random-intercept Cox model based on separate models 1 and 2

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \mu_{0i} + \varepsilon_{ij}$$

$$h_{(t,x)} = h_{0(t)} e^{(\sum_{i=1}^{p} \beta_i x_i + y\mu_i)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(5)}$$

The random intercepts μ_i are now shared between the two models and depend on the longitudinal CD4 cell count and survival outcome of the time to develop TB. Before developing both separate linear mixed and survival models, a univariable model was fitted for each of the predictor variables, and significant variables (P-value<0.25) were considered for the multivariable analysis. Multivariable analysis helps to control potential confounders and analyze the effect of a factor in the presence of other factors. The longitudinal and survival submodels were fitted jointly, using the current value parameterization. Model with the smallest AIC and BIC was considered the best-fit.

The association parameter (alpha value) from the fitted joint model was used to assess the association between longitudinal CD4 cell count changes and the time to develop TB among PLHIV under ART. This value indicates the value of the CD4 cell count at which TB disease occurs, and for any change in the CD4 cell count, the time to develop TB disease was predicted. In the multivariable joint modeling, a covariate with a P-value of < 0.05 that was contained in the 95% confidence interval was considered a significant covariate. In the final Cox proportional hazard model, the hazard ratio was interpreted as follows: for any change or unit increase in the predictor, the hazard of TB development increased or decreased with increasing hazard ratio from the baseline. For categorical predictors, it was interpreted by comparing the hazard of TB development in one category of the variable to the reference category, while holding others constant.

4.11. Ethical considerations

This research was conducted after ethical approval was obtained from the institutional review board of Mekelle University, College of Health Sciences. Since the study relied upon secondary data that were available at both ART clinics, the need for participant’s consent was waived by institutional review board of Mekelle University, College of Health Sciences. A permission letter was also obtained from both MGH and ACSH hospitals. The data were kept confidential and used for this purpose only, and the study subjects were anonymous.

4.12. Plan for dissemination of findings

After completion of the thesis, the report will be disseminated to Mekelle University, College of Health Science, School of Public Health; Tigray Regional Health Bureau, MGH; ACSH; and reputable journals for publication.

5. RESULTS

5.1. Descriptive Statistics

The study was conducted among 449 adult HIV patients on ART, of which 300 were from MGH and 149 were from ACSH. The participants were followed for a minimum of 1 year and a maximum of 6 years. Descriptive statistics such as frequency, mean, standard deviation, and range were used to describe the study population.

Approximately 62% of the total study participants were female, and 91% were orthodox Christians. In terms of marital status, half (51%) of the study participants were married. With respect to education, the majority (84%) of the study participants had primary, secondary, or tertiary education. The majority (74%) of the study participants resided in urban areas. More than two-thirds (69%) of the study participants had a job (for more see Table 2).

Table 2: Sociodemographic characteristics of PLHIV under ART follow-up at ACSH and MGH, Mekelle, Ethiopia, 2018–2024

Characteristics	**Category**	**TB status**		**Total (%)**
		Event (%)	**Censored (%)**	
Sex	Female	60(55.56)	219(64.22)	279(62.14)
	Male	48(44.44)	122(35.78)	170(37.86)
Health facility	ACSH	38 (35.2)	111 (32.55)	149 (33.18)
	MGH	70 (64.8)	230 (67.45)	300(66.82)
Religion	Catholic	3(2.78)	6(17.76)	9(2.00)
	Muslim	8(7.41)	22(6.45)	30(6.68)
	Orthodox	97(89.81)	313(91.79)	410(91.31)
Marital Status	Divorced	24 (22.22)	55 (16.13)	79(17.59)
	Married	58(53.70)	171 (50.15)	229(51.00)
	Single	16(14.81)	92(26.98)	108(24.05)
	Widowed	10(9.26)	23(6.74)	33(7.35)
Educational Status	No formal education	17(15.74)	48(14.08)	65(14.48)
	Primary school (1-8 grade)	62(57.41)	155(45.45)	217(48.33)
	Secondary school (9-12 grade)	22(20.37)	102(29.91)	124(27.62)
	Tertiary school	7(6.48)	36(10.56)	43(9.58)
Job	Yes	71(65.74)	238(69.79)	309(68.82)
	No	37(34.26)	103(30.21)	140(31.18)
Residency	Rural	26 (24.07)	89(26.01)	115(25.61)

	Urban	82(75.93)	252(73.9)	334(74.39)

Regarding the ART regimens, 76% of the study participants were on the 1E [1]regimen. More than half (57%) of the study participants had good ART medication adherence. While almost all (97%) of the participants had no previous history of TB infection, the majority (73%) had a working functional status. While 83% of the study participants had a history of CPT utilization, 60% had no history of use of isoniazid preventive therapy (IPT). During admission, more than half (60%) of the study participants were in WHO clinical stages one and two (early stage). Only approximately one-third (33%) of the study participants had baseline opportunistic infections (for more see Table 3 below).

Table 3: Clinical characteristics of PLHIV under ART follow-up at ACSH and MGH, Mekelle, Ethiopia, 2018--2024

Characteristics	**Category**	**TB status**		**Total (%)**
		Event (%)	**Censored (%)**	
Type of ART medication (Regimen)	1E	75(69.40)	265 (77.71)	340(75.72)
	1J[2]	31(28.70)	75(21.99)	106(23.61)
	1C[3]	2(1.85)	1(0.29)	3(0.67)
ART Medication Adherence level	Fair	28(25.93)	47(13.78)	75(16.70)
	Good	53(49.07)	204(59.82)	257(57.24)
	Poor	27(25.00)	90(26.39)	117(26.06)
Previous History of TB infection	No	1(0.93)	1(0.29)	2(0.45)
	Yes	1(0.93)	12(3.52)	13(2.90)
CPT	No	23(21.30)	53(15.54)	76(16.93)
	Yes	85(78.70)	288(84.46)	373(83.07)
Isoniazid Preventive Therapy (IPT)	No	78(72.22)	189(55.43)	267(59.47)
	Yes	30(27.78)	152(44.57)	182(40.53)
Baseline WHO clinical stage	Stage I	22(20.37)	185(54.25)	207(46.10)
	Stage II	25(23.15)	35(10.26)	60(13.36)
	Stage III	29(26.85)	65(19.06)	94(20.94)
	Stage IV	32(29.63)	56(16.42)	88(19.60)

[1] 1E is a fixed-dose combination of Tenofovir disoproxil fumarate (TDF) + Lamivudine (3TC)+ Efavirenz (EFV)

[2] 1J is a fixed-dose combination of Tenofovir disoproxil fumarate (TDF) + Lamivudine (3TC)+ Dolutegravir (DTG)

[3] 1C is a fixed-dose combination of Zidovudine (AZT) + Lamivudine (3TC) + Nevirapine (NVP)

Baseline functional status	Ambulatory	31(28.70)	58(17.01)	89(19.82)
	Bedridden	17(15.74)	16(4.69)	33(7.35)
	Working	60(55.56)	267(78.30)	327(72.83)
Baseline presence of opportunistic infection	No	40(37.04)	260(76.25)	300(66.82)
	Yes	68(62.96)	81(23.75)	149(33.18)

The median age and BMI of the participants at admission were 39 years and 16.8, respectively. The study participants had a median hemoglobin level of 12.7 g/dL, with an IQR of 11.9–13.5 g/dL. The median baseline CD4 cell count of the study participants was 452 cells/$mm^{3,}$ with an IQR of 268–601 cells/mm^3 (for more see Table 4).

Table 4: Descriptive statistics for continuous variables among PLHIV under ART follow-up at ACSH and MGH, Mekelle, Ethiopia, 2018–2024

Variables	Median	IQR
Age	39	14
BMI	16.77	2.62
Hemoglobin	12.7	1.6
CD4 baseline	452	333

5.2. Survival submodel

Incidence of TB disease

A total of 449 patients were followed for a minimum of 12 months and a maximum of 72 months, with a median follow-up time of 24 months [IQR: 30]. Among the total 449 study participants, 108 (24.05%) were diagnosed with TB disease. The TB incidence rate was 6.77 cases/100 person-years of observation. In this study, since majority of the study participants haven't developed TB in the follow-up period, the overall median survival time was not observed. Instead, by using the truncation time of 72 months, the overall restricted mean survival time was calculated as 60 months (see Figure 2).

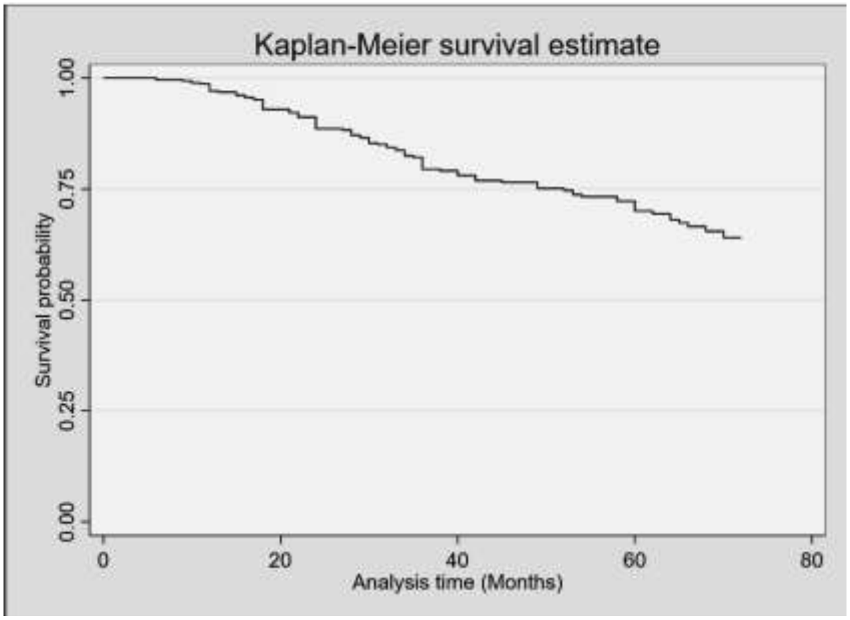

Figure 2: Overall Median survival time to develop TB in PLHIV under ART follow-up at ACSH and MGH, Mekelle, Ethiopia, 2018–2024

Kaplan–Meier curve and Log-rank test

Graphically, the Kaplan–Meier hazard estimate indicates that there seems a significant difference in the survival functions of sex, history of IPT utilization, history of CPT utilization, previous history of TB, baseline functionality, and baseline presence of opportunistic infection.

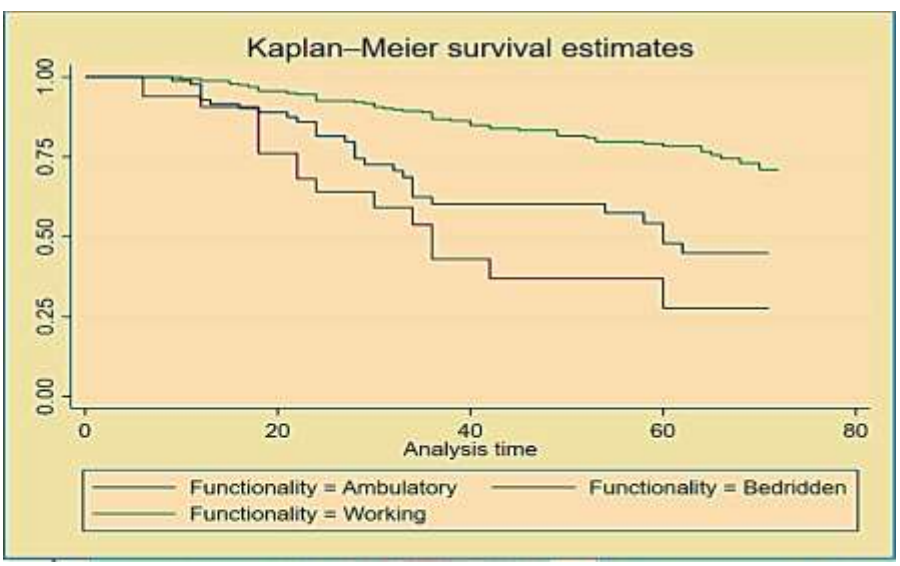

Figure 3: Kaplan–Meier survival curve by functionality level of PLHIV under ART in ACSH & MGH, Mekelle, Ethiopia, 2018-2024

Statistically using logrank test, there was a significant difference between the female and male patient groups (X^2=6.63, P-value=0.01). These findings suggest that males have a shorter survival time than females do. In terms of marital, even though the Kaplan–Meier curve does not show a difference in the survival function of PLHIV, the log-rank test revealed a significant difference in the survival distributions across different categories of marital status (X^2=7.94, P-value=0.0473). Even though the Kaplan–Meier curve does not show a difference in the survival function of PLHIV with regard to educational status, the log-rank test revealed a significant difference in the survival distributions of the educational level (X^2=7.85, P-value=0.0493).

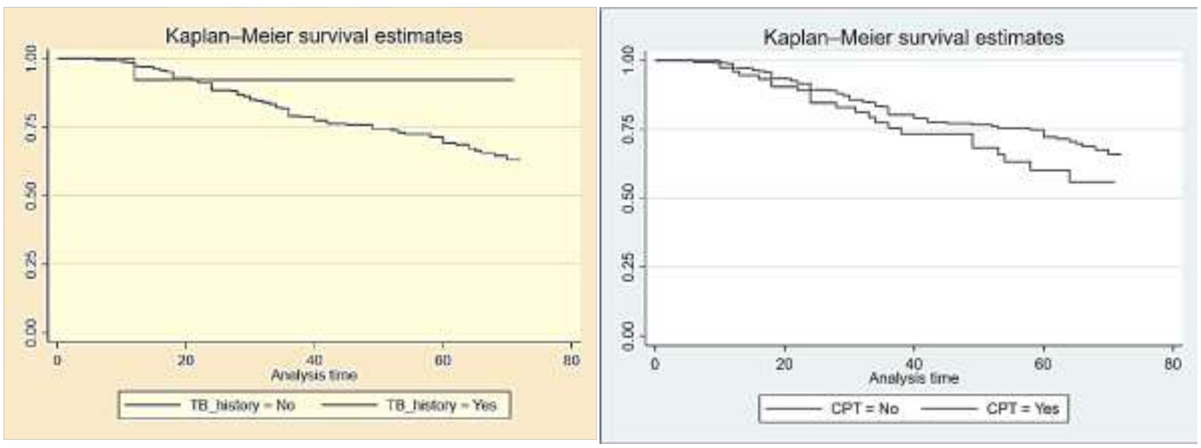

Figure 4: Kaplan–Meier survival curve by previous TB history and CPT utilization history of PLHIV under ART in ACSH & MGH, Mekelle, Ethiopia, 2018-2024

The log-rank test indicated that there was a statistically significant difference in the survival functions across the adherence levels (between good and fair and between fair and poor) (X^2=34.61, P-value=0.00) and functionality levels (between working and ambulatory; and between working and bedridden) (X^2=45.07, P-value=0.00), history of IPT utilization (X^2=15.14, P-value=0.0001), WHO clinical stage (between stage I and II, between stage I and III, and between stage I and IV) (X^2=50.13, P-value=0.00), and baseline history of opportunistic infection (X^2=80.07, P-value=0.00) (see Figure 5).

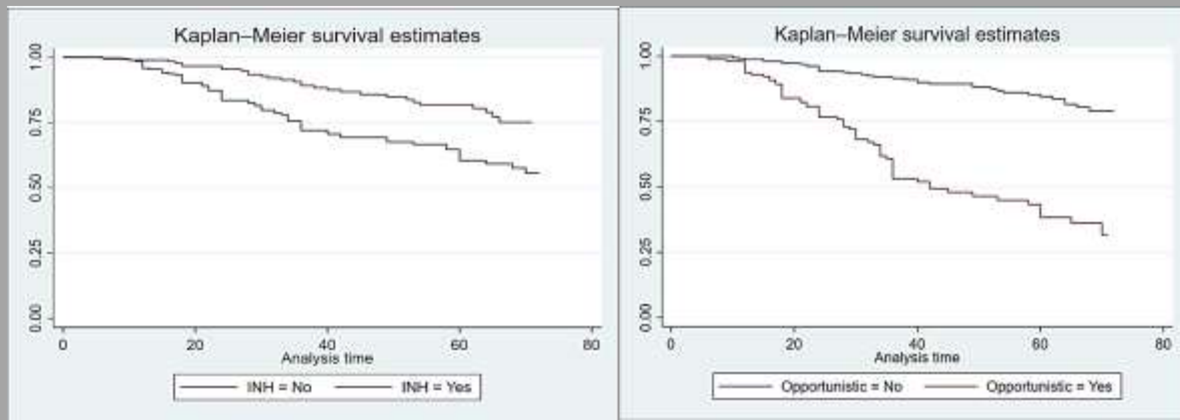

Figure 5: Kaplan–Meier survival curve by IPT utilization and presence of opportunistic infection of PLHIV under ART in ACSH & MGH, Mekelle, Ethiopia, 2018-2024

However, there was no statistically significant difference in survival time among the different groups in terms of religion (X^2=2.43, P-value=0.2963), job availability (X^2=1.17, P-value=0.28), residence (X^2=0.00, P-value=0.9513), ART regimen (X^2=4.71, P-value=0.0951), previous TB history (X^2=2.09, P-value=0.1486), and CPT utilization (X^2=2.6, P-value=0.1069).

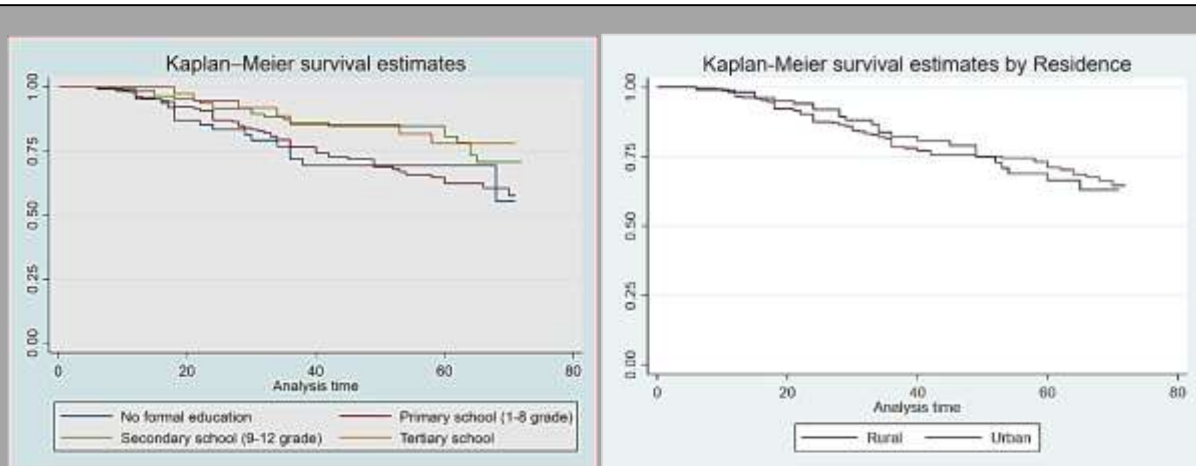

Figure 6: Kaplan–Meier survival curve by educational status and residence of PLHIV under ART follow-up in ACSH & MGH, Mekelle, Ethiopia, 2018-2024

Cox-Proportional Hazard Assumption Checking

The PHA was checked for all the independent variables both graphically via the cumulative hazard plot and statistically via the Schoenfeld residual test before conducting the Cox-regression analysis.

Graphically (as depicted in Figure 7), the PHA for the variables sex, CPT, and opportunistic infection was fulfilled. Statistically, the Schoenfeld residual test confirmed that PHAs were satisfied for all the independent variables, except adherence (P-value=0.02).

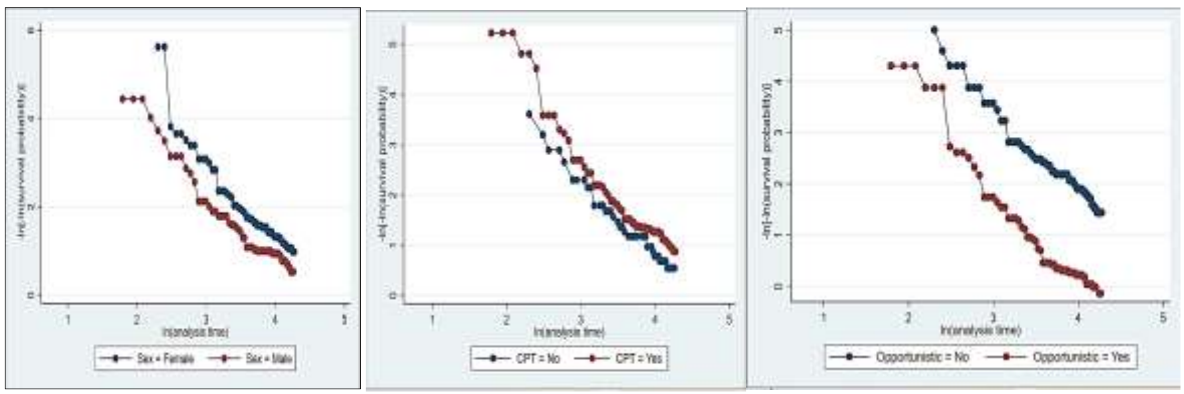

Figure 7: Cumulative hazard plot of Sex, CPT utilization, and presence of opportunistic infection among PLHIV under ART in ACSH and MGH, Mekelle, Ethiopia, 2018-2024

Using the Schonefeld residual test, all the variables listed in the Table 5 were found to satisfy the PHA.

Table 5: Schonefeld PHA testing of independent variables among PLHIV under ART follow-up at ACSH and MGH, Mekelle, Ethiopia, 2018–2024

Variables	x^2	P-value	Variables	x^2	P-value
sex	0.8	0.37	BMI	0.02	0.89
Age	0.07	0.79	Religion	1.79	0.41
Hemoglobin	<0.001	0.98	Marital status	0.28	0.96

Baseline CD4	2.98	0.08	Educational status	2.04	0.56
ART type	2.99	0.22	Residence	1.93	0.17
Occupation	13.10	0.07	CPT	0.09	0.77
TB history	1.59	0.21	IPT	1.88	0.17
WHO clinical stage	5.75	0.12	Functional status	1.44	0.49
Baseline presence of OI	0.43	0.51			

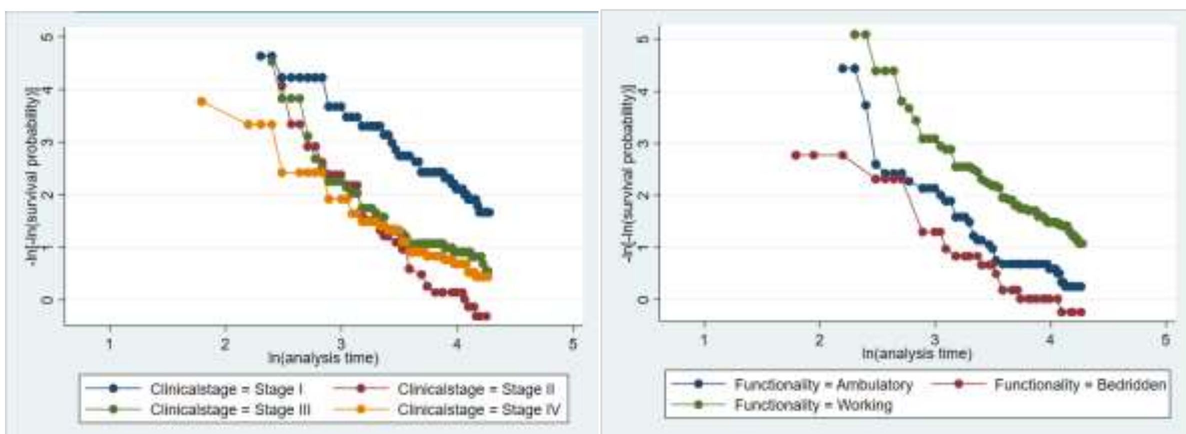

Figure 8: Graphical plot of PHA checking of WHO clinical stage and functionality level among PLHIV under ART in ACSH and MGH, Mekelle, Ethiopia, 2018-2024

Time-Varying Covariates

After the Schoenfeld residual test, adherence was found to violate the PHA assumption (P-value=0.0183). To solve this failure of the assumption, a time-interaction test for adherence was performed.

Univariable Cox-PH Model

Univariable Cox–PH regression models were employed to assess the impact of each independent variable on the time to develop TB among adult HIV patients on ART. Variables that had a P-value of less than or equal to 0.25 for the Wald Z-statistic of the log-hazard ratio in the univariable analysis were considered in the multivariable Cox model. .

On the basis of the univariable analysis, age, BMI, hemoglobin, baseline CD4, sex, ART type, adherence, religion, marital status, educational status, previous history of TB, use of CPT, use of IPT, WHO clinical stage, functional status, and baseline presence of opportunistic infection variables were included in the multivariable Cox analysis.

Table 6: Uni-variable analyses of predictors of time to develop TB among adult PLHIV under ART at ACSH and MGH, 2018--2024

Variable	Category	HR	Std.err	P>|z|	95% CI	
Age		1.01578	.0088442	0.072	.9985927	1.033263
BMI		.867	.036	.001	.799	.941
Hemoglobin		.751	.039	<0.001	.678	.832
CD4_baseline		.995	.001	<0.001	.994	.996
Sex	Female	.607	.119	.011	.412	.892
	Male (Ref.)	1				
ART regimen type	1E	.241	.173	.048	.059	.986
	1J	.244	.179	.054	.058	1.026
	1C (Ref)	1				
Adherence level	Fair	4.331	1.005	<0.001	2.749	6.824
	Poor	4.675	1.294	<0.001	2.718	8.042
	Good (Ref)	1				
Religion	Muslim	.64	.434	.51	.17	2.414
	Orthodox	.467	.274	.195	.147	1.478
	Catholic (Ref)	1				
Marital Status	Divorced	1.117	.272	.649	.693	1.801
	Single	.479	.143	.014	.267	.86
	Widowed	1.113	.399	.766	.551	2.248
	Married (Ref)	1				
Educational status	No formal education	2.095	.943	.1	.867	5.061
	Primary school	2.047	.82	.074	.934	4.489

	Secondary school	1.169	.511	.72	.497	2.752
	Tertiary school (Ref)	1				
Residence	Urban	1.009	.235	.969	.639	1.593
	Rural	1				
Previous history of TB	No	3.783	3.802	.186	.528	27.118
	Yes (Ref)	1				
Use of CPT	No	1.454	.344	.114	.915	2.311
	Yes (Ref)	1				
Use of IPT	No	2.284	.501	<0.001	1.486	3.51
	Yes (Ref)	1				
WHO clinical stage	Stage II	5.449	1.569	<0.001	3.099	9.58
	Stage III	6.779	1.909	<0.001	3.903	11.772
	Stage IV	10.718	3.027	<0.001	6.162	18.642
	Stage I (Ref)	1				
Functional Status	Ambulatory	2.759	.623	<0.001	1.773	4.295
	Bedridden	4.674	1.327	<0.001	2.68	8.153
	Working (Ref)	1				
Baseline Presence of OI	No	.193	.04	<0.001	.129	.289
	Yes (Ref)	1				

Multivariable Cox Model

In the multivariable analysis at the 5% significance level, the baseline CD4 count, sex, use of CPT, WHO clinical stage, adherence, and presence of opportunistic infection were the variables that were found to be significantly associated with the time to develop TB among adult HIV patients on ART. Multicollinearity was assessed at each stage, and variables with a variance inflation factor (VIF) greater than 10 were excluded from the model. The variable with the smallest Z-value was removed first, leading to the exclusion of age, BMI, hemoglobin, religion, educational status, marital status, use of IPT, ART type, functional status, and previous TB history from the model (see Table 7 below).

With respect to sex, being female was associated with a 42.6% reduction in the hazard of developing TB compared with being male (HR=0.57, 95% CI: 0.32, 0.82). On the other hand, PLHIV with no history of opportunistic infection had a 51.8% lower hazard of developing TB than their counterparts did [HR=0.482, 95% CI: 0.28–0.82], while other factors remained constant.

The other factor that showed a significant association was the baseline WHO clinical stage. Those PLHIV on ART with baseline WHO stage II disease had an approximately 2-fold [HR=2.28, 95% CI: 1.12--4.66] greater risk of developing TB than those with WHO clinical stage one disease. Patients who had a history of as good ART drug adherence was interacting with were observed to have a 1.1% lower hazard of developing TB [HR=0.989; 95% CI: 0.978–0.999] than their counterparts, keeping other factors constant. Moreover, those patients who had no history of CPT utilization at baseline had an approximately 2.3-fold greater hazard of developing TB than did those with a history of CPT utilization [HR=2.33, 95% CI: 1.42--3.83].

Table 7: Multivariable Cox analysis of predictors of time to develop TB among adult PLHIV on ART at ACSH and MGH, Mekelle, Ethiopia, 2018--2024

Variable	**Category**	**HR**	**Std.err**	**P>|z|**	**95% CI**	
CD4_base_sqrt		.901	.017	<0.001	.869	.934
Sex (Female)	Female	.574	.119	.008	.382	.863
	Male (Ref)	1				
Use of CPT	No	2.33	.59	.001	1.42	3.83
	Yes (Ref)	1				
Use of IPT	No	1.33	.309	.212	.848	2.099
	Yes (Ref)					
WHO clinical stage	Stage II	2.28	.832	.024	1.117	4.662
	Stage III	1.64	.58	.164	.817	3.28
	Stage IV	1.53	.578	.258	.731	3.211
	Stage I (Ref)	1				
ART drug Adherence (tvc)		.989	.005	.047	.979	1
Baseline presence of OI	No	.482	.131	.007	.283	.821
	Yes (Ref)	1				

Model Comparison

When the overall model fitness was compared, the reduced model with time interaction by adherence was most fitted, where the AIC and BIC values were 1043.136 and 1097.714, respectively. Therefore, the reduced model with time interaction was found to be the best-fitting model since it had the lowest AIC and BIC values.

Table 8: Model Comparison of the survival model

COX-PH Model	AIC	BIC
Reduced Model with time interaction	1043.136	1097.714
Full Model	1043.4	1195

5.3. Longitudinal submodel (linear mixed effect model)

Exploratory Analysis

The normality assumption for the longitudinal record of the CD4 cell count was visualized via a histogram, which revealed a positively skewed distribution with outliers. To satisfy the normality assumption, a square root transformation was applied, as it resulted in the lowest Chi² value among the tested transformation options, indicating improved normality.

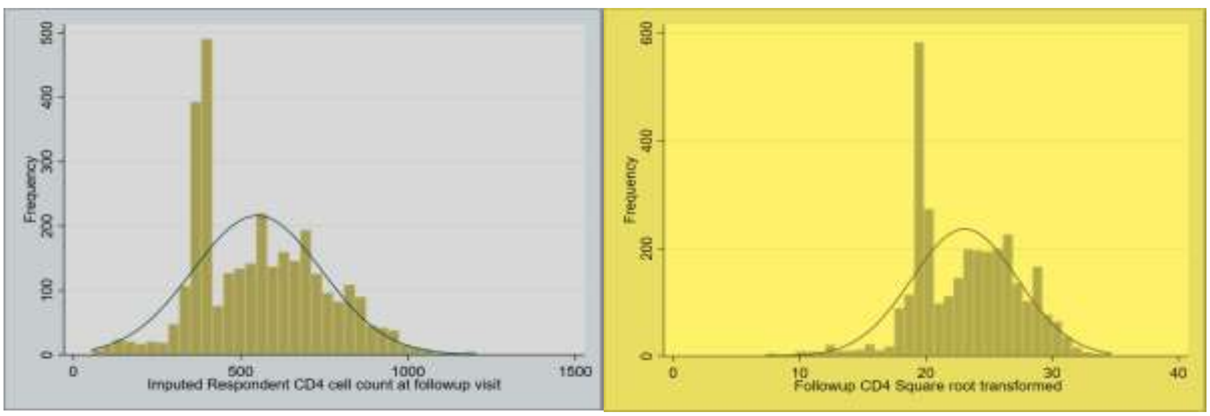

Figure 9: Histogram plot for normality assumption test of CD4 cell count before and after square root transformation

The residual versus fitted values plot (rvfplot) was used to assess the linearity and constancy of the variance. Since the residuals are concentrated aound the zero residuals line, it fulfills the linearity and constancy of variance assumption and is not funnel shaped in the distributions of the residuals. The normality assumption was visualized via a quantile–quantile (Q-Q) plot and histogram. The Q–Q plot indicated that the normality assumption was fulfilled with heavy tails, meaning that the distribution has a high probability on both tails. The Breusch–Pagan/Cook–Weisberg test for heteroskedasticity showed that the equality of variance assumption is fulfilled

(x^2=2.73, P-value=0.0987). Missing values were managed via the MICE method. The pattern of Missingness was diagnosed using the logistic regression, and was found to be missing at random. On the other hand, multicollinearity was checked at every step, and variables with a VIF>10 were removed from the model.

,

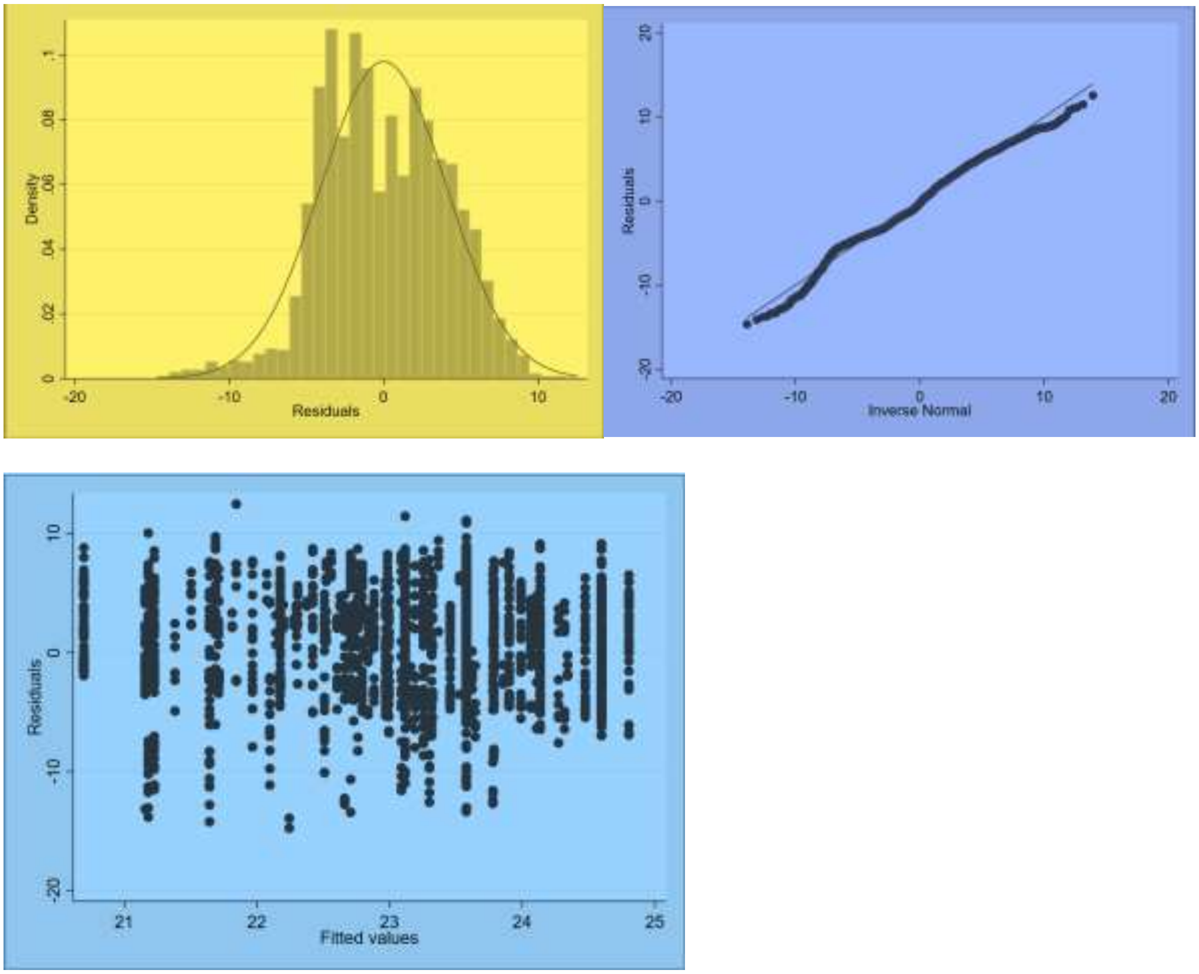

Figure 10: Histogram, Q–Q plot, and rvf plot for equality of variance and normality of variance assumption test of longitudinal CD4 cell count

Table 9: Summary of longitudinally measured CD4 at every visit among PLHIV under ART at ACSH and MGH, Mekelle, 2018–2024

time_point	N	Mean	SD
1	449	546.0958	193.1931
2	449	552.1359	188.2614
3	413	546.431	189.846
4	377	557.1379	203.2499
5	318	532.3836	196.5985
6	271	531.8487	181.6782
7	220	551.3727	189.1788
8	191	560.4084	202.246
9	165	556.5636	184.8552
10	144	578.2361	196.1002
11	112	557.2321	202.4417
12	69	554.1304	175.9155
Total	3178	549.5079	192.626

In summary, the mean CD4 cell count slightly increased over time, which indicated that there might be an improvement in the CD4 cell count because the patients were receiving treatment.

Graphs were used to visualize the patterns of individual profile plots and mean profile plots. This graphical visualization is used to show how individual CD4 levels vary from the overall mean CD4 value. Individuals varied more at baseline from the mean than through time because the individual profile plots were parallel through time, as we can see from the graph below. The mean CD4 value at the start of follow-up was 546 cells/mm^3.

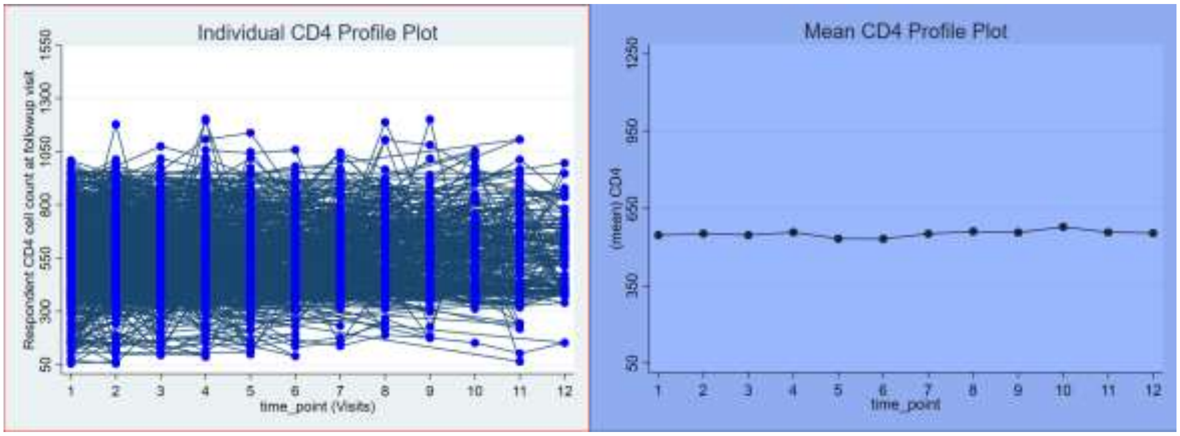

Figure 11: Individual and mean profile plot of CD4 levels of PLHIV under ART in ACSH and MGH, Mekelle, Ethiopia, 2018-2024

Univariable Longitudinal Analysis

Variables with P- values < 0.25 in the univariable longitudinal analysis were included in the multivariable analysis. Those were age, CD4_base_sqrt, hemoglobin, sex, ART medication type, adherence, opportunistic infection, religion, WHO clinical stage, and availability of job.

Multivariable Longitudinal Analysis

The standard deviation of the random intercept across different subjects was 3.2 [95% CI: 2.94--3.42], which showed that there was significant variability in the baseline CD4 counts between individuals. On the other hand, the standard deviation of the random slope (effect of time on CD4 count) across different subjects was 0.00000004, which means that the effect of time on CD4 count variation between individuals is insignificant. Since the longitudinal CD4 cell count was square root transformed, caution must be taken in directly interpreting the result as the effect of the predictors on the original CD4 cell count is not a constant additive difference, but rather changes depending on the level of CD4 cell count.

From the multivariable analysis, occupational status was found to be statistically significant. On average, the mean $\sqrt{\text{CD4 cell count}}$ for PLHIV with no job was 0.7 units higher than for PLHIV with job, after accounting for the effect of other variables. In addition, individuals with poor adherence, compared to those with good ART adherence, had significantly lower mean CD4 cell count measurement. This means that the mean $\sqrt{\text{CD4 cell count}}$ for PLHIV with poor adherence was decreased by 1.6 units, as compared to PLHIV with good adherence, keeping others constant.

Table 10: Multivariable longitudinal analysis of CD4 cell count change among adult PLHIV on ART at ACSH and MGH, Mekelle, Ethiopia, 2018-2024

Variable	**Category**	**Coef.**	**St.Err.**	**P-value**	**95% CI**	
Observed Time (Months_visit)		.002	.003	.467	-.004	.008
Sex	Female	.532	.339	.117	-.134	1.497
	Male (Ref)					
Adherence level	Fair	-.299	.475	.53	-1.23	.632
	Poor	-1.583	.479	.001	-2.522	-.645
	Good (Ref)	1				
Baseline WHO stage	Stage II	-.376	.572	.511	-1.497	.745
	Stage III	-.491	.558	.379	-1.584	.602
	Stage IV	-1.056	.608	.083	-2.247	.136
	Stage I (Ref)					
Baseline OI	No	-.135	.451	.764	-1.02	.749
	Yes (Ref)					
Availability of Job	No	-.706	.36	.05	-1.411	-.001
	Yes (Ref)					

Intercepts

Random effect	**Estimate**
Sd (Intercept)	3.167799
Sd (Slope)	.0.00000004
Sd (Residual) error	2.731261

Model comparison

Model comparison was performed between the random intercept only model and random intercept and slope model. The random intercept only model was found to be the best-fit model since it had lower AIC and BIC values. The AIC and BIC of the reduced model were 16438.44 and 16511.21, respectively (for more see Table 11 Table 13 Table 14).

Table 11: Model comparison of the multivariable longitudinal analysis

Model	AIC	BIC
Random intercept model	16438.44	16511.21
Random intercept and slope model	16440.44	16519.27

5.4. Joint Modeling Analysis

In the joint modeling analysis, the longitudinal measure of the CD4 count was jointly fitted with time to develop TB of the Cox-PH model. In this analysis, the associations between the longitudinal and time–to-event models were assessed, and there was a mean change in the CD4 count in patients with TB compared with patients without TB (censored), as shown in the graph below. This means that as the mean CD4 count increased, the hazard of TB infection decreased. However, in the censored group, there was no change in the mean CD4 count, and no events occurred, as shown in the figure below. Therefore, joint modeling analysis was better than a separate analysis of survival and a longitudinal submodel.

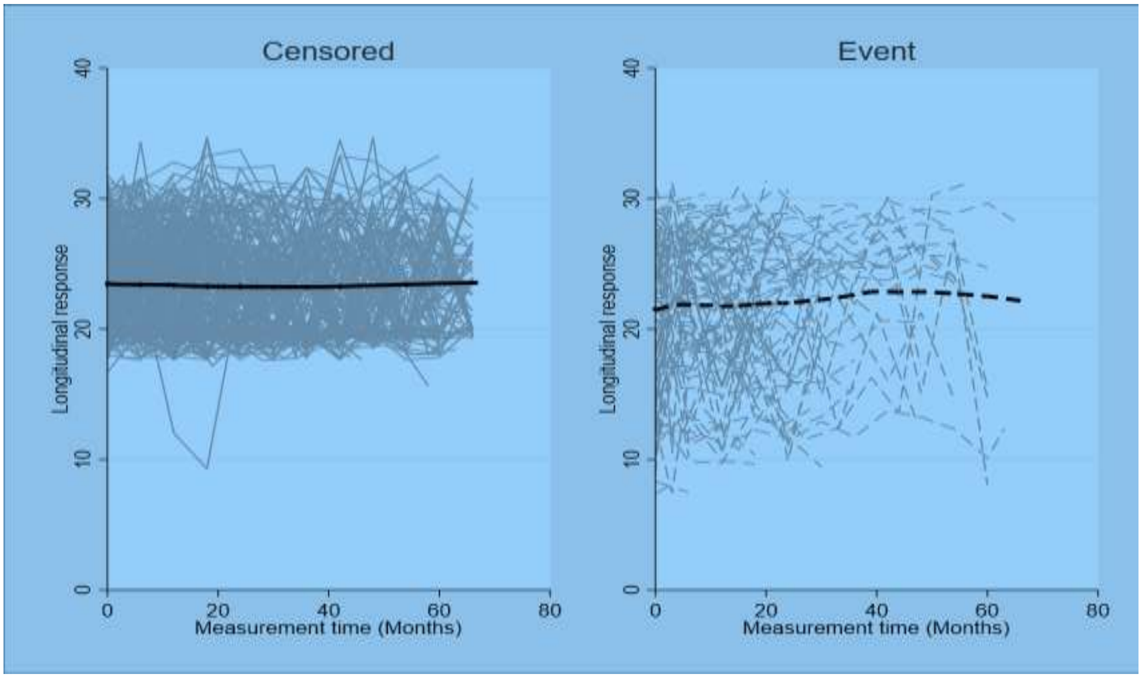

Figure 12: The association between longitudinal CD4 cell count and time to develop TB among PLHIV under ART in ACSH and MGH, Mekelle, Ethiopia, 2018-2024

Result Interpretation in Joint Model Analysis

The results were interpreted via the HR and 95% CI, including the P-value. This analysis revealed an association parameter, alpha, with a HR of 0.854 (95% CI: 0.8, 0.91), indicating a statistically significant association between $\sqrt{\text{CD4 cell count}}$ and the time to develop TB. This means that for a unit increase in the average $\sqrt{\text{CD4 cell count}}$, the hazard of TB infection decreased by 14.6%, while keeping others constant. The lower the CD4 count is, the greater the probability of TB infection. The restricted cubic spline function indicates the variation in CD4 counts within a patient throughout the longitudinal measurement time. From a sample of three measurements of CD4 count the two had significant differences from the individual mean CD4 count, and this difference had a statistically significant effect on the hazard of TB infection.

The calculated hazard proportion of the WHO clinical stages of HIV was statistically significant, with HR = 1.024 (95% CI: 1.017–1.033), meaning that patients in advanced WHO clinical stages had a 2.4% greater hazard of TB infection than those in stage I HIV, keeping others constant. The calculated hazard ratio for sex was significantly associated with the time to TB infection, with a HR of 1.62 (95% CI: 1.09, 2.4). This means that in female patients, the hazard of TB infection is 2.4 times greater than that in male patients, holding others constant.

The calculated hazard ratio of CPT intake was significantly associated with the time to TB infection, with a HR of 0.55 (95% CI: 0.35, 0.89). This means that not taking CPT prophylaxis decreases the hazard of TB infection by 45% compared with those who take CPT while keeping others constant. The calculated hazard ratio for adherence was significantly associated with the time to TB infection, with a HR of 0.38 (95% CI: 0.28, 0.52). This means that having good adherence decreases the hazard of TB infection by 62% compared with those who have poor ART drug adherence, keeping others constant.

Random Effect Part

The random effect parameters in the joint modeling described how the variation in the longitudinal CD4 count contributed to the hazard of TB infection. The random intercept model indicated that the greater variation in CD4 counts at baseline contributed strongly to the hazard of time to develop TB. The value of the standard deviation in the random intercept was 3.2. Therefore, individual variation at baseline leads to variation in the time to TB infection.

The standard deviation for the random slope model was 0.09e-10. This finding indicated a very small variation in the time to TB infection, which was attributed to the variation in the CD4 count over time between patients. Therefore, the random intercept model was the best-fit model that contributed to the greater variation in the hazard of TB infections. The correlation coefficient between the random intercept and slope was -0.497, indicating a moderate negative correlation between the random intercept and slope. This means that CD4 counts vary at baseline and do not vary linearly over time. The standard deviation of the residual was 2.74, indicating a variation between the observed CD4 count and the predicted value.

Model Comparison

The hazard of developing TB among PLHIV under ART was assessed via three models, namely, survival, longitudinal, and joint modeling of longitudinal survival data. Among these models, joint modeling analysis was the best-fit model because the associations between the longitudinal trajectories and the time to develop TB were strong. Therefore, the final model was fitted via joint modeling.

Final Model

The final model was a joint random intercept Cox-PH model and was built from a joint model analysis of the Cox-PH submodel and longitudinal submodel. The model included association parameters, random slopes, random intercepts, and regression coefficients for independent variables from the joint model. It is written as follows in the table below:

$$\text{CD4} = 24.96 - 0.738\text{Adherence} + 0.061\text{CD4\_base\_sqrt} - 0.363\text{WHOClinical Stage} + 3.2$$

$$\text{h(TB)(t)} = \text{h}_{0(t)}\text{e}^{(0.024\text{WHOclinical stage} + 0.483\text{Sex} - 0.588\text{CPT} - 0.953\text{Adherence} - 0.158\sqrt{\text{CD4 cell count}})}$$

Table 12: Joint model analysis of time to develop TB associated with longitudinal CD4 cell count among PLHIV under ART at ACSH and MGH, Mekelle, Ethiopia, 2018--2024

Longitudinal part	**Category**	**Coeff.**	**Std. err**	**P>|Z|**	**95% CI**	
Observed time		0.000	<0.001	0.367	-0.000	0.000
_time_1_BMI		-0.000	<0.001	0.406	-0.000	0.000
CD4_base_sqrt		0.061	0.030	0.040	0.003	0.120
Availability of job	Yes	0.683	0.359	0.057	-0.021	1.387
	No (Ref)					
Adherence	Poor	-0.738	0.277	0.008	-1.281	-0.195
	Good (Ref)					
Sex	Female	-0.548	0.346	0.113	-1.226	0.130
	Male (Ref)					
BMI		-0.084	0.074	0.255	-0.228	0.061
WHO Clinical Stage	Stage IV	-0.363	0.168	0.030	-0.692	-0.035
	Stage I (Ref)					
CPT	No	0.245	0.433	0.571	-0.604	1.094
	Yes (Ref)					
Intercept		24.961	1.706	0.000	21.617	28.305

Random effect Parameters	**Estimates**	**Std.err**	**95% CI**	
Sd (Slope, time)	7.09e-10	3.75e-10	2.51e-10	2.00e-09
sd(Intercept)	3.202	0.124	2.969	3.454
Correlation	-0.497	0.357	-0.900	0.366
sd(Residual)	2.741	0.038	2.668	2.815

Survival Part							
Parameter		**HR**	**Coeff**	**Std. err**	**P>|Z|**	**95% CI**	
Who Clinical Stage	Stage IV	1.024	0.024	0.004	0.000	0.017	0.032
	Stage I (Ref)	1					
Association(α)		0.854	-0.158	0.031	0.000	-0.219	-0.097
Sex	Female	1.62	0.483	0.201	0.016	0.090	0.876
	Male (Ref)	1					
CPT	No	0.55	-0.588	0.239	0.014	-1.056	-0.119
	Yes (Ref)	1					
Adherence	Poor	0.38	-0.953	0.150	0.000	-1.246	-0.660
	Good (Ref)	1					
Rcs1		2.29	0.831	0.190	0.000	0.459	1.203
Rcs2		1.46	0.380	0.149	0.011	0.087	0.673
Rcs3		0.87	-0.135	0.110	0.223	-0.351	0.082
Constant		0.53	-0.637	0.918	0.488	-2.436	1.162

Loglikelihood value = -8806.4669

6. DISCUSSION

The main objective of this study was to predict the associations between the time to develop TB and longitudinal CD4 cell count and associated factors among PLHIV under ART follow-up at ACSH and MGH from 2018–2024.

The study revealed that the incidence density of TB disease among PLHIV under ART was 6.77 cases/100 person-years. This result is higher than that of a study performed in Addis Ababa, Ethiopia, which reported a 3.08 cases/1000 person-year incidence rate of TB disease (25), and it is slightly lower than those reported in studies performed in Addis-Ababa and Afar, which reported rates of 6.82 and 8.6 per 100 person-years of observation, respectively (10,19). This variation may be due to the improved HIV case management of the find-and-treat strategy.

Moreover, this study revealed that the restricted mean survival time to develop TB disease was found as 60 months. This finding is similar to a study performed in Amhara, Ethiopia, which reported that the mean survival time to develop TB was 60.8 months (22). This similarity may be attributed to the similar study period of six years. This was higher than a study done in Northeastern Ethiopia (54 months) (9). This may be attributed to a difference in study period.

The significant factors associated with the risk of TB disease were WHO clinical stage, sex, history of CPT intake, adherence, and longitudinal CD4 cell count. This study revealed that for a unit increase in the average square root CD4 cell count, the hazard of TB disease development decreased by 14.6%, keeping others constant. This was lower than a study performed in Iran which reported an increasing trajectory of CD4 cell count was associated with a 68% decreased risk of TB co-infection (21). This variation may be attributed to a difference in study type, setting, and period.

In this study, female patients had a 62% greater risk of TB disease than their male counterparts did. This finding contrasts with those of a systematic review and meta-analysis performed in SSA and a study performed in Thailand, which revealed that male patients were approximately 1.43 and 1.4 times more likely to develop TB disease than their female counterparts, respectively (8,25). This may be due to variations in the study setting and period.

Moreover, this study revealed that PLHIV with an advanced WHO clinical stage had a 2.4% greater likelihood of having TB disease than those with WHO clinical stage one. This result is lower than a systematic review and meta-analysis performed in SSA, in which PLHIV with advanced WHO clinical stages were twice as likely to be TB coinfected (8). Moreover, it is also lower than several studies performed in Ethiopia, in which PLHIV under ART with an advanced WHO clinical stage (stages III and IV) were approximately seven and three times more likely to be coinfected with TB, respectively (19,24). This may be due to variation in analysis methodology, the study setting and period.

In this study, a history of CPT utilization was found to be an important predictor of TB disease among PLHIV under ART. The PLHIV with no history of CPT utilization had a 45% decreased hazard of TB disease. This was similar to a study done in South Africa where no use of CPT was associated with a 42% decrease in TB incidence (27). This may be because PLHIV with advanced WHO clinical stages are more likely to utilize CPT than those with early WHO clinical stages are. Moreover, this study revealed that PLHIV who had good ART adherence had a 62% lower hazard of TB disease than did those with poor ART drug adherence. This is similar to a study performed in northwestern Ethiopia which revealed that PLHIV with good ART medication adherence had a 77.3% lower risk of TB disease (25).

7. STRENGTHS AND LIMITATIONS

This study used a joint modeling analysis of survival and longitudinal data types, which is a robust method of data analysis. This modeling handles informative drop-out, assesses dependency and association between the longitudinal CD4 count and the time to develop TB simultaneously. In addition, this study employed the Kobo toolbox tool for data collection, which increases the accuracy and efficiency of the study, ensuring reliable and comprehensive data for analysis. Despite its strengths, the study also has limitations; since the data was collected secondarily from the patient registry and followup, important factors like household income were not studied. Moreover, longitudinal factors like viral load status were not studied.

8. CONCLUSION

This research highlights the complex interplay between various predictors and their impact on both the time to develop TB disease and longitudinal CD4 cell count levels among PLHIV under ART. PLHIV with a decreasing trajectory of CD4, advanced WHO clinical stage, female sex, and poor adherence, and history of CPT intake have a greater risk of developing TB disease. This study revealed that CD4 cell count measurement is an important surrogate indicator of risk of developing TB disease among PLHIV in a ART follow-up.

9. RECOMMENDATIONS

On the basis of these findings, it is strongly recommended that the government and relevant health actors working on TB/HIV should intensify activities that improve patient adherence and a regular CD4 cell measurement. Moreover, health care workers working on TB/HIV area advised to focus on early initiation of ART medication for all HIV-infected individuals to prevent advanced WHO clinical stage and to consequently decrease TB disease among PLHIV. It is also recommended that the regional health bureau give a special focus to female clients. Finally, the PLHIV are advised to take improve their adherence and take regular CD4 measurement. Finally, researcher are recommended to study on the factors behind the highest TB development among female patients.

10.ACKNOWLEDGMENTS

First and foremost, I would like to thank GOD with his mom St. Mary for giving me the strength and patience to reach this time. Next, I am very grateful to Mekelle University, the College of Health Sciences, the School of Public Health, and the Department of Biostatistics for allowing me to perform my thesis work on this topic. After that, I am grateful to my advisors Mr. Kidanemariam Alem (Assistant Professor) and Mr. Goitom Halefom (MPH) for their support throughout the full process of conducting my research thesis. Additionally, I would like to express my gratitude to the administration bodies, electronic medical records, and card rooms staff members of MGH and ACSH's, as well as the data collectors who actively took part in the data collection process.

11.REFERENCES

1. World Health Organization (WHO). HIV_AIDS Key facts. 2024;86(July 2023):1–6.
2. CDC. Clinical Care for People with TB and HIV _ Tuberculosis (TB) _ CDC [Internet]. 2025. Available from: https://www.cdc.gov/tb/hcp/clinical-care/hiv.html?CDC_AAref_Val=https://www.cdc.gov/tb/topic/treatment/tbhiv.htm
3. World Health Organization (WHO). Consolidated Guidelines on HIV Prevention, Testing, Delivery and Monitoring: Treatment, Service Recommendations for a Public Health Approach. Geneva: WHO; 2021.
4. WHO. Tuberculosis Key facts [Internet]. WHO. 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/tuberculosis
5. FMOH. Guidelines for Clinical and Programmatic Management of TB , TB / HIV , DR-TB and Leprosy in Ethiopia. 2021. 1–249 p.
6. (UNAIDS) JUNP on H. The path that ends AIDS, UNAIDS Global AIDS Update. Geneva; 2023.
7. WHO. Global TB report 2023 [Internet]. January. 2023. 75 p. Available from: https://iris.who.int/.%0Ahttps://www.who.int/teams/global-tuberculosis-programme/tb-reports
8. Wondmeneh TG, Mekonnen AT. The incidence rate of tuberculosis and its associated factors among HIV-positive persons in Sub-Saharan Africa: a systematic review and meta-analysis. BMC Infect Dis [Internet]. 2023;23(1):1–24. Available from: https://doi.org/10.1186/s12879-023-08533-0
9. Ahmed A, Mekonnen D, Kindie M. Acquired Immuno Deficiency Syndrome Incidence and Predictors of Tuberculosis among Adult People Living. 2018;1(1):1–8.
10. Alemu A, Yesuf A, Zerihun B, Getu M, Worku T, Bitew ZW. Incidence and determinants of tuberculosis among HIV-positive individuals in Addis Ababa, Ethiopia: A retrospective cohort study. Int J Infect Dis [Internet]. 2020;95:59–66. Available from: https://doi.org/10.1016/j.ijid.2020.02.053

11. Tesfaye B, Alebel A, Gebrie A, Zegeye A, Tesema C, Kassie B. The twin epidemics: Prevalence of TB/HIV coinfection and its associated factors in Ethiopia; A systematic review and meta-analysis. PLoS One. 2018 Oct 1;13(10).

12. Martins R, Silva GL, Andreozzi V. Bayesian joint modeling of longitudinal and spatial survival AIDS data. Vol. 35, Statistics in Medicine. 2016.

13. Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. Vol. 28, Journal of Clinical Oncology. 2010. p. 2796–801.

14. WHO. HIV data and statistics Global data People living with HIV. 2024;1–8. Available from: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics

15. Suwanpimolkul G, Gatechompol S, Kawkitinarong K, Ueaphongsukkit T, Sophonphan J, Siriyakorn N, et al. Incidence of active tuberculosis among people living with HIV receiving long-term antiretroviral therapy in high TB/HIV burden settings in Thailand: implication for tuberculosis preventive therapy. J Int AIDS Soc. 2022;25(4):1–8.

16. Mollel EW, Maokola W, Todd J, Msuya SE, Mahande MJ. Incidence Rates for Tuberculosis Among HIV Infected Patients in Northern Tanzania. Front Public Heal. 2019;7(October):1–9.

17. Van Halsema CL, Okhai H, Hill T, Sabin CA. Incidence of and risk factors for tuberculosis among people with HIV on antiretroviral therapy in the United Kingdom. Aids. 2020;34(12):1813–21.

18. Azanaw MM, Derseh NM, Yetemegn GS. Incidence and predictors of tuberculosis among HIV patients after initiation of antiretroviral treatment in Ethiopia : a systematic review and meta-analysis. 2021;2.

19. Ahmed A, Mekonnen D, Shiferaw AM, Belayneh F, Yenit MK. Incidence and determinants of tuberculosis infection among adult patients with HIV attending HIV care in north-east Ethiopia: A retrospective cohort study. BMJ Open. 2018;8(2):1–14.

20. Ayana GM, Akalu TY, Ayele TA. Joint modeling of incidence of tuberculosis and change

in viral load over time among adult hiv/aids patients on anti-retroviral therapy at zewditu memorial hospital in Addis Ababa, Ethiopia. HIV/AIDS - Res Palliat Care. 2021;13:239–49.

21. Alafchi B, Mahjub H, Tapak L, Poorolajal J, Roshanaei G. Modeling the trajectory of cd4 cell count and its effect on the risk of aids progression and tb infection among hiv-infected patients using a joint model of competing risks and longitudinal ordinal data. Epidemiol Biostat Public Heal. 2019;16(4):1–11.

22. Temesgen E. Incidence and predictors of tuberculosis occurrence among adults on antiretroviral therapy at Debre Markos referral hospital, Northwest Ethiopia: Retrospective follow-up study. BMC Infect Dis. 2020;20(1):1–10.

23. Arefaine ZG, Abebe S, Bekele E, Adem A, Adama Y, Brockmeyer NH, et al. Incidence and predictors of HIV related opportunistic infections after initiation of highly active antiretroviral therapy at Ayder Referral Hospital, Mekelle, Ethiopia: A retrospective single centered cohort study. PLoS One [Internet]. 2020;15(4):1–11. Available from: http://dx.doi.org/10.1371/journal.pone.0229757

24. Umeta AK, Yermosa SF, Dufera AG. Bayesian parametric modeling of time to tuberculosis co-infection of HIV/AIDS patients at Jimma Medical Center, Ethiopia. Sci Rep. 2022;12(1):1–30.

25. Agegn SB. Joint Modeling in Detecting Predictors of Cd4 Cell Count and Status of Tuberculosis Among HIV Infected Patients Under HAART at Felege Hiwot Teaching and Specialized Hospital, North-west Ethiopia. 2020; Available from: https://doi.org/10.21203/rs.3.rs-89429/v1

26. Beshir MT, Beyene AH, Tlaye KG, Demelew TM. Incidence and predictors of tuberculosis among HIV-positive children at Adama Referral Hospital and Medical College, Oromia, Ethiopia: A retrospective follow-up study. Epidemiol Health. 2019;41:1–8.

27. Hoffmann CJ, Chaisson RE, Martinson NA. Cotrimoxazole Prophylaxis and Tuberculosis Risk among People Living with HIV. 2014;9(1):1–4.

28. Federal Ministry of Health EF. National Comprehensive HIV Prevention , Care and Treatment Training for Healthcare Providers. 2020;

29. Ayder Hospital - Wikipedia [Internet]. [cited 2025 May 21]. Available from: https://en.wikipedia.org/wiki/Ayder_Hospital

30. Dimitris Rizopoulos M. Title Joint Modeling of Longitudinal and Survival Data [Internet]. 2022. Available from: http://jmr.r-forge.r-project.org/

31. FMOH Ethiopia. National Consolidated Guidelines for Comprehensive Hiv Prevention , Care and Treatment. FMOH. 2018. 1–238 p.

32. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, recommendations for a public health approach second edition. WHO Press. 2013;(June):1–251.

12.ANNEXES

Missing Data Diagnosis

```
. gen missing_CD4=missing( CD4_visit )

. logit missing_CD4 Sex Age BMI Religion Marital Education Occupation_group Residence Haemoglobin CD4_
> baseline Adherence Clinicalstage ART TB_history CPT INH Functionality Opportunistic

Iteration 0:   log likelihood = -1013.6416
Iteration 1:   log likelihood = -989.87716
Iteration 2:   log likelihood = -989.05872
Iteration 3:   log likelihood = -989.05689
Iteration 4:   log likelihood = -989.05689

Logistic regression                                     Number of obs =  3,178
                                                        LR chi2(18)   =  49.17
                                                        Prob > chi2   = 0.0001
Log likelihood = -989.05689                             Pseudo R2     = 0.0243
```

missing_CD4	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Sex	.0578799	.1371217	0.42	0.673	-.2108737	.3266334
Age	.0065922	.0065366	1.01	0.313	-.0062193	.0194038
BMI	.0401077	.0293034	1.37	0.171	-.0173258	.0975412
Religion	-.0182996	.1852291	-0.10	0.921	-.3813419	.3447428
Marital	.275412	.0773777	3.56	0.000	.1237544	.4270695
Education	.2100402	.0852352	2.46	0.014	.0429823	.377098
Occupation_group	.1218267	.1431895	0.85	0.395	-.1588195	.402473
Residence	-.2345148	.146933	-1.60	0.110	-.5224983	.0534686
Haemoglobin	.0429711	.0410359	1.05	0.295	-.0374579	.1234
CD4_baseline	.0011563	.0003637	3.18	0.001	.0004434	.0018691
Adherence	-.3060464	.1290476	-2.37	0.018	-.5589751	-.0531178
Clinicalstage	.1970546	.0731859	2.69	0.007	.0536129	.3404964
ART	-.0054737	.1389851	-0.04	0.969	-.2778795	.2669322
TB_history	.3388324	.3398204	1.00	0.319	-.3272034	1.004868
CPT	.3262798	.1854411	1.76	0.078	-.037178	.6897377
INH	.2026786	.1345537	1.51	0.132	-.0610418	.466399
Functionality	.0654387	.0944805	0.69	0.489	-.1197397	.250617
Opportunistic	.3127325	.1854187	1.69	0.092	-.0506815	.6761464
_cons	-6.720391	1.26785	-5.30	0.000	-9.205332	-4.235449

Survival final model output

```
No. of subjects =     449                       Number of obs =  3,178
No. of failures =     104
Time at risk    = 18,427
                                                LR chi2(9)    = 138.64
Log likelihood = -512.56879                     Prob > chi2   = 0.0000
```

_t	Haz. ratio	Std. err.	z	P>|z|	[95% conf.	interval]
main						
CD4_base_sqrt	.9010832	.0165931	-5.66	0.000	.8691413	.934199
Sex						
Female	.5740922	.1194188	-2.67	0.008	.3818743	.8630639
CPT						
No	2.33168	.5902565	3.34	0.001	1.419679	3.829548
INH						
No	1.334214	.3085685	1.25	0.212	.8479398	2.099355
Clinicalstage						
Stage II	2.281926	.8319015	2.26	0.024	1.116829	4.662475
Stage III	1.637377	.5804243	1.39	0.164	.8173555	3.280093
Stage IV	1.53221	.5783604	1.13	0.258	.7311671	3.210848
Opportunistic						
No	.4819401	.1308885	-2.69	0.007	.2830207	.820669
tvc						
Adherence	.9893659	.005336	-1.98	0.047	.9789626	.9998798

Note: Variables in tvc equation interacted with _t.

. estat ic

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	3,178	-581.8906	-512.5688	9	1043.138	1097.714

Variable	VIF	1/VIF
CD4_base_s~t	8.88	0.112629
1.Sex	2.72	0.367199
1.CPT	1.29	0.776662
1.INH	2.09	0.478242
Clinicalst~e		
2	1.42	0.704019
3	2.21	0.453080
4	2.21	0.452969
1.Opportun~c	6.22	0.160675
Adherence		
1	1.30	0.767604
3	2.72	0.367120

```
. corr CD4_base_sqrt Sex CPT INH Clinicalstage Opportunistic Adherence //To assess intercation between independent variables//
(obs=3,178)

             | CD4_ba~t      Sex      CPT      INH Clinic~e Opport~c Adhere~e
-------------+---------------------------------------------------------------
CD4_base_s~t |   1.0000
         Sex |  -0.0322   1.0000
         CPT |  -0.1126   0.1294   1.0000
         INH |   0.2103  -0.0881   0.0006   1.0000
Clinicalst~e |  -0.4135   0.0426   0.0476  -0.1098   1.0000
Opportunis~c |  -0.5042  -0.0240   0.1080  -0.1182   0.4881   1.0000
   Adherence |   0.0373   0.0572  -0.0215   0.0714   0.4395  -0.0827   1.0000
```

Longitudinal Analysis Final Output

Table 13: Random Intercept only longitudinal Model with model comparison

```
. asdoc xtmixed CD4_sqrt  Months_visit ib2. Sex ib2. Opportunistic ib2. Adherence ib1. Clinicalstage i
> b2. Occupation_group || id :
(File Myfile.doc already exists, option append was assumed)

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0:   log likelihood = -8207.2204
Iteration 1:   log likelihood = -8207.2204

Computing standard errors:

Mixed-effects ML regression                     Number of obs     =      3,178
Group variable: id                              Number of groups  =        449
                                                Obs per group:
                                                              min =          2
                                                              avg =        7.1
                                                              max =         12
                                                Wald chi2(9)      =      43.17
Log likelihood = -8207.2204                     Prob > chi2       =     0.0000
```

CD4_sqrt	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Months_visit	.0021984	.0030244	0.73	0.467	-.0037293	.0081262
Sex						
Female	.5316786	.3394425	1.57	0.117	-.1336166	1.196974
Opportunistic						
No	-.1352838	.4513174	-0.30	0.764	-1.01985	.749282

Adherence						
Fair	-.2986761	.4750838	-0.63	0.530	-1.229823	.6324711
Poor	-1.58339	.4788375	-3.31	0.001	-2.521894	-.6448855
Clinicalstage						
Stage II	-.3758258	.5718692	-0.66	0.511	-1.496669	.7450173
Stage III	-.4909856	.557596	-0.88	0.379	-1.583854	.6018825
Stage IV	-1.055553	.6080613	-1.74	0.083	-2.247331	.1362252
Occupation_group						
Yes	-.7058146	.359735	-1.96	0.050	-1.410882	-.0007469
_cons	24.05563	.6444858	37.33	0.000	22.79246	25.3188

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
id: Identity				
sd(_cons)	3.167799	.1217503	2.937939	3.415643
sd(Residual)	2.731261	.0370496	2.659602	2.804851

LR test vs. linear model: chibar2(01) = 1524.82 Prob >= chibar2 = 0.0000

. estat ic

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	3,178	.	-8207.22	12	16438.44	16511.21

Note: BIC uses N = number of observations. See [R] BIC note.

Table 14: Random Intercept and Slope longitudinal Model with model comparison

```
. asdoc xtmixed CD4_sqrt  Months_visit ib2. Sex ib2. Opportunistic ib2. Adherence ib1. Clinicalstage i
> b2. Occupation_group || id : Months_visit
(File Myfile.doc already exists, option append was assumed)

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0:   log likelihood = -8213.3904
Iteration 1:   log likelihood = -8207.2596
Iteration 2:   log likelihood = -8207.2205
Iteration 3:   log likelihood = -8207.2204

Computing standard errors:

Mixed-effects ML regression                     Number of obs     =      3,178
Group variable: id                              Number of groups  =        449
                                                Obs per group:
                                                              min =          2
                                                              avg =        7.1
                                                              max =         12
                                                Wald chi2(9)      =      43.17
Log likelihood = -8207.2204                     Prob > chi2       =     0.0000
```

CD4_sqrt	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Months_visit	.0021984	.0030244	0.73	0.467	-.0037293	.0081262
Sex						
Female	.5316786	.3394425	1.57	0.117	-.1336165	1.196974

Opportunistic						
No	-.1352838	.4513173	-0.30	0.764	-1.019849	.7492819
Adherence						
Fair	-.2986761	.4750838	-0.63	0.530	-1.229823	.632471
Poor	-1.58339	.4788375	-3.31	0.001	-2.521894	-.6448856
Clinicalstage						
Stage II	-.3758259	.5718692	-0.66	0.511	-1.496669	.7450171
Stage III	-.4909856	.5575959	-0.88	0.379	-1.583853	.6018824
Stage IV	-1.055553	.6080613	-1.74	0.083	-2.247331	.1362251
Occupation_group						
Yes	-.7058146	.3597349	-1.96	0.050	-1.410882	-.000747
_cons	24.05563	.6444857	37.33	0.000	22.79246	25.3188

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
id: Independent				
sd(Months_visit)	4.00e-08	.0000228	0	.
sd(_cons)	3.167799	.1217504	2.937939	3.415643
sd(Residual)	2.731261	.0370498	2.659602	2.804852

LR test vs. linear model: chi2(2) = 1524.82   Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

```
. estat ic
```

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	3,178	.	-8207.22	13	16440.44	16519.27

Note: BIC uses N = number of observations. See [R] BIC note.

Final Joint Model STATA Output

```
. asdoc stjm CD4_sqrt CD4_base_sqrt Occupation_group Adherence Sex BMI Clinicalstage CPT , panel(id) s
> urvmodel(rcs) df(3) rfp(5) timeinterac( BMI ) survcov( Sex CPT Adherence ) assoccov( Clinicalstage )
(File Myfile.doc already exists, option append was assumed)
-> gen double _time_1 = X^(5)
-> gen double _time_1_BMI = BMI * _time_1
(where X = _t0)

Obtaining initial values:

Fitting full model:

-> Conducting adaptive Gauss-Hermite quadrature

Iteration 0:   log likelihood = -8815.2479  (not concave)
Iteration 1:   log likelihood = -8808.5917
Iteration 2:   log likelihood = -8808.5035  (backed up)
Iteration 3:   log likelihood = -8807.5705
Iteration 4:   log likelihood = -8806.4714

->No longer updating quadrature node locations

Iteration 5:   log likelihood = -8806.4675
Iteration 6:   log likelihood = -8806.4669

Joint model estimates                            Number of obs.     =     3178
Panel variable: id                               Number of panels   =      449
                                                 Number of failures =      104

Log-likelihood = -8806.4669
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Longitudinal						
_time_1	1.48e-09	1.64e-09	0.90	0.367	-1.74e-09	4.70e-09
_time_1_BMI	-7.66e-11	9.23e-11	-0.83	0.406	-2.57e-10	1.04e-10
CD4_base_s~t	.0613377	.0299262	2.05	0.040	.0026834	.119992
Occupation~p	.6829746	.359415	1.90	0.057	-.0214658	1.387415
Adherence	-.7377784	.2770802	-2.66	0.008	-1.280846	-.1947111
Sex	-.5476498	.3460008	-1.58	0.113	-1.225799	.1304993
BMI	-.083968	.073735	-1.14	0.255	-.2284859	.0605499
Clinicalst~e	-.363273	.1676416	-2.17	0.030	-.6918446	-.0347015
CPT	.2453332	.4331717	0.57	0.571	-.6036677	1.094334
_cons	24.96098	1.706131	14.63	0.000	21.61703	28.30494
Survival						
assoc:value						
Clinicalst~e	.0244344	.0038821	6.29	0.000	.0168255	.0320432
_cons	-.1581845	.0309688	-5.11	0.000	-.2188822	-.0974869
xb						
Sex	.4826747	.2005146	2.41	0.016	.0896732	.8756762
CPT	-.5877958	.2390716	-2.46	0.014	-1.056367	-.1192241
Adherence	-.9531779	.1495681	-6.37	0.000	-1.246326	-.6600299
_rcs1	.8310083	.1895488	4.38	0.000	.4594994	1.202517
_rcs2	.3800177	.1493483	2.54	0.011	.0873005	.6727349
_rcs3	-.1345037	.1102896	-1.22	0.223	-.3506673	.08166
_cons	-.6368448	.9178275	-0.69	0.488	-2.435754	1.162064

Random effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
id: Unstructured				
sd(_time_1)	7.09e-10	3.75e-10	2.51e-10	2.00e-09
sd(_cons)	3.20244	.1235958	2.969132	3.454081
corr(_time_1,_cons)	-.4971483	.3569525	-.9004904	.3659894
sd(Residual)	2.740887	.037511	2.668344	2.815402

```
Longitudinal submodel: Linear mixed effects model
    Survival submodel: Restricted cubic spline hazard model
   Integration method: Adaptive Gauss-Hermite quadrature using 5 nodes
    Cumulative hazard: Gauss-Kronrod quadrature using 15 nodes
```

Data Extraction Sheet

Age (in years) ______________________________

Gender	☐ Male	☐ Female		
Residency of the client	☐ Urban	☐ Rural		
Marital status	☐ Single	☐ Married	☐ Divorced	☐ Widowed
Religion	☐ Orthodox	☐ Muslim	☐ Catholic	☐ Protestant
Educational level	☐ No education	☐ Primary school (1-8 grade)	☐ Secondary school (9-12 grade)	☐ Tertiary school
Occupational status	Farmer	Merchant	Government employee	Daily laborer
	Housewife	Other		
Weight (in kg)	______________________________			
Height (in cm)	______________________________			
ART enrollment date	______________________________			
Baseline CD4 cell count (in cells/mm^3)	______________________________			
ART adherence status	☐ Good	☐ Fair		☐ Poor
Functional status	☐ Bedridden	☐ Ambulatory		☐ Working

Baseline WHO clinical stage	☐ Stage I	☐ Stage II	☐ Stage III	☐ Stage IV
Baseline Hemoglobin level (in g/dl)	______________________________			
Cotrimoxazole Preventive Therapy (CPT)	☐ **Yes**		☐ **No**	
Isoniazid Preventive Therapy (IPT)	☐ **Yes**		☐ **No**	
Type of ART medication (Regimen)	______________________________			
Previous History of TB infection	☐ **Yes**	☐ **No**		
TB screening result	☐ **Positive**	☐ **Negative**		
History of opportunistic infections	☐ **Yes**	☐ **NO**		
If yes to the above question, specify the type of opportunistic infection encountered.	______________________________			
Previous history of TB infection	☐ **Yes**	☐ **NO**		
Censoring Status	Transfer-Out	Lost to follow-up/Withdrawal	Death	End of study period

CD4 cell count measurement (in cells/mm^3)	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**
(repeated follow-up period)	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**
	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**	☐ **CD4** ☐ **Date**