

Predictors of Virologic Failure among Adults with Low-Level Viremia: A Retrospective Cohort Study in Oromia, Ethiopia

*Lamessa Tadesse Amante

School of Public Health, Texila American University, Georgetown, Guyana

Muluembet Abera Wordofa

Jimma University, School of Public Health

*Corresponding author

Abstract

Background

Low-level viremia (LLV), defined as an HIV-1 RNA level of 50 to 999 copies/mL, is a common but clinically challenging state among adults receiving antiretroviral therapy (ART). Although LLV does not meet the conventional threshold for virologic failure, it may identify individuals at increased risk of subsequent treatment failure. We evaluated predictors of progression from LLV to virologic failure and characterized the time-dependent probability of progression in a large Ethiopian ART cohort.

Methods

We conducted a retrospective cohort study among adults with at least one LLV episode documented between October 2023 and June 2025 across 105 health facilities in Oromia, Ethiopia. The primary outcome was progression from LLV to virologic failure, defined as an HIV-1 RNA level of ≥ 1000 copies/mL, within 18 months. Predictors were evaluated using generalized estimating equations (GEE) with patient-level clustering. A risk stratification model was derived using L1-penalized logistic regression with a patient-level train/test split and validated using grouped cross-validation. Time to virologic failure was evaluated using Kaplan-Meier methods, with a sensitivity analysis based on alternative censoring definitions.

Results

Among 6713 LLV observations contributed by 6167 unique patients, 189 virologic failure events occurred over 18 months of follow-up. The cumulative incidence of virologic failure was 0.76% at 6 months, 2.22% at 12 months, and 3.02% at 18 months at the patient level. In multivariable GEE analysis, receipt of second- or third-line ART was strongly associated with increased odds of virologic failure (adjusted odds ratio [aOR], 3.93; 95% CI, 2.85-5.43; $P < .001$), whereas multi-month dispensing (MMD) of ≥ 90 days was associated with lower odds of virologic failure (aOR, 0.54; 95% CI, 0.38-0.79; $P = .001$). Increasing age was independently associated with a modest reduction in the odds of virologic failure (aOR per year, 0.99; 95% CI, 0.98-1.00; $P = .046$). Sex, adherence status, prior visit frequency, prior viral load documentation, and prior viral load magnitude were not independently associated with virologic failure after adjustment. The risk-stratification model demonstrated moderate discrimination (area under the receiver operating characteristic curve [AUC-ROC], 0.79; precision-recall area under the curve [PR-AUC], 0.09) but poor calibration, with a mean predicted probability of 38.9% compared with an observed event prevalence of 2.23%. Kaplan-Meier analysis estimated a cumulative probability of virologic failure of 1.27% at 6 months, 8.10% at 12 months, and 34.51% (95% CI, 26.31%-44.40%) at 18 months in the primary cohort; estimates were similar in the sensitivity cohort (35.14%; 95% CI, 26.95%-44.95%).

Conclusions and Relevance

Among adults with LLV receiving ART in Oromia, Ethiopia, MMD of ≥ 90 days and ART regimen line were the principal independent predictors of subsequent virologic failure. The risk-stratification model demonstrated moderate discrimination but poor absolute calibration, indicating that recalibration and external validation are necessary before clinical application. Kaplan-Meier analysis estimated a cumulative probability of virologic failure of 34.51% at 18 months, although this estimate should be interpreted cautiously because only 12 patients remained at risk at that time point. These findings provide programmatically relevant evidence regarding MMD and the management of patients

receiving second- or third-line ART and support further prospective evaluation of these factors within differentiated HIV service-delivery models.

Keywords: *HIV; Low-Level Viremia; Virologic Failure; Ethiopia; Antiretroviral Therapy; Generalized Estimating Equations; Risk Stratification.*

I. INTRODUCTION

1.1 Background

The global scale-up of antiretroviral therapy (ART) has produced substantial improvements in HIV treatment outcomes. By 2023, approximately 29.8 million people were receiving ART worldwide [1]. Nevertheless, a proportion of patients receiving treatment continue to have detectable HIV RNA levels despite not meeting the conventional definition of virologic failure.

Low-level viremia (LLV), commonly defined as an HIV-1 RNA concentration of 50–999 copies/mL, has been reported among approximately 3–20% of people receiving ART [2–4]. LLV represents a clinically heterogeneous state. Some individuals subsequently re-suppress without treatment modification, whereas others progress to higher-level viremia and virologic failure [5,6]. Persistent LLV has also been associated with the emergence of antiretroviral drug resistance and adverse clinical outcomes [2–4].

The management of LLV is therefore an important challenge for HIV programs. Decisions regarding enhanced adherence support, differentiated service delivery, repeat viral-load monitoring, regimen review, and potential treatment modification must be made in settings where laboratory and clinical resources may be constrained.

1.2 The Prediction Challenge

The clinical course following LLV is heterogeneous. Not all patients with LLV progress to virologic failure, and the timing of progression may differ substantially between individuals [5,6]. Identifying characteristics associated with progression could therefore support more targeted clinical and programmatic responses.

This issue is particularly relevant in resource-constrained settings, where routine viral-load monitoring and clinical follow-up may be limited and where unnecessary clinical interventions can place additional burdens on both patients and health systems. A useful risk-stratification approach should distinguish patients at different levels of risk while maintaining appropriate calibration of predicted probabilities. Importantly, predictive discrimination and calibration represent different properties: a model may distinguish relatively higher- and lower-risk patients while still producing poorly calibrated absolute-risk estimates.

1.3. Study Objectives

1. This study had four objectives:
2. To identify predictors of progression from LLV to virologic failure.
3. To derive and evaluate a clinically applicable risk-stratification tool.
4. To characterize the time-dependent probability of progression to virologic failure.
5. To describe variation in virologic outcomes across health facilities.

II. METHODS

2.1 Study Design and Setting

This retrospective cohort study was conducted among adults with at least one documented LLV episode between October 2023 and June 2025 across 105 health facilities in Oromia, Ethiopia. The study was approved by the Texila American University Institutional Review Board and the Oromia Regional Health Bureau.

2.2 Data Cleaning and Preparation

The initial electronic medical record extract contained 647 109 records. A structured data cleaning procedure was implemented before construction of the analytic cohort. First, exact duplicate records were identified using patient, date, category, and count/value fields and removed. This step eliminated substantial redundancy in the raw extract. Second, records with conflicting viral-load categories recorded on the same date were adjudicated. Where the failure-to-LLV ratio was approximately 10-fold, suggesting a possible decimal-placement error, the record was reclassified as LLV. In other same-day conflicts, the more severe virologic category was retained. Third, negative viral-load values were recorded to zero. Following these procedures, the cleaned dataset contained 127 166 records from 85 261 patients.

2.3 Study Population

The final analytic cohort consisted of 6713 LLV observations contributed by 6167 unique patients. Patients were eligible for analysis if they had at least one documented LLV observation during the study period. The analytic unit for the longitudinal predictor analysis was LLV observation, while patient-level clustering was accounted for in the GEE analysis.

2.4 Outcome Definitions

Virologic outcomes were defined as follows:

- Viral suppression: HIV-1 RNA <50 copies/mL.
- Low-level viremia: HIV-1 RNA 50 to 999 copies/mL.
- Virologic failure: HIV-1 RNA ≥ 1000 copies/mL.

The primary outcome was progression from an LLV observation to virologic failure within 18 months.

2.5 Predictor Variables

The principal predictors evaluated were:

- Multi-month dispensing (MMD): dispensing of ART for ≥ 90 days, categorized as a binary variable.
- ART regimen line: first-line versus second- or third-line regimen.
- Age: age in years, centered for modeling.
- Sex: male or female.
- Adherence level: poor versus fair/good.
- Visit history: number of viral-load measurements recorded before the index LLV observation.
- Prior viral load: represented using a two-part specification consisting of an indicator for whether a prior viral-load measurement was documented and the $\log_{10}$ magnitude of the prior viral-load value among observations with a documented prior measurement.

The distinction between absence of a documented prior measurement and the magnitude of an observed prior viral load was maintained because many first LLV observations did not have a preceding viral-load measurement.

2.6 Statistical Analysis

GEE Analysis: Generalized estimating equations were used as the primary inferential model to account for repeated LLV observations contributed by the same patient. An exchangeable working correlation structure was specified, with patient-level clustering and robust standard errors. Adjusted odds ratios (aORs), 95% CIs, and P values were reported.

Risk-Score Development and Validation: A separate predictive model was developed using L1-penalized logistic regression. The data were divided into training and held-out test sets using a patient-level 80%/20% split to prevent observations from the same patient from appearing in both datasets.

Model validation incorporated five-fold stratified grouped cross-validation. Predictive performance was evaluated using the area under the receiver operating characteristic curve (AUC-ROC), precision-recall area under the curve (PR-AUC), Brier score, and calibration diagnostics.

The risk score was considered a complementary predictive model rather than an alternative inferential model to the GEE analysis.

Survival Analysis: Time to virologic failure was evaluated using Kaplan-Meier methods. End-of-follow-up was determined using a predefined hierarchy based on actual observation dates, including the date of virologic failure, the last documented viral-load measurement, the last documented clinical encounter, and administrative censoring.

A sensitivity analysis restricted the cohort to patients with documented clinical follow-up. The two survival distributions were compared using the log-rank test. All analyses were conducted using Python version 3.11. Statistical significance was defined as $P < .05$.

III RESULTS

3.1 Cohort Assembly and Baseline Characteristics

Among 647 109 viral load records identified from 85 261 patients, exact deduplication and same-day adjudication yielded 127 166 eligible records. Application of study eligibility criteria resulted in a final analytic cohort comprising 6713 low-level viremia (LLV) observations contributed by 6167 unique patients (Figure 1).

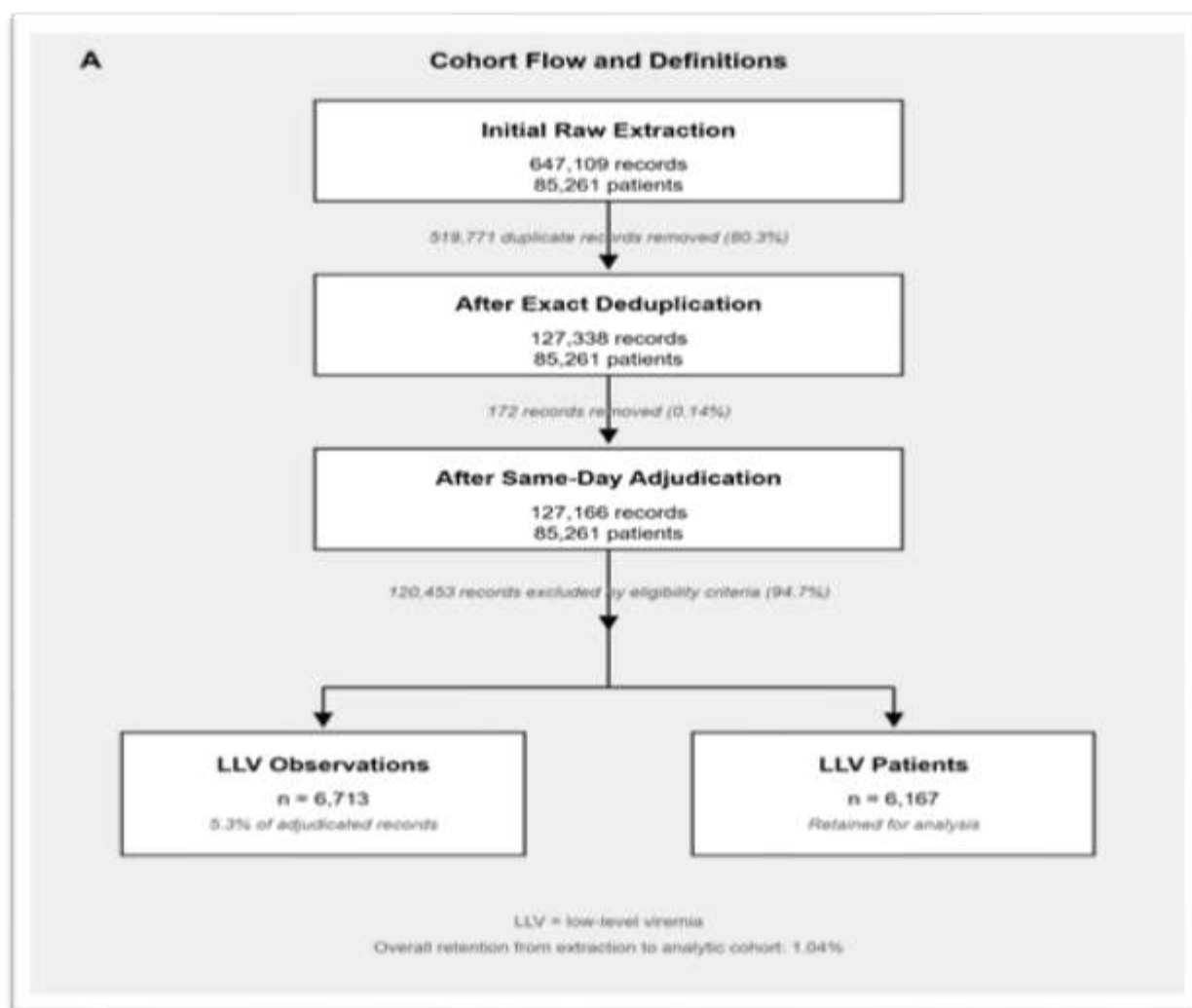


Figure 1A. Cohort flow and derivation of the analytic low-level viremia (LLV) cohort

The median age at the index LLV observation was 41 years (interquartile range [IQR], 32-49 years), and 57.4% of patients were female. Patients had received antiretroviral therapy (ART) for a median of 127 months (IQR, 69-175 months). Most patients were receiving first-line ART (87.0%), were enrolled in multi-month dispensing (MMD) programs with dispensing intervals ≥ 90 days (89.3%), and had documented good or fair adherence (96.6%). The median index viral load was 130 copies/mL (IQR, 75-251 copies/mL), with 68.4% of LLV observations occurring in the 50-199 copies/mL category (Table 1).

Table 1. Baseline Characteristics of Patients With Low-Level Viremia at Index Observation (N = 6167)

Characteristic	Value
Demographic Characteristics	
Age at LLV index, y, median (IQR)	41 (32-49)
Age group, No. (%)	

Characteristic	Value
15-24 y	661 (10.7)
25-34 y	996 (16.2)
35-44 y	2014 (32.7)
45-54 y	1532 (24.8)
≥55 y	694 (11.3)
Sex, No. (%)	
Female	3538 (57.4)
Male	2629 (42.6)
Clinical Characteristics	
Time on ART at LLV index, mo, median (IQR)	127 (69-175)
Regimen line, No. (%)	
First-line	5363 (87.0)
Second-/third-line	804 (13.0)
MMD status, No. (%)	
MMD ≥90 d	5508 (89.3)
No MMD (<90 d)	659 (10.7)
Adherence, No. (%)	
Good/fair	5957 (96.6)
Poor	210 (3.4)
Documented prior viral load, No. (%)	
Yes	2362 (38.3)
No (first LLV observation)	3805 (61.7)
Laboratory Characteristics	
Index viral load, copies/mL, median (IQR)	130 (75-251)
Index viral load category, No. (%)	
50-199 copies/mL	4594 (68.4)
200-499 copies/mL	1411 (21.0)
500-999 copies/mL	708 (10.5)
LLV observations per patient, mean (SD)	1.09 (0.30)
1 observation, No. (%)	5643 (91.5)
2 observations, No. (%)	502 (8.1)
3 observations, No. (%)	22 (0.4)

Notes: Values are presented as No. (%), median (IQR), or mean (SD), as appropriate. Patient-level characteristics were summarized using the first qualifying LLV observation (index LLV). Laboratory characteristics were summarized across all 6713 LLV observations.

Abbreviations: ART, antiretroviral therapy; IQR, interquartile range; LLV, low-level viremia; MMD, multi-month dispensing; SD, standard deviation.

3.2 Virologic Failure Events

The frequency of virologic failure increased progressively with longer follow-up duration (Table 2). Within 6 months of the index LLV episode, 47 virologic failure events were observed, corresponding to an observation-level event rate of 0.70% and a patient-level event rate of 0.76%. By 12 months, 138 virologic failure events had occurred, yielding observation-level and patient-level rates of 2.06% and 2.22%, respectively. At 18 months, a total of 189 virologic failure events were observed among 186 patients, corresponding to observation-level and patient-level rates of 2.82% and 3.02%, respectively (Table 2).

These findings indicate that progression from LLV to virologic failure was uncommon during short-term follow-up but accumulated steadily over time. Notably, approximately three-fourths of all observed virologic failure events occurred after the first 6 months of follow-up, emphasizing the importance of continued longitudinal monitoring following an LLV episode.

Table 2. Virologic Failure Events by Prediction Horizon Following Low-Level Viremia

Horizon	Event Observations, No.	Event Patients, No.	Observation-Level Rate, %	Patient-Level Rate, %
6 months	47	47	0.70	0.76
12 months	138	137	2.06	2.22
18 months	189	186	2.82	3.02

Notes: Virologic failure was defined as the first viral load measurement ≥ 1000 copies/mL occurring within the specified interval following the index LLV observation. Observation-level rates were calculated using 6713 LLV observations as the denominator, whereas patient-level rates were calculated using 6167 unique patients.
Abbreviation: LLV, low-level viremia.

3.3 Predictors of Virologic Failure

In multivariable generalized estimating equation (GEE) analysis, receipt of second- or third-line ART was strongly associated with increased odds of virologic failure (adjusted odds ratio [aOR], 3.93; 95% CI, 2.85-5.43; $P < .001$). Similarly, patients without MMD (<90 -day dispensing interval) had significantly higher odds of virologic failure than those receiving MMD ≥ 90 days (aOR, 1.85; 95% CI, 1.27-2.65; $P = .001$).

Increasing age was independently associated with a modest reduction in the odds of virologic failure (aOR per year, 0.99; 95% CI, 0.98-1.00; $P = .046$). Sex, adherence status, prior visit frequency, prior viral load documentation, and prior viral load magnitude were not independently associated with virologic failure after adjustment (**Table 3; Figure 2**).

Table 3. Multivariable Predictors of Virologic Failure Following Low-Level Viremia (GEE Model)

Variable	aOR	95% CI	P Value
Age (per year, centered)	0.99	0.98-1.00	.046
Male sex	0.96	0.71-1.29	.77
Second-/third-line ART	3.93	2.85-5.43	$<.001$
No MMD (<90 d)	1.85	1.27-2.65	.001

Notes: Adjusted odds ratios were estimated using generalized estimating equations with an exchangeable working correlation structure to account for repeated LLV observations within patients. All prespecified covariates were retained in the final model regardless of statistical significance.

Model characteristics: 6713 LLV observations; 6167 patients; 189 virologic failure events.

Abbreviations: aOR, adjusted odds ratio; CI, confidence interval; GEE, generalized estimating equations; LLV, low-level viremia; MMD, multi-month dispensing.

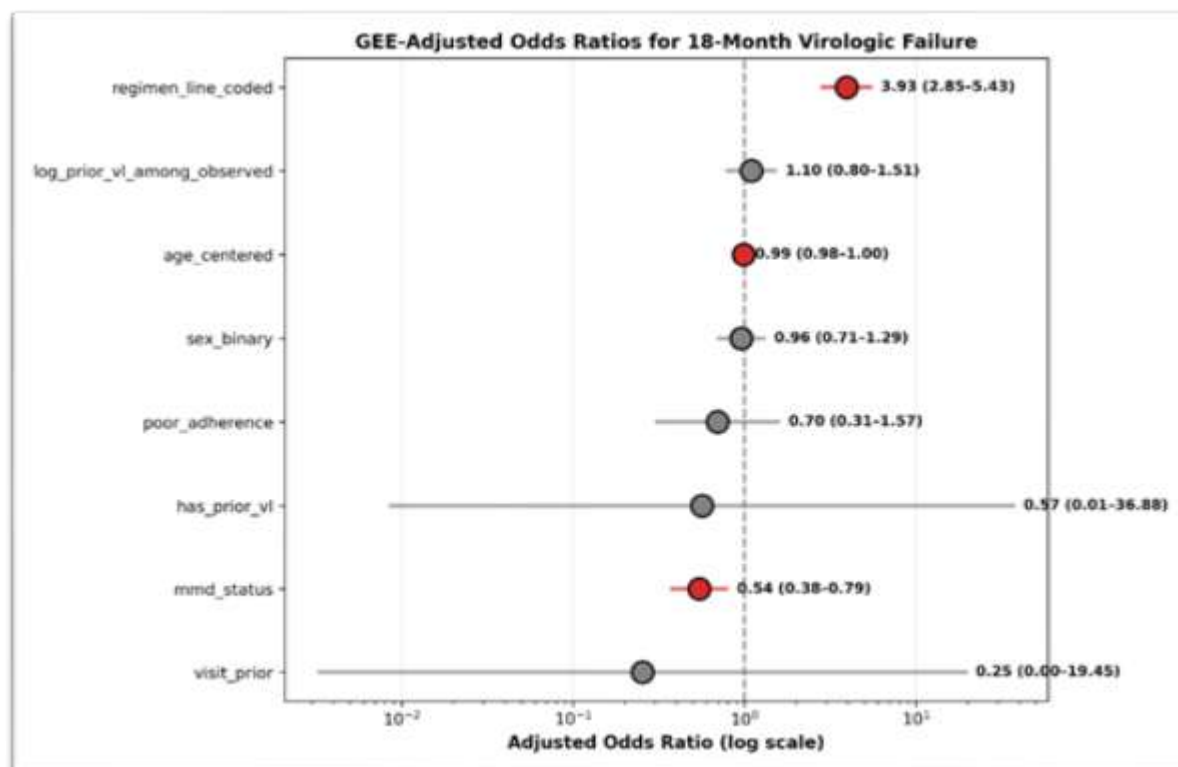


Figure 2. Forest plot of GEE-adjusted odds ratios for 18-month virologic failure.

Adjusted odds ratios and corresponding 95% confidence intervals are presented. The dashed vertical line represents the null value (OR = 1). aOR indicates adjusted odds ratio; CI, confidence interval; GEE, generalized estimating equations; LLV, low-level viremia; MMD, multi-month dispensing.

3.4 Risk Score Performance

Performance of the LLV risk score in the held-out test dataset is summarized in Table 4 and illustrated in Figure 3.

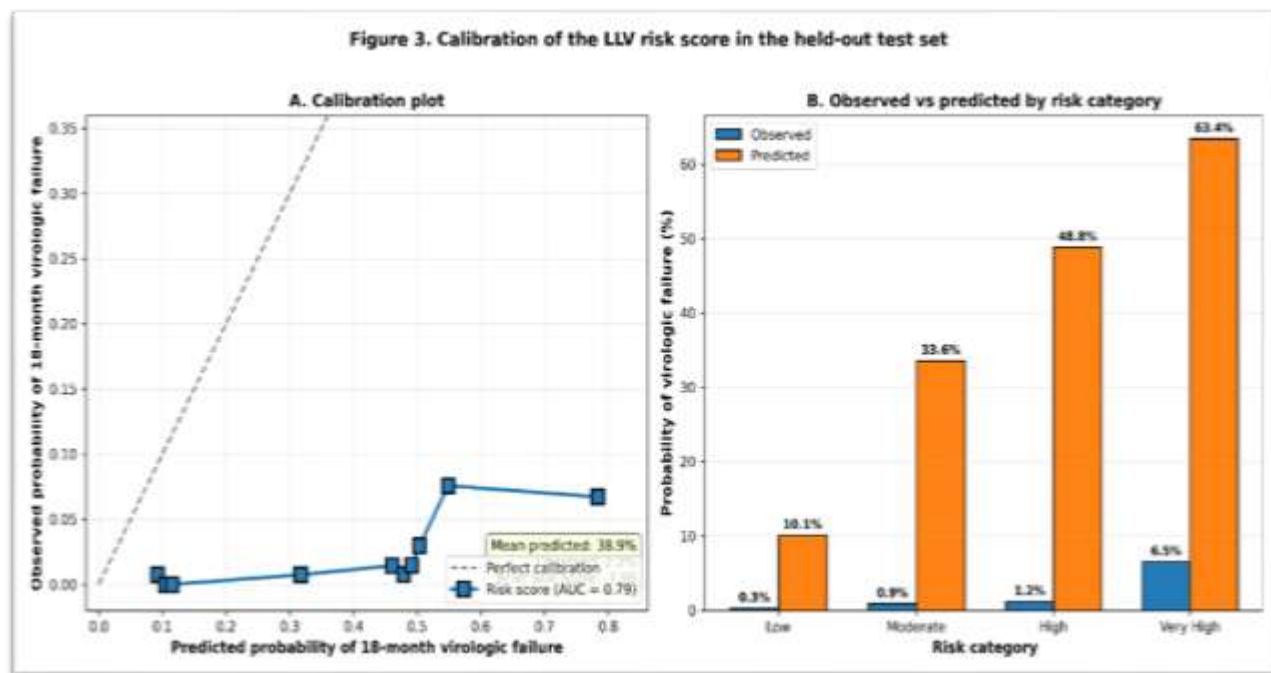


Figure 3. Calibration of the LLV risk prediction model in the held-out test dataset.

Observed and predicted probabilities of virologic failure are compared across risk strata. The model demonstrated moderate discrimination (AUC-ROC = 0.79) but poor calibration, with substantial overestimation of absolute risk. The mean predicted probability was 38.9%, compared with an observed event prevalence of 2.23%. The calibration slope was 0.92 and the calibration intercept was -3.71.

The prediction model demonstrated moderate discrimination, with an area under the receiver operating characteristic curve (AUC-ROC) of 0.79 and a precision-recall area under the curve (PR-AUC) of 0.09. However, calibration performance was poor. The mean predicted probability of virologic failure was 38.9%, substantially exceeding the observed event prevalence of 2.23%.

Calibration metrics further indicated systematic overprediction of absolute risk. The Brier score was 0.196, compared with a null-model Brier score of 0.022, yielding a strongly negative Brier Skill Score (-7.98). Although the calibration slope (0.92) was reasonably close to the ideal value of 1.0, the calibration intercept (-3.71) deviated substantially from the ideal value of 0, indicating poor agreement between predicted and observed probabilities (Table 4; Figure 3).

Collectively, these findings suggest that the model provides useful discrimination and relative risk ranking but does not produce reliable estimates of absolute risk in its current form. Accordingly, the score should be interpreted as a risk-stratification tool rather than a calibrated clinical prediction model and would require recalibration before implementation in independent populations.

Table 4. Risk Score Validation in the Held-Out Test Set

Metric	Value
Discrimination	
AUC-ROC	0.79
PR-AUC	0.09
Calibration	
Brier score	0.196

Null Brier score	0.022
Brier Skill Score	-7.98
Calibration slope	0.92
Calibration intercept	-3.71
Prediction summary	
Mean predicted probability, %	38.9
Observed prevalence, %	2.23

Notes: Model performance was evaluated in a held-out test set comprising **1345** LLV observations. Discrimination was assessed using the area under the receiver operating characteristic curve (AUC-ROC) and precision-recall area under the curve (PR-AUC). Calibration was assessed using the Brier score, calibration slope, and calibration intercept. A negative Brier Skill Score indicates poorer probabilistic prediction than a prevalence-based null model. Ideal calibration values are a slope of 1.0 and an intercept of 0.

Abbreviations: AUC-ROC, area under the receiver operating characteristic curve; LLV, low-level viremia; PR-AUC, precision-recall area under the curve.

3.5 Survival Analysis

Kaplan-Meier estimates of cumulative virologic failure are summarized in **Table 5** and displayed in **Figure 2**.

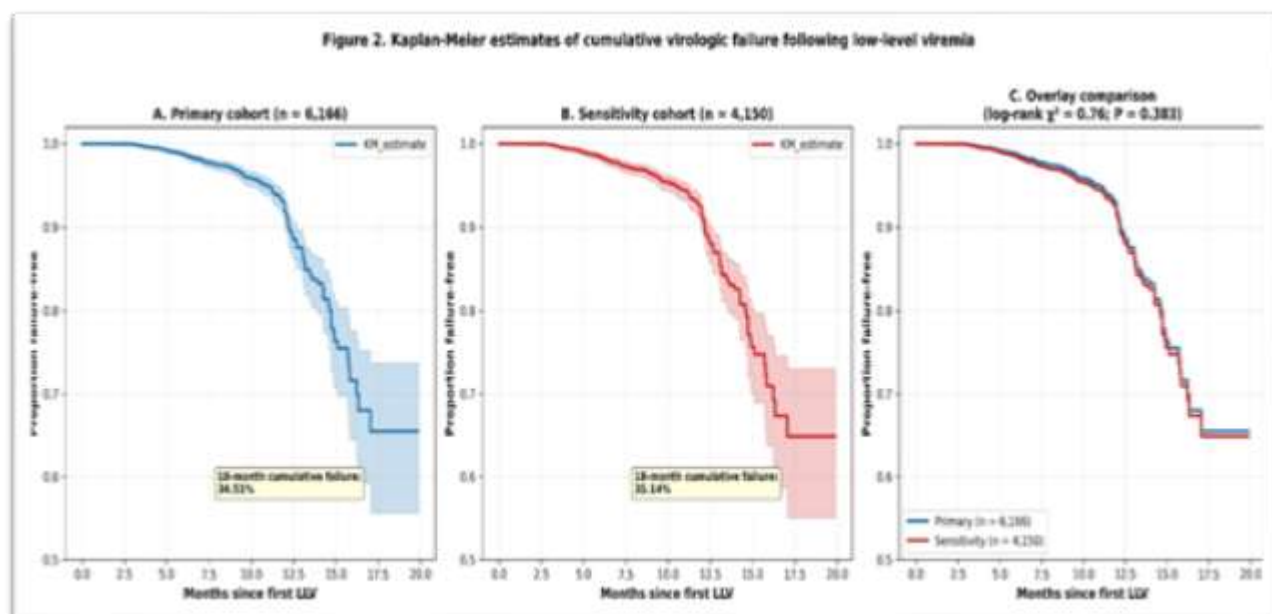


Figure 2. Kaplan-Meier estimates of cumulative virologic failure following low-level viremia. The primary cohort and sensitivity cohort demonstrated similar cumulative probabilities of virologic failure over follow-up. Shaded regions indicate 95% confidence intervals. Survival distributions did not differ significantly between cohorts (log-rank $P = 0.383$).

In the primary survival cohort ($n = 6166$), which excluded one patient with zero follow-up time, the estimated cumulative probability of virologic failure was 1.27% at 6 months, 8.10% at 12 months, and 34.51% (95% CI, 26.31%-44.40%) at 18 months (**Table 5, Figure 2**).

In the sensitivity cohort ($n = 4150$), corresponding estimates were 1.59%, 8.66%, and 35.14% (95% CI, 26.95%-44.95%), respectively (Table 5). The absolute difference between cohorts at 18 months was only 0.63 percentage points, and survival distributions did not differ significantly (log-rank $\chi^2 = 0.76$; $P = .38$).

Median follow-up was 5.3 months (IQR, 2.6-9.5 months) in the primary cohort and 9.8 months (IQR, 6.3-12.3 months) in the sensitivity cohort. Only 12 patients remained at risk at the 18-month time point in each cohort.

The directly observed cumulative incidence of virologic failure was 3.02% (186/6166), substantially lower than the corresponding 18-month Kaplan-Meier estimate. This discrepancy reflects extensive right censoring and the limited number of patients remaining under observation at later time points. Whereas observed cumulative incidence describes the proportion of

patients who were directly observed to fail, Kaplan-Meier methods estimate cumulative failure probability while accounting for unequal follow-up duration and censoring. Accordingly, the tail of the survival curve should be interpreted cautiously.

Despite these limitations, the close agreement between the primary and sensitivity cohorts supports the robustness of the estimated progression risk following LLV (Figure 2; Table 5).

Table 5. Kaplan-Meier Estimates of Cumulative Probability of Virologic Failure

Time Point	Primary Cohort (n = 6166)	Sensitivity Cohort (n = 4150)
6 months	1.27%	1.59%
12 months	8.10%	8.66%
18 months	34.51% (95% CI, 26.31%-44.40%)	35.14% (95% CI, 26.95%-44.95%)

Notes: Values represent cumulative probabilities of virologic failure estimated using Kaplan-Meier methods. Values at 18 months are presented with 95% CIs. The absolute difference between cohorts at 18 months was 0.63 percentage points. Survival distributions did not differ significantly between cohorts (log-rank $\chi^2 = 0.76$; $P = .38$). Median follow-up was 5.3 months (IQR, 2.6-9.5) in the primary cohort and 9.8 months (IQR, 6.3-12.3) in the sensitivity cohort.

Abbreviations: CI, confidence interval; IQR, interquartile range; KM, Kaplan-Meier

3.6 Sensitivity Analysis

Results were highly consistent in Cox proportional hazards models conducted among patients with valid post-index follow-up. Receipt of second- or third-line ART remained the strongest predictor of virologic failure (adjusted hazard ratio [aHR], 2.99; 95% CI, 2.15-4.18; $P < .001$), while MMD ≥ 90 days was associated with approximately 50% lower hazard of virologic failure (aHR, 0.50; 95% CI, 0.35-0.72; $P < .001$). Older age remained protective (aHR per year, 0.98; 95% CI, 0.97-0.99; $P = .002$) (Figure 5).

The consistency between GEE and Cox models strengthens the inference that ART regimen line and MMD status are the dominant predictors of progression from LLV to virologic failure.

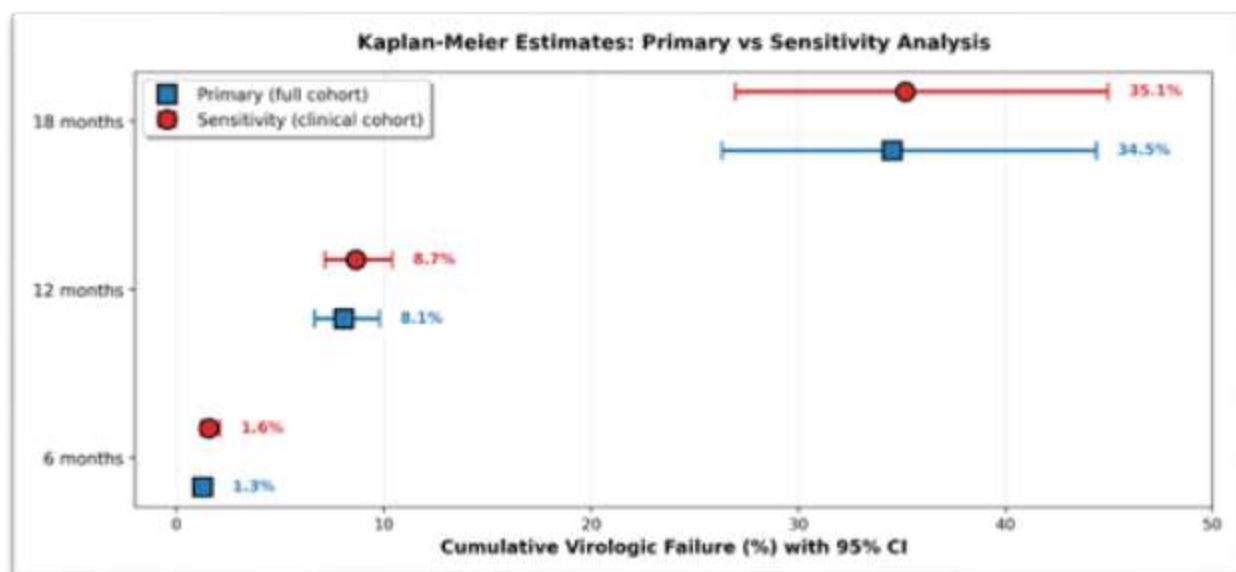


Figure 5. Sensitivity analysis of Kaplan-Meier estimates. Kaplan-Meier curves compare the primary cohort with the clinical-follow-up sensitivity cohort under the alternative censoring specification.

3.7 Facility-Level Variation

Facility-level characteristics are summarized in Supplementary Table S6. Among 105 facilities included in the dataset, 102 (97.1%) contributed at least one patient with low-level viremia (LLV). Facility LLV burden ranged from 0.0% to 20.0%, with a median burden of 4.1% (IQR, 2.4%-8.1%).

Patient-level virologic failure proportions stratified by facility LLV burden are presented in Table 6. Facilities classified as having high LLV burden ($>10\%$; $n = 17$) contributed 2045 patients, among whom 51 experienced virologic failure, corresponding to a failure proportion of 2.49%. Facilities with lower LLV burden ($\leq 10\%$; $n = 85$) contributed 4122 patients, among whom 135 experienced virologic failure, corresponding to a failure proportion of 3.28%.

Across all participating facilities, 186 of 6167 patients experienced virologic failure during follow-up, yielding an overall patient-level cumulative incidence of 3.02%. Although virologic failure proportions appeared slightly lower among patients receiving care at high-burden facilities than among those attending low-burden facilities, these comparisons were descriptive and were not adjusted for patient-level characteristics, facility-level clustering, or other potential confounding factors.

Accordingly, these findings should not be interpreted as evidence of a causal relationship between facility LLV burden and risk of virologic failure. Rather, they suggest that facility-level LLV prevalence alone may not adequately explain subsequent virologic outcomes. Future investigations using hierarchical or multilevel modeling approaches are warranted to more rigorously evaluate the contribution of facility-level factors to progression from LLV to virologic failure.

Table 6. Patient-Level Virologic Failure by Facility LLV-Burden Stratum

Facility LLV-Burden Group	Facilities, No.	Patients, No.	Virologic Failures, No.	Failure Proportion, %
High LLV burden (>10%)	17	2045	51	2.49
Low LLV burden ($\leq$ 10%)	85	4122	135	3.28
Overall	102	6167	186	3.02

Notes: Results are presented at the patient level. Facility LLV burden was defined as the proportion of patients within a facility meeting the LLV definition during the study period. Comparisons are descriptive and do not account for patient clustering or facility-level confounding.

Abbreviation: LLV, low-level viremia.

IV. DISCUSSION

4.1 Principal Findings

In this routine-care cohort from Oromia, later-line ART and MMD status were the clearest adjusted correlates of virologic failure after LLV. The consistency of the GEE and Cox estimates supports the robustness of these associations across analytic approaches, but the observational design does not establish causality.

MMD and ART regimen line were the principal independent predictors identified in the multivariable GEE analysis. MMD of ≥ 90 days was associated with lower odds of virologic failure, whereas receipt of second- or third-line ART was associated with substantially higher odds.

Kaplan-Meier analysis produced an estimated cumulative probability of virologic failure of 34.51% at 18 months, although this estimate should be interpreted cautiously because only 12 patients remained at risk at that time point. The estimate was similar under the tested sensitivity specification.

The risk-stratification model demonstrated moderate discrimination (AUC, 0.74) but poor calibration of absolute risk. Thus, although the model may have potential for relative risk stratification, recalibration using an independent cohort would be required before its predicted probabilities could be considered suitable for clinical application.

4.2 Regimen Line and Progression to Virologic Failure

The association with second- or third-line ART is consistent with evidence that LLV is associated with subsequent failure and that risk may be greater among patients with more complex treatment histories.¹⁻⁴ Later-line status may capture accumulated treatment experience, previous virologic failure, resistance, or other clinical complexity not fully measured in the routine dataset.^{1,2,5-6} The finding therefore supports closer clinical review after LLV among patients receiving later-line therapy, but it should not be interpreted as evidence that the regimen itself causes failure or that switching therapy solely because of regimen line will improve outcomes.

This finding is clinically plausible because patients receiving later-line regimens may represent a population with a more complex treatment history, including previous treatment failure, exposure to multiple antiretroviral regimens, or accumulated drug resistance.^{4,6,10} Persistent LLV in such patients may therefore require careful clinical assessment and continued virologic monitoring.^{12,13}

The present analysis cannot establish the mechanism underlying the observed association because detailed information on historical resistance testing, cumulative drug exposure, and previous regimen changes was not incorporated into the model.

4.3 MMD and Virologic Outcomes

MMD of at least 90 days was associated with lower odds of failure. This is directionally consistent with observational evidence and randomized evidence that differentiated dispensing can maintain treatment outcomes among clinically stable patients.^{7,8} However, MMD allocation is not random in routine care and may reflect eligibility, engagement, provider assessment, or other characteristics associated with outcomes.⁷ The observed association should therefore not be interpreted as a causal treatment effect. Target-trial emulation or other causal methods could better address confounding by indication.

Patients receiving MMD may differ from those receiving shorter dispensing intervals in ways that are not completely captured by the available electronic medical record variables. For example, eligibility for MMD may reflect clinical stability, prior treatment engagement, or facility-level implementation practices.

Nevertheless, the finding provides a programmatically relevant signal that warrants further investigation, particularly as differentiated service delivery models continue to expand.^{7,8}

4.4 Other Predictors

Age demonstrated a statistically significant but small inverse association with virologic failure. The effect size was close to the null, suggesting that the magnitude of the association was modest.

Sex, adherence category, visit history, documentation of a prior viral-load measurement, and the magnitude of the prior viral load were not statistically significant in the multivariable model.¹¹

The interpretation of the prior viral-load variables requires particular caution. More than 60% of LLV observations had no documented preceding viral-load measurement. This pattern largely reflected first observed LLV episodes rather than conventional missing-data processes. Accordingly, the analysis treated prior viral-load documentation as a two-part variable rather than simply treating the absence of a measurement as missingness.

The wide confidence intervals for visit history and prior viral-load variables indicate limited precision. Consequently, the absence of statistical significance should not be interpreted as evidence that these variables have no relationship with virologic outcomes.

4.5 Risk Stratification

The risk score achieved an AUC of 0.74 in the held-out test set, indicating moderate ability to distinguish observations with different levels of risk.

However, discrimination and calibration measure different aspects of prediction. In this analysis, the model substantially overestimated absolute risk: the mean predicted probability was 39.8%, whereas the observed event rate was 2.8%.

An AUC of 0.79 indicates moderate ability to rank patients, but the mean predicted probability (38.9%) was far above the observed prevalence (2.23%), with a negative Brier Skill Score. The model therefore is not suitable for individual probability estimates or treatment decisions in its current form. Recalibration and validation in temporally or geographically independent populations are needed before implementation, consistent with prediction-model reporting standards.⁹

This discrepancy has important implications for potential implementation. A model can have useful discriminatory performance while producing probabilities that are inappropriate for direct clinical interpretation.⁸ The present risk score should therefore be regarded as a preliminary risk-stratification model rather than a calibrated prediction model.

4.6 Time-Dependent Risk

The Kaplan-Meier analysis demonstrated increasing cumulative probability of virologic failure over the observed follow-up period. The estimated probability was 1.27% at 6 months, 8.10% at 12 months, and 34.51% at 18 months.

The 18-month estimate was similar in the sensitivity analysis, which yielded an estimate of 35.14%. The small difference between the two estimates suggests that the point estimate was relatively stable under the tested alternative censoring specification.

The difference between the directly observed 18-month failure proportion (3.02%) and the Kaplan-Meier estimate (34.51%) is important for interpretation. The observed proportion describes failures actually documented during available follow-up, whereas the Kaplan-Meier estimate models the probability of failure by 18 months under a non-informative-censoring assumption. Because median follow-up was only 5.3 months and only 12 patients remained at risk at 18 months, the tail of the curve is highly uncertain. The estimate should not be used as an individual risk estimate or interpreted as an extrapolation beyond the observed follow-up.

The sensitivity analysis supports stability of the point estimate under the tested specification, but it does not establish that the censoring mechanism was non-informative.

4.7 Facility-Level Variation

Facility-level LLV burden varied substantially, but the descriptive difference in subsequent failure between high- and low-burden facilities should not be interpreted causally. Facility burden may reflect differences in patient case mix, testing intensity, treatment mix, clinical staffing, or other unmeasured factors. Multilevel analyses that account for facility clustering and include facility-level characteristics would be more informative in future work.

Accordingly, the observed difference should not be interpreted as evidence that facility LLV burden itself influences virologic outcomes. Instead, the finding identifies a potential area for future multilevel investigation.

4.8 Strengths and Limitations

This study has several strengths, including a large routine-care cohort, explicit data-cleaning procedures, patient-level clustering in the primary inferential analysis, and sensitivity analyses using time-to-event methods.

Limitations include reliance on routine documentation, incomplete and unequal follow-up, sparse late risk sets, potential residual confounding, structural missingness of prior viral-load measurements, and lack of external validation of the prediction model. Some adjusted estimates were also imprecise, particularly for prior visits and prior viral-load documentation.

The retrospective design limits causal interpretation of the observed associations. Prior viral-load measurements were not documented for 63.6% of LLV observations, although this largely reflected the absence of a preceding measurement rather than conventional missing data.

The risk score has not undergone external validation and demonstrated poor absolute calibration. Accordingly, its predicted probabilities should not be used directly for clinical decision-making.⁹

The facility-level analysis was descriptive and did not adjust for clustering or potential confounding.

Finally, the 18-month Kaplan-Meier estimate was based on only 12 patients remaining at risk at that time point. The estimate should therefore be interpreted cautiously as a model-based cumulative probability. Although the sensitivity analysis produced a similar point estimate, this does not establish that the censoring mechanism was non-informative.

4.9 Future Directions

Future research should validate these associations in other Ethiopian and sub-Saharan African cohorts using harmonized definitions of single and persistent LLV.^{2,3,14} Studies should examine whether associations differ by LLV magnitude and regimen type, evaluate the causal effect of MMD using formal causal-inference methods, externally validate and recalibrate the prediction model, and use multilevel methods to quantify facility-level contributions to virologic outcomes.^{7,9,15}

Multilevel modeling could further examine facility-level variation while accounting simultaneously for patient- and facility-level characteristics.

Prospective studies could also evaluate whether expansion of MMD is associated with changes in virologic outcomes under routine program conditions and identify the mechanisms through which differentiated dispensing may influence treatment continuity.^{7,8}

V. CONCLUSION

In conclusion, later-line ART and absence of MMD were the clearest adjusted correlates of virologic failure following LLV in this Oromia cohort. The findings support risk-aware clinical review and continued viral-load monitoring but do not establish a causal benefit of MMD, a precise 18-month population risk, or readiness for clinical use of the prediction model.

Among adults with LLV receiving ART in Oromia, Ethiopia, MMD of ≥ 90 days and ART regimen line were the principal independent predictors of subsequent virologic failure in the multivariable GEE analysis.

The risk-stratification model demonstrated moderate discrimination but poor absolute calibration, indicating that recalibration and external validation are necessary before clinical application.⁹

Kaplan-Meier analysis estimated a cumulative probability of virologic failure of 34.51% at 18 months, with a similar estimate under the tested sensitivity specification. Because few patients remained at risk at 18 months, this estimate should be

interpreted as a model-based cumulative probability rather than as a direct proportion of the original cohort experiencing failure.

The findings provide programmatically relevant evidence regarding MMD and the management of patients receiving second- or third-line ART and support further prospective evaluation of these factors within differentiated HIV service-delivery models.^{7, 8}

Implications for Policy & Practice

- Patients with LLV who are receiving second- or third-line ART may warrant closer clinical review and follow-up because later-line therapy was associated with higher odds of subsequent virologic failure.^{1-4, 12, 13}
- MMD of at least 90 days was associated with lower odds of failure. This should be interpreted as an observational association rather than a causal effect; local eligibility criteria and potential confounding should be considered.^{7, 8}
- Routine viral-load monitoring should continue beyond the immediate post-LLV period because documented failures accumulated over time.^{10, 13} Late Kaplan-Meier estimates require caution because of sparse follow-up.
- The current prediction model should not be used to generate individual absolute-risk estimates until recalibrated and externally validated. Program use should prioritize validated tools and clinical judgment.⁹

Author Contributions

Concept and design: Amante, and Prof. Muluembet Abera. Acquisition, analysis, or interpretation of data: Amante. Drafting of the manuscript: Amante. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Amante. Administrative, technical, or material support: Amante. Supervision: Prof. Muluembet Abera.

Access to Data: Dr Amante had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of Interest Disclosures

The authors declare no conflicts of interest.

Funding/Support

No specific grant from any funding agency in the public, commercial, or not-for-profit sectors supported this study.

Data Sharing Statement

The data analyzed in this study were obtained from the Oromia Regional Health Bureau and participating health facilities. Restrictions apply to the availability of these data, which were used under institutional approval for the current study. De-identified data may be available from the corresponding author upon reasonable request and with permission from the Oromia Regional Health Bureau.

Disclaimer

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Oromia Regional Health Bureau or the participating health facilities.

Additional Contributions

We thank the Oromia Health Bureau and the 105 participating health facilities for providing access to the data used in this study. We also acknowledge the Texila American University research committee for methodological guidance. No compensation was received for these contributions.

Additional Information

This study was approved by the Texila American University Institutional Review Board and the Oromia Regional Health Bureau.

REFERENCES

- [1]. UNAIDS. Global HIV & AIDS statistics—Fact sheet. 2024. Accessed September 15, 2025. <https://www.unaids.org/en/resources/fact-sheet>
- [2]. Zhao S, Wang W, Li S, et al. The prevalence of low-level viraemia and its association with virological failure in people living with HIV: a systematic review and meta-analysis. *Emerg Microbes Infect.* 2024;13(1):2447613. doi:10.1080/22221751.2024.2447613
- [3]. Chun HM, Abutu AA, Milligan K, et al. Low-level viraemia among people living with HIV in Nigeria: a retrospective longitudinal cohort study. *Lancet Glob Health.* 2022;10(12):e1815-e1824. doi:10.1016/S2214-109X(22)00413-2
- [4]. Hermans LE, Moorhouse M, Carmona S, et al. Effect of HIV-1 low-level viraemia during antiretroviral therapy on treatment outcomes in WHO-guided South African treatment programmes: a multicentre cohort study. *Lancet Infect Dis.* 2018;18(2):188-197. doi:10.1016/S1473-3099(17)30681-3
- [5]. Gaifer Z, Boulassel MR. Low-level viremia predicts virological failure in HIV-infected Omani patients receiving antiretroviral therapy. *J Int Assoc Provid AIDS Care.* 2020;19:2325958220979817. doi:10.1177/2325958220979817
- [6]. Zaçe D, Rindi LV, Compagno M, et al. Managing low-level HIV viraemia in antiretroviral therapy: a systematic review and meta-analysis. *Sex Transm Infect.* 2024;100(7):460-468. doi:10.1136/sextrans-2024-056198
- [7]. Blanco N, Lavoie MC, Ngeno C, et al. Effects of multi-month dispensing on clinical outcomes: retrospective cohort analysis conducted in Kenya. *AIDS Behav.* 2023;27(11):3554-3563. doi:10.1007/s10461-023-04247-1
- [8]. Hoffman RM, Moyo C, Balakasi KT, et al. Multimonth dispensing of up to 6 months of antiretroviral therapy in Malawi and Zambia (INTERVAL): a cluster-randomised, non-blinded, non-inferiority trial. *Lancet Glob Health.* 2021;9(5):e628-e638. doi:10.1016/S2214-109X(21)00039-5
- [9]. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med.* 2015;13:1. doi:10.1186/s12916-014-0241-z
- [10]. Esber A, Polyak C, Kiweewa F, et al. Persistent low level viremia predicts subsequent virologic failure: is it time to change the third 90? *Clin Infect Dis.* 2019;69(5):805-812. doi:10.1093/cid/ciy949
- [11]. Mesic A, Spina A, Mar H, et al. Predictors of virological failure among people living with HIV receiving first line antiretroviral treatment in Myanmar: retrospective cohort analysis. *AIDS Res Ther.* 2021;18(1):16. doi:10.1186/s12981-021-00343-1
- [12]. Bareng O, Moyo S, Choga W, et al. Low-level viremia among adults living with HIV on dolutegravir-based first-line antiretroviral therapy is a predictor of virological failure in Botswana. *Viruses.* 2024;16(5):720. doi:10.3390/v16050720
- [13]. Aoko A, Pals S, Ngugi E, et al. Retrospective longitudinal analysis of low-level viremia among HIV-1 infected adults on antiretroviral therapy in Kenya. *EClinicalMedicine.* 2023;63:102169. doi:10.1016/j.eclinm.2023.102169
- [14]. Leon Z, Edun O, Toska E, et al. Prevalence of low-level viremia and its association with virological failure among adolescents living with HIV in Eastern Cape, South Africa. *HIV Med.* 2026;27(3):245-254. doi:10.1111/hiv.13621
- [15]. Sodeke O, Milligan K, Vallabhaneni S, et al. Longitudinal viral load outcomes of adults with HIV after detectable viremia on tenofovir, lamivudine, and dolutegravir. *AIDS.* 2024;38(10):1503-1512. doi:10.1097/QAD.0000000000003892

Figure 2. Calibration of the risk-stratification model. Observed and predicted probabilities of virologic failure are presented across risk deciles in the held-out test set. The dashed diagonal line represents perfect calibration. The substantial deviation between observed and predicted probabilities demonstrates poor absolute-risk calibration despite moderate discrimination (AUC=0.74).