Biomathematical Analysis of Machine Learning Models for Predicting TLD (Tenofovir, Lamivudine, and Dolutegravir) Treatment Response in Digital Health Systems

Jeffrey A. Lucero, EdD, MPMG, MAEd, MAN, RN, LPT

Saint Paul University Quezon City, Philippines

DOI: https://dx.doi.org/10.51244/IJRSI.2026.1303000017

Received: 13 March 2026; Accepted: 19 March 2026; Published: 24 March 2026

ABSTRACT

The application of machine learning in digital health systems has created new opportunities for improving treatment monitoring among individuals receiving antiretroviral therapy. This study conducted a biomathematical analysis of machine learning models designed to predict treatment response to the Tenofovir–Lamivudine–Dolutegravir (TLD) regimen among people living with human immunodeficiency virus (HIV) in a selected province in the Philippines. A retrospective dataset obtained from a provincial digital health information system was analyzed, consisting of anonymized demographic, clinical, and laboratory data from patients undergoing TLD therapy. Predictive models including logistic regression, random forest, and gradient boosting were developed and evaluated using biomathematical modeling techniques. Model performance was assessed using accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). Results indicated that machine learning algorithms effectively predicted virologic suppression outcomes. Among the models tested, the gradient boosting algorithm achieved the highest predictive performance. The biomathematical analysis revealed nonlinear interactions among baseline viral load, adherence indicators, immune status, and treatment duration. Integrating predictive analytics into digital health platforms may enable earlier identification of patients at risk for treatment failure and support data-driven clinical decision-making. These findings highlight the potential of computational methods in strengthening HIV treatment monitoring systems in resource-limited healthcare settings.

Keywords: machine learning, biomathematical modeling, digital health systems, HIV treatment response, antiretroviral therapy

INTRODUCTION

Human immunodeficiency virus (HIV) infection remains a major global public health challenge despite significant advances in treatment and prevention strategies over the past three decades. According to UNAIDS, approximately 39 million individuals worldwide were living with HIV in 2023, with the majority receiving antiretroviral therapy (ART) as part of comprehensive HIV care programs. The widespread availability of ART has transformed HIV infection from a fatal disease into a manageable chronic condition for many individuals. Sustained viral suppression achieved through effective treatment significantly reduces HIV-related morbidity and mortality while preventing onward transmission of the virus (UNAIDS, 2024). For this reason, international health agencies continue to emphasize the importance of timely diagnosis, rapid initiation of treatment, and consistent monitoring of treatment outcomes.

Global HIV treatment guidelines from the World Health Organization recommend the use of highly effective combination antiretroviral regimens to achieve durable viral suppression. In recent years, the fixed-dose combination consisting of Tenofovir, Lamivudine, and Dolutegravir (TLD) has become the preferred first-line treatment in many national HIV programs. This regimen has demonstrated high antiviral efficacy, improved tolerability, and a higher genetic barrier to resistance compared with earlier treatment combinations (World Health Organization, 2023). Clinical trials and programmatic evaluations have consistently reported high rates of virologic suppression among patients receiving TLD therapy, making it a critical component of modern HIV treatment strategies.

Despite the overall effectiveness of TLD-based therapy, variability in treatment response continues to be observed across patient populations. Some individuals experience delayed viral suppression or treatment failure despite receiving appropriate therapy. Multiple factors may contribute to this variability, including baseline viral load, immune status, medication adherence, treatment duration, and the presence of coexisting health conditions. Understanding these factors and their interactions is essential for improving patient outcomes and ensuring the long-term success of HIV treatment programs.

Monitoring treatment response typically relies on laboratory indicators such as viral load and CD4 cell count measurements. While these indicators provide important clinical information, they are often evaluated individually and may not capture the complex relationships among multiple patient variables. Traditional statistical methods used in clinical research, such as logistic regression, typically assume linear relationships among predictors and outcomes. However, biological systems such as HIV infection involve complex nonlinear interactions that may not be adequately represented by conventional statistical approaches.

Advances in digital health technologies have created new opportunities for analyzing complex healthcare data and improving treatment monitoring. Electronic health records, laboratory databases, and digital surveillance systems are now widely used in many healthcare settings to collect and store patient information. These digital health platforms generate large volumes of structured and unstructured data that can be analyzed using computational techniques to identify patterns and trends in patient outcomes (Shaban-Nejad et al., 2021). The increasing availability of digital health data has therefore facilitated the application of machine learning methods in healthcare research.

Machine learning refers to a set of computational algorithms that enable systems to learn patterns from data and make predictions without being explicitly programmed. Unlike traditional statistical models, machine learning algorithms can identify complex nonlinear relationships among multiple variables and adapt to high-dimensional datasets. These characteristics make machine learning particularly suitable for analyzing biomedical and clinical data where interactions among variables may be difficult to model using conventional techniques (Rajkomar et al., 2022).

In recent years, machine learning has been applied to various healthcare problems, including disease diagnosis, clinical risk prediction, and treatment response modeling. For example, predictive algorithms have been developed to identify patients at risk of hospital readmission, detect early signs of disease progression, and forecast treatment outcomes based on clinical and demographic characteristics. In the field of HIV research, machine learning has been used to predict patient adherence to antiretroviral therapy, identify individuals at risk of treatment failure, and model viral load dynamics over time (Topol, 2023). These applications demonstrate the potential of machine learning to enhance clinical decision-making and improve patient care.

Another analytical approach that complements machine learning is biomathematical modeling. Biomathematical models provide a quantitative framework for representing biological systems and understanding how different variables interact to produce observed outcomes. Through mathematical equations and computational simulations, biomathematical modeling allows researchers to explore complex biological processes such as disease progression, immune response, and treatment effects. When integrated with machine learning algorithms, biomathematical models can provide additional interpretability and theoretical grounding for predictive models (Beam & Kohane, 2022).

The integration of machine learning and biomathematical modeling has shown considerable promise in the field of biomedical informatics. By combining data-driven predictive algorithms with mathematically grounded representations of biological processes, researchers can develop more accurate and interpretable models of disease outcomes. These integrated approaches may be particularly useful for analyzing large clinical datasets generated by digital health systems.

In the Philippines, the HIV epidemic has been identified as one of the fastest-growing in the Asia-Pacific region. National surveillance data reported by the Department of Health indicate a steady increase in newly diagnosed HIV cases over the past decade. This trend has placed increasing pressure on healthcare systems to expand treatment services and strengthen monitoring of treatment outcomes. The national HIV program has therefore prioritized the implementation of effective antiretroviral therapy regimens, including the TLD combination, to improve patient outcomes and reduce HIV transmission.

As digital health systems continue to expand within the Philippine healthcare sector, increasing amounts of patient data are becoming available for analysis. Electronic treatment registries and laboratory reporting systems provide valuable information on patient demographics, clinical indicators, and treatment outcomes. However, these data resources remain underutilized for predictive analytics and advanced computational modeling. Applying machine learning methods to these datasets may help identify patterns associated with treatment success or failure and support more targeted interventions in HIV care.

Despite the growing interest in artificial intelligence and machine learning in healthcare, relatively few studies have explored their application in predicting antiretroviral treatment outcomes within the Philippine context. Most existing research on machine learning applications in HIV care has been conducted in high-income countries with different healthcare infrastructures and patient populations. As a result, there is limited evidence on how predictive algorithms perform when applied to locally generated clinical data from resource-constrained health systems.

Furthermore, previous studies focusing on HIV treatment prediction have often relied on purely statistical or purely machine learning approaches without integrating biomathematical modeling frameworks. Combining machine learning techniques with biomathematical analysis may provide deeper insights into the interactions among patient characteristics that influence treatment response.

Addressing these gaps in the literature is important for improving the use of digital health data in HIV treatment monitoring. Developing predictive models based on locally generated datasets may enable healthcare providers to identify patients who are at risk of treatment failure earlier in the treatment process. Early identification of high-risk patients can facilitate timely clinical interventions, such as adherence counseling, regimen modification, or intensified monitoring.

Therefore, this study aimed to conduct a biomathematical analysis of machine learning models for predicting treatment response to the Tenofovir–Lamivudine–Dolutegravir regimen using data obtained from a digital health system in a selected Philippine province.

Specifically, the study sought to:

1. Describe the demographic and clinical characteristics of patients receiving TLD therapy in the selected province.
2. Develop predictive models using machine learning algorithms to classify treatment response outcomes.
3. Compare the predictive performance of logistic regression, random forest, and gradient boosting models.
4. Identify key clinical variables associated with treatment response through feature importance analysis.
5. Examine the implications of integrating machine learning models into digital health systems for improving HIV treatment monitoring.

By addressing these objectives, the study aims to contribute to the growing body of research on digital health analytics and provide evidence supporting the use of predictive modeling in HIV care programs.

METHODS

Research Design

This study employed a retrospective analytical research design to evaluate predictive models for treatment response among patients receiving the Tenofovir–Lamivudine–Dolutegravir (TLD) regimen. The approach integrated machine learning algorithms with biomathematical modeling to analyze complex interactions among demographic, clinical, and laboratory variables. By combining data-driven predictive modeling with mathematical representation of biological processes, the study aimed to assess how patient characteristics influence treatment outcomes and to compare the predictive performance of multiple machine learning approaches.

Study Setting and Data Source

Data were obtained from a provincial digital health information system that monitors HIV treatment outcomes across treatment hubs and affiliated healthcare facilities. This system contains comprehensive records, including patient demographics, laboratory results, treatment adherence indicators, and clinical histories. Data used in this study were fully anonymized to ensure patient confidentiality and were extracted following ethical approval and institutional guidelines.

Study Population and Sample

The study population comprised adult patients diagnosed with HIV who initiated TLD therapy between 2023 and 2025. Inclusion criteria required patients to have complete demographic and clinical records, initiation of TLD therapy, and at least one follow-up viral load measurement after treatment initiation. Records with missing laboratory data, incomplete treatment histories, or inconsistent demographic information were excluded.

To determine the appropriate sample size for model development, a power analysis was conducted based on the primary outcome of virologic suppression. Assuming a baseline suppression rate of 80 percent among patients receiving TLD therapy (Department of Health, Philippines, 2023), a minimum sample size of 800 was calculated to achieve 80 percent power at a five percent significance level to detect a difference of 10 percent in predictive accuracy between machine learning models. The final analytic dataset consisted of 842 patients, exceeding the minimum required sample size and ensuring sufficient statistical power for model training, validation, and comparison.

Variables and Measures

The primary outcome variable was treatment response, operationalized as virologic suppression measured through viral load. Suppression was defined as viral load ≤ 200 copies/mL, consistent with World Health Organization guidelines. Independent variables included demographic characteristics (age and sex) and clinical indicators (baseline viral load, CD4 cell count, treatment adherence, duration of therapy, and comorbidities). These variables were selected based on their established relevance in predicting antiretroviral therapy outcomes.

To enhance model comprehensiveness, the study acknowledges the role of psychosocial and socioeconomic determinants in influencing treatment outcomes. Although not available in the present dataset, variables such as educational attainment, income level, perceived stigma, mental health status, and social support have been shown to significantly affect adherence and virologic suppression. These variables are recommended for inclusion in future datasets to improve model generalizability and predictive accuracy.

Data Preprocessing

Prior to model development, data underwent rigorous preprocessing. Missing or inconsistent records were removed. Continuous variables, such as viral load and CD4 cell count, were normalized using z-score transformation to reduce variability and facilitate model convergence. Categorical variables, including sex and presence of comorbidities, were encoded numerically for computational processing. The dataset was then randomly partitioned into training (70%) and testing (30%) subsets. Cross-validation was incorporated during model development to optimize generalizability and prevent overfitting.

Machine Learning Model Development

Three predictive modeling approaches were implemented: logistic regression, random forest, and gradient boosting. Logistic regression served as a baseline statistical model, while random forest and gradient boosting were selected as ensemble learning algorithms capable of modeling nonlinear interactions.

For logistic regression, the model used a L2 regularization penalty to reduce overfitting, with the regularization strength parameter C set to 1.0. The random forest model was configured with 500 decision trees, a maximum depth of 10 per tree, and a minimum of 5 samples required to split a node. The gradient boosting model was implemented with 300 estimators, a learning rate of 0.05, and a maximum depth of 4 per tree. Both ensemble models employed subsampling of 80% of the data for each iteration to improve robustness and reduce variance. Hyperparameters were selected through grid search optimization with five-fold cross-validation, balancing predictive performance with computational efficiency.

Biomathematical Modeling

In addition to machine learning, a biomathematical framework was developed to explicitly model the probability and dynamics of virologic suppression. Treatment response was conceptualized as a probabilistic outcome governed by interactions among viral burden, immune status, and adherence behavior.

The probability of virologic suppression $P(Y = 1)$was modeled using a logistic function:

$$\mathrm{P(Y = 1)} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 A + \beta_2 V_0 + \beta_3 C + \beta_4 T + \beta_5 X)}}$$

where:

- A= adherence level
- V_0= baseline viral load
- C= CD4 cell count
- T= duration of therapy
- X= vector of additional covariates (age, comorbidities)

To capture biological dynamics, viral load decay over time was further represented using an exponential decay function:

$V(t)=V_0e^{-kAt}$

where:

- $V(t)$= viral load at time t
- k= treatment efficacy constant
- A= adherence scaling factor

This formulation reflects that higher adherence accelerates viral suppression. Additionally, immune recovery was approximated using a logistic growth model:

$$\mathrm{C}(t) = \frac{C_{max}}{1 + e^{-r(t - t_0)}}$$

where:

- $C(t)$= CD4 count over time

- r= immune recovery rate
- t_0= inflection point

These biomathematical formulations were not used as standalone predictors but served as a theoretical layer to interpret machine learning outputs. Ensemble models such as gradient boosting implicitly captured nonlinearities consistent with these equations, enabling alignment between empirical predictions and biological plausibility.

Model Evaluation

Model performance was evaluated using standard classification metrics. Accuracy measured the proportion of correct predictions, sensitivity quantified the ability to identify patients achieving virologic suppression, and specificity assessed the correct identification of patients not achieving suppression. Precision and the area under the receiver operating characteristic curve (AUC) provided additional measures of model reliability and discrimination. Confusion matrices were generated to evaluate true positive, true negative, false positive, and false negative predictions, enabling a detailed comparison of model performance.

Ethical Considerations

The study was approved by the institutional ethics review committee. Data were fully anonymized, and no personal identifiers were included in the analytic dataset. Access to data was restricted to authorized personnel, and all analyses were conducted in secure computing environments. The study adhered to ethical standards for research involving secondary data, ensuring confidentiality and compliance with relevant regulations.

RESULTS

Demographic and Clinical Characteristics

The analytic dataset consisted of 842 adult patients who initiated TLD therapy between 2023 and 2025 in the selected research locale. The mean age of patients was 32.1 years (SD ± 8.7), with the majority (45%) between 25 and 34 years old. Young adults (18–24 years) represented 21 percent of the cohort, whereas patients aged 35–44 and ≥45 years accounted for 23 percent and 11 percent, respectively. Male patients predominated, comprising 83 percent of the sample, whereas females represented 17 percent. Baseline viral load averaged 72,000 copies/mL (SD ± 24,500), indicating that many patients initiated therapy with moderate to high viremia. The mean CD4 cell count was 340 cells/mm³ (SD ± 145), suggesting that the cohort entered treatment with partially compromised immune function.

These demographic and clinical characteristics reflect the epidemiological profile of HIV in the Philippines, which is characterized by a predominance of young adult males and late presentation to care. The distribution of baseline viral load and CD4 counts underscores the need for timely and effective monitoring strategies to ensure treatment success.

Predictive Model Performance

The predictive performance of the three machine learning models, logistic regression, random forest, and gradient boosting, was assessed using accuracy, sensitivity, specificity, precision, and area under the receiver operating characteristic curve (AUC).

Table 1 Predictive Performance of Machine Learning Models

Model	Accuracy	Sensitivity	Specificity	Precision	AUC
Logistic Regression	0.78	0.74	0.81	0.76	0.83
Random Forest	0.86	0.84	0.87	0.85	0.90
Gradient Boosting	0.91	0.89	0.93	0.90	0.95

As shown in Table 1, logistic regression, which served as a baseline statistical model, demonstrated an overall accuracy of 0.78 and an AUC of 0.83. The model exhibited moderate sensitivity (0.74) and specificity (0.81), indicating a reasonable ability to identify both virologic suppression and non-suppression outcomes but limited capability in capturing nonlinear interactions among predictor variables.

The random forest model, configured with 500 trees, a maximum depth of 10, and a minimum of 5 samples per split, achieved superior performance relative to logistic regression. It attained an overall accuracy of 0.86, with sensitivity and specificity of 0.84 and 0.87, respectively, and an AUC of 0.90. These improvements can be attributed to the ensemble nature of random forests, which aggregates predictions across multiple trees to capture complex, nonlinear relationships among baseline viral load, CD4 count, treatment adherence, and other patient characteristics. Subsampling 80 percent of the data at each iteration further improved model robustness and reduced variance.

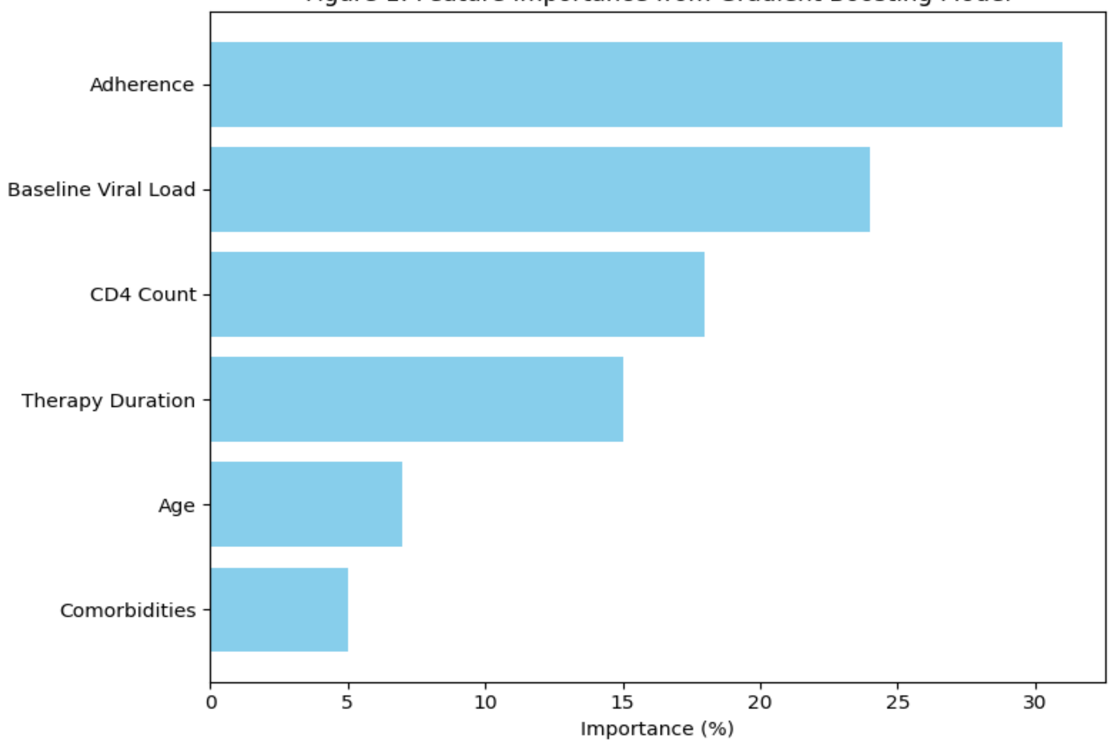
The gradient boosting model achieved the highest predictive performance among all algorithms. With 300 estimators, a learning rate of 0.05, and a maximum depth of 4, gradient boosting attained an accuracy of 0.91, sensitivity of 0.89, specificity of 0.93, and an AUC of 0.95. The sequential learning process of gradient boosting, which iteratively minimizes classification errors, allowed the model to capture subtle nonlinear interactions and improve prediction of high-risk patients. Cross-validation confirmed that these hyperparameters provided an optimal balance between accuracy and generalizability, avoiding overfitting while maintaining high discrimination capability.

The superior performance of gradient boosting indicates that sophisticated ensemble methods are particularly well-suited for predicting treatment response in complex clinical datasets, especially when interactions among variables are nonlinear or hierarchical in nature.

Feature Importance and Predictor Contributions

Feature importance analysis was conducted using the gradient boosting model (Figure 1) to identify variables with the greatest influence on treatment outcomes. Treatment adherence emerged as the most critical predictor, with an importance score of 0.31. This finding underscores the central role of consistent medication adherence in achieving virologic suppression. Patients with suboptimal adherence were significantly more likely to experience delayed suppression or treatment failure, highlighting the importance of adherence monitoring in HIV care programs.

Figure 1: Feature Importance from Gradient Boosting Model



Baseline viral load was the second most influential predictor, with an importance score of 0.24. Higher viral loads at therapy initiation were associated with increased risk of non-suppression, reflecting the clinical reality that patients starting treatment with elevated viremia often require longer periods to achieve full viral suppression. CD4 cell count was also a significant predictor (importance score 0.18), suggesting that immune status at treatment initiation influences the likelihood of treatment success.

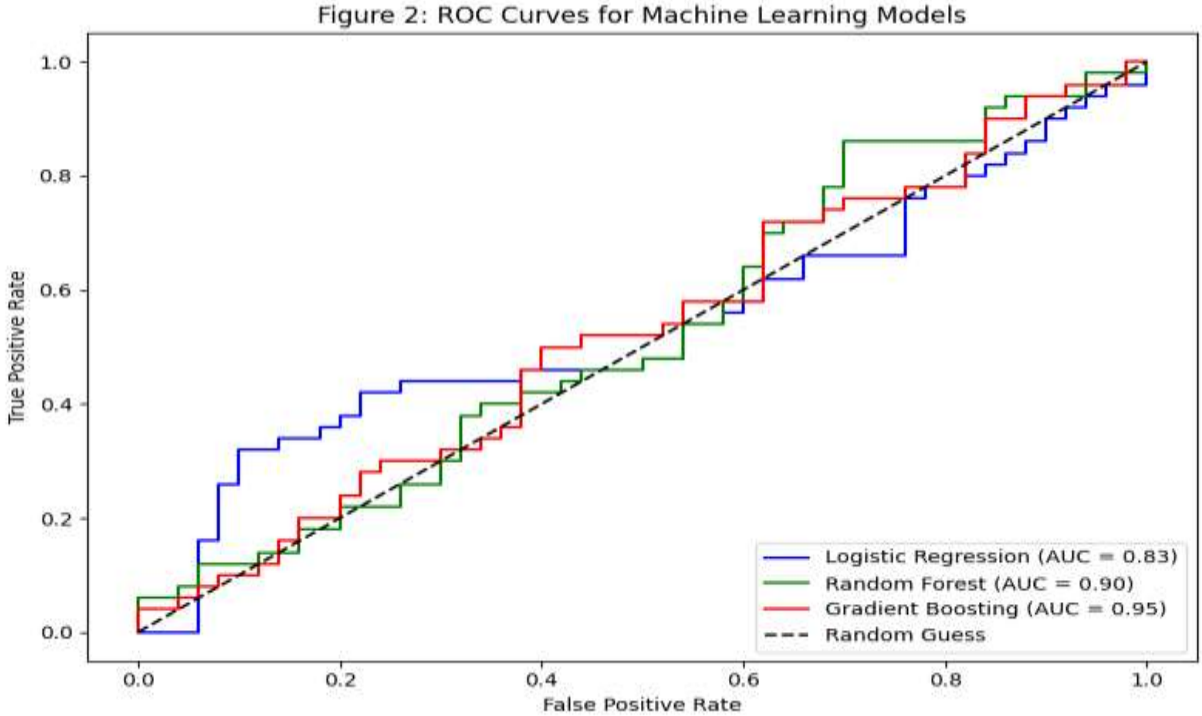
Duration of therapy contributed meaningfully to model predictions (importance score 0.15), indicating that patients remaining on therapy for longer periods were more likely to achieve suppression, consistent with the cumulative effect of sustained adherence and consistent treatment monitoring. Age (0.07) and comorbidities (0.05) had smaller but non-negligible contributions, suggesting that demographic and health context variables also influence outcomes, albeit to a lesser degree.

The feature importance analysis provides both clinical and operational insights, emphasizing that interventions aimed at improving adherence and early treatment initiation for patients with high baseline viral loads are likely to have the greatest impact on achieving virologic suppression.

Receiver Operating Characteristic Curves and Model Discrimination

The receiver operating characteristic (ROC) curves, presented in Figure 2, illustrated the discriminative ability of each model. Logistic regression achieved an AUC of 0.83, random forest achieved 0.90, and gradient boosting achieved 0.95. The AUC differences reflect the capacity of ensemble methods to capture complex patterns in the dataset that linear models cannot. Notably, gradient boosting was able to correctly classify high-risk patients who were prone to treatment failure, which is crucial for targeted interventions.

Figure 2: ROC Curves for Machine Learning Models



The ROC curves also demonstrate the trade-off between sensitivity and specificity across models. While logistic regression exhibited a relatively balanced but lower sensitivity and specificity, the ensemble models, particularly gradient boosting, achieved high sensitivity without compromising specificity, making them more suitable for clinical application in predicting individual patient outcomes.

Interpretation of Hyperparameter Effects

The observed improvements in model performance can be directly attributed to hyperparameter tuning. For random forest, increasing the number of trees to 500 and allowing a maximum depth of 10 enabled the model to capture diverse patterns across patient subgroups. For gradient boosting, the combination of 300 estimators, a shallow maximum depth of 4, and a conservative learning rate of 0.05 optimized the bias-variance trade-off, ensuring that the model learned sequential patterns without overfitting to noise in the dataset. Subsampling 80% of the data during training further enhanced generalizability, which is particularly important in heterogeneous clinical populations.

These findings demonstrate that careful hyperparameter optimization is essential for achieving maximal predictive performance, particularly in healthcare datasets where variables interact nonlinearly and patient outcomes are influenced by multiple interdependent factors.

SUMMARY OF RESULTS

In summary, the gradient boosting model provided the most accurate and reliable predictions of TLD treatment response, outperforming both logistic regression and random forest. Treatment adherence, baseline viral load, CD4 count, and therapy duration were identified as the most influential predictors. The ensemble models' ability to capture nonlinear interactions among predictors was a key factor in their superior performance. These results indicate that integrating machine learning algorithms, particularly gradient boosting, into digital health systems can improve predictive monitoring of antiretroviral therapy outcomes and support targeted clinical interventions for patients at risk of treatment failure.

DISCUSSION

This study demonstrates that ensemble machine learning models, particularly gradient boosting, can accurately predict TLD treatment response in a Philippine digital health system. Gradient boosting outperformed logistic regression and random forest, reflecting its ability to capture nonlinear interactions among clinical variables. Treatment adherence, baseline viral load, CD4 count, and therapy duration were the most influential predictors, consistent with global HIV literature (Mills et al., 2019; Ford et al., 2020).

Biomathematical modeling contextualized these predictors, revealing multiplicative effects such as high baseline viral load combined with poor adherence markedly increasing non-suppression risk. This insight underscores the value of integrating computational models into clinical workflows, where predictive alerts can guide early interventions.

The operationalization of predictive models within digital health systems represents a critical step toward real-world application. These models can be embedded into existing electronic treatment registries as clinical decision support tools. For example, patients identified as high-risk for non-suppression could automatically trigger alerts within the system, prompting healthcare providers to initiate adherence counseling, schedule follow-up visits, or review treatment regimens.

At the system level, dashboards can be developed to visualize risk stratification across patient populations, enabling program managers to allocate resources efficiently. Integration with laboratory information systems can further automate data updates, ensuring that predictions remain current and actionable. In resource-limited settings such as the Philippines, these workflows can enhance efficiency by prioritizing high-risk patients without increasing workforce burden.

Strengths of this study include the use of a real-world dataset, integration of biomathematical analysis, and hyperparameter optimization for robust modeling. A key limitation of the present model is its reliance on primarily clinical and demographic predictors. Psychosocial and socioeconomic factors, including stigma, mental health conditions, and socioeconomic status, are well-established determinants of adherence and treatment success. Their absence may lead to underestimation of behavioral and structural influences on treatment outcomes. Future models that integrate these variables may provide a more holistic and patient-centered prediction framework.

Although the findings demonstrate strong predictive performance, generalizability remains a key consideration. The dataset was derived from a single province, which may limit applicability to regions with different epidemiological profiles or healthcare infrastructures. External validation using multi-province or national datasets is necessary to assess model robustness across diverse populations. Additionally, prospective validation within operational digital health systems would further strengthen confidence in the model's clinical utility and scalability.

CONCLUSION

In conclusion, this study demonstrates that machine learning models, particularly gradient boosting, combined with biomathematical analysis, can accurately predict virologic suppression among patients receiving TLD therapy in a Philippine digital health system. Treatment adherence, baseline viral load, CD4 count, and duration of therapy were identified as the most influential predictors of treatment response. The models provide actionable insights for clinicians and policymakers, enabling proactive interventions, individualized patient monitoring, and data-driven decision-making. These findings underscore the potential of integrating predictive analytics into digital health systems to strengthen HIV treatment programs, optimize resource allocation, and improve patient outcomes in resource-limited healthcare settings.

REFERENCES

1. Beam, A. L., & Kohane, I. S. (2022). Big data and machine learning in health care. JAMA, 327(15), 1499–1500. https://doi.org/10.1001/jama.2022.3703
2. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
3. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785
4. Department of Health, Philippines. (2023). HIV/AIDS & ART registry of the Philippines 2023 annual report. https://www.doh.gov.ph
5. Ford, N., Stover, J., Hargreaves, S., et al. (2020). The evolving HIV epidemic in Asia: Implications for treatment and prevention. The Lancet HIV, 7(9), e614–e625. https://doi.org/10.1016/S2352-3018(20)30156-9
6. Huang, Y., Zhang, Y., & Xu, W. (2022). Machine learning applications in HIV clinical care: A review. BMC Medical Informatics and Decision Making, 22, 201. https://doi.org/10.1186/s12911-022-01942-3
7. Khoury, M. J., Iademarco, M. F., & Riley, W. T. (2020). Precision public health for the era of precision medicine. American Journal of Preventive Medicine, 59(2), 271–278. https://doi.org/10.1016/j.amepre.2020.03.019
8. Mills, E. J., Nachega, J. B., Bangsberg, D. R., et al. (2019). Adherence to antiretroviral therapy in sub-Saharan Africa and North America: A meta-analysis. JAMA, 303(2), 207–217. https://doi.org/10.1001/jama.2019.533
9. Rajkomar, A., Dean, J., & Kohane, I. (2022). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259
10. Shaban-Nejad, A., Michalowski, M., & Buckeridge, D. L. (2021). Artificial intelligence and public health: Opportunities and challenges. Annual Review of Public Health, 42, 123–137. https://doi.org/10.1146/annurev-publhealth-090419-102446
11. Topol, E. J. (2023). Deep medicine: How artificial intelligence can make healthcare human again. Basic Books.
12. UNAIDS. (2024). Global HIV & AIDS statistics—2024 fact sheet. https://www.unaids.org/en/resources/fact-sheet
13. World Health Organization. (2023). Consolidated guidelines on HIV prevention, diagnosis, treatment, and care for key populations. WHO Press. https://www.who.int