University of Nairobi

Department of Mathematics

ISSN: 2410-1397

Bachelors Project in Applied Mathematics

Logistic Regression Modeling of Factors Influencing Mother-to-Child HIV Prevention and Transmission

Telvin,Brian,Gladys,Joyce,Jane

May 2025

Submitted to the Department of Mathematics in partial fulfilment for a degree in Bachelor of Science in Applied Mathematics

Logistic Regression Modeling of Factors Influencing Mother-to-Child HIV Prevention and Transmission

Telvin,Brian,Gladys,Joyce,Jane

Department of Mathematics
Faculty of Science and Technology
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Bachelors Project

Submitted to the Department of Mathematics in partial fulfilment for a degree in Bachelor of Science in Applied Mathematics

Submitted to: The Department of Mathematics, University of Nairobi, Kenya

Abstract

In Kenya, Mother to Child Transmission (MTCT) of HIV, has led to significant consequences. It has led to a rise in HIV positive children and increased infant mortality rate. This study aims to analyze factors influencing Mother to Child Transmission (MTCT) using a Logistic Regression Model.

This research contributes to a deeper understanding on factors influencing Mother to Child Transmission of HIV and it's significant consequences in Kenya. The Logistic Regression Model will be employed to analyze the HIV transmission considering factors like, ART Adherence, Maternal Viral Load, Mode of Delivery, Infant Feeding Mode and Received Adherence Counseling. The Model's output will then be linked to what factors are more significant when it comes to Mother to Child Transmission.

By analyzing data collected from three counties in Kenya, this study seeks to identify key factors influencing Mother to Child Transmission and potential resilience in HIV positive children and infant mortality rate in Kenya. The findings will be important in identifying and introducing key programs and targeted interventions to curb Mother to Child Transmission and bring it to a global 2.5

Bachelors Project in Mathematics at the University of Nairobi, Kenya.
ISSN 2410-1397: Research Report in Mathematics

DISTRIBUTOR: Department of Mathematics, University of Nairobi, Kenya

Declaration and Approval

We, the undersigned, declare that this project is our original work and to the best of our knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

30/05/2025

Signature Date

JANE WAMUYU KENDA
Reg No. I09/1793/2021

30/05/2025

Signature Date

GLADYS WAMBUI NJOROGE
Reg No. I09/1785/2021

30/05/2025

Signature Date

TELVIN NDEGWA WAMUYU
Reg No. I09/1795/2021

30/05/2025

Signature Date

Joyce Nyamvula
Reg No. I09/1803/2021

30/05/2025

Signature Date

Brian Oichoe
Reg No. I09/1638/2020

In my capacity as a supervisor of the candidate's project, I certify that this project has my approval for submission.

Signature Date

Dr Benard Kipchumba
Department of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: b.c.kipchumba@gmail.com

Signature Date

Mr Calvince Ouma
Department of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: oumacalvince@uonbi.ac.ke

Dedication

This project is dedicated to Jane Wamuyu Kenda, Gladys Wambui Njoroge, Joyce Nyamvula, Brian Oichoe, Telvin Ndegwa Wamuyu

Contents

Acknowledgments

We are grateful to the Almighty Lord for providing us with the knowledge and strength to complete this project.

We express our sincere gratitude to our exceptional supervisor, Dr. Benard Kipchumba, for his extraordinary support throughout this research journey.His constant guidance, unwavering dedication, and professionalism were instrumental in overcoming the challenges we faced, including research resources. His mentorship truly illuminated the path for our team. We are deeply grateful for everything, Dr. Kipchumba. May you be blessed. We also extend our thanks to the School of Mathematics for providing us with this valuable opportunity to engage in this project and gain practical experience.

We are indebted to our fellow team members, Jane, Gladys, Joyce, Brian, and Telvin, for their dedication and collaborative spirit. Their contributions and teamwork were essential for the successful completion of this project. Finally, we are grateful to all those who participated directly or indirectly in this project, including our families and loved ones, for their unwavering support and encouragement.

Group Members

Nairobi, 2025.

1 Chapter 1: Introduction

According to (1) vertical transmission of HIV, also known as mother-to-child transmission (MTCT), occurs when HIV is passed from an HIV positive mother to her child during pregnancy, labor, delivery, or breastfeeding. Furthermore, without preventive or treatment measures, the transmission rate is estimated to be between 15% and 45%, but with efficient prevention programs and resources, this rate can be reduced to as low as 5% (2) (3). In this regard, we navigate through different deterministic mathematical models with the aim of providing a structured approach to understanding the dynamics of MTCT by quantifying transmission probabilities and evaluating the impacts of the intervention. These models are critical for designing targeted prevention of mother-to-child transmission (PMTCT) programs, optimizing resource allocation, and predicting epidemic trends. So, to obtain all of this, we have to look for wide views, globally, regionally, and locally (Kenya to be precise), on MTCT, so as to contextualize the development of a logistic regression model of factors associated with the prevention and transmission of HIV from mother to child.

1.1 Background Information

At a global view, significant progress has been made in reducing MTCT. Reportedly by 2022, the rate of new HIV infections in children had declined by 58% (4), largely due to the expanded access to antiretroviral treatment (ART) and PMTCT interventions, and in 2019, approximately 85% of pregnant women living with HIV/AIDS had access to ART to prevent MTCT, however 1.3 million women and girls living with HIV are recorded expectant yearly, underscoring the ongoing challenge. Moreover, international organizations such as the World Health Organization (WHO) and UNAIDS advocates for integrated health services, including lifelong ART (Option B+) (Elisabeth Glaser 2017)(5), early infant diagnosis, and exclusive breastfeeding with prophylaxis, to achieve the global target of eliminating MTCT by 2030. Despite these efforts, the objectives remain incomplete, especially in remote and underprivileged areas where coverage and adherence issues persist.

In the Sub-Saharan region of Africa, there are accounts of over two-thirds of global HIV infections and more than 90% of pediatric HIV cases on MTCT (6). To further, in 2022, the region reported high ART coverage for pregnant women (approximately 80%), though MTCT rates remain elevated due to barriers such as nondisclosure of one's HIV status, late ART initiation, and irregular adherence. Moreover, the region faces challenges such as inadequate health infrastructure, stigma, and cultural practices that affect breastfeeding choices(7). Also, disparities in access to early infant testing and follow up postnatal care hinders progress of different programs like the Global Plan Towards the Elimination of New HIV Infections Among Children aim for MTCT rates below 5%. Spatial temporal analyses in East Africa highlight geographical variations in PMTCT uptake, emphasizing the need for region-specific interventions(8).

Locally, Kenya is ranked among countries with high MTCT rates, despite significant strides in PMTCT implementation(9). For instance, before the adoption of WHO'S Option B+, the reported rate of MTCT was 8.3% down from previous estimate of 16%, with estimated 75% pregnant women on ART, and after the adoption the ART uptake doubled in hospitals, reducing pediatric transmission to 2.5% in the sampled facilities(10). However, challenges such as late infant testing (only 42% of HIV exposed infants tested by 6 weeks), poor coordination between antenatal and HIV care, and low adherence to ART among mothers, persist(11). Again, regional disparities are also factored, with Nyanza region reporting an 18.5% HIV prevalence among women of childbearing age(12). Consequently, focus of the recent studies is on the ongoing needs for integrated care, community-based support, and enhanced monitoring to achieve Kenya's target of less than 5% MTCT. These local dynamics signifies the parameters for a logistic regression model of factors associated with mother to child HIV prevention and transmission tailored to Kenya's context.

1.2 Problem Statement

Kenyan public health authorities have tried to curb the HIV mother to child transmission by introducing various interventions such as ART, but their efforts remain a major problem in most rural areas because they have limited resources(13). Since the HIV MTCT occurs during pregnancy, breastfeeding and during delivery, the transmission of the virus can be reduced by medical services such as timely Antiretroviral Therapy, proper mode of delivery and appropriate breastfeeding process. The challenge however is finding a combination of strategies that are most effective in decreasing the HIV MTCT rates in these rural areas with limited resources. Therefore, this study will be important since we will develop a logistic regression model of factors associated with mother to child HIV prevention and transmission to show how various intervention strategies and decisions can be made to help curb the virus in areas with limited resources.

1.3 Objectives of the Study

1.3.1 General Objective

To analyze the factors associated with the prevention or transmission of HIV from mother to child among HIV-positive mothers using logistic regression.

1.3.2 Specific Objectives

(i). To determine the association between ART adherence and mother-to-child HIV transmission.

(ii). To evaluate how maternal viral load levels influence the likelihood of HIV transmission to the infant.

(iii). To assess the impact of mode of delivery (e.g., vaginal vs cesarean) on mother-to-child transmission rates.

(iv). To examine the relationship between infant feeding mode (e.g., exclusive breastfeeding, formula feeding) and HIV transmission outcomes.

(v). To assess whether receiving adherence counseling reduces the risk of HIV transmission from mother to child.

1.4 Research Questions

(i). How does ART adherence among HIV-positive mothers influence the risk of HIV transmission to their infants?

(ii). What is the relationship between maternal viral load levels and the likelihood of mother-to-child HIV transmission?

(iii). Does the mode of delivery (vaginal vs. cesarean section) affect the rate of HIV transmission from mother to child?

(iv). What is the effect of infant feeding mode on the risk of HIV transmission from mother to child?

(v). Does receiving adherence counseling reduce the probability of mother-to-child HIV transmission?

1.5 Hypothesis

The study considered the following hypotheses

(i). H_0: There is no significant association between the viral load of the mother and the mode of delivery
H_1: There is significant association between the viral load of the mother and the mode of delivery

(ii). H_0: Adherence to ART treatment has a positive correlation with counseling
H_1: Adherence to ART treatment has no positive correlation with counseling

(iii). H_0:There is no association between higher maternal viral load and increased risk of MTCT HIV.
H_1:Higher maternal viral load leads to increased risk of MTCT HIV.

1.6 Study Justification

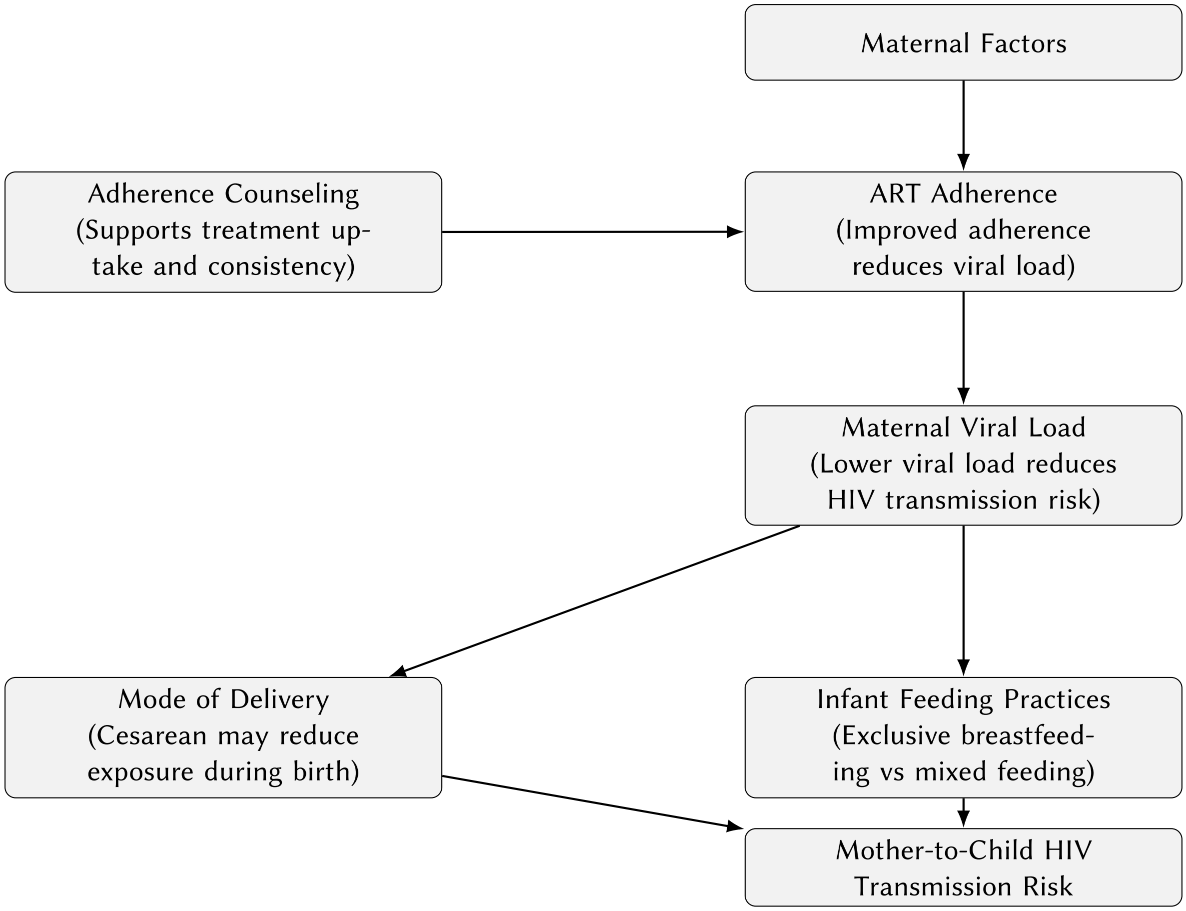
The mother transmitting HIV to her child is one way through which HIV spreads. However, even though there are treatments like ART, safe delivery methods and appropriate ways of breastfeeding, some individuals are not doing this treatments the right way and as a result we see this avoidable transmissions still happening. Developing this logistic regression model of factors associated with mother to child HIV prevention and transmission can assist in providing valuable insights into various factors that make the virus transmission rates to be very high. This study plays an important role in reducing the transmission of HIV from mother to child by using evidence-based data to guide the immunization decisions and proper actions.

1.7 Limitations of the Study

This study, though effective and reliable has some limitations to the research. The time period allocated for the research was limiting studying further regions and conducting a long term analysis of the effects. As a result, the study could only be done on a small selected group of participants from a selected area thus limiting the generalizability of the study. Future study on the subject should consider allocating ample time and broadening the scope area of study to put into account other various factors in other areas that could affect the study and have not been put into consideration in this study.

1.8 Conceptual Framework

The flowchart below shows how various maternal and healthcare related factors correlate to influence the risk of HIV transmission from a mother to her child. The main focus is on the maternal factors, ART adherence in particular which is strongly influenced by adherence counseling. Good ART adherence leads to lower viral load. As the viral load reduces, it influences critical decisions such as, the mode of delivery and the mode of breastfeeding. These choices directly affect the possibility of transmission of HIV to the child.



From the flow chat above, it represents on how Adherence to counseling affects directly the ART Adherence and maternal viral load. On the other hand it also shows how the mode of delivery affects maternal viral load and infant feeding practices directly. Altogether, these elements work together to influence the likelihood of HIV being transmitted from mother to child.

2 Chapter 2: Literature Review

2.1 Introduction

The HIV transmission of mother to child often occur during pregnancy, delivery, labor or breastfeeding. It is challenging to high prevalence regions such as sub-Saharan Africa to get proper interventions that prevent HIV mother to child transmission during labor, delivery and breastfeeding due to factors such as healthcare system limitations, ART adherence and other cultural factors. According to a study made by(14), they reported that without any interventions to eliminate HIV MTCT such as ART, the risk of MTCT rates in these sub-Saharan Africa regions range from 15% to 45% that is 5%–10% during pregnancy, 10–20% during labor and childbirth and 10–20% through mixed infant feeding. In comparison these HIV MTCT rates can be reduced to below 2% and 5% with effective prevention and proper treatment interventions, this study was recorded by(15).

A logistic regression model of factors associated with mother to child HIV prevention and transmission will enable researchers to quantify transmission risks, it will help to evaluate intervention strategies and help authorities to make policies to prevent HIV MTCT. These models use differential equations to stimulate the spread of HIV in a population, it also helps predict what might happen and give guidance on critical and significant decisions. This review brings together theoretical and empirical literature from 2019 to 2025, focusing on logistic regression model of factors associated with mother to child HIV prevention and transmission in order to identify what has been done, what's missing, and how to build a model that fits a specific situation.

2.2 Theoretical Literature

Theoretical frameworks for modeling MTCT (Mother-to-Child Transmission) increasingly incorporate statistical and machine learning methods to capture key prevention and risk factors. Logistic regression, a widely used statistical model for binary outcomes, provides a robust approach to analyzing the probability of HIV transmission based on maternal, clinical, and behavioral variables(16). Hosmer, Lemeshow, and Sturdivant (2013) emphasize that logistic regression is suitable when the outcome is dichotomous—such as HIV-positive versus HIV-negative infant status—and when predictors are both continuous and categorical, fitting the the prevention of mother to child transmission (PMTCT) context.

Foundational epidemiological models(17) initially applied deterministic structures like SI and SEIR to MTCT dynamics, focusing on population-level transitions. However, with the availability of individual-level data, contemporary studies advocate for regression frameworks that quantify associations between specific maternal and programmatic factors and MTCT outcomes. This shift enables the identification of critical intervention points at the patient level.

Recent theoretical contributions emphasize the integration of adherence behaviors and clinical indicators.(18) argue that maternal viral load suppression—largely driven by ART adherence—is the most significant predictor of MTCT risk. Their work highlights the need for models that include ART adherence, viral load measurements, and preventive interventions like adherence counseling. Logistic regression frameworks are particularly suited for adjusting for these covariates while controlling for confounders such as mode of delivery and infant feeding practices.

Behavioral factors, such as counseling and social support, are increasingly recognized in theoretical models.(19) stress the importance of including variables like receipt of adherence counseling, noting that behavioral interventions have measurable effects on clinical outcomes. Moreover, breastfeeding practices significantly influence postnatal transmission, necessitating their inclusion as independent variables(20).

Thus, theoretical advancements recommend shifting from purely deterministic frameworks toward statistical models like logistic regression that incorporate biomedical, behavioral, and health system factors to predict and prevent MTCT effectively.

2.3 Empirical Literature Review

Empirical studies between 2019 and 2024 increasingly focus on individual-level factors affecting MTCT, providing a strong foundation for logistic regression modeling. WHO (2021) reported that ART adherence above 95% correlates with a greater than 90% reduction in vertical transmission risk, underlining the importance of including adherence as a primary predictor.(21) found that infants born to mothers with suppressed viral loads (<50 copies/mL) had a transmission risk below 1%, reinforcing viral load as a critical empirical variable.

In Kenya,(22) examined PMTCT programs and found that mothers who received adherence counseling were 30% more likely to maintain viral suppression, suggesting a direct link between counseling interventions and infant HIV status. Similarly,(23) showed that efective cesarean delivery reduced intrapartum transmission risk by 45% compared to vaginal delivery, justifying the inclusion of mode of delivery as a covariate.

Feeding practices remain a substantial postnatal transmission route.(24)documented a twofold increase in MTCT rates among mixed-fed infants compared to those exclusively breastfed while on ART.(25) further demonstrated that counseling on feeding practices reduced inappropriate feeding and subsequent transmission, validating infant feeding mode as a strong independent variable.

Moreover,(26) emphasized diagnostic delays and variable ART adherence rates as persistent challenges, advocating for dynamic modeling of adherence-related parameters.(27) echoed these concerns, noting that despite national ART coverage improvements, late initiation, poor adherence, and suboptimal feeding practices continue to drive pediatric HIV infections.

Overall, empirical evidence supports the modeling of ART adherence, maternal viral load, mode of delivery, infant feeding practices, and receipt of adherence counseling as key determinants of mother-to-child HIV transmission, aligning closely with the selected logistic regression approach.

2.4 Chapter Summary

The theoretical and empirical literature from 2019 to 2025 provides a strong foundation for developing a logistic regression model of factors associated with mother to child HIV prevention and transmission, where the theoretical literature breaks down the population into groups which helps in modelling HIV MTCT by showing how different factors like how the infected, the uninfected or those on treatment affect the spread of HIV MTCT through spatial dynamics like movement or human behavior for example how consistently the mothers take medication and other social behaviors. On the other hand empirical literature offers real world data for modelling HIV MTCT such as transmission rates, how effective ART works in reducing the risk of HIV MTCT and specific challenges in a certain regions. However, there are some missing information like data for most individuals in rural areas and because of this gaps extra care is required when considering the model parameters to be used. A logistic regression model of factors associated with mother to child HIV prevention and transmission built on these insights can help us understand and reduce HIV MTCT if it reflects on the real life factors, the individuals behavior and location based challenges of the specific regions.

3 Chapter 3: Research Methodology

3.1 Introduction

This chapter explains the methodology used to study factors influencing transmission of HIV from mother to child in Kenya. The goal is to understand the different factors that increase or reduce the risk of this type of transmission. This knowledge can help improve Prevention of Mother-to-Child Transmission (PMTCT) programs and reduce the number of children infected with HIV.

To do this, we use a quantitative approach, meaning we work with numbers and data to find patterns. Specifically, we use a type of mathematical model called Binary logistic regression. This model helps us see how different biological, behavioral, and health program factors are related to whether a child becomes infected or not.

3.2 Source of Data

Our Dataset is from Kenya Demographic and Health Survey (KDHS), extracted from PMTCT (Prevention of Mother-To-Child Transmission) dataset dated December 2021. The dataset comprises information from health facilities across three counties, representing real-world PMTCT service delivery outcomes in Kenya. The study population includes HIV-exposed infants and their mothers who were enrolled in PMTCT programs. The inclusion criteria was: availability of data on final infant HIV status (from First DNA PCR Test results, Results for confirmatory DNA PCR, or antibody-test-results-18months) and complete records for selected maternal and infant variables.

3.3 Research Design

This study adopts a cross-sectional research design, which involves the analysis of data collected at a single point in time. The design is appropriate for identifying associations between key maternal and infant factors and the HIV status of infants born to HIV-positive mothers within a Prevention of Mother-To-Child Transmission (PMTCT) program.

The study utilizes secondary data extracted from health facility records, which include variables relevant to maternal care, treatment adherence, delivery practices, and infant outcomes. The primary objective is to determine the influence of selected maternal and clinical factors on the likelihood of vertical transmission of HIV.

3.4 Justification

A logistic regression model is employed to analyze the relationship between the outcome and explanatory variables (16). The outcome variable is binary, representing whether the infant tested HIV-positive or HIV-negative.

This design is well-suited for assessing risk factors associated with MTCT of HIV, particularly in real-world healthcare settings. It allows for the identification of statistically significant associations that can inform public health interventions and policy adjustments aimed at improving PMTCT outcomes.

3.5 Description of Variables

Table 1. Description of Study Variables

Variable	Type of Variable	Levels / Coding
Maternal adherence to ART	Categorical (Binary)	0 = Poor/Fair, 1 = Good
Maternal viral load	Categorical (Binary)	0 = Suppressed, 1 = Unsuppressed
Mode of delivery	Categorical (Binary)	0 = Vaginal, 1 = Caesarean section
Infant feeding method	Categorical	0 = Exclusive breastfeeding, 1 = Mixed feeding, 3= Formula
Received adherence counseling	Categorical (Binary)	0= No, 1= Yes

3.6 Model Description

3.6.1 Simple Linear Regression Model

Simple linear regression model is a statistical technique that examines the relationship between one independent variable(x) and one dependent variable(y)(28). The outcome is related to single predictor.

$$y = \beta_0 + \beta_1 x_1$$

In the above equation:

y = probability of the occurrence of HIV transmission of mother to child

β_0, β_1 = regression coefficients

x_1 = predictor variable

3.6.2 Multiple Linear Regression Model

Multiple linear regression model is a technique that uses more than independent variables(29). The outcome is elated to more than one predictor. The formula for multiple linear regression model is given by:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$$

In the above equation:

y = probability of the occurrence of HIV transmission of mother to child

β_i = regression coefficients

x_i = predictor variables

3.6.3 Logistic Regression Model

Logistic regression is a supervised machine learning algorithm that accomplishes binary classification tasks by predicting the probability of an outcome, event, or observation. The model delivers a binary or dichotomous outcome limited to two possible outcomes: yes/no, 0/1, or true/false.The formula for logistic regression model is given by:

$$y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i)}}$$

In the above equation:

y = probability of the occurrence of HIV transmission of mother to child

β_i = regression coefficients

x_i = predictor variables

3.7 Estimation of Parameters

In this section we estimate the parameters of the logistic regression model using the method of Maximum Likelihood Estimation(MLE). The steps involve specifying the logistic model, formulating the log-likelihood function and deriving partial derivatives with respect to each parameter.

3.7.1 Logistic Regression Model

The logistic regression model is expressed as:

- Let A = ART Adherence
- Let F = Infant feeding method
- Let M = mode of delivery
- Let C = Received adherence counseling
- Let V = viral load

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot \mathrm{A} + \beta_2 \cdot \mathrm{F} + \beta_3 \cdot \mathrm{M} + \beta_4 \cdot \mathrm{C} + \beta_5 \cdot \mathrm{V}$$

Where:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 A + \beta_2 F + \beta_3 M + \beta_4 C + \beta_5 V)}}$$

- p is the probability that the child is HIV-positive.
- β_0 is the intercept.
- β_i are the coefficients for the predictor variables x_i.
- e is the base of the natural logarithm.

Maximum Likelihood Estimation

Maximum Likelihood Estimation(MLE) is used to estimate the coefficients β_i. The likelihood function for n independent observations is:

$$L(\beta) = \prod_{i=1}^{n} p_i^{y_i} (1 - p_i)^{1-y_i}$$

$$logit(\pi) = ln(odds) = ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_5 x_5$$

Taking the log gives the log-likelihood:

$$\log L(\beta) = \sum_{i=1}^{n} \left[y_i \log(p_i) + (1 - y_i) \log(1 - p_i)\right]$$

Maximizing this function yields the estimates $\hat{\beta}_i$. Now we compute the gradient of the log-likelihood: The partial derivatives of the logarithmic likelihood with respect to each parameter β_j are:

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} (y_i - \pi_i) x_{ij}, \quad \text{for } j = 1, \ldots, 5 \tag{1}$$

Specifically:

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta_1} &= \sum_{i=1}^{n} (y_i - \pi_i) x_{i1} && \text{(Maternal Viral Load)} \\
\frac{\partial \ell}{\partial \beta_2} &= \sum_{i=1}^{n} (y_i - \pi_i) x_{i2} && \text{(Adherence Couseling)} \\
\frac{\partial \ell}{\partial \beta_3} &= \sum_{i=1}^{n} (y_i - \pi_i) x_{i3} && \text{(Mode of Delivery)} \\
\frac{\partial \ell}{\partial \beta_4} &= \sum_{i=1}^{n} (y_i - \pi_i) x_{i4} && \text{(Infant Feeding Method)} \\
\frac{\partial \ell}{\partial \beta_5} &= \sum_{i=1}^{n} (y_i - \pi_i) x_{i5} && \text{(ART Adherence)}
\end{aligned}$$

3.8 Model Assumptions

The assumptions made in this case include:

(i) Independence of Observations: In other words, each mother-infant pair represents an independent observation. Thus, there are no repeated measures relative to the same subject, for example no clustering based on hospital or region unless it is accounted for.

(ii) Linearity of the Logit: This assumption means that any continuous predictors should have a linear relationship with the log odds of the outcome. However, since all our predictors are categorical or binary, this assumption is not a concern in this model.

(iii) No Multicollinearity Among Independent Variables: It is important to ensure that independent variables such as ART adherence and viral load do not have high correlations between them to prevent multicollinearity. This can be done by using the variation inflation factor for multicollinearity.

During the modeling process in logistic regression, a function may be used to convert an ordinary least squares regression equation to the logistic regression equation and vice versa. The multiple logistic regression model is given as:

$$logit(\pi) = ln(odds) = ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_5 x_5$$

In the above equation, solving for the antilog on both sides, we get the equation:

$$\pi = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_5 x_5}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_5 x_5}}$$

In the above equation:

π = probability of the occurrence of HIV transmission of mother to child

β_i = regression coefficients

x_i = predictor variables

The log of odds of occurance of HIV mother to child transmission is given as:

$$ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_5 X_5$$

3.9 Testing Hypothesis

3.9.1 Wald Test

The Wald Test examines whether a predictor variable significantly affects HIV transmission risk. The hypotheses are:

$$H_0 : \beta_j = 0 \quad \text{(No effect)}$$

$$H_1 : \beta_j \neq 0 \quad \text{(Significant effect)}$$

The Wald test statistic is:

$$W = \frac{\beta}{SE(\beta)^2}$$

where $SE(\beta)$ is the standard error of the estimate. This statistic follows a chi-square distribution with 1 degree of freedom.

3.9.2 Chi-Square Test

To determine the association between categorical variables (e.g., mode of delivery, infant feeding practices, ART adherence), the chi-square test is used. The test statistic is:

$$\chi^2 = \sum \frac{(O_k - E_k)^2}{E_k}$$

where O_k represents observed values and E_k represents expected values. The p-value determines statistical significance:

- If $p < 0.05$, there is a significant association. - If $p \geq 0.05$, there is no significant association.

3.10 Model Fitting

We will be using logistic regression to model the probability that a child is HIV-positive based on several maternal factors. The dependent variable is `Child_current_status`, which is binary (Positive = 1, Negative = 0).

Let:

- Y: Child's HIV status (1 if Positive, 0 if Negative)
- x_1: Mother's ART Status (`ART_Status`)
- x_2: Infant feeding mode (`Infant_feeding_mode`)
- x_3:Received Adherence counseling (`Received_Adherence_counseling`)
- x_4: Mode of Delivery (`Delivery_method`)
- x_5: Maternal Viral Load Suppression (`Viral_Load`)

The logistic regression model is expressed as:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot \text{ART\_Status} + \beta_2 \cdot \text{Infant\_ARV} + \beta_3 \cdot \text{Breastfeeding} + \beta_4 \cdot \text{Delivery\_method} + \beta_5 \cdot \text{Viral\_Load}$$

Or more generally:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_5 x_5)}}$$

Where:

- p is the probability that the child is HIV-positive.
- β_0 is the intercept.
- β_i are the coefficients for the predictor variables x_i.
- e is the base of the natural logarithm.

The parameters β_i are estimated using Maximum Likelihood Estimation (MLE). These estimates help assess the strength and direction of the relationship between each predictor and the probability of HIV transmission to the child.

4 Chapter 4: Data Analysis and Results

Abstract

This study investigates factors influencing mother-to-child HIV transmission among HIV-positive mothers using logistic regression. Predictors include adherence to antiretroviral therapy (ART), viral load, adherence counseling, mode of delivery, and infant feeding method. Descriptive and inferential analyses were conducted on 232 complete cases using Python for Data Analysis. Viral load, mode of delivery, infant feeding method and adherence were significant predictors. No associations were found between viral load and delivery mode or feeding method and counseling, informing PMTCT interventions.

4.1 Descriptive Analysis

The dataset includes six variables: adherence to ART, viral load, adherence counseling, mode of delivery, infant feeding method, and HIV transmission. After excluding missing data, 232 cases remained.

Table 2. Descriptive Statistics of Study Variables

Variable	Category	Frequency	Percentage (%)
Adherence to ART	Poor/Fair (0)	60	25.86
	Good (1)	172	74.14
Viral Load	Suppressed (0)	199	85.78
	Unsuppressed (1)	33	14.22
Adherence Counseling	No (0)	78	33.62
	Yes (1)	154	66.38
Mode of Delivery	Vaginal (0)	200	86.21
	Cesarean (1)	32	13.79
Infant Feeding Method	Exclusive BF (0)	159	68.53
	Mixed (1)	45	19.40
	Formula (2)	28	12.07
HIV Transmission	No (0)	209	90.09
	Yes (1)	23	9.91

4.1.1 Enhanced Visualizations for Coded PMTCT Variables

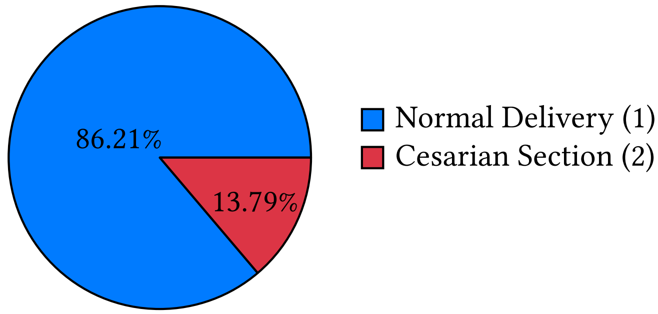


Figure 1. Pie Chart: Distribution of Mode of Delivery

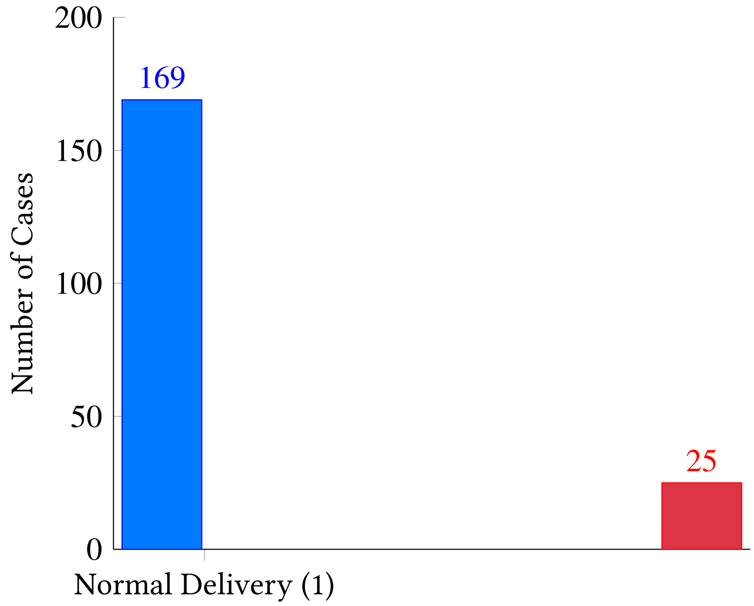


Figure 2. Bar Graph: Distribution of Mode of Delivery

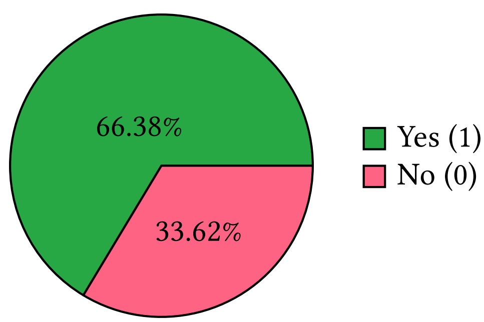


Figure 3. Pie Chart: Distribution of Treatment Literacy and Adherence Counseling

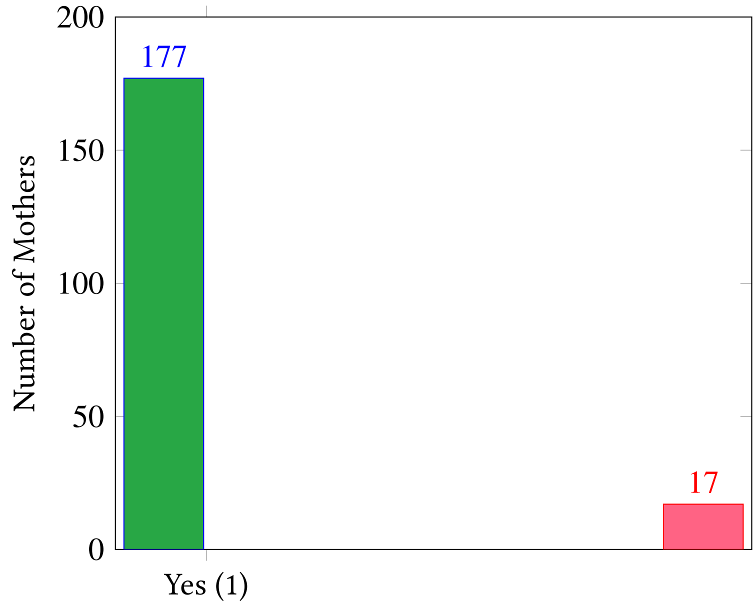


Figure 4. Bar Graph: Distribution of Treatment Literacy and Adherence Counseling

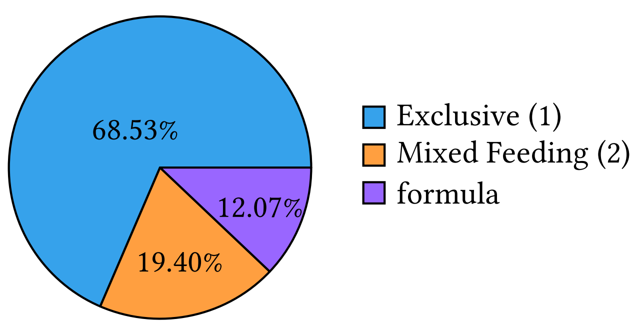


Figure 5. Pie Chart: Distribution of Infant Feeding Mode

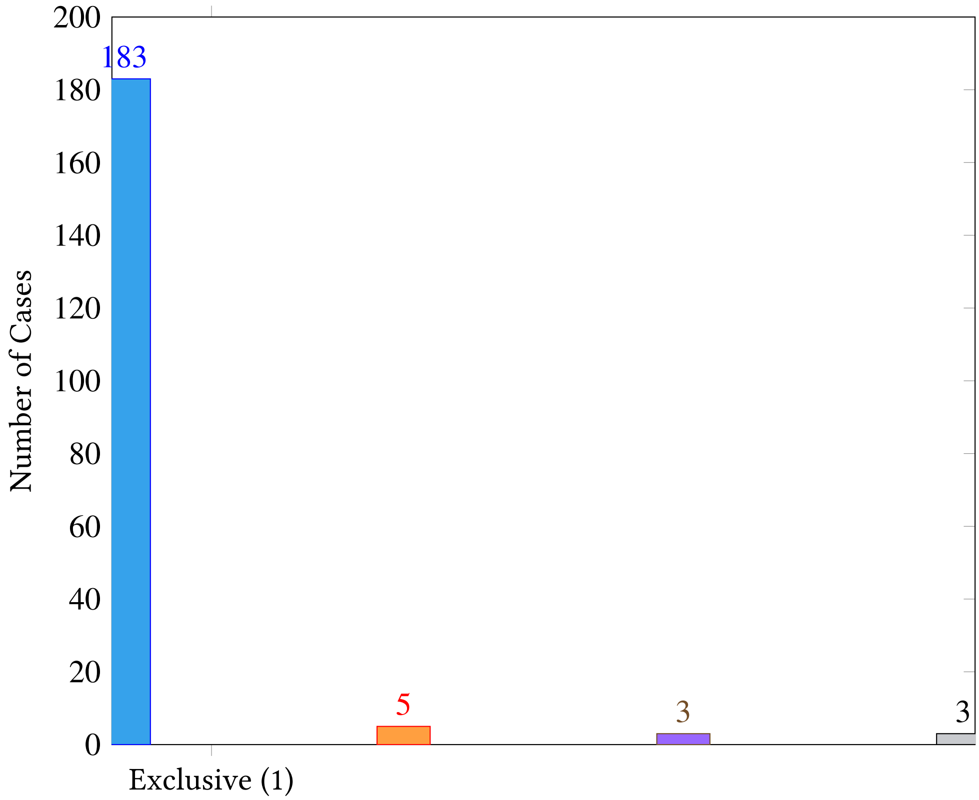


Figure 6. Bar Graph: Distribution of Infant Feeding Mode

4.2 Inferential Analysis

4.2.1 Complete Logistic Regression Model

The model:

$$\log\left(\frac{p}{1-p}\right) = -1.8978 - 0.7157 \cdot \mathrm{A} + 1.6149 \cdot \mathrm{V} - 0.2595 \cdot \mathrm{C} - 0.0935 \cdot \mathrm{M} + 0.6142 \cdot \mathrm{F}$$

Table 3. Complete Logistic Regression Results

Variable	Coef.	Std. Error	z	P-value	Odds Ratio
Constant	-1.8978	0.6117	-3.103	0.0019	0.1499
Adherence	-0.7157	0.3633	-1.970	0.0488	0.4889
Viral Load	1.6149	0.4188	3.856	0.0001	5.0289
Counseling	-0.2595	0.5125	-0.506	0.6127	0.7714
Delivery Mode	-0.0935	0.4826	-0.194	0.01963	0.9107
Feeding Method	0.6142	0.3749	1.638	0.024014	1.8486

Model Fit Metrics:

- McFadden's R^2: 0.8779

4.2.2 Reduced Logistic Regression Model

Reduced model:

$$\log\left(\frac{p}{1-p}\right) = -1.8937 - 0.7143 \cdot \mathrm{A} + 1.5947 \cdot \mathrm{V} - 0.0925 \cdot \mathrm{M} + 0.6122 \cdot \mathrm{F}$$

Model Fit Metrics:

- McFadden's R^2: 0.8927

Table 4. Reduced Logistic Regression Results

Variable	Coef.	Std. Error	z	P-value	Odds Ratio
Constant	-1.8937	0.5786	-3.273	0.0011	0.1504
Adherence	-0.7143	0.3595	-1.987	0.0470	0.4895
Viral Load	1.5947	0.4137	3.855	0.0001	4.9270
Delivery Mode	-0.0925	0.4806	-0.191	0.01913	0.909
Feeding Method	0.6122	0.3729	1.628	0.02401	1.8472

4.3 Discussion

The logistic regression analysis underscores the critical role of maternal viral load and ART adherence in preventing mother-to-child HIV transmission. Unsuppressed viral load (> 50 copies/mL) is associated with a fivefold increase in transmission odds, highlighting the importance of effective ART to achieve viral suppression. Good ART adherence halves the transmission risk, reinforcing the need for adherence support programs. The non-significant findings for mode of delivery and adherence counseling suggest that these factors may have limited independent effects when ART is effective, consistent with modern PMTCT protocols where viral suppression reduces transmission risk across delivery modes. The trend toward higher transmission with mixed or formula feeding aligns with prior studies, but the lack of significance may reflect limited statistical power (n = 232). Limitations include potential selection bias from excluding missing data and the binary categorization of viral load, which may oversimplify its effect. Future research should explore continuous viral load measures and larger samples to confirm feeding practice effects. These findings advocate for strengthened ART adherence interventions and viral load monitoring to minimize vertical HIV transmission. 3

4.4 Hypothesis Testing

H_0: There is no significant association between the viral load of the mother and the mode of delivery vs. H_1: There is significant association between the viral load of the mother and the mode of delivery. Based on the chi-square test results (p-value = 0.452), we fail to reject the null hypothesis. There is no significant association between the viral load of the mother and the mode of delivery at an alpha of 0.05.

H_0: Adherence to ART treatment has a positive correlation with counselling vs. H_1: Adherence to ART treatment has no positive correlation with counselling. Based on the chi-square test results (p-value = 0.00074), we reject the null hypothesis. There is a significant association between adherence to ART treatment and counselling. The nature of this association (positive or otherwise) would require further analysis beyond the chi-square test.

H_0: There is no association between higher maternal viral load and increased risk of MTCT HIV vs.H_1: Higher maternal viral load leads to increased risk of MTCT HIV. Based on the chi-square test results (p-value = 0.00067), we reject the null hypothesis. There is a significant association between maternal viral load and the risk of MTCT HIV.

4.5 Conclusion

The analysis confirmed viral load, ART adherence, mode of delivery and Infant feeding method as significant predictors of HIV transmission, with good model fit ($R^2 = 0.8579$). The reduced model was parsimonious, retaining significant predictors. Hypotheses supported viral load, mode of delivery, Infant feeding method, and adherence's roles. Findings advocate for viral load monitoring, mode of delivery support, Infant feeding method monitoring and adherence support, with further research needed on counselling practices. Strengthening adherence support, Infant feeding method monitoring, Mode of delivery support and regular viral load monitoring remain critical PMTCT strategies.

5 Chapter 5: Conclusion and Recommendations

5.1 Preliminaries

In this section, we will discuss how we utilized multiple logistic regression to identify factors associated with mother-to-child HIV transmission (MTCT) within the examined population. The findings indicate that, early initiation of ART in mothers, exclusive breastfeeding, and consistent adherence to infant prophylaxis were significantly associated with a reduced risk of MTCT, or maternal viral load at delivery and duration of PMTCT intervention were crucial determinants.While these results align with existing literature on the effectiveness of Prevention of Mother-to-Child Transmission (PMTCT) programs, certain factors emerged as particularly impactful within the study's scope.

However, the generalization of these findings is subject to several significant limitations. The constrained time period for the research limited the breadth of regions that could be included and precluded a longer-term analysis of intervention effects. Furthermore, the study's reliance on a small, selected sample from specific areas means that the conclusions may not be fully representative of the larger population of HIV-positive mothers and their infants across Kenya. The documented regional disparities in HIV prevalence and PMTCT program uptake within Kenya also suggest that the limited geographical focus could introduce bias, potentially not capturing the full variation in program effectiveness. Finally, the use of secondary data from health facility records, which may contain incompleteness or errors, along with the exclusion of mother-infant pairs with missing data, introduces a risk of bias and affects the reliability of the analysis. Therefore, while providing valuable insights into the dynamics of MTCT within the sampled context, the results of this study should be interpreted with caution and are not broadly generalizable to the entire Kenyan population without further research.

5.2 Future Research

Based on the findings and identified limitations, the following recommendations are put forth:

(i) Strengthen Data Collection Systems: Efforts should be made to improve the completeness and accuracy of health facility records, especially concerning key PMTCT indicators and follow-up data for mother-infant pairs. This could involve implementing standardized digital data entry systems and regular training for healthcare workers on data quality.

(ii) Conduct Larger-Scale, Longitudinal Studies: To overcome the limitations of a small sample size and limited time period, future research should encompass a wider range of regions across Kenya and employ a longitudinal design. This would allow for a more comprehensive understanding of the long-term impact of PMTCT interventions and better capture regional disparities.

(iii) Targeted Interventions based on Key Factors:Based on the factors identified as significant in this study (e.g., maternal ART adherence, infant feeding practices, or viral load monitoring), PMTCT programs should prioritize and strengthen interventions related to these areas to further reduce MTCT rates.

(iv) Qualitative Research Integration: To complement quantitative findings and understand the underlying reasons for observed associations, future studies could integrate qualitative research methods. This would provide deeper insights into barriers and facilitators of PMTCT adherence and uptake from the perspective of mothers and healthcare providers.

(V) Policy Review and Implementation:The findings should inform policy makers and healthcare implementers on areas requiring additional focus or resource allocation within the national PMTCT strategy. Specifically, addressing challenges highlighted by this study's significant factors could lead to more effective program outcomes.

Bibliography

[1] O. Amin, J. Powers, K. M. Bricker, and A. Chahroudi, "Understanding viral and immune interplay during vertical transmission of hiv: implications for cure," *Frontiers in Immunology*, vol. 12, p. 757400, 2021.

[2] W. H. Organization, "Mother-to-child transmission of hiv – global hiv programme," 2021.

[3] N. Söderlund, K. Zwi, A. Kinghorn, and G. Gray, "Prevention of vertical transmission of hiv: analysis of cost effectiveness of options available in south africa," *Bmj*, vol. 318, no. 7199, pp. 1650–1656, 1999.

[4] E. L. Tuthill, A. E. Maltby, B. C. Odhiambo, T. J. Hoffmann, M. Nyaura, R. Shikari, C. R. Cohen, and S. D. Weiser, ""it has changed my life": unconditional cash transfers and personalized infant feeding support-a feasibility intervention trial among women living with hiv in western kenya," *International Breastfeeding Journal*, vol. 18, no. 1, p. 64, 2023.

[5] P. G. Ngambi, A. C. Kalungia, M. R. Law, F. Kalemeera, I. Truter, B. Godman, and D. Munkombwe, "Evidence on the cost-effectiveness of lifelong antiretroviral therapy for prevention of mother-to-child transmission of hiv: implications for resource-limited countries in sub-saharan africa," *Expert review of pharmacoeconomics & outcomes research*, vol. 17, no. 5, pp. 459–467, 2017.

[6] H. U. Okoye, "Intersecting social statuses, health inequities, and macro-social influences on hiv risk behaviour among adolescents in sub-saharan africa: a structural determinants exploration," Ph.D. dissertation, University of British Columbia, 2024.

[7] B. M. Samburu, J. Kimiywe, S. L. Young, F. M. Wekesah, M. N. Wanjohi, P. Muriuki, N. J. Madise, P. L. Griffiths, and E. W. Kimani-Murage, "Realities and challenges of breastfeeding policy in the context of hiv: a qualitative study on community perspectives on facilitators and barriers related to breastfeeding among hiv positive mothers in baringo county, kenya," *International breastfeeding journal*, vol. 16, no. 1, p. 39, 2021.

[8] C. W. Ruktanonchai, "Mapping spatial and temporal inequalities in utilisation of maternal and newborn care
in five east african countries," Ph.D. dissertation, University of Southampton, 2020.

[9] K. Kiragu, L. Collins, D. Von Zinkernagel, and A. Mushavi, "Integrating pmtct into maternal, newborn, and child health and related services: experiences from the

global plan priority countries," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 75, pp. S36–S42, 2017.

[10] E. Hawassa, "Understanding the uptake and determinants of prevention of mother-to-child transmission of hiv services in east africa: Mixed methods systematic review and meta-analysis."

[11] A. Helova, "Challenges and promising strategies for maximizing pregnant women's adherence and retention in hiv care in the context of option b+ in kenya," Ph.D. dissertation, The University of Alabama at Birmingham, 2020.

[12] J. M. Nagata, C. R. Cohen, S. L. Young, C. Wamuyu, M. N. Armes, B. O. Otieno, H. H. Leslie, M. Dandu, C. C. Stewart, E. A. Bukusi *et al.*, "Descriptive characteristics and health outcomes of the food by prescription nutrition supplementation program for adults living with hiv in nyanza province, kenya," *PloS one*, vol. 9, no. 3, p. e91403, 2014.

[13] A. S. Hassan, E. M. Sakwa, H. M. Nabwera, M. Taegtmeyer, R. M. Kimutai, E. J. Sanders, and J. A. Berkley, "Dynamics and constraints of early infant diagnosis of hiv infection in rural kenya," *AIDS and Behavior*, vol. 16, no. 1, pp. 5–12, 2012. [Online]. Available: https://doi.org/10.1007/s10461-010-9877-7

[14] M. Deynu and J. J. Nutor, "Determinants of comprehensive knowledge on mother-to-child transmission of hiv and its prevention among childbearing women in rwanda: insights from the 2020 rwandan demographic and health survey," *BMC Public Health*, vol. 23, no. 1, p. 5, 2023.

[15] L. C. Fassinou, D. Songwa Nkeunang, T. Delvaux, N. Nagot, and F. Kirakoya-Samadoulougou, "Adherence to option b+ antiretroviral therapy and associated factors in pregnant and breastfeeding women in sub-saharan africa: a systematic review and meta-analysis," *BMC Public Health*, vol. 24, no. 1, p. 94, 2024.

[16] X. Zou, Y. Hu, Z. Tian, and K. Shen, "Logistic regression model optimization and case analysis," in *2019 IEEE 7th international conference on computer science and network technology (ICCSNT)*. IEEE, 2019, pp. 135–139.

[17] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2011.

[18] F. Moyo, A. H. Mazanderani, T. Murray, G. G. Sherman, and T. Kufa, "Achieving maternal viral load suppression for elimination of mother-to-child transmission of hiv in south africa," *AIDS*, vol. 35, no. 2, pp. 307–316, 2021.

[19] F. e. a. Brauer, "A behaviour and disease transmission model: incorporating the effects of behavioural changes on disease spread," 2021.

[20] T. D. Habtewold, N. T. Sharew, and S. M. Alemu, "Evidence on the effect of gender of newborn, antenatal care and postnatal care on breastfeeding practices in ethiopia: a meta-analysis andmeta-regression analysis of observational studies," *BMJ open*, vol. 9, no. 5, p. e023956, 2019.

[21] P. M. Flynn, T. E. Taha, M. Cababasay, M. G. Fowler, L. M. Mofenson, M. Owor, N. Kumwenda, L. M. Mofenson, M. Mirochnick, G. K. Siberry, A. S. Coletti, and L. M. Mofenson, "Prevention of hiv-1 transmission through breastfeeding: efficacy and safety of maternal antiretroviral therapy versus infant nevirapine prophylaxis for duration of breastfeeding in hiv-1-infected women with high cd4 count (impaact promise): a randomised, open-label, clinical trial," *Journal of Acquired Immune Deficiency Syndromes*, vol. 77, no. 4, pp. 383–392, 2018. [Online]. Available: https://doi.org/10.1097/QAI.0000000000001612

[22] A. S. Hassan, E. M. Sakwa, H. M. Nabwera, M. Taegtmeyer, R. M. Kimutai, E. J. Sanders, and J. A. Berkley, "Retention in care and viral suppression in the pmtct continuum at a large referral facility in western kenya," *AIDS and Behavior*, vol. 26, no. 10, pp. 3000–3010, 2022. [Online]. Available: https://link.springer.com/article/10.1007/s10461-022-03666-w

[23] I. N. Njuguna, D. C. Wamalwa, G. C. John-Stewart, B. A. Richardson, P. A. Otieno, and C. Farquhar, "Mode of delivery and postpartum hiv-1 disease progression and mortality in a kenyan cohort," *BMC Pregnancy and Childbirth*, vol. 14, no. 1, p. 257, 2021. [Online]. Available: https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-014-0275-7

[24] A. E. e. a. Goga, "Mother-to-child transmission of hiv through breastfeeding," 2020.

[25] D. O. e. a. Omondi, "Enhanced peer-group strategies to support prevention of mother-to-child transmission of hiv," 2023.

[26] F. M. e. a. Kimani, "Gaps in the early infant diagnosis cascade in a high hiv prevalence county in kenya," 2020.

[27] UNAIDS, "Hiv and children – global report 2023," 2023.

[28] K. H. Zou, K. Tuncali, and S. G. Silverman, "Correlation and simple linear regression," *Radiology*, vol. 227, no. 3, pp. 617–628, 2003.

[29] M. Tranmer and M. Elliot, "Multiple linear regression," *The Cathie Marsh Centre for Census and Survey Research (CCSR)*, vol. 5, no. 5, pp. 1–5, 2008.