**CA-IeDEA HIV dynamic transmission model: Documentation of approach**

Prepared by

Zhongzhe ‘Portia’ Pan, Deo Mujwara, and April Kimmel

Department of Health Behavior and Policy, Virginia Commonwealth University

for the Central Africa-International Epidemiologic Databases on AIDS (CA-IeDEA) Consortium

**Contents**

[List of Figures 4](#_Toc82592637)

[List of Tables 4](#_Toc82592638)

[1 Overview 6](#_Toc82592639)

[2 Motivating research question 6](#_Toc82592640)

[3 Model structure 7](#_Toc82592641)

[4 Sub-populations and mixing 9](#_Toc82592642)

[5 System of equations 12](#_Toc82592643)

[6 Force of infection 15](#_Toc82592644)

[6.1 Assumptions and expression 15](#_Toc82592645)

[6.2 Probability of HIV transmission per sex act 17](#_Toc82592646)

[6.3 Condom effectiveness 18](#_Toc82592647)

[6.4 Proportion with consistent condom use 19](#_Toc82592648)

[6.5 Effectiveness of viral suppression in reducing HIV transmission 21](#_Toc82592649)

[6.6 Number of sexual acts 22](#_Toc82592650)

[7 Parameter inputs 23](#_Toc82592651)

[7.1 Initial population 23](#_Toc82592652)

[7.1.1 The size of sub-populations, by HIV status (Table 6 ): 23](#_Toc82592654)

[7.1.2 The size of sub-populations living with HIV across compartments (Table 7): 25](#_Toc82592656)

[7.2 Susceptible population growth rate 29](#_Toc82592657)

[7.3 HIV diagnosis 31](#_Toc82592658)

[7.4 Natural history 34](#_Toc82592659)

[7.5 Linked to care 37](#_Toc82592660)

[7.6 Lost to follow-up 38](#_Toc82592661)

[7.7 On ART, Suppressed 42](#_Toc82592748)

[7.8 On ART, Not Suppressed 45](#_Toc82592749)

[7.9 Death 48](#_Toc82592750)

[8 Calibration approach 51](#_Toc82592751)

[8.1 Calibration targets 51](#_Toc82592752)

[8.2 Approach: Parameters used in the calibration process 56](#_Toc82592753)

[8.3 Approach: Goodness of fit 56](#_Toc82592754)

[8.4 Approach: Method for identifying appropriate values for parameters inputs 57](#_Toc82592755)

[8.5 Approach: Defining convergence criteria 57](#_Toc82592756)

[8.6 Approach: Stopping rule (i.e. when calibration process is complete) 58](#_Toc82592757)

[8.7 Calibration approach: Probabilistic distributions for model parameters 58](#_Toc82592758)

[9 Implementation (in R) 65](#_Toc82592759)

[10 References 95](#_Toc82592760)

[11 Appendix A 100](#_Toc82592761)

[11.1 Datasets 100](#_Toc82592762)

[11.2 Codes to derive parameter inputs 100](#_Toc82592763)

[11.3 System of Differential Equations 100](#_Toc82592764)

[11.4 Methods to estimate CD4 strata and time-at-risk patient living with HIV 102](#_Toc82592765)

[11.5 Average monthly population growth rate, by sub-populationa 103](#_Toc82592766)

[11.6 Comparing different for measuring the goodness of fit in the calibration process. 104](#_Toc82592767)

[11.7 Methods for searching for parameters. 110](#_Toc82592768)

[11.8 Probabilistic distribution assigned to model parameters 117](#_Toc82592769)

# List of Figures

[Figure 1. Simplified schematic for the CA-IeDEA model](#_Toc82592770)

[Figure 2. Sub-populations and mixing patterns in the CA-IeDEA model](#_Toc82592771)

# List of Tables

[Table 1. Model sub-populations: model implementation, justification, and supporting evidence](#_Toc82599260)

[Table 2. Model compartments: a legend for the system of equations](#_Toc82599261)

[Table 3. CA-IeDEA model parameters](#_Toc82599262)

[Table 4: Probabilities of HIV transmission per sex act, by sub-group and direction](#_Toc82599263)

[Table 5: Proportion of consistent condom use, by sub-population](#_Toc82599264)

[Table 6. Initial population size by sub-population and HIV status](#_Toc82599265)

[Table 7. Size of initial sub-populations, by CD4 stratum and engagement in care](#_Toc82599266)

[Table 8. Average monthly population growth ratea](#_Toc82599267)

[Table 9. Monthly probability of HIV diagnosis, by sub-population and CD4 stratum](#_Toc82599268)

[Table 10. Sub-population monthly probability of disease progression, by CD4 stratum and engagement in care](#_Toc82599269)

[Table 11: Monthly probability of sub-population linkage to HIV care, by CD4 stratum](#_Toc82599270)

[Table 12: Monthly probability of sub-population LTFU, by CD4 stratum and engagement in care](#_Toc82599271)

[Table 13: IeDEA variables for estimating model parameter inputsa](#_Toc82599272)

[Table 14: Monthly probability of sub-population On ART, Suppressed, by CD4 stratum](#_Toc82599273)

[Table 15: Monthly probability of sub-population virologic failure](#_Toc82599274)

[Table 16: Monthly probability of sub-population death, by CD4 stratum and engagement in care](#_Toc82599275)

[Table 17. Calibration targets: Sub-population HIV prevalence over time](#_Toc82599276)

[Table 18: Calibration targets: Sub-population HIV incidence over time](#_Toc82599277)

[Table 19: Calibration targets: Number of people on ART over time](#_Toc82599278)

[Table 20: Calibration targets: Percentage at Different Steps Along the HIV Care Continuum based on 2018-2019 Rwanda Population-based HIV Impact Assessment Survey19](#_Toc82599279)

[Table 21. Calibration: Parameter estimates for beta distribution assigned to probabilities of HIV diagnosis](#_Toc82599280)

[Table 22. Calibration: Parameter estimates for multiplier applied to probability of HIV diagnosis and linkage to care](#_Toc82599281)

[Table 23. Calibration: Parameter estimates for beta distribution assigned to probabilities of linkage to care](#_Toc82599282)

[Table 24. Calibration: Parameter estimates for distributions assigned to the probabilities of HIV transmission](#_Toc82599283)

[Table 25. Calibration: Parameter estimates for beta distribution assigned to proportion of condom use](#_Toc82599284)

**List of Equations**

[Equation 1 Force of infection](#_Assumptions_and_expression)

[Equation 2 Force of infection: Estimating the effectiveness of viral suppression](#_Effectiveness__)

[Equation 3 Initial population: Estimating the proportion of lost to follow-up (pre-ART)](#_The_size_of)

[Equation 4 Initial population: Estimating the proportion of lost to follow-up (on ART)](#_The_size_of)

[Equation 5 Initial population: Estimating the proportion on ART](#_The_size_of)

[Equation 6 Initial population: Estimating the number On ART, Not Suppressed](#_The_size_of)

[Equation 7 Susceptible population: population growth](#_Susceptible_population_growth)

[Equation 8 Disease progression parameterization: Weighted competing risks model](#_Natural__)

[Equation 9 Disease progression parameterization: Calculating probability weights](#_Natural__)

[Equation 10 Lost to follow-up parameterization: Weighted competing risks model](#_Lost_to_follow-up)

[Equation 11 Lost to follow-up parameterization: Calculating probability weights](#_Lost_to_follow-up)

[Equation 12 On ART, Suppressed parameterization: Weighted competing risks model](#_On_ART,_Suppressed)

[Equation 13 On ART, Suppressed parameterization: Calculating probability weights](#_On_ART,_Suppressed)

[Equation 14 On ART, Not Suppressed parameterization: Weighted competing risks model](#_On_ART,_Not)

[Equation 15 On ART, Not Suppressed parameterization: Calculating probability weights](#_On_ART,_Not)

[Equation 16 Death parameterization: Weighted competing risks model](#_Death)

[Equation 17 Death parameterization: Calculating probability weights](#_Death)

[Equation 18 Probabilistic distributions for model parameters: Beta distribution weighted mean](#_Calibration_approach:_Probabilistic)

[Equation 19 Probabilistic distributions for model parameters: Beta distribution standard deviation](#_Calibration_approach:_Probabilistic)

[Equation 20 Probabilistic distributions for model parameters: Beta distribution shape parameter](#_Calibration_approach:_Probabilistic)

[Equation 21 Probabilistic distributions for model parameters: Beta distribution scale parameter](#_Calibration_approach:_Probabilistic)

# Overview

This document details and justifies analytic decisions for the CA-IeDEA HIV dynamic transmission model, including model structure, parameter inputs, model implementation, and model calibration procedures. This documentation of approach is a living record of model-related analytic decisions, which are presented in nine sections. Section 2 outlines the research questions motivating this initial work. Section 3 discusses the model structure and justifies the inclusion/exclusion of individual compartments. Section 4 describes model sub-populations, assumptions regarding sub-population mixing (i.e., sexual interaction). Section 5 presents the systems of equations that represent how population move across compartments. Section 6 introduces an expression for the force of infection, which calculates risk of HIV acquisition. Section 7 describes the estimation procedures for parameter inputs, while Section 8 describes the model calibration approach and current targets. Finally, Section 9 shows implementation of the model in R and the output. Data sources and statistical procedures to derive parameter inputs are shown in the Appendix.

# Motivating research question

This model was motivated by the following research question: What improvements are needed along the HIV care continuum to achieve Treat All and the 95-95-95 UNAIDS targets in Rwanda?

***Purpose of the model:*** The model’s purpose is to examine the spectrum of HIV policies and response programs, including for different sub-populations and sub-national units, needed in Rwanda to realize universal ART and propel movement toward ending the epidemic.

***The audience:*** The target audience includes the variety of stakeholders supporting the HIV response, particularly the Rwandan Ministry of Health and donors. For the current model version, stakeholder interest has influenced decisions related to model structure, including for high-risk sub-populations (e.g., female sex workers) and for different sub-national units (e.g., urban vs rural).

***Longer term plan:*** The current model will serve as a foundation for future work, such as the efficiency and affordability of alternative strategies to achieve 95-95-95 UNAIDS targets in Rwanda by 2030. It will also serve as the foundation for similar country-specific models in the IeDEA Central Africa region.

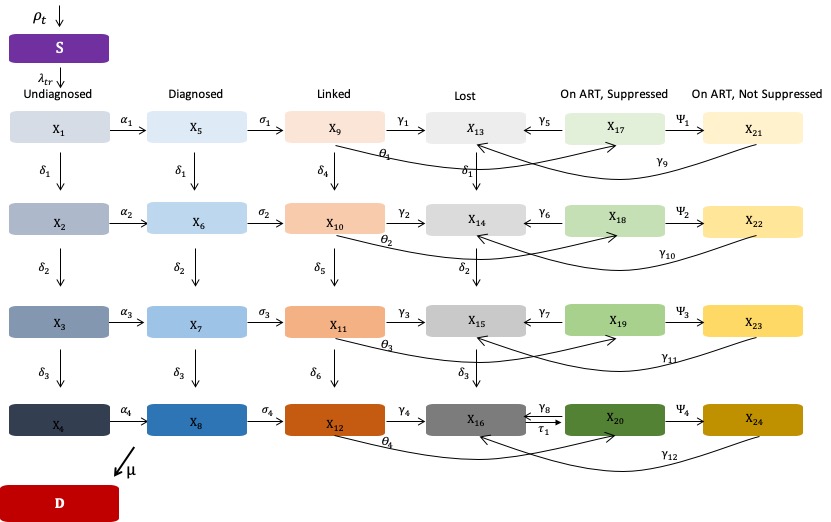
# Model structure

The model is an open systems model, specifically a dynamic compartmental model. These types of models are also referred to as a dynamic transmission model or dynamic epidemiologic transmission model. ‘Dynamic’ refers to the mathematical approximation of HIV disease spread over time. The model is a simple SIR model, where *S* is susceptible, *I* is infectious, and *R* is removed (i.e. death). The infectious compartment is further adapted so that it contains 26 compartments, which are defined by disease progression and engagement in HIV care (**Figure 1**) The model captures 25 different sub-populations defined by sex, age group, risk of HIV acquisition, and urbanicity. Each of the model’s 25 sub-populations are represented by the compartments described, resulting in a total of 650 model compartments (25 sub-populations x 26 compartments = 650 sub-population compartments). The model is deterministic to reduce computational time and has a one-month time step to reflect real-world clinical decision making and disease progression. The model will be calibrated to historical data between 2003 and 2020, with model projections over a 30-year time horizon until 2050. The 30-year analytic time horizon captures both short- and longer-term decision making.

We describe analytic decision regarding the model’s infectious compartments below.

**Figure 1. Simplified schematic for the CA-IeDEA model**

Legend: The arrows represent the movement of population across compartments. The definition of compartments and parameters are described in Table 2 and Table 3. For simplicity, the subscripts indicating time is not presented. The probability of death in each compartment is presented in aggregated form, despite that it varies by compartments.

****

***Disease progression:*** Disease progression in the absence of antiretroviral therapy is defined according to 4 CD4 strata (see below). These strata were selected to reflect historical World Health Organization (WHO) guidelines for ART initiation,1–4 which informs model calibration and validation. Their inclusion also captures empirical differences in HIV testing uptake, linkage to care, ART initiation and mortality based on baseline CD4 cell count.5,6

* Asymptomatic early (CD4 >500)
* Asymptomatic late (CD4 >350 - 500)
* Symptomatic (CD4 >200-350)
* AIDS (CD4 200)

We emphasize that acute HIV was not explicitly modeled. This is because: 1) acute HIV has a small window of detection (approximately 14 days) which will be captured in a single model cycle, 2) most commonly used HIV diagnostic tests in settings like Rwanda are unable to detect HIV at this stage,7 3) evidence on its impact on disease spread is limited, and 4) this stage does not require different programmatic approaches for epidemic control.8

We assume no explicit disease progression among individuals on ART given evidence suggesting CD4 at ART initiation as a primary predictor of health outcomes (e.g., time to AIDS or death).9

***Engagement in care:*** We model distinct stages of engagement in HIV care and the HIV care continuum, namely:

* Undiagnosed
* Diagnosed
* Linked to care
* Lost from care
* On ART (both virally suppressed and not virally suppressed)

Diagnosis and Linked to care*:* The analytic rationale to explicitly model (and decouple) diagnosis and linkage to care is based on current evidence. First, recent evidence suggests that not all Rwandans diagnosed with HIV are linked to care and initiated on ART within 30 days,10,11 despite that timely linkage to care (i.e., within 30 days after diagnosis) is recommended for HIV treatment and care.4 Additionally, evidence suggests that linkage to care varies by type of HIV testing strategy11,12 and study population.13 Therefore, we explicitly model linkage to care, including separate compartments for “**diagnosed”** and **“linked”** to account for differences in probability of linkage to care and ART initiation.

Lost to follow-up*:* We model the HIV individuals that are lost from care have a natural history disease progression and return to care at the AIDS stage when sick.5,14

Viral suppression***:*** We model viral suppression as “**On ART and viral suppressed**” and “**On ART and not viral suppressed**” to estimate the number of individuals that are virally suppressed and account for the health and transmission benefits of being virally suppressed.15,16 While viral suppression is not modeled for each line of therapy (e.g., 1st, 2nd, 3rd, etc.), future iterations of the model can be adapted to include viral suppression by ART regimen.

# Sub-populations and mixing

The study population is stratified into 25 sub-populations stratified by age, sex, risk, and urbanicity based on the differences in the risk of HIV acquisition and HIV prevalence. **Table 1** (next page) shows the rationale for sub-populations considers in this study.

We make the following assumptions to describe the sub-populations and/or their mixing pattern:

* High-risk women are not stratified by urbanicity given evidence suggesting that 80% of high-risk women (female sex workers) reside in urban areas in Rwanda.17
* High-risk women engage sexually with both urban and rural men. Evidence from the Rwanda Demographic and Health Survey suggests men residing in both rural and urban areas pay for sex, with the majority of new infections occurring in rural area.18
* For low-risk sub-populations (i.e., men and women in the general population), no migration occurs between urban and rural areas.
* Heterosexual mixing occurs among all age groups (**Figure 2**).

**Figure 2. Sub-populations and mixing patterns in the CA-IeDEA model**

Legend: Arrows indicate mixing across sub-populations. The probabilities of mixing differ depending on the direction of the arrows. Sub-populations are numbered from 1 to 25 so that they can be identified in model implementation.

**A screenshot of a cell phone

Description automatically generated**

**Table 1.** **Model sub-populations: model implementation, justification, and supporting evidence**

|  |  |  |
| --- | --- | --- |
| **Domain** | **Model implementation and justification** | **Supporting evidence** |
| Age | Modeled as 10-year intervals.  10-year intervals capture differences in HIV prevalence, risk of HIV acquisition, and engagement in care by age, while limiting model computational complexity. | HIV prevalence is currently 3.0% in Rwanda,19 but varies across age groups.18,20 Although HIV prevalence is low (<1.5%) among young adults in Rwanda, it is increasing, particularly among young girls.18 HIV prevalence is higher among people above 25 years and increases with age from 3.1% among 25-29 to 7.1% among 45-49.18  Risk of HIV acquisition is higher among young people due in part to risky sexual behaviors. In sub-Saharan Africa, young girls have sexual relationships with older people and marry at younger ages, which increases risk HIV acquisition.21  Engagement in HIV care, including viral suppression. The probability of HIV testing is lower among young people and older people.18,20,22 Younger age is a risk factor for loss to follow-up both pre-ART and on ART, while older age is a risk factor for mortality.5 Viral suppression is less likely among young people (15-24) compared to older people (49+).23 |
| Sex | Modeled as men and women.  Gender identity is not captured by these sub-populations. We do not explicitly model gender identity due to limited data availability. | HIV prevalence. Women are disproportionately affected by the HIV epidemic in Rwanda. The majority (62%) of people living with HIV in Rwanda are women, with HIV prevalence among women is higher (3.6%) relative to men (2.3%).18  Risk of HIV acquisition is higher among women, especially young women. Biologically, women have a higher risk of acquiring HIV than men because they are more likely to be exposed to injury and pathogens during sexual intercourse.24 Contextually, young women are more likely to be forced to have sex, compared to their male counterpart.25 Behaviorally, young women are more likely to start sexual experiences early, have sex partners that are much older, or exchange for money.25  Engagement in HIV care. Men are more likely to be lost to follow-up or die compared to women.5 |
| Risk of HIV acquisition | Modeled as high- and low-risk.  Among women, the high-risk sub-population represents female sex workers. Men are modeled as low-risk only. | HIV prevalence. HIV prevalence varies by sub-population, with female sex workers having HIV prevalence of 51% in 201026,27 and 45.8% in 201528 compared to 3.6% for women in the low-risk population (15-49 years).18 Men are modeled as low-risk only since HIV prevalence for low- vs high-risk men is similar (3% vs 5%).29  Risk of HIV acquisition. The government of Rwanda classifies female sex workers as a high-risk group, given high HIV prevalence and low utilization of HIV preventive services.30  Engagement in HIV care. While the percentage HIV testing is similar among female risk populations;31 the percentage initiating ART among female sex workers is higher than among low-risk women.32 |
| Sub-national (Urbanicity) | Modeled as urban and rural.  The urban-rural classification would capture differences in HIV prevalence and HIV testing at regional level, while limiting the model complexity. | HIV prevalence. In Rwanda, HIV prevalence is more than double in urban (6%) compared to rural (2.2%) areas.18 In Kigali city, the capital city of Rwanda and home to half of the urban population in the country, HIV prevalence is 6.2% vs 2-3%18 in the other provinces that are predominantly rural.  Risk of HIV acquisition. Populations residing in urban areas are more likely to acquire HIV, with HIV incidence in urban areas higher than in rural areas (0.65 vs 0.22 per 100 person-years; adjusted hazard ratio: 3.1, 1.3 – 7.0).33  Engagement in HIV care. HIV testing rates vary by urbanicity,18,20 with more people likely to test for HIV in urban vs rural areas. |

# System of equations

Epidemic dynamics and disease progression in the current model are captured by a system of equations. Systems of equations in dynamic models are typically expressed using ordinary differential equations, which are represented below. Similar to other modeling groups,34,35 however, we implement the system using a discrete approximate of the derivatives in a system of difference equations. This approach decreases computational burden relative to use of differential equations and allows the model to project outcomes at discrete time points, versus measuring rates of change over time. We emphasize that the systems of equations shown apply to each modeled sub-population, such that the system below is applied to each of the 25 individual sub-populations resulting in 650 model compartments. The systems of difference equations are shown below, while the system of differential equations is in **Section 11.3**. Model compartments in the system of equations are described in **Table 2**.

*Susceptible Population*

*Infected, Undiagnosed*

*Infected, Diagnosed*

*Infected, Linked to Care*

*Infected, Lost from Care*

*Infected, On ART and Virally Suppressed*

*Infected, On ART and Not Virally Suppressed*

*Dead*

**Table 2.** **Model compartments: a legend for the system of equations**

|  |  |  |  |
| --- | --- | --- | --- |
| **Compartment** | **Description** | | |
| S | Susceptible | At risk of acquiring HIV | |
| D | Dead | Death (absorbing compartment) | |
|  | Infected | Undiagnosed | CD4>500 |
|  | CD4 >350 – 500 |
|  | CD4 >200 – 350 |
|  | CD4 <200 |
|  | Diagnosed, not in care | CD4>500 |
|  | CD4 >350 – 500 |
|  | CD4 >200 – 350 |
|  | CD4 <200 |
|  | Linked to care\* | CD4>500 |
|  | CD4 >350 – 500 |
|  | CD4 >200 – 350 |
|  | CD4 <200 |
|  | Lost to follow-up† | CD4>500 |
|  | CD4 >350 – 500 |
|  | CD4 >200 – 350 |
|  | CD4 <200 |
|  | On ART, virally suppressed‡ | CD4 >500 |
|  | CD4 >350 – 500 |
|  | CD4 >200 – 350 |
|  | CD4 <200 |
|  | On ART, not virally suppressed‡ | CD4 >500 |
|  | CD4 >350 – 500 |
|  | CD4 >200 – 350 |
|  | CD4 <200 |

\* Defined as HIV care engagement after HIV diagnosis. † Defined as not continuously engaged in HIV care for at least 12 months (pre-ART) or 6 months (on ART). ‡CD4 strata is assumed to be the same as the CD4 at ART initiation.

# Force of infection

## Assumptions and expression

The force of infection () of the susceptible population in sub-population r is defined as the probability of the susceptible population acquiring HIV at time t.34

***Assumptions made for estimating the force of infection***

* All contacts in the model are heterosexual, reflecting the majority heterosexual transmission (i.e., 65% in stable heterosexual relationship, 20% in female sex workers and their clients, 10% in casual heterosexual relationship) in this setting.36
* There is random mixing across individuals in different sub-populations since all individuals in the model are assumed to be sexually active. That is, all individuals in sub-population *r* have equal probability of engaging in sexual behaviors with individuals in other sub-populations.

***Force of infection expression:*** Equation 1 shows the force of infection of sub-population at time period. Parameters used in the force of infection are defined in **Table 3.**

(1)

Where

Note:

* denotes the probability of sub-population r partnering with sub-population j among all possible partners
* denotes HIV prevalence in sub-population
* denotes the probability of HIV acquisition when partnering with individuals in sub-population j with sex acts per partnership

**Table 3.** **CA-IeDEA model parameters**

|  |  |  |
| --- | --- | --- |
| **Subscript** | **Description** | **Data source** |
| *r* | Sub-population\* | N/A |
|  | Partners’ sub-population | N/A |
|  | Compartment | N/A |
|  | Time period | N/A |
| **Parameters (force of infection)** | |  |
|  | Probability of HIV transmission per unprotected sex contact when not suppressed for sub-population *r* | Literature review |
|  | Average number of sexual acts per month for sub-population *r* | Literature review |
|  | Percentage of individuals consistently using a condom for sub-population | DHS (2005, 2010, 2015) |
|  | Reduction in probability of HIV transmission when using a condom | Literature review |
|  | Reduction in probability of HIV transmission when on ART and virally suppressed | Literature review |
|  | Probability of individuals living with HIV who are on ART and viral suppressed at time period for sub-population | IeDEA |
|  | Probability of individuals living with HIV who are not viral suppressed at time period for sub-population | IeDEA |
|  | Number of individuals for sub-population *r* | World Bank |
|  | HIV prevalence at time period t for sub-population *r* | DHS (2005); Estimated |
|  | Multiplier applied for condom use for sub-population *r* | Calculated |
|  | Multiplier applied for viral suppressed individuals | Calculated |
| **Parameters (transition probabilities)†** | |  |
|  | Force of infection in sub-population at time | Calculated |
|  | Probability of HIV diagnosis for compartment at time t | IeDEA |
|  | Probability of HIV disease progression for compartment | IeDEA |
|  | Probability of death for compartment | IeDEA |
|  | Probability of linkage to care for compartment | Literature review |
|  | Probability of lost to follow-up for compartment | IeDEA |
|  | Probability of on ART and virally suppressed for compartment | IeDEA |
|  | Probability of return to ART and virally suppressed for compartment | IeDEA |
|  | Probability of failure to maintain viral suppression for compartment | IeDEA |
|  | Multiplier applied to adjust probabilities of diagnosis and linkage for population with CD4<200 | Literature review |
| **Compartments** | |  |
|  | Number of susceptible individuals at time | Calculated |
|  | Number of infected individuals in compartment at time | Calculated |
| D | Number of individuals who died at time | Calculated |

*Abbreviates: DHS = Demographic Health Survey; IeDEA=International epidemiology Databases to Evaluate AIDS;*

\* Sub-populations are defined by age, sex, urbanicity, and HIV acquisition risk. † Parameters are for the implemented system of difference equations.

## Probability of HIV transmission per sex act

The per sex act probability of acquiring HIV in the force of infection equation is defined as the probability of HIV transmission per sex act when an individual not living with HIV will become sero-positive after partnering with non-viral suppressed partner, without using a condom.

***Data source(s)***: A systematic review of probability of HIV transmission per sex act is used.37

* Among low-risk women, the systematic review found probabilities of male-to-female transmission is 0.30%.
* Among high-risk women, the probabilities of acquiring HIV per sex act due to partner living with HIV differ significantly compared to low-risk women. The probability of male-to-female transmission in commercial sex, which is 0.05%.
* In particular, men who were engaging in commercial sex were found to have a higher per-act probability of acquiring HIV of 2.44% and men regardless of the type of sex partner were found to have a per-act probability of acquiring HIV of 0.87%. Since we did not model men who are the clients of female sex workers independently, the probability of acquiring HIV regardless of sex partner is used (0.87%).

***Assumptions***

* Probability of HIV transmission is assumed to be the same irrespective of age and urbanicity since a study suggested that probability of HIV transmission does not statistically differ by susceptible population’s age group.38
* We did not account for differences in probabilities of HIV transmission for disease stage since we did not stratify the sub-population by disease stage.

***Analytic decisions***

* The probabilities of HIV transmission are reported in **Table 4**.

**Table 4:** **Probabilities of HIV transmission per sex act, by sub-group and direction**

|  |  |  |
| --- | --- | --- |
| **Sub-population** | **Direction** | **Probabilities (95% CI)** |
| High-risk women | Male-to-female | 0.0005 (0.0002 – 0.0013)a |
| Low-risk women | Male-to-female | 0.0030 (0.0014 – 0.0063) |
| Men | Female-to-male  (Pooled estimate for all men) | 0.0087 (0.0028 – 0.0270) |

*Note: 95% CI is extracted from literature.*

a The probability of client-to-female-sex workers transmission is low possibly due to imprecise data or low prevalence of other sexually transmitted diseases among female sex workers in the country studied. Given the uncertainty in the data, the baseline probability might be chosen outside the confidence interval during model calibration.

## Condom effectiveness

The effectiveness of condom in the force of infection equation is defined as the reduced probability of HIV transmission per sex act when an individual not living with HIV is consistently using condom during sex behaviors with sero-positive, non-viral suppressed partner.

***Data source***:

* A systematic review found that ‘always’ using a condom (i.e. use condom for all sex acts) is effective is reducing the HIV incidence by 77.6% - 82.9%, depending on the characteristics of ‘never’ using a condom group. The effectiveness of condom could range between 34.5% and 94.2%.39
* A more recent quantitative study among African suggested a similar outcome that condom use could reduce the per-act HIV risk by 78%, with a confidence interval of 58% to 89%.40

***Assumptions***

* We assumed that only male condom will be used based on the Demographic and Health Survey 2014-2015 that almost none of the women aged 19-45 were using female condom.18
* We assume that the effectiveness of condom is applied on a per-act bases since the systematic review on population-level analysis is comparable to the study on per-act effectiveness.

***Analytic decision:***

* The effectiveness of condom is 80% based on the systematic review, which is commonly cited by other modeling studies.41,42 The range of the condom effectiveness is from 34.5% to 94.2% for sensitivity analysis.39

## Proportion with consistent condom use

Consistent condom use is defined as individuals who reported always using a condom. The proportion of consistent condom use in the force of equation is measured as the proportion of individuals in a sub-population that reports consistent condom use in the past 30 days.

***Data sources:***

* Among low-risk population, we will use Demographic and Health Survey (2005, 2010, 2015) to estimate the proportion of consistent condom use by age, sex, and urbanicity.22,29,43 The DHS interviewed on “How long ago that you had your last sexual relations with a woman (men)?” “The last time you had sexual intercourse with this (second/third) person, was a condom used?” and “Did you use a condom every time you had sexual intercourse with this person in the last 12 months, if used condom last time?”
* Among high-risk women, the proportion of women consistently used condom will come from the Behavioral and Biological Survey, 2010 in Rwanda, which is the only publicly available survey data that stratified consistent condom use by age.44

***Assumptions***

* We assume the proportion of individuals who uses condoms for each sex act is uniform over a year.
* We assume the same proportion of consistent condom use for individuals aged 25+ since the estimated proportion from DHS does not differ significantly.

***Analytic decision:***

* The proportion of consistent condom use is presented in **Table 5**.

Table 5: Proportion of consistent condom use, by sub-population

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sub-group** | **Years of data applied** | **Age group (Years)** | **Proportion of consistent condom use** | **Reference** |
| Low-risk urban women | 2004-2005 | 15-24 | 0.134 (0.091– 0.194) | DHS 200522 |
| 25+a | 0.043 (0.030 – 0.063) |
| 2006-2010 | 15-24 | 0.165 (0.123 – 0.218) | DHS 201043 |
| 25+a | 0.066 (0.050 – 0.085) |
| 2011-End of projection period | 15-24 | 0.266 (0.221 – 0.316) | DHS 201529 |
| 25+a | 0.090 (0.075 – 0.108) |
| Low-risk rural women | 2004-2005 | 15-24 | 0.017 (0.010 – 0.027) | DHS 200522 |
| 25+a | 0.008 (0.006 – 0.012) |
| 2006-2010 | 15-24 | 0.066 (0.053 – 0.082) | DHS 201043 |
| 25+a | 0.032 (0.028 – 0.038) |
| 2011-End of projection period | 15-24 | 0.098 (0.082 – 0.117) | DHS 201529 |
| 25+a | 0.039 (0.034 – 0.045) |
| Urban men | 2004-2005 | 15-24 | 0.408 (0.290 – 0.536) | DHS 200522 |
| 25+a | 0.120 (0.090 – 0.159) |
| 2006-2010 | 15-24 | 0.512 (0.407 – 0.617) | DHS 201043 |
| 25+a | 0.121 (0.096 – 0.153) |
| 2011-End of projection period | 15-24 | 0.622 (0.527 – 0.708) | DHS 201529 |
| 25+a | 0.133 (0.109 – 0.161) |
| Rural men | 2004-2005 | 15-24 | 0.089 (0.060 – 0.131) | DHS 200522 |
| 25+a | 0.014 (0.009 – 0.020) |
| 2006-2010 | 15-24 | 0.210 (0.173 – 0.253) | DHS 201043 |
| 25+a | 0.040 (0.034 – 0.049) |
| 2011-End of projection period | 15-24 | 0.276 (0.230 – 0.325) | DHS 201529 |
| 25+a | 0.050 (0.042 – 0.058) |
| High-risk urban women | 2004-End of projection period | 15-24b | 0.351 (0.298 – 0.404) | BSS 201044 |
| 25+a,b | 0.276 (0.235 – 0.317) |

a We assume the proportion of consistent condom use is the same for individuals aged 25+ given the existing data.

b Due to lack of data, we assume that the confidence interval would be a relative of the baseline value.

## Effectiveness of viral suppression in reducing HIV transmission

The effectiveness of viral suppression on transmission in the force of infection equation is defined as the percent reduction in the probability of HIV transmission per sex act when an individual living with HIV viral suppressed but not using condom consistently. Viral suppression status is defined as having viral load < 1000 copies/ml.

***Data source (s)***:

* A systematic review on per-act HIV transmission risk for among individuals living with HIV off ART and under combination ART (cART) for more than 6 months suggested that the per-act probabilities of HIV transmission for unprotected sex are 0.0014 (0.0010 – 0.0018) for those off-ART and range between 0.000024 (0.0000006 – 0.000087) and 0.000047 (0.0000057 – 0.00013) for those under cART for 6 months depending on whether the transmission occur after 6 months of cART.45
  + The reviewed articles have inconsistent viral suppression threshold ranging from 50 copies/ml to 400 copies/ml, with a majority of studies defining viral suppression by 400 copies/ml. Our model defines a lower level of HIV RNA load for viral suppression suggesting that the estimates might underestimate the effectiveness of viral suppression.

***Assumptions:***

* For model structure simplicity, we assume equal risk of HIV transmission overtime when on ART before virological failure and all transmission occurs after 6 months of cART. The estimate might underestimate the effectiveness of viral suppression since transmission are more likely to occur among those with higher viral load level.

***Analytic decision:***

* We derive estimates on the effectiveness of viral suppression in reducing HIV transmission based on studies reporting the per-act probability of HIV transmission,45 as the following equation.

(2)

Where denotes overall per-act transmission probability off ART and denotes per-act transmission probability when virally suppressed.

* The effectiveness of viral suppression is 96.6% with the confidence interval of (92.8%, 99.4%). The upper bound and lower bound are calculated using the upper bound and lower bound of the probability of HIV transmission.

## Number of sexual acts

The number of sex acts in the force of infection equation represents the number of vaginal sex in the past 30 days for both men and women regardless of their partner’s HIV status.

***Data source(s):***

* A study published in 2011 that examined risk behaviors among female sex workers (FSWs) and female who received voluntary HIV counseling and testing in Rwanda found that the median numbers of vaginal sex acts per month are 40 and 1 for FSWs and female clients, respectively.46

***Assumptions:***

* Due to the lack of data, we assume that the number of sex acts per months for low-risk women and men are the same as female who received voluntary HIV counseling and testing regardless of age and urbanicity.
* We assume that no anal sex occurs during heterosexual behavior since study suggested that only a small proportion of FSWs had anal sex in the past months.47

***Analytic decision:***

* For low-risk men and women, the number of sex acts is 1 and the confidence interval is (0, 8).
* For high-risk women, the number of sex acts is 40 and the confidence interval is (20, 64)

# Parameter inputs

In this section, we describe the process used and analytic decisions made to parameterize the model. The data sources are summarized in **Section 11.1.**

## Initial population

To instantiate the model, we use the data from multiple sources to estimate the population size for each sub-population across compartments. We first describe all data source used to estimate the population size. In section 7.1.1, we estimate the size of each sub-population that stratified by age, sex, risk, and urbanicity by HIV infected status. In section 7.1.2, we estimate the distribution of population living with HIV across Infected compartments to further divide the population living with HIV within each sub-population by compartments.

***Data sources:***

* ***Low-risk populations (with and without HIV).*** Weuse World Bank population estimates from 2003 for individuals aged 15-64 years old.48 The population estimates data are reported by age group and sex, with percent of population living in urban and rural areas.
* ***High-risk populations (with and without HIV).*** We use the population estimate of female sex workers from the Joint United Nations Programme on HIV/AIDS (UNAIDS) and age-distribution estimates from 2010 Behavioral and Biological Surveillance (BSS) Survey in Rwanda.44,49 The population estimate for female sex workers comes from the 2012 Sex Worker Size Estimation Survey conducted by Rwanda Biomedical Center/Institute of HIV/AIDS Disease Prevention and Control.50 It defines female sex workers as women who have sex for money regardless of the place where they sell sex (e.g. home based, street based, or venue based). The age-distribution estimates come from BSS 2010 for female sex workers.44
* ***HIV prevalence.*** We use HIV prevalence estimates from the 2005 Demographic and Health Survey and estimates of HIV prevalence from BSS 2010 to calculate the size of low- and high-risk sub-populations living with HIV, respectively.
  + ***Distribution of population living with HIV across compartments.*** We estimate the size of populations in each infected compartment by sub-population using multiple sources, including DHS 2005,22 International epidemiology Databases to Evaluate AIDS (IeDEA),51 and published literature. The process for deriving the distribution across compartments is described later in this document.

### The size of sub-populations, by HIV status (Table 6 ):

* ***Size of sub-population.*** We estimate the size of each sub-population stratified by age, sex, risk, and urbanicity by multiplying the total size of population and the distribution reported by age groups, sex, and urbanicity in low-risk population and by age groups in high-risk population.
  + For low-risk populations, the World Bank reported the population data by age group and sex together with the percent of population living in urban and rural areas. We assume similar age and sex distributions in urban and rural areas to estimate the size of low-risk sub-populations, which is applied to all age-groups and sexes.
  + For high-risk populations, the BSS 2011 data reported the population size by age-groups. While the age-group stratification differs between BSS 2010 and our model, we assume uniform distributions within each age-group in BSS 2010 to derive our own estimates.
* ***Size of people living with HIV in each sub-population.*** We estimate the size of each sub-population living with HIV by multiplying the size of each sub-population and the HIV prevalence.
* ***Size of susceptible population in each sub-population.*** We estimate the size of each susceptible population by subtracting the people living with HIV from the size of total population in each sub-population.

**Table 6.** **Initial population size by sub-population and HIV status**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sub-population** | **Total population (2003), n** | **HIV prevalence,a %** | **People living with HIV, n** | **Susceptible population, n** |
| Low-risk urban women (15-24) | 174,420 | 4.04 | 7,047 | 167,373 |
| Low-risk urban women (25-34) | 96,551 | 9.43 | 9,105 | 87,446 |
| Low-risk urban women (35-44) | 69,640 | 19.13 | 13,322 | 56,318 |
| Low-risk urban women (45-54) | 45,588 | 7.57 | 3,451 | 42,137 |
| Low-risk urban women (55-64)b | 25,313 | 7.57 | 1,916 | 23,397 |
| Low-risk rural women (15-24) | 857,407 | 0.99 | 8,488 | 848,919 |
| Low-risk rural women (25-34) | 474,621 | 3.49 | 16,564 | 458,057 |
| Low-risk rural women (35-44) | 342,332 | 4.46 | 15,268 | 327,064 |
| Low-risk rural women (45-54) | 224,101 | 3.58 | 8,023 | 216,078 |
| Low-risk rural women (55-64)b | 124,433 | 3.58 | 4,455 | 119,978 |
| Urban men (15-24) | 163,133 | 1.1 | 1,794 | 161,339 |
| Urban men (25-34) | 89,208 | 7.33 | 6,539 | 82,669 |
| Urban men (35-44) | 72,332 | 12.35 | 8,933 | 63,399 |
| Urban men (45-54) | 41,887 | 7.98 | 3,343 | 38,544 |
| Urban men (55-64) | 19,325 | 0 | 0 | 19,325 |
| Rural men (15-24) | 801,925 | 0.27 | 2,165 | 799,759 |
| Rural men (25-34) | 438,524 | 1.62 | 7,104 | 431,420 |
| Rural men (35-44) | 355,569 | 3.61 | 12,836 | 342,733 |
| Rural men (45-54) | 205,905 | 2.79 | 5,745 | 200,160 |
| Rural men (55-64) | 94,995 | 0 | 0 | 94,995 |
| High-risk women (15-24)c | 5,648 | 39.5 | 2,231 | 3,417 |
| High-risk women (25-34)c | 4,543 | 53 | 2,408 | 2,135 |
| High-risk women (35-44)c | 1,498 | 62 | 929 | 569 |
| High-risk women (45-54)c | 295 | 63 | 186 | 109 |
| High-risk women (55-64)b,c | 295 | 63 | 186 | 109 |
| a. HIV prevalence data are based on the DHS 2005 data for low-risk population and BSS 2010 for high-risk women. | | | | | |
| b. Given limited data, we assume that women aged 55+ have the same HIV prevalence as women aged 45-54. | | | | | |
| c. We assume the total population of female sex worker is 12,278 based on the 2012 Sex Worker Size Estimation Survey. The distribution of FSWs by age is estimated based on 2010 BSS data. Given 2010 BSS data has different age groups over 30, we assume a uniform distribution within age groups. | | | | | |

### The size of sub-populations living with HIV across compartments (Table 7):

The distribution of people living with HIV in each sub-population across compartments are estimated first along the HIV care continuum and, then, by CD4 stratum. The number in each compartment (i.e., the size of each sub-population) are then calculated based on the proportions derived.

* ***Estimating the distribution of at key steps along the HIV care continuum***

*Estimating the proportion diagnosed among all people living with HIV*:

Data:

* + We used Demographic and Health Surveys (DHS) for Rwanda in 200552, a national household survey, to estimate the proportion of people diagnosed with HIV. In DHS (2005), a random sub-sample of the respondents received an HIV test, regardless of prior test history, to detect their HIV status. Respondents were also asked if they had previously received an HIV test and if so, whether they received the result of their last HIV test based on the following DHS (2005) questions in individual’s questionnaire:
    - I don’t want to know the results, but have you ever been tested to see if you have the AIDS virus
    - I don’t want to know the results, but did you get the results of the test? (YES/NO)

Assumptions:

* We assume individuals are diagnosed if they were found living with HIV and ever received HIV testing per DHS data. This might overestimate the proportion of individuals who are diagnosed since individuals currently living with HIV might acquire HIV after receiving the previous testing result.
* We assume all individuals who are tested will know their result since approximately 92% of the individuals ever tested have received their results.
* We assume that the proportion of individuals diagnosed does not differ by sub-population due to the small sample size.

Estimation:

* The sample is restricted to individuals living with HIV since the proportion diagnosed is conditional on people living with HIV. The proportion of diagnosed among all individuals living with HIV is 44.6% (39.3%, 50.5%)

*Estimating the proportion linked to care among diagnosed individuals*

Data:

* + We used a prospective cohort study in 2007 among female sex workers, where 85% were linked to care within 3 months of diagnosis and referral.53

Assumptions:

* + We assumed that the proportion of individuals diagnosed did not differ by sub-population.

Estimation:

* + The proportion of individuals who are linked is 85%.

*Estimating proportion of individuals lost to follow-up (LTFU)*

Data:

* + We will use the 2004 IeDEA-Rwanda data to estimate the proportion of individuals LTFU among all individuals linked to care.51 LTFU is defined as at least 365 days between last clinical visit and censor date before ART initiation or at least 6 months between last clinical visit and censor date when on ART based on Rwanda national HIV treatment guidelines and are consistent with other studies examining LTFU in sub-Saharan Africa.57–60
  + The proportion of individuals LTFU will be adjusted to account the misclassification of LTFU. We will extract the multipliers from multiple literature including:
    - A retrospective study from Kenya (2009-2011) estimating the misclassification of loss to follow-up (LTFU) for pre-ART patients. The proportion of pre-ART patients who have died but coded as LTFU is 17%.54
    - A prospective study from Uganda (2007-2011) tracking the pre-ART LTFU patients for their HIV treatment and care status. Approximately 13% of the pre-ART patients who considered LTFU received HIV care in a clinic other than the original one.55
    - A meta-analysis on misclassification of LTFU for individuals on ART in sub-Saharan region. The proportions of patients coded as LTFU but have died or transferred are 20.8% and 35.9% respectively.56

Estimation:

* The estimation is done by sub-populations.
* For pre-ART patients,

(3)

* For on ART patients,

(4)

*Estimating the proportion On ART*

Data:

* For each sub-population, we will use 2004 IeDEA-Rwanda data to estimate the proportion of individuals on ART among those who are linked to care and not LTFU, by sub-population.51

Assumption:

* We assume that individuals who are on ART without evidence of virological failure are virally suppressed.

Estimation:

* The estimation is done by sub-populations.

(5)

*Estimating the number On ART and not suppressed*

*Data:*

* We will use 2004 IeDEA-Rwanda data to estimate the proportion of individuals on ART and not suppressed among those who are on ART and not LTFU, by sub-population.51 We define virologic failure as viral suppressed patients with at least 1 viral load tests of > 1000 HIV RNA copies/ml in the follow-up tests.

Estimation:

* The estimation is done by sub-populations.

(6)

* ***Estimating the distribution of people living with HIV, by CD4 stratum***

Data:

* For each sub-population, we will use IeDEA data (Rwanda, 2004) to estimate the distribution of people living with HIV, by CD4 stratum.

Assumptions:

* We assume the CD4 distribution of individuals living with HIV does not differ among individuals who are undiagnosed, diagnosed, linked, or LTFU.
* We assume that CD4 distribution at enrollment is representative of the overall CD4 distribution of those living with HIV.

Estimation:

* *Undiagnosed, diagnosed, linked, LTFU****:*** The CD4 cell count distributions for those undiagnosed, diagnosed, and linked are based on the estimated CD4 cell count at enrollment date. The CD4 cell count distribution for those LTFU is based on the CD4 estimate at the date of LTFU.
* *On ART and suppressed, On ART not suppressed:* The CD4 distribution for those On ART is based on CD4 cell count at ART initiation. The process for estimation of CD4 cell count is in **Section 11.4**.

**Table 7.** **Size of initial sub-populations, by CD4 stratum and engagement in care**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sub-population | Diagnosis, n | | | | Lost, n | | | | On ART & suppressed, n | | | | On ART & not suppressed, n | | | |
| CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 | CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 | CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 | CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 |
| Low-risk urban women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

## Susceptible population growth rate

We apply a time-step-specific (in this case, monthly) growth rate for each sub-population.

***Data source***: Data come from World Bank annual population estimates and projections, 2004–2045.48 The population estimates data are reported by age group and sex, with percent of population living in urban and rural areas.

***Assumptions***:

* The growth rate is exponentially distributed, i.e., constant over time.
* The susceptible population growth rate for the low-risk sub-population reflects Rwanda’s overall population growth rate.
* The population growth rate in high-risk urban women is the same as low-risk urban women.
* The growth rate for the susceptible population aged 24-54 years with similar sex, risk status, and residency does not vary by age since the growth rates for susceptible population within this age group are not statistically different (**Section 11.5**).

***Estimation:***

The population growth rate is estimated in three steps.

* We estimate the size of each sub-population using the World Bank data from 2003 – 2045 following the same procedure described earlier in **Section 7.1**.
* We calculate the monthly population growth for each year from 2003 - 2045 using the following equation:

(7)

where, is the population estimate in the next time period in sub-population , is the population in the current time period in sub-population , is the number of months in a year, and is the monthly population growth rate in sub-population .

* The overall monthly population growth rate is estimated as the average monthly growth rate over time and the confidence interval is estimated using the average and standard deviation of the population growth rates. The monthly growth rate and the confidence interval is reported in **Table 8**.

**Table 8.** **Average monthly population growth ratea**

|  |  |  |
| --- | --- | --- |
| **Sub-population (age group)** | **Monthly growth rate** | **95% Confidence Interval** |
| Low-risk urban women (15-24) | 0.20% | (0.17%, 0.23%) |
| Low-risk urban women (25-54)b | 0.33% | (0.32%, 0.34%) |
| Low-risk urban women (55-64) | 0.45% | (0.40%, 0.50%) |
| Low-risk rural women (15-24) | 0.08% | (0.05%, 0.11%) |
| Low-risk rural women (25-54)b | 0.21% | (0.19%, 0.24%) |
| Low-risk rural women (55-64) | 0.33% | (0.29%, 0.37%) |
| Urban men (15-24) | 0.21% | (0.18%, 0.24%) |
| Urban men (25-54)b | 0.33% | (0.31%, 0.34%) |
| Urban men (55-64) | 0.48% | (0.42%, 0.53%) |
| Rural men (15-24) | 0.09% | (0.06%, 0.12%) |
| Rural men (25-54)b | 0.21% | (0.19%, 0.24%) |
| Rural men (55-64) | 0.36% | (0.30%, 0.41%) |
| High-risk urban women (15-24)c | 0.20% | (0.17%, 0.23%) |
| High-risk urban women (25-54)b,c | 0.33% | (0.32%, 0.34%) |
| High-risk urban women (55-64)c | 0.45% | (0.40%, 0.50%) |
| 1. The monthly growth rate is estimated based on the World Bank population estimates and projection.48 The monthly population growth rate is an average of the monthly growth rate between 2003 and 2045. 2. We adopt the same population growth rate for individuals aged 25-34, 35-44, and 45-54 years because previous analysis showed that the growth rate within the three age groups did not differ significantly. 3. We assume that the population growth rate in high-risk urban women is the same as low-risk urban women. | | |

## HIV diagnosis

***Data sources:***

* For low-risk population, we used Demographic and Health Surveys (DHS) for Rwanda in 200552, 201020 and 201518, a national household survey, to estimate the probability of HIV diagnosis for low-risk women and men in both urban and rural areas. In the DHS, respondents were asked if they had previously received an HIV test and if so, the timing of the tests. They were also asked whether they received the result of their last HIV test. A random sub-sample of these respondents also received an HIV test, regardless of prior test history. Our estimates were based on the following DHS questions:
* In 2005. Individual’s questionnaire No.715, No. 716, No.718.
  + I don’t want to know the results, but have you ever been tested to see if you have the AIDS virus
  + When was the last time you were tested? (LESS THAN 12 MONTHS/12-23 MONTHS/2 YEARS OR MORE)
  + I don’t want to know the results, but did you get the results of the test? (YES/NO)
* In 2010 and 2015. Women’s questionnaire No.926, No. 927, No.928; men’s questionnaire No.712, No. 713, No. 714.
  + I don’t want to know the results, but have you ever been tested to see if you have the AIDS virus? (YES/NO)
  + How many months ago was your most recent HIV test? (\_\_MONTHS AGO/TWO OR MORE YEARS)\
  + I don’t want to know the results, but did you get the results of the test? (YES/NO)
* The Behavioral and Biological Surveillance Survey (BBSS) (2010), a survey of female sex workers in Rwanda,44 was used to estimate the probability of HIV diagnosis among high-risk women in urban areas. In the BBSS, respondents were asked if they have received an HIV test within the past 12 months and if so, had the result been received.

***Assumptions:***

* The probability of HIV diagnosis among low-risk populations does not differ by urbanicity, given no statistically significant difference in our estimates for urban versus rural populations (2010 and 2015 only).
* The probabilities of HIV diagnosis for men in age groups 15-24 years and 25-34 years do not differ; and the probabilities of HIV diagnosis for men and women in age groups 35-44, 45-44, and 55+ do not differ. The assumption is made given small sub-population sample sizes and similar estimates across select sub-populations.
* The probability of HIV diagnosis does not vary by CD4 stratum for individuals with CD4>200, given the evidence suggesting those with CD4 count ≤200 are 1.56 times (95% CI 1.11, 2.20) more likely to be linked compared to those with CD4>200.57 The evidence is applicable to HIV diagnosis since it is comparable with the estimates used in other modelling studies.58
* For low-risk sub-populations: Respondents who received their last HIV test result more than 12 months prior to the DHS interview date knew their HIV status.
* For high-risk sub-populations: The probability of HIV diagnosis is the same regardless of HIV status, since high-risk sub-populations are targeted by HIV testing promotion policies and campaigns.59
* For high-risk sub-population: The probability of HIV diagnosis does not vary based on age and urbanicity, given similar estimates in the sample by age and region.44

***Estimation:***

Low-risk women and men in both urban and rural areas. We estimated the probability of HIV diagnosis estimated using the proportion of DHS respondents living with HIV who received an HIV test result in the past 12 months, given previously unknown HIV status. We restricted the sample to individuals living with unknown HIV status, given individuals living with HIV are more likely to have high-risk behaviors and thus get tested.60 To adjust for the time step in our model (monthly), the annual probability of HIV diagnosis will be converted to annual rate and then the monthly rate. The monthly rate will then be converted back to monthly probability of HIV diagnosis.

High-risk women in urban areas. We used the proportion receiving an HIV test within the past 12 months, which is consistent with other modelling study,61 adjusted for the time step in our model (monthly), as a proxy for probability of HIV diagnosis for female sex workers. We used the proportion of those testing for HIV and receiving their result within the past 12 months to represent the annual probability of HIV diagnosis. The annual probability of HIV diagnosis is converted to annual rate of HIV diagnosis, and then the monthly rate of HIV diagnosis. The monthly rate of HIV diagnosis will be further concerted to monthly probability of diagnosis.

**Table 9.** **Monthly probability of HIV diagnosis, by sub-population and CD4 stratum**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sub-population** | **Years of data applied** | **Age group (Years)** | **Monthly probabilities (95% CI)** | | **Reference** |
| **CD4 >200\*** | **CD4 ≤200** |
| Low-risk women | 2004-2005 | 15-24 | 0.016 (0.007 – 0.037) | 0.025 (0.010 – 0.057) | DHS 200522 |
| 25-34 | 0.027 (0.016 – 0.043) | 0.041 (0.025 – 0.067) |
| 35-44 | 0.027 (0.018 – 0.042) | 0.043 (0.027 – 0.065) |
| 45-54 | 0.027 (0.018 – 0.042) | 0.043 (0.027 – 0.065) |
| 55+ | 0.027 (0.018 – 0.042) | 0.043 (0.027 – 0.065) |
| 2006-2010 | 15-24 | 0.145 (0.058 – 0.246) | 0.226 (0.091 – 0.383) | DHS 201043 |
| 25-34 | 0.146 (0.076 – 0.225) | 0.228 (0.118 – 0.351) |
| 35-44 | 0.111 (0.069 – 0.160) | 0.173 (0.108 – 0.250) |
| 45-54 | 0.111 (0.069 – 0.160) | 0.173 (0.108 – 0.250) |
| 55+ | 0.111 (0.069 – 0.160) | 0.173 (0.108 – 0.250) |
| 2011- End of projection period | 15-24 | 0.126 (0.065 – 0.197) | 0.196 (0.101 – 0.307) | DHS 201529 |
| 25-34 | 0.133 (0.081 – 0.191) | 0.207 (0.127 – 0.297) |
| 35-44 | 0.115 (0.066 – 0.173) | 0.179 (0.103 – 0.269) |
| 45-54 | 0.115 (0.066 – 0.173) | 0.179 (0.103 – 0.269) |
| 55+ | 0.115 (0.066 – 0.173) | 0.179 (0.103 – 0.269) |
| Low-risk men | 2004-2005 | 15-24 | 0.018 (0.008 – 0.038) | 0.028 (0.012 – 0.059) | DHS 200522 |
| 25-34 | 0.018 (0.008 – 0.038) | 0.028 (0.012 – 0.059) |
| 35-44 | 0.021 (0.012 – 0.037) | 0.033 (0.018 – 0.058) |
| 45-54 | 0.021 (0.012 – 0.037) | 0.033 (0.018 – 0.058) |
| 55+ | 0.021 (0.012 – 0.037) | 0.033 (0.018 – 0.058) |
| 2006-2010 | 15-24 | 0.071 (0.024 – 0.226) | 0.111 (0.038 – 0.353) | DHS 201043 |
| 25-34 | 0.071 (0.024 – 0.226) | 0.111 (0.038 – 0.353) |
| 35-44 | 0.139 (0.076 – 0.210) | 0.217 (0.119 – 0.328) |
| 45-54 | 0.139 (0.076 – 0.210) | 0.217 (0.119 – 0.328) |
| 55+ | 0.139 (0.076 – 0.210) | 0.217 (0.119 – 0.328) |
| 2011- End of projection period | 15-24 | 0.045 (0.037 – 0.111) | 0.071 (0.058 – 0.173) | DHS 201529 |
| 25-34 | 0.045 (0.037 – 0.111) | 0.071 (0.058 – 0.173) |
| 35-44 | 0.146 (0.098 – 0.198) | 0.228 (0.153 – 0.309) |
| 45-54 | 0.146 (0.098 – 0.198) | 0.228 (0.153 – 0.309) |
| 55+ | 0.146 (0.098 – 0.198) | 0.228 (0.153 – 0.309) |
| High-risk urban women | 2004- End of projection period | 15-24 | 0.168 (0.088 – 0.241) | 0.241 (0.129 – 0.339) | DHS 200522 |
| 25-34 | 0.168 (0.088 – 0.241) | 0.241 (0.12 – -0.339) |
| 35-44 | 0.168 (0.088 – 0.241) | 0.241 (0.129 – 0.339) |
| 45-54 | 0.168 (0.088 – 0.241) | 0.241 (0.129 – 0.339) |
| 55+ | 0.168 (0.088 – 0.241) | 0.241 (0.129 – 0.339) |

\* CD4 >200 represents the following CD4 strata: >500, >350–500, and >200–350.

## Natural history

The natural history of HIV is defined as advancing through CD4 count strata (>500, >350-500, >200-350, <200). Progression to the next disease stage occurs if the estimated CD4 cell count falls below the threshold of a given CD4 count stratum.

***Data source.***

* To estimate natural history of HIV, we will use the International epidemiology databases to evaluate AIDS (IeDEA) consortium in Rwanda from 2004 to most current data.51
* To estimate the weights applied to the competing risks model, we will use a retrospective study from Kenya (2009-2011), which describes the underestimation of death due to the misclassification of loss to follow-up (LTFU) in the IeDEA data. The proportion of patients who have died but coded as LTFU is 17%.54

***Assumptions.***

* We assume a linear trend in the decrease of CD4 level for individuals living with HIV.

***Estimation***

The probability of disease progression will be estimated using a weighted competing risk model for each sub-population (**Table 10)**. This model is selected because the probability of disease progression will be affected depending on the whether the patient has initiated ART or died. ART initiation and death will be modeled as the competing risks. The competing risk model is presented as follows.

(8)

Where, *p* = disease progressed (i.e., reaching the CD4 threshold), w = weight to account for underestimation of disease progression stopped (i.e. death or ART initiation), *S* = event-free function (survival from competing events, or disease progression), = weighted probability of being disease progression before time, ***X*** = vector of predictor variables (e.g., site) for patient , = cause-specific weighted hazard function for disease progression, and = time-dependent weighted probability that the patient with factors ***X*** is event-free.

In the weighted competing risk model, we will calculate time at risk for each individual to estimate the hazard function of disease progression, which is based on the residence time within a CD4 cell count stratum. The process for estimating the time-at-risk can be found in **Section 11.4.** We will control for site-level factors only (i.e. clinics) since clinic- or site-level factors can be associated with disease progression (e.g. nutrition support programs).62 We did not control for demographic characteristics or disease stage since the model is applied to each sub-population by CD4 stratum. The variables created for the model and IeDEA data used is presented in **Table 13**.

The model will adjust for misclassification of competing risks (due to death or ART initiation) and censoring events (e.g. LTFU). Evidence from sub-Saharan Africa suggests that some patients recorded as LTFU had died,56,63–66 which suggests a potential underestimation of the competing risks. Thus, we adjust for the probability of competing risks and the probability of censoring considering patient who is misclassified as LTFU. The function for deriving weights is adapted from Haas et al.,56 with the probability weight calculated as:

(9)

We parameterize this expression as follows:

* Proportion of death among LTFU. The proportion of patients LTFU but who have died is based on a retrospective study from Kenya that evaluated outcomes of people living with HIV in the general population and LTFU (data from 2009-2011); the study found that 17% of pre-ART patients classified as LTFU had died.54

**Table 10.** **Sub-population monthly probability of disease progression, by CD4 stratum and engagement in care**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sub-population | Linked | | | Undiagnosed, Diagnosed, and Lost to follow-up | | |
| CD4 >500 | CD4 350-500 | CD4 200-350 | CD4 >500 | CD4 350-500 | CD4 200-350 |
| Low-risk urban women (15-24) |  |  |  |  |  |  |
| Low-risk urban women (25-34) |  |  |  |  |  |  |
| Low-risk urban women (35-44) |  |  |  |  |  |  |
| Low-risk urban women (45-54) |  |  |  |  |  |  |
| Low-risk urban women (55-64) |  |  |  |  |  |  |
| Low-risk rural women (15-24) |  |  |  |  |  |  |
| Low-risk rural women (25-34) |  |  |  |  |  |  |
| Low-risk rural women (35-44) |  |  |  |  |  |  |
| Low-risk rural women (45-54) |  |  |  |  |  |  |
| Low-risk rural women (55-64) |  |  |  |  |  |  |
| Urban men (15-24) |  |  |  |  |  |  |
| Urban men (25-34) |  |  |  |  |  |  |
| Urban men (35-44) |  |  |  |  |  |  |
| Urban men (45-54) |  |  |  |  |  |  |
| Urban men (55-64) |  |  |  |  |  |  |
| Rural men (15-24) |  |  |  |  |  |  |
| Rural men (25-34) |  |  |  |  |  |  |
| Rural men (35-44) |  |  |  |  |  |  |
| Rural men (45-54) |  |  |  |  |  |  |
| Rural men (55-64) |  |  |  |  |  |  |
| High risk women (15-24)a |  |  |  |  |  |  |
| High risk women (25-34)a |  |  |  |  |  |  |
| High risk women (35-44)a |  |  |  |  |  |  |
| High risk women (45-54)a |  |  |  |  |  |  |
| High risk women (55-64)a |  |  |  |  |  |  |

a The natural history for high-risk women will be estimated using the same estimated for low-risk urban women since most high-risk women resides in urban area and the natural history does not differ by risk of HIV acquisition.

## Linked to care

***Data sources.*** Two studies were used as sources for linkage to care estimates.

* For low-risk women and men in both urban and rural areas. We will use a cross-sectional study that collects HIV care and treatment information from individuals who participated national HIV care programme in 2013-2014 in Rwanda.5 It is estimated that 90% of individuals diagnosed with HIV were linked to care in a year.5
* For high-risk urban women, a cohort study examined the rate of linkage and treatment since diagnosis were conducted in 2007-2008.53 It is estimated that 77% of individuals diagnosed with HIV will be linked to care based on a study among female sex workers in Rwanda.53

***Assumptions***

* The probabilities of linkage to care is the same for individuals in CD4 count strata >200-350, >350-500, and >500. Linkage to care for those with CD4≤200 are 1.56 times more likely to be linked to care compared to those with CD4>200. This assumption is based on a randomized controlled trial in Kenya that examined testing and linkage to care based on CD4, with 1.56 (95% CI 1.11-2.20) times vs those with CD4>200.57
* Linkage to care is the same irrespective of sex, age and urbanicity due to the limited data.

***Estimation***

For low-risk population, the annual probability of linkage to care is 90%. And for the high-risk population, the annual probability of linkage to care is 77%.

The annual probability of linkage to care will be converted to monthly probability. Since the original data does not provide the confidence interval, we will use the 0.5 and 1.5 times the baseline value as the lower and upper bound for the confidence interval. (**Table 11)**

**Table 11:** **Monthly probability of sub-population linkage to HIV care, by CD4 stratum**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sub-population** | **Monthly probability (95% CI)** | | **Reference** |
| **CD4>500, CD4>350–500, CD4>200–350** | **CD4≤200** |
| High-risk women | 0.115 (0.059 – 0.168) | 0.174(0.091 – 0.249) | 5 |
| Low-risk women and men | 0.175 (0.092 – 0.250) | 0.259 (0.139 – 0.362) | 53 |

## Lost to follow-up

LTFU is defined as at least 365 days between last clinical visit and censor date before ART initiation or at least 6 months between last clinical visit and censor date when on ART. The definition is based on Rwanda national HIV treatment guidelines and are consistent with other studies examining LTFU in sub-Saharan Africa.57–60 In the sensitivity analysis, we will test the impact of definition LTFU based 120 days as stated in the Rwanda guidelines.67 Parameter values for LTFU will vary based on disease progression, engagement in HIV care, and sub-groups. Model compartments that apply to LTFU include linkage to care; On ART and suppressed; On ART and not suppressed.

***Data source.***

* To estimate LTFU, we will use the International epidemiology databases to evaluate AIDS (IeDEA) consortium in Rwanda from 2004 to most current data.51
* To estimate the weights applied to the competing risks model, we will extract data from multiple literature including:
  + A retrospective study from Kenya (2009-2011) estimating the misclassification of loss to follow-up (LTFU) for pre-ART patients. The proportion of pre-ART patients who have died but coded as LTFU is 17%.54
  + A prospective study from Uganda (2007-2011) tracking the pre-ART LTFU patients for their HIV treatment and care status. Approximately 13% of the pre-ART patients who considered LTFU received HIV care in a clinic other than the original one.55
  + A meta-analysis on misclassification of LTFU for individuals on ART in sub-Saharan region. The proportions of patients coded as LTFU but have died or transferred are 20.8% and 35.9% respectively.56

***Assumptions.***

* We assume a linear trend in the decrease of CD4 level for individuals living with HIV.
* We assume that all patients who initiate ART are virally suppressed unless there is evidence of not being viral suppressed.68 To calculate the time-at-risk for individuals on ART and virally suppressed, the starting date until LTFU would be the date of ART initiation.

***Estimation.***

The probability of LTFU will be estimated using a weighted competing risk model for each sub-population (**Table 12)**. This model is selected against the traditional Kaplan-Meier model because the probability of event (i.e., LTFU) is not independent to the censor date(e.g., the probability of LTFU will be zero when patient has died but would not be zero if they initiated ART). Death will be modeled as the competing risk. The competing risk model is presented as follows.

(10)

Where, *p* = LTFU, w = weight to account for underestimation of death, *S* = event-free function (survival from death, or LTFU), = weighted probability of LTFU before time, ***X*** = vector of predictor variables (e.g., site) for patient , = cause-specific weighted hazard function for LTFU, and = time-dependent weighted probability that the patient with factors ***X*** is event-free.

In the weighted competing risk model, we will calculate time at risk within each CD4 cell count stratum for each individual to estimate the hazard function in the model. The process for estimating the time-at-risk can be found in **Section 11.4.** We will control for site-level factors only (i.e. clinics) since clinic- or site-level factors can be associated with disease progression (e.g. nutrition support programs).62 We did not control for demographic characteristics or disease stage since the model is applied to each sub-population by CD4 stratum. The variables created for the model and IeDEA data used is presented in **Table 13**.

The model will adjust for misclassification of competing risks (due to death) and censoring events (e.g. database closer, transfer, ART initiation, or virological failure). Evidence from sub-Saharan Africa suggests that some patients recorded as LTFU had died or transferred to another clinic.56,63–66 When estimating the cumulative probability of LTFU, we will adjust for mortality and self-transfer among patients misclassified as LTFU using probability weights.56 The probability weight is calculated as:

(11)

Where P(LTFU who die) is the proportion of LTFU who died and P(LTFU who self-transfer) is the proportion of LTFU who self-transferred to other sites.

We parameterize this expression as follows:

* Weights for Linked to Care. The weight is **0.70** {1-(0.17-0.13)}. The proportion of patients LTFU but who have died is based on a retrospective study from Kenya that evaluated outcomes of people living with HIV in the general population and LTFU (data from 2009-2011); the study found that 17% of pre-ART patients classified as LTFU had died.54 The proportion of patients LTFU but who have transferred is based on a retrospective study from Uganda that traced people living with HIV in the general population who were on Pre-ART and LTFU (data from 2007-2011); the study found that 13% of patients classified as LTFU before initiating ART had transferred to another clinic.55
* Weights for On ART. The weight is **0.43**{(1-0.208-0.359)}. The proportion of patients LTFU having died or transferred to another site is based on a meta-analysis on LTFU among patients in ART programs in sub-Saharan Africa; the study found that 20.8% and 35.9% of patients recorded as LTFU had died or self-transferred to another ART clinic.56

**Table 12:** **Monthly probability of sub-population LTFU, by CD4 stratum and engagement in care**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sub-population | Linked to Care | | | | On ART, Suppressed | | | | On ART, Not Suppressed | | | |
| CD4 >500 | CD4 350-500 | CD4 200-350 | CD4 < 200 | CD4 >500 | CD4 350-500 | CD4 200-350 | CD4 < 200 | CD4 >500 | CD4 350-500 | CD4 200-350 | CD4 < 200 |
| Low-risk urban women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |

**Table 13:** **IeDEA variables for estimating model parameter inputsa**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variablesb** | | **IeDEA data usedb** | |
| ***Variable*** | ***Definition*** | ***Variable*** | ***Definition*** |
| Patient ID | Unique patient ID | Patient ID | Unique patient ID |
| LTFU | Patient classified as LTFU | ENROL\_D | Date of enrollment |
| HIV\_POS\_D | Date of first positive HIV test |
| VIS\_D | Date of last clinic visit |
| Dead | Patient classified as dead | DEATH\_D | Date of death |
| Transfer | Patient classified as transferred | DROP\_RS | Reason for drop |
| DROP\_D\_A | Date of drop out |
| CD4\_count | Patient’s CD4 cell count stratum based on predicted CD4 | CD4\_D | Date of CD4 measurement |
| CD4\_V | Value of CD4 measurement |
| CD4\_U | Unit of measurement (cells/mm3) |
| Viral suppression | Patient’s level of viral load | HIV RNA\_D | Date of viral load measurement |
| HIV RNA\_V | HIV-RNA measurement value in copies/mL |
| Age | Patient age | BIRTH\_D | Date of birth |
| VIS\_D | Date of last visit |
| ENROL\_D | Date of enrollment |
| Sex | Patient sex | SEX | Sex at birth (1 = Male; 2 = Female; 9 = Unknown) |
| Urbanicity | Residence of the patient | RURAL | Clinic location (1=urban, 2=mostly urban, 3=mostly rural, 4=rural, 9=unknown |
| Clinic | Clinic the patient is receiving treatment | CENTER | Clinic code |

a The model parameter inputs include the probability of disease progression, LTFU, On ART\_ suppressed, On ART\_not suppressed, and death.

b Variables are the variables we use in the regression model to estimate parameter inputs, which will be informed or constructed by the variables in the IeDEA databased listed.

## On ART, Suppressed

Patients who are on ART and virally suppressed are assumed to be those on ART and without evidence of virological failure.68 Virological failure is defined as HIV RNA >1000 copies/ml based on Rwandan HIV guidelines to account for differences in the probability of HIV transmission.68 Parameter values for probability of on ART and viral suppressed will vary based on disease progression and sub-group. Model compartments that apply to probability of on ART and viral suppression include only linkage to care.

***Data source.***

* To estimate the probability of being on ART and viral suppressed among individuals linked to care, we will use the International epidemiology databases to evaluate AIDS (IeDEA) consortium in Rwanda from 2004 to most current data.51
* For patients on ART, the weight applied to adjust for the misclassification of patients is extracted from previous study examining LTFU using IeDEA data, which is calculated based on a meta-analysis in sub-Saharan area. The proportion of patients LTFU having died is 20.8%.56

***Assumptions.***

* We assume a linear trend in the decrease of CD4 level for individuals living with HIV.
* We assume that all patients who initiate ART are virally suppressed unless there is evidence of virological failure.68
* We assume that a single HIV RNA test results of >1000 copies/ml will indicate the virological failure.

***Estimation***

The probability of on ART and viral suppression (**Table 14)** will be estimated using a weighted competing risk model for each sub-population. This model is selected against the traditional Kaplan-Meier model because the probability of event (i.e., ART and suppressed) is not independent to the censor date (e.g., the probability of viral suppression will be zero when patient has died or LTFU). LTFU and death will be modeled as the competing risks. The competing risk model is presented as follows.

(12)

Where, *p* = viral suppression, w = weight to account for underestimation of on ART and viral suppression, *S* = event-free function (survival from death, or on ART and suppressed), = weighted probability of being on ART and suppressed before time, ***X*** = vector of predictor variables (e.g., site) for patient , = cause-specific weighted hazard function for on ART and suppressed, and = time-dependent weighted probability that the patient with factors ***X*** is event-free.

In the weighted competing risk model, we will calculate time at risk for each individual to estimate the hazard function of on ART and suppressed, which is based on the residence time within a CD4 cell count stratum. The process for estimating the time-at-risk can be found in **Appendix A3.** We will control for site-level factors only (i.e. clinics) since clinic- or site-level factors can be associated with viral suppression.62 We did not control for demographic characteristics or disease stage since the model is applied to each sub-population by CD4 stratum. The variables created for the model and IeDEA data used is presented in **Table 13**.

The model will adjust for misclassification of competing risks (due to death) and censoring events (e.g. LTFU). Evidence from sub-Saharan Africa suggests that some patients recorded as LTFU had died or transferred to another clinic.56,63–66 When estimating the cumulative probability of viral suppression, we will adjust for the probability of competing risks and the probability of censoring considering that patients who were transferred misclassified as LTFU. The patients who are dies but misclassified as death will be ignored in this analysis since both events are competing events of viral suppression. The function for deriving weights are adapted from the Haas paper.56 The probability weight is calculated as:

(13)

We parameterize this expression as follows:

* Proportion of transfer among LTFU. The proportion of patients LTFU having transferred to another site is based on a meta-analysis on LTFU among patients in ART programs in sub-Saharan Africa; the study found that 35.9% of patients recorded as LTFU had self-transferred to another ART clinic.56

**Table 14:** **Monthly probability of sub-population On ART, Suppressed, by CD4 stratum**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sub-population | On ART & suppressed, Linkage | | | |
| CD4 >500 | CD4 >350-500 | CD4 >200-350 | CD4 <200 |
| Low-risk urban women (15-24) |  |  |  |  |
| Low-risk urban women (25-34) |  |  |  |  |
| Low-risk urban women (35-44) |  |  |  |  |
| Low-risk urban women (45-54) |  |  |  |  |
| Low-risk urban women (55-64) |  |  |  |  |
| Low-risk rural women (15-24) |  |  |  |  |
| Low-risk rural women (25-34) |  |  |  |  |
| Low-risk rural women (35-44) |  |  |  |  |
| Low-risk rural women (45-54) |  |  |  |  |
| Low-risk rural women (55-64) |  |  |  |  |
| Urban men (15-24) |  |  |  |  |
| Urban men (25-34) |  |  |  |  |
| Urban men (35-44) |  |  |  |  |
| Urban men (45-54) |  |  |  |  |
| Urban men (55-64) |  |  |  |  |
| Rural men (15-24) |  |  |  |  |
| Rural men (25-34) |  |  |  |  |
| Rural men (35-44) |  |  |  |  |
| Rural men (45-54) |  |  |  |  |
| Rural men (55-64) |  |  |  |  |
| High risk women (15-24) |  |  |  |  |
| High risk women (25-34) |  |  |  |  |
| High risk women (35-44) |  |  |  |  |
| High risk women (45-54) |  |  |  |  |
| High risk women (55-64) |  |  |  |  |

## On ART, Not Suppressed

We define virologic failure as viral suppressed patients with at least 1 viral load tests of > 1000 HIV RNA copies/ml in the follow-up tests. This study will assume viral suppression as 1000 HIV RNA copies/ml based on Rwandan HIV guidelines to account for differences in the probability of HIV transmission.68 Parameter values of probabilities of on ART and not viral suppression (or virological failure) will vary based on disease progression and sub-group. Model compartments that apply to probability of virological failure include only On ART and viral suppression.

***Data source.***

* To estimate probability of On ART and not suppressed, we will use the International epidemiology databases to evaluate AIDS (IeDEA) consortium.51
* For patients on ART, the weight applied to adjust for the misclassification of patients is extracted from previous study examining LTFU using IeDEA data, which is calculated based on a meta-analysis in sub-Saharan area. The proportion of patients LTFU having died is 20.8%.56

***Assumptions.***

* We assume a linear trend in the decrease of CD4 level for individuals living with HIV.
* We assume that all patients who initiate ART are virally suppressed unless there is evidence of virological failure.68
* We assume that a single HIV RNA test results of >1000 copies/ml will indicate the virological failure.

***Estimation***

The probability of on ART and not viral suppression will be estimated using a weighted competing risk model for each sub-population **(Table 15)**. This model is selected against the traditional Kaplan-Meier model because the probability of event (i.e., ART and not suppressed) is not independent to the censor date (e.g., the probability of viral suppression will be zero when patient has died). Death will be modeled as the competing risks. The competing risk model is presented as follows.

(14)

Where, *p* = on ART and not viral suppression, w = weight to account for underestimation of on ART and not viral suppression, *S* = event-free function (survival from death, or on ART and not suppressed), = weighted probability of being on ART and not suppressed before time, ***X*** = vector of predictor variables (e.g., site) for patient , = cause-specific weighted hazard function for on ART and not suppressed, and = time-dependent weighted probability that the patient with factors ***X*** is event-free.

In the weighted competing risk model, we will calculate time at risk for each individual to estimate the hazard function of on ART and not suppressed, which is based on the residence time within a CD4 cell count stratum. The process for estimating the time-at-risk can be found in **Section 11.4.** We will control for site-level factors only (i.e. clinics) since clinic- or site-level factors can be associated with on ART and not suppressed (e.g. nutrition support programs).62 We did not control for demographic characteristics or disease stage since the model is applied to each sub-population by CD4 stratum. The variables created for the model and IeDEA data used is presented in **Table 13**.

The statistical model will also adjust for misclassification of death. Evidence from sub-Saharan Africa suggests that some patients recorded as LTFU had died or transferred to another clinic.56,63–66 When estimating the cumulative probability of virological failure, we will adjust for mortality misclassified as LTFU using probability weights.56 The function for deriving weights are adapted from the Haas paper.56 The probability weight is calculated as:

(15)

We parameterize this expression as follows:

* Proportion of died or transfer among LTFU. The proportion of patients LTFU having transferred to another site is based on a meta-analysis on LTFU among patients in ART programs in sub-Saharan Africa; the study found that 20.8% and 35.9% of patients recorded as LTFU had died or self-transferred to another ART clinic.56

**Table 15:** **Monthly probability of sub-population virologic failure**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sub-population | On ART & not suppressed, Linkage | | | |
| CD4, >500 | CD4, 350-500 | CD4, 200-350 | CD4, < 200 |
| Low-risk urban women (15-24) |  |  |  |  |
| Low-risk urban women (25-34) |  |  |  |  |
| Low-risk urban women (35-44) |  |  |  |  |
| Low-risk urban women (45-54) |  |  |  |  |
| Low-risk urban women (55-64) |  |  |  |  |
| Low-risk rural women (15-24) |  |  |  |  |
| Low-risk rural women (25-34) |  |  |  |  |
| Low-risk rural women (35-44) |  |  |  |  |
| Low-risk rural women (45-54) |  |  |  |  |
| Low-risk rural women (55-64) |  |  |  |  |
| Urban men (15-24) |  |  |  |  |
| Urban men (25-34) |  |  |  |  |
| Urban men (35-44) |  |  |  |  |
| Urban men (45-54) |  |  |  |  |
| Urban men (55-64) |  |  |  |  |
| Rural men (15-24) |  |  |  |  |
| Rural men (25-34) |  |  |  |  |
| Rural men (35-44) |  |  |  |  |
| Rural men (45-54) |  |  |  |  |
| Rural men (55-64) |  |  |  |  |
| High risk women (15-24) |  |  |  |  |
| High risk women (25-34) |  |  |  |  |
| High risk women (35-44) |  |  |  |  |
| High risk women (45-54) |  |  |  |  |
| High risk women (55-64) |  |  |  |  |

## Death

Patients are classified as death when marked dead in the IeDEA data. The parameter estimates for probability of death will apply to all compartments in infectious stage, including Undiagnosed; Diagnosed; Linkage to care; Loss-to follow up (LTFU); On ART and suppressed; and On ART and not suppressed.

***Data source.***

* To estimate probability of death, we will use the International epidemiology databases to evaluate AIDS (IeDEA) consortium in Rwanda from 2004 to most current data.51
* To adjust for the misclassification of loss to follow-up (LTFU) in the IeDEA data, a retrospective study from Kenya (2009-2011) is used, in which the proportion of patients who have died but coded as LTFU is 17%.54
* For patients on ART, the weight applied to adjust for the misclassification of patients is extracted from previous study examining LTFU using IeDEA data, which is calculated based on a meta-analysis in sub-Saharan area. The proportion of patients LTFU having died is 20.8%.56

***Assumptions.***

* We assume a linear trend in the decrease of CD4 level for individuals living with HIV.
* We assume that the probability of death for undiagnosed, diagnosed, and LTFU compartments are similar to the probability of death for linked due to the data limitation that IeDEA data is not feasible to examine individuals who are not linked to care. In sensitivity analysis, a multiplier (>1) might be applied to account for increased probability of death when not linked.

***Estimation.***

The probability of death will be estimated using a weighed cox proportional regression model for each sub-population (**Table 16)**. The model is selected against the Kaplan-Meier model because cox proportional hazard regression model allows us to control for the clinic- or site-level predictors. The weighted cox proportional regression model is presented as follows.

(16)

Where,w = weight to account for misclassification of death, *h* = hazard, = weighted probability of death at time, ***X*** = vector of predictor variables (e.g., site) for patient , = time-dependent hazard function for death.

In the cox proportional regression model, we will calculate the time-at-risk for each individual to estimate the hazard function of death. The process for estimating the time-at-risk can be found in **Appendix A3.** We will control for site-level factors only (i.e. clinics) since clinic- or site-level factors can be associated with death (e.g. nutrition support programs). We did not control for demographic characteristics or disease stage since the model is applied to each sub-population by CD4 stratum. The variables created for the model and IeDEA data used is presented in **Table 13.**

The statistical model will also adjust for misclassification of death. Evidence from sub-Saharan Africa suggests that some patients recorded as LTFU had died or transferred to another clinic.56,63–66 When estimating the probability of death, we will adjust for the estimates among patients misclassified as LTFU using probability weights.56 The probability weight is calculated as:

(17)

We parameterize this expression as follows:

* Proportion of death among LTFU. The proportion of patients LTFU but who have died is based on a retrospective study from Kenya that evaluated outcomes of people living with HIV in the general population and LTFU (data from 2009-2011); the study found that 17% of pre-ART patients classified as LTFU had died.54
* Proportion of died among LTFU when on ART. The proportion of patients LTFU having transferred to another site is based on a meta-analysis on LTFU among patients in ART programs in sub-Saharan Africa; the study found that 20.8% of patients recorded as LTFU had died.56

**Table 16:** **Monthly probability of sub-population death, by CD4 stratum and engagement in care**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sub-population | Linkeda | | | | On ART & suppressed | | | | On ART & not suppressed | | | |
| CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 | CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 | CD4,  >500 | CD4, 350-500 | CD4, 200-350 | CD4,  < 200 |
| Low-risk urban women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk urban women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Low-risk rural women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Urban men (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| Rural men (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (15-24) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (25-34) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (35-44) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (45-54) |  |  |  |  |  |  |  |  |  |  |  |  |
| High risk women (55-64) |  |  |  |  |  |  |  |  |  |  |  |  |

1. The probability of death for individuals undiagnosed, diagnosed, and LTFU will be the same as the probability of death for individuals linked to care. In the sensitivity analysis, a multiplier will be applied to capture the difference in probability of death due to medical care (e.g., nutrition support)

# Calibration approach

External model calibration is a process to identify unobserved parameters or to inform parameter inputs with significant uncertainty so that the model outputs are reasonably consistent with data from other sources (external data) which is not used to parameterize the model.69

## Calibration targets

We identified 6 calibration targets based on the surveys and Rwanda national reports. The targets include: 1) HIV prevalence; 2) HIV incidence; 3) number of individuals on ART; 4) percentage of individuals diagnosed with HIV on ART; 5) percent virally suppressed; and 6) percent virally suppressed conditional on ART. The calibration targets are presented below.

***HIV prevalence***

*Data source:*The HIV prevalence data is derived from Demographic Health Survey in Rwanda in years 2005, 2010, and 2015 by linking DHS HIV diagnosis data with individual level characteristics estimating the proportion of those diagnosed with HIV by age group, sex and urbanicity.22,29,43 In year 2019, the HIV prevalence data for the overall population is derived from Rwanda Population-based HIV Impact Assessment.19 For female sex worker, the HIV prevalence data comes from the Behavioral Surveillance Survey in year 2010, 2015.26,44**(Table 17)**

***HIV incidence***

*Data source:* The HIV incidence data is derived from Rwanda AIDS Indicator and HIV Incidence Survey (2013) and Rwanda Population-based HIV Impact Assessment in year 2019 for the overall population.19,70 HIV incidence by sub-groups (i.e. sex-disaggregated HIV incidence) will not be used as a calibration target since the survey is not powered to generate such estimates. **(Table 18)**

***Number of individuals on ART***

*Data source:* The number of individuals on antiretroviral therapy is collected by TRACnet or Rwanda Health Management Information System.28,71 The calibration target is critical at population-level. **(Table 19)**

***Percent of individuals on antiretroviral therapy (ART)***

*Data source:* The percent on treatment is defined as the proportion of population diagnosed with HIV who is on ART, which is derived from Rwanda Population-based HIV Impact Assessment in year 2019 for the overall population.19 The percent of population on ART among those diagnosed with HIV is one of the major indicators for HIV care continuum. **(Table 20)**

***Percent viral suppression (conditional or not conditional on ART)***

*Data source:* The percent viral suppression is defined in two ways: 1) the proportion of population living with HIV that are virally suppressed regardless of HIV status diagnosis and treatment status; and 2) the proportion of population who are on antiretroviral therapy that are virally suppressed. Both indicators were extracted from Rwanda Population-based HIV Impact Assessment in year 2019.19 The viral suppression status is both one of the major indicators of HIV care continuum and the one of the goals of providing HIV treatment. When viral suppressed, population living with HIV have much lower risk of HIV transmission and improved health outcomes such as life-expectancy. **(Table 20)**

**Table 17.** **Calibration targets: Sub-population HIV prevalence over time**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sub-population** | **Age Group** | **2005a** | | | **2010a** | | | **2015a** | | | **2019b** | | |
| **Sample size** | **Prevalence (95% CI)** | **Sample size** | | **Prevalence (95% CI)** | **Sample size** | | **Prevalence (95% CI)** | **Sample size** | | **Prevalence (95% CI)** |
| Urban men | 15-24 | 340 | 1.10 (0.39 – 3.01) | 387 | | 1.57 (0.70 – 3.44) | 454 | | 1.79 (0.90 – 3.53) | — | | — |
| 25-34 | 217 | 7.33 (4.52 – 11.69) | 352 | | 4.72 (2.93 – 7.51) | 434 | | 3.49 (2.12 – 5.70) | — | | — |
| 35-44 | 138 | 12.35 (7.77 – 19.07) | 127 | | 12.98 (8.11 – 20.11) | 212 | | 8.36 (5.29 – 12.97) | — | | — |
| 45-54 | 86 | 7.98 (3.77 – 16.09) | 96 | | 13.79 (8.15 – 22.40) | 105 | | 18.60 (12.18 – 27.35) | — | | — |
| 55+ | 11 | 0.00 (0.00 – 0.00) | 25 | | 0.00 (0.00 – 0.00) | 23 | | 11.67 (3.26 – 34.13) | — | | — |
| 15 – 59 | 793 | 5.49 (4.11 – 7.32) | 987 | | 5.31 (4.07 – 6.90) | 1,229 | | 5.15 (4.04 – 6.53) | — | | — |
| Rural men | 15-24 | 1,721 | 0.27 (0.11 – 0.67) | 2,261 | | 0.24 (0.11 – 0.56) | 1,880 | | 0.57 (0.31 – 1.03) | — | | — |
| 25-34 | 893 | 1.62 (0.97 – 2.69) | 1,385 | | 1.88 (1.23 – 2.75) | 1,427 | | 1.43 (0.93 – 2.20) | — | | — |
| 35-44 | 696 | 3.61 (2.45 – 5.29) | 791 | | 4.31 (3.09 – 5.97) | 825 | | 2.22 (1.39 – 3.47) | — | | — |
| 45-54 | 547 | 2.79 (1.69 – 4.57) | 696 | | 4.31 (3.03 – 6.10) | 628 | | 5.65 (4.09 – 7.75) | — | | — |
| 55+ | 86 | 0.00 (0.00 – 0.00) | 183 | | 1.72 (0.56 – 5.12) | 203 | | 2.22 (0.88 – 5.51) | — | | — |
| 15 – 59 | 3,942 | 1.51 (1.17 – 1.94) | 5,315 | | 1.86 (1.53 – 2.26) | 4,963 | | 1.79 (1.46 – 2.21) | — | | — |
| Men | 15 – 49 | — | — | — | | — | — | | — | 12,167 | | 1.8 (1.5 – 2.1) |
| 15 – 64 | — | — | — | | — | — | | — | 13,780 | | 2.2 (1.9 – 2.6) |
| Low-risk urban women | 15-24 | 435 | 4.04 (2.54 – 6.37) | 472 | | 3.70 (2.32 – 5.85) | 520 | | 2.74 (1.64 – 4.57) | — | | — |
| 25-34 | 300 | 9.43 (6.59 – 13.32) | 335 | | 7.45 (5.07 – 10.81) | 466 | | 9.59 (7.22 – 12.63) | — | | — |
| 35-44 | 164 | 19.13 (13.78 – 25.95) | 178 | | 20.75 (15.38 – 27.39) | 222 | | 13.88 (9.89 – 19.12) | — | | — |
| 45-54 | 43 | 7.57 (2.46 – 21.02) | 63 | | 17.75 (10.02 – 29.49) | 66 | | 13.87 (7.28 – 24.82) | — | | — |
| 55+ | — | — | — | | — | — | | — | — | | — |
| 15 – 49 | 942 | 8.55 (6.92 – 10.52) | 1,049 | | 8.65 (7.09 – 10.51) | 1,273 | | 7.77 (6.42 – 9.37) | — | | — |
| Low-risk rural women | 15-24 | 2, 037 | 0.99 (0.64 – 1.53) | 2,446 | | 1.11 (0.77 – 1.62) | 2,068 | | 0.98 (0.64 – 1.51) | — | | — |
| 25-34 | 1,345 | 3.49 (2.63 – 4.62) | 1,825 | | 3.41 (2.66 – 4.34) | 1,795 | | 2.77 (2.10 – 3.64) | — | | — |
| 35-44 | 931 | 4.46 (3.30 – 5.99) | 1,156 | | 4.97 (3.85 – 6.38) | 1,197 | | 4.83 (3.75 – 6.20) | — | | — |
| 45-54 | 426 | 3.58 (2.17 – 5.84) | 473 | | 4.21 (2.73 – 6.45) | 429 | | 4.15 (2.62 – 6.52) | — | | — |
| 55+ | — | — | — | | — | — | | — | — | | — |
| 15 – 49 | 4,739 | 2.61 (2.19 – 3.12) | 5,900 | | 2.83 (2.43 – 3.28) | 5,489 | | 2.65 (2.26 – 3.11) | — | | — |
| High-risk womenc | 15-24 | — | — |  | | 42.1 | — | | — | — | | — |
| 25-34 | — | — |  | | 55.6 | — | | — | — | | — |
| 35-44 | — | — |  | | 63 | — | | — | — | | — |
| 45-54 | — | — |  | | (63) | — | | — | — | | — |
| 55+ | — | — |  | | (63) | — | | — | — | | — |
| 15 + | — | — | 1112 | | 51 | 1967 | | 45.8 | — | | — |
| Women | 15 – 49 | — | — | — | | — | — | | — | 14,659 | | 3.3 (2.9 – 3.8) |
| 15 – 64 | — | — | — | | — | — | | — | 16, 857 | | 3.7 (3.3 – 4.1) |
| Overall | 15 – 49 | 10,416 | 2.95 (2.64 – 3.29) | 13,251 | | 3.08 (2.80 – 3.39) | 12,955 | | 3.07 (2.78 – 3.38) | 26,826 | | 2.6 (2.3 - 2.9) |
| 15 – 64 | — | — | — | | — | — | | — | 30,637 | | 3.0 (2.7 - 3.3) |
| a Data source: Demographic Health Survey, 2005, 2010, and 201518,20,22 | | | | | | | | | | | | | | |
| b Data source: Population-based HIV Impact Assessment Survey, 201919 | | | | | | | | | | | | | | |
| c Data source: Behavioral Surveillance Survey among Female Sex Workers, 2010 and 201526,44 | | | | | | | | | | | | | | |

**Table 18:** **Calibration targets: Sub-population HIV incidence over time**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Age Group** | **2014a** | | **2019b** | |
| **Sub-population** | **Sample size** | **Incidence (%, 95% CI)** | **Sample size** | **Incidence (%, 95% CI)** |
| Men | 15 – 49 | — | — | — | 0.10 (0.00 - 0.20) |
| 15 – 64 | — | — | — | 0.09 (0.00 - 0.17) |
| Women | 15 – 49 | — | — | — | 0.06 (0.00 - 0.13) |
| 15 – 64 | — | — | — | 0.07 (0.00 - 0.15) |
| Overall | 15 – 49 | — | — | — | 0.08 (0.02-0.14) |
| 15 – 64 | >13000 | 0.27 | — | 0.08 (0.02-0.14) |

1. The HIV incidence estimation comes from Rwanda AIDS Indicator and HIV Incidence Survey 2013.70 The HIV incidence is reported in Rwanda Annual Report for HIV. Micro-data is not available.
2. The HIV incidence estimation comes from Population-based HIV Impact Assessment Survey 2019.19 The estimates by sex are based on a small number of new infections and may be interpreted with caution due to the lack of power for sex-disaggregated estimates of the survey.

**Table 19:** **Calibration targets: Number of people on ART over time**

|  |  |
| --- | --- |
| **Yeara** | **# of people living with HIV on ART** |
| 2004 | 6,356 |
| 2005 | 19,058 |
| 2006 | 34,136 |
| 2007 | 48,069 |
| 2008 | 63,149 |
| 2009 | 76,726 |
| 2010 | 83,041 |
| 2011 | 96,123 |
| 2012 | 107,938 |
| 2013 | 123,499 |
| 2014 | 130,406 |
| 2015 | — |
| 2016 | 166,818 |
| 2017 | 175,359 |
| 2018 | 182,987 |

1. From 2004-2013, the numbers of people living with HIV on ART come from TRACnet, a national phone-based and internet-based HIV reporting system in Rwanda.71 The data is reported in Rwanda Annual Report for HIV. From 2014-2018, the numbers come from Rwanda Health Information System and are reported in Rwanda Country Operational Plan (COP).28

**Table 20:** **Calibration targets: Percentage at Different Steps Along the HIV Care Continuum based on 2018-2019 Rwanda Population-based HIV Impact Assessment Survey**19

|  |  |  |  |
| --- | --- | --- | --- |
| Domain | Female  Estimate (95% CI) | Male  Estimate (95% CI) | Total  Estimate (95% CI) |
| **Diagnosed (%)** |  |  |  |
| Ages 15-64 years | 85.6 | 80.4 | 83.8 |
| **On Treatment (%)a** |  |  |  |
| Ages 15-64 years | 97.6 | 97.2 | 97.5 |
| **Viral Load Suppression (%) b** |  |  |  |
| Ages 15-24 years | 62.3 (50 – 75) | 55.9 (35 – 75) | 60.6 (50 – 70) |
| Ages 25-34 years | 73.5 (65 – 80) | 45.9 (30 – 60) | 66.2 (55 – 75) |
| Ages 35-44 years | 85.2 (75 – 90) | 75.0 (65 – 85) | 81.4 (76 – 86) |
| Ages 45-54 years | 84.8 (74 – 94) | 75.7 (65 – 85) | 81.1 (74 – 88) |
| Ages 55-64 years | 81.7 (71 – 91) | 84.9 (72 – 98) | 83.1 (75 – 92) |
| Ages 15-49 years | 78.6 (74.2 – 83.0) | 65.7 (57.2 – 74.2) | 74.3 (69.7 – 78.8) |
| Ages 15-64 years | 79.1 (74.9 – 83.2) | 70.5 (63.8 – 77.2) | 76.0 (72.0 – 80.0) |
| **Viral Load Suppression conditional on ART (%)** | | | |
| Ages 15-64 years | 92.4 | 85.4 | 90.1 |

1. On treatment is defined as proportion of people diagnosed with HIV that were on antiretroviral therapy.
2. Viral load suppression is defined as HIV ribonucleic acid (RNA) <1,000 copies per milliliter (mL) of plasma among HIV-positive individuals regardless of knowledge of HIV status and treatment stat. Confidence bounds for age groups were not reported in the report. These bounds were visually estimated from error bars in the graph.

## Approach: Parameters used in the calibration process

We will use all parameters to estimate through calibration process since all parameters will contribute to the outcomes of the model. For each sub-group, we will have approximately 60 parameter inputs to calibrate and we will have 25 sub-groups, suggesting that the number of parameter inputs to calibrate could be more than 1500. An alternative option is that we will use parameters that we have less confidence in when calibrating the model (e.g. parameters extracted from literature) so that the calibration process is less computational intense.

## Approach: Goodness of fit

The goodness of fit is to evaluate how close the model predictions are to the target data. In statistics, the most commonly used measure for goodness of fit are least squares, weighted least squares, chi-square, the likelihood method, and multiple goodness of fit estimation approaches.

In our analysis, we will use weighted sum of least squares and weighted percentage deviation to measure the goodness of fit for multiple targets similar to other mathematical modeling studies.72,73 The methods are intuitive and empirically feasible given the calibration targets we have since we have limited micro-level data and lack variation in some targets. We will assume equal weight in the initial attempts but apply weights to adjust difference in the importance of each targets. The comparison between methods can be found in **Section 11.6**.

## Approach: Method for identifying appropriate values for parameters inputs

We will search for parameters or a set of parameter values that produce model outputs that match specified calibration targets from external data most closely. A number of alternative strategies that can be used to search for parameters, which can be classified according to the existence of constraints, the nature of the design variables, the physical structure of the problem, the nature of the equations involved, the deterministic nature of the variables, the number of objective functions, and others.

In our analysis, we will use mixed approaches that includes random search and grid search approaches for out model in the external calibration process.Random search method, which assigns a distribution to each parameter input, is better since it is intuitive and commonly found in literatures that calibrate epidemiological model. Grid search will be used for the parameters describing initial distribution of population across CD4 stratum since the probabilities are correlated with each other (i.e., sum up to 1) and it is difficult to find a joint probability distribution to describe the pattern of probabilities in random search. Grid search will not be used for other parameter inputs since the number of combinations of parameter values increases quickly with more parameter inputs. Given that we have thousands of parameter inputs, the calibration process will become extremely computational intense.

Other methods have major limitations in identifying appropriate values for parameter inputs. The Bayesian calibration method might not be applicable given the limited information on our calibration targets identified. The optimization methods (i.e. reduced gradient method, Nelder-Mead method, Simulated Annealing method) are not used since they generate only one set of parameters that best-fit the objective function(s) based on the initial value chosen to search for the optimum. In our calibration process, we would like to identify multiple sets of parameters that can fit our model. In addition, the optimization methods are seen in engineering modeling but less commonly used in calibrating epidemiological models. The comparison between different methods to identify parameter values are presented in **Section 11.7**.

## Approach: Defining convergence criteria

Convergence criteria describes the process of defining acceptable sets of input parameter values. The possible approaches to define convergence criteria include following.

* Optimization of the GOF. The predicted output will minimize (or maximize) the goodness of fit measures so that the set of parameters are optimal. The method ignores the potential uncertainty of the parameters.
* Visual observe the model fit. The predicted output parameters of many parameter sets are plotted and the worst fitting set that is acceptable is identified.
* Targeted ranges. The targeted ranges are set based on the data informing the calibration target(s) and select those parameter sets that produce model output within those ranges.
* Define confidence interval around the GOF. Confidence interval will be defined around the goodness of fit of the best-fit parameter set and to deem acceptable all the parameter sets with GOF estimates within that interval.

I would propose to use the combination of optimization of goodness of fit method and the targeted ranges method given that not all of our calibration targets would have a range or confidence interval. The acceptable sets of input parameter values will include the sets of parameter inputs that produces outputs within the range of calibration targets and minimize the GOF when calibration targets are not available.

## Approach: Stopping rule (i.e. when calibration process is complete)

The stopping rule determines whether the calibration process is complete. There are two broad criteria that can be used:

* Acceptability of the convergence of the model outputs to the observed calibration targets, and/or
* Completion of a specified number of searches (or iterations within the parameter space).

Given the convergence criterion chosen, I would propose to use the Completion of a specified number of searches. We will complete a specific number of searches. Among the outputs generates, the first 100 set of parameter inputs that have lowest sum of least squares will be identified.

## Calibration approach: Probabilistic distributions for model parameters

As we have decided to use a random search approach to identify parameter inputs, we will estimate distributions to each parameter input based on the information collected from the survey or literature. The assigned distribution as well as the justification and the data availability for each parameter input of the model is described in **Section 11.8.**

***Probability of HIV diagnosis***

We assigned a beta distribution to the parameter inputs for probability of HIV diagnosis in the Rwanda model with a range from 0 to 1.

*Low-risk population:* Since the original data is a survey data that includes weighting, we cannot estimate the beta distribution simply by multiplying sample size and the number of HIV diagnosis. Instead, we will use the weighted mean and standard error to estimate the shape ( and scale parameters for the beta distribution. The annual probability will be transformed to monthly probability before we estimate the distribution.

Since the mean and standard error of the beta distribution can be calculated as follows.

Thus, the parameters for the beta distribution can be estimates as:

*High-risk population:*  No original data is available for high-risk population. According to the report where we extract data, the sample size and the proportion of high-risk population who received HIV testing in past 12 month is available. We will first estimate the shape and scale parameter for the distribution of annual probability assume that the annual probability of HIV testing follows beta distribution as well and estimate the standard error for annual probability of HIV testing using the equation (2) above. Then, the annual probability and standard deviation will be converted to monthly probability and standard error to estimate the shape ( and scale parameters for the beta distribution assigned to monthly probabilities.

*Population with CD4 <200:* For individuals with CD4 < 200, the probability of HIV diagnosis is calculated based on the baseline probability of HIV diagnosis and a multiplier extracted from literature. Since the multiplier comes from cox proportional hazard regression, the logarithm of the multiplier follows a normal distribution, which suggests that the multiplier follows a log-normal distribution. With the baseline value and confidence interval, we can calculate the mean and standard error for the log-normal distribution.

Note: We will calibrate the baseline probability of HIV diagnosis and the multiplier so that the calibration process is less computational intense. However, it is possible that a single multiplier does not apply to all parameter inputs. Thus, we may further calibrate the parameters independently based on the calculated baseline and confidence interval independently.

The estimates are presented in **Table 21** and **Table 22**. The process for calculating parameters for beta distributions in **Table 21** is stored in [the excel.](file:///Users/portia/Dropbox/VCU_PhD_Year%203&4/GRA/Rwanda-Model/Calibration/Calibration%20documentation/Beta%20distribution-probability%20of%20HIV%20testing_new.xlsx)

***Probability of link to care***

We assigned a beta distribution to the parameter inputs for probability of linkage to care in the Rwanda model with a range from 0 to 1.

*Low-risk /high-risk population:* No original data is available for linkage to care. According to the report where we extract data, the sample size and the proportion of population who linked to care is available. Without further information, we will estimate the beta distribution parameters for the annual probability of linkage to care. The standard error will then be estimated based on sample size and the probability of linkage to care. The annual probabilities and standard error will be converted to monthly probability and standard error to estimate the shape ( and scale parameters for the beta distribution assigned to monthly probabilities. **(Table 23)**

*Population with CD4 <200:* For individuals with CD4 < 200, the probability of linkage is calculated based on the baseline probability of HIV diagnosis and a multiplier extracted from literature. Since the multiplier comes from cox proportional hazard regression, the logarithm of the multiplier follows a normal distribution, which suggests that the multiplier follows a log-normal distribution. With the baseline value and confidence interval, we can calculate the mean and standard error for the log-normal distribution. (**Table 22**)

Note: We will calibrate the baseline probability of HIV diagnosis and the multiplier so that the calibration process is less computational intense. However, it is possible that a single multiplier does not apply to all parameter inputs. Thus, we may further calibrate the parameters independently based on the calculated baseline and confidence interval independently.

***Probability of HIV transmission per sex act***

The original data source for extracting probability of HIV transmission per sex act shows that the probability is estimated based on the assumption that the logarithm of the probability is normally distributed. Thus, we assume that it follows a log-normal distribution.

*Low-risk/high-risk population*: For each sub-population, the data has presented the baseline probability of HIV transmission and the 95% confidence interval. According to the methods supplement, the reported estimates will be log-transformed to derive the pooled estimated as well as the 95% confidence interval. The mean and standard error can be calculated and used in the log-normal distribution for the probability of HIV transmission. **(Table 24)**.

***Proportion of condom use***

We assigned a beta distribution to the parameter inputs for proportion of condom use in the Rwanda model with a range from 0 to 1.

*Low-risk population:* Since the original data is a survey data that includes weighting, we cannot estimate the beta distribution simply by sample size and the number of consistent condom use. Instead, we will use the weighted mean and standard error to estimate the shape ( and scale parameters for the beta distribution. **(Table 25)**

Since the mean and standard deviation of the beta distribution is calculated using the same function as for HIV diagnosis.

*High-risk population:*  No original data is available for high-risk population. According to the report where we extract data, the sample size and the proportion of high-risk population who consistently used condom is available. Since we assume consistent proportion of consistent condom use over year. The annual proportion will be used as an approximation for the monthly proportion of population consistently used condom. The sample size and the proportion of high-risk population who consistently used condom will be used to estimate for the beta distribution. (**Table 25)**

The process for calculating parameters for beta distributions in **Table 25** is stored in the [excel.](file:///Users/portia/Dropbox/VCU_PhD_Year%203&4/GRA/Rwanda-Model/Calibration/Calibration%20documentation/Beta%20distribution%20for%20condom%20use.xlsx)

**Table 21.** **Calibration: Parameter estimates for beta distribution assigned to probabilities of HIV diagnosis**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sub-group** | **Year of data collection** | **Age group (Years)** | **Monthly probabilities (SE)** | | **Parameters for beta distribution (** | |
| **CD4>500, >350–500, and >200–350** | **CD4≤200** | **CD4>500, >350–500, and >200–350** | **CD4≤200** |
| Low-risk women | 2005 | 15-24 | 0.0163 (0.0059) | 0.0243 (0.0089) | (7.43, 448.48) | (7.32, 293.45) |
| 25-34 | 0.0266 (0.0051) | 0.0396 (0.0076) | (26.60, 974.38) | (25.95, 629.37) |
| 35+ | 0.0274 (0.0045) | 0.0408 (0.0068) | (35.89, 1273.64) | (34.98, 821.66) |
| 2010 | 15-24 | 0.1446 (0.0087) | 0.2089 (0.0130) | (236.99, 1401.57) | (204.03, 772.64) |
| 25-34 | 0.1460 (0.0068) | 0.2109 (0.0102) | (396.00, 2315.55) | (340.11, 1272.89) |
| 35+ | 0.1112 (0.0064) | 0.1621 (0.0096) | (265.06, 2117.76) | (236.62, 1222.84) |
| 2015 | 15-24 | 0.1255 (0.0078) | 0.1822 (0.0117) | (224.03, 1561.39) | (196.99, 884.27) |
| 25-34 | 0.1327 (0.0059) | 0.1923 (0.0089) | (436.71, 2853.76) | (380.60, 1598.39) |
| 35+ | 0.1147 (0.0072) | 0.1670 (0.0108) | (223.19, 1722.77) | (198.55, 990.30) |
| Men | 2005 | 15-34 | 0.0180 (0.0059) | 0.0269 (0.0089) | (9.09, 495.51) | (8.94, 323.56) |
|  | 35+ | 0.0214 (0.0049) | 0.0319 (0.0074) | (18.43, 843.77) | (18.07, 548.56) |
| 2010 | 15-34 | 0.0711 (0.0138) | 0.1047 (0.0207) | (24.41, 318.92) | (22.81, 194.97) |
|  | 35+ | 0.1392 (0.0067) | 0.2013 (0.0100) | (372.25, 2302.73) | (322.20, 1278.37) |
| 2015 | 15-34 | 0.0683 (0.0087) | 0.1007 (0.0131) | (56.95, 776.70) | (53.28, 475.84) |
|  | 35+ | 0.1463 (0.0045) | 0.2113 (0.0068) | (888.01, 5180.73) | (761.69, 2843.94) |
| High-risk urban women | 2010 | 15-55+ | 0.154 (0.077 – 0.231) | 0.222 (0.111–0.333) | (31967.81, 175326.00) | (27127.63, 94980.61) |

**Table 22.** **Calibration: Parameter estimates for multiplier applied to probability of HIV diagnosis and linkage to care**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sub-population** | **Multiplier (95% CI)** | **Log-transformation** | **Log-normal distribution**  **(mean, standard error)** |
| Any | 1.56 (1.11 – 2.20) | 0.445 (0.104, 0.788) | (0.445, 0.175) |

**Table 23.** **Calibration: Parameter estimates for beta distribution assigned to probabilities of linkage to care**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sub-population** | **Year of data collection** | **Monthly probabilities (CIs)** | | **Parameters for beta distribution (** | |
| **CD4>500, >350–500, and >200–350** | **CD4≤200** | **CD4>500, >350–500, and >200–350** | **CD4≤200** |
| Low-risk women and men | 2013 | 0.175 (0.092 – 0.250) | 0.259 (0.139 – 0.362) | (7699609.36, 37156115.76) | (6401713.31, 19611835.49) |
| High-risk urban women | 2007-2008 | 0.1153 (0.059 – 0.168) | 0.1678(0.091 – 0.249) | (1313.32, 10080.12) | (1165.43, 5779.00) |

**Table 24.** **Calibration: Parameter estimates for distributions assigned to the probabilities of HIV transmission**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sub-population** | **Direction** | **Probabilities (95% CI)** | **Log-transformation** | **Log-normal distribution**  **(mean, standard error)** |
| High-risk women | Male-to-female | 0.0005 (0.0002 – 0.0013) | -7.600 (-8.517, -6.645) | (-7.600, 0.478) |
| Low-risk women | Male-to-female | 0.0030 (0.0014 – 0.0063) | -5.809 (-6.571, -5.067) | (-5.809, 0.384) |
| Men | Female-to-male  (Pooled estimate for all men) | 0.0087 (0.0028 – 0.0270) | -4.744 (-5.878, -3.612) | (-4.744, 0.578) |

**Table 25.** **Calibration: Parameter estimates for beta distribution assigned to proportion of condom use**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sub-group** | **Year of data collection** | **Age group (Years)** | **Sample sizea** | **Proportion of consistent condom use (SE)** | **Parameters for beta distribution (** |
| Low-risk urban women | 2005 | 15-24 | 175 | 0.1392 (0.0262) | (24.16, 149.40) |
| 25+b | 598 | 0.0445 (0.0084) | (26.77, 574.83) |
| 2010 | 15-24 | 243 | 0.1651 (0.0239) | (39.68, 200.64) |
| 25+b | 814 | 0.0660 (0.0087) | (53.69, 759.74) |
| 2015 | 15-24 | 335 | 0.2656 (0.0242) | (88.20, 243.87) |
| 25+b | 1087 | 0.0903 (0.0085) | (316.21, 591.92) |
| Low-risk rural women | 2005 | 15-24 | 952 | 0.0190 (0.0044) | (18.27, 943.48) |
| 25+b | 3531 | 0.0088 (0.0016) | (29.98, 3376.28) |
| 2010 | 15-24 | 1150 | 0.0659 (0.0073) | (76.06, 1078.08) |
| 25+b | 5229 | 0.0324 (0.0024) | (176.31, 5265.44) |
| 2015 | 15-24 | 1087 | 0.0978 (0.0090) | (106.44, 981.88) |
| 25+b | 5217 | 0.0392 (0.0027) | (202.49, 4962.96) |
| Urban men | 2005 | 15-24 | 63 | 0.4074 (0.0626) | (49.37, 302.26) |
| 25+b | 349 | 0.1199 (0.0174) | (65.23, 171.27) |
| 2010 | 15-24 | 88 | 0.5124 (0.0536) | (118.03, 272.68) |
| 25+b | 511 | 0.1213 (0.0145) | (161.49, 195.64) |
| 2015 | 15-24 | 111 | 0.6218 (0.0463) | (150.29, 302.65) |
| 25+b | 647 | 0.1329 (0.0134) | (212.49, 227.36) |
| Rural men | 2005 | 15-24 | 253 | 0.0892 (0.0180) | (22.28, 227.47) |
| 25+b | 1798 | 0.0136 (0.0.0027) | (25.01, 1814.18) |
| 2010 | 15-24 | 404 | 0.2101 (0.0203) | (84.40, 317.32) |
| 25+b | 2766 | 0.0406 (0.0038) | (109.48, 2587.00) |
| 2015 | 15-24 | 343 | 0.2752 (0.0242) | (93.46, 246.14) |
| 25+b | 2800 | 0.0497 (0.0041) | (139.59, 2669.04) |
| High-risk urban women | 2010 | 15-24c | 612 | 0.351 (0.298 – 0.404) | (214.81, 397.19) |
|  | 25+b,c | 689 | 0.276 (0.235 – 0.317) | (190.16, 498.84) |

# Implementation (in R)

The model is implemented using R software. Sample code (v5.2, May 22nd, 2021) for operationalize and calibrating the model is attached below.

**R code:**

# CODE FOR A DYNAMIC MODEL OF HIV TRANSMISSION FOR RWANDA #  
  
# # START OF THE CODE # #  
  
# Clean global environment (clear previous data and results)  
rm(list=ls())  
  
# Specify the number of trials that alter parameter values   
trials <-1  
# Specify the number of sub-groups we are going to run  
sgroup <-25  
# Command specifing the number of monthly time periods in the model (432 monthly time periods from 2004 - 2040)  
generations <- 432  
  
# This command sets up variables to store the value for the force of infection  
# Note that the initial value of lambda shall set to 0 otherwise error may occur during operation

lambda\_t\_r <- array(0,dim=c(trials, sgroup, generations))

# This command sets up variables to store the value for each compartment

# The initial value is set as null.

S <- array(NA,dim=c(trials, sgroup, generations))

S\_y <- array(NA,dim=c(trials, sgroup, generations/12))

S\_total <- array(0,dim=c(trials, 1, generations))

S\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

D <- array(NA,dim=c(trials, sgroup, generations))

X1 <- array(NA,dim=c(trials, sgroup, generations))

X2 <- array(NA,dim=c(trials, sgroup, generations))

X3 <- array(NA,dim=c(trials, sgroup, generations))

X4 <- array(NA,dim=c(trials, sgroup, generations))

X5 <- array(NA,dim=c(trials, sgroup, generations))

X6 <- array(NA,dim=c(trials, sgroup, generations))

X7 <- array(NA,dim=c(trials, sgroup, generations))

X8 <- array(NA,dim=c(trials, sgroup, generations))

X9 <- array(NA,dim=c(trials, sgroup, generations))

X10 <- array(NA,dim=c(trials, sgroup, generations))

X11 <- array(NA,dim=c(trials, sgroup, generations))

X12 <- array(NA,dim=c(trials, sgroup, generations))

X13 <- array(NA,dim=c(trials, sgroup, generations))

X14 <- array(NA,dim=c(trials, sgroup, generations))

X15 <- array(NA,dim=c(trials, sgroup, generations))

X16 <- array(NA,dim=c(trials, sgroup, generations))

X17 <- array(NA,dim=c(trials, sgroup, generations))

X18 <- array(NA,dim=c(trials, sgroup, generations))

X19 <- array(NA,dim=c(trials, sgroup, generations))

X20 <- array(NA,dim=c(trials, sgroup, generations))

X21 <- array(NA,dim=c(trials, sgroup, generations))

X22 <- array(NA,dim=c(trials, sgroup, generations))

X23 <- array(NA,dim=c(trials, sgroup, generations))

X24 <- array(NA,dim=c(trials, sgroup, generations))

# This command sets up variables to store the number of individuals infected.

N <- array(0,dim=c(trials, sgroup, generations))

N\_y <- array(0,dim=c(trials, sgroup, generations/12))

N\_total <- array(0,dim=c(trials, 1, generations))

N\_total\_y <- matrix(0,nrow=trials, ncol=generations/12)

#This command sets up variables to store the number ofindividuals virally not suppresed.

Z <- array(0,dim=c(trials, sgroup, generations))

Z\_total <- array(0,dim=c(trials, 1, generations)

#This command sets up variables to store the number of individuals virally suppressed.

V <- array(0,dim=c(trials, sgroup, generations))

V\_total <- array(0,dim=c(trials, 1, generations))

# This command sets up variables to store the percent of individuals virally not suppressed.

p\_r <- array(0,dim=c(trials, sgroup, generations)

# This command sets up variables to store the percent of individuals virally suppressed.

p\_r\_v <- array(0,dim=c(trials, sgroup, generations))

# This command sets up variables to store HIV treatment engagement status in the cohort.

Undiagnosed <- array(NA,dim=c(trials, sgroup, generations))

Undiagnosed\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Undiagnosed\_total <- array(0,dim=c(trials, 1, generations))

Undiagnosed\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Diagnosed <- array(NA,dim=c(trials, sgroup, generations))

Diagnosed\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Diagnosed\_total <- array(0,dim=c(trials, 1, generations))

Diagnosed\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Diagnosed\_total\_py <- matrix(NA,nrow=trials, ncol=generations/12)

New\_Diagnosed <- array(NA,dim=c(trials, sgroup, generations))

New\_Diagnosed\_y <- array(NA,dim=c(trials, sgroup, generations/12))

New\_Diagnosed\_total <- array(0,dim=c(trials, 1, generations))

New\_Diagnosed\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Linked <- array(NA,dim=c(trials, sgroup, generations))

Linked\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Linked\_total <- array(0,dim=c(trials, 1, generations))

Linked\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Lost <- array(NA,dim=c(trials, sgroup, generations))

Lost\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Lost\_total <- array(0,dim=c(trials, 1, generations))

Lost\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Suppressed <- array(NA,dim=c(trials, sgroup, generations))

Suppressed\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Suppressed\_total <- array(0,dim=c(trials, 1, generations))

Suppressed\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Target\_VS <- matrix(NA,nrow=trials, ncol=generations/12)

Suppressed\_onART\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Suppressed\_onART\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Target\_VSonART <- matrix(NA,nrow=trials, ncol=generations/12)

Not\_Suppressed <- array(NA,dim=c(trials, sgroup, generations))

Not\_Suppressed\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Not\_Suppressed\_total <- array(0,dim=c(trials, 1, generations))

Not\_Suppressed\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

On\_ART <- array(NA,dim=c(trials, sgroup, generations))

On\_ART\_y <- array(NA,dim=c(trials, sgroup, generations/12))

On\_ART\_total <- array(0,dim=c(trials, 1, generations))

On\_ART\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

On\_ART\_total\_ny <- matrix(NA,nrow=trials, ncol=generations/12)

Target\_onART\_N <- matrix(NA,nrow=trials, ncol=generations/12)

Target\_onART\_P <- matrix(NA,nrow=trials, ncol=generations/12)

# This command sets up variables to store HIV incidence

Incidence <- array(NA,dim=c(trials, sgroup, generations))

Incidence\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Incidence\_total <- array(0,dim=c(trials, 1, generations))

Incidence\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Target\_incidence <- matrix(NA,nrow=trials, ncol=generations/12)

# This command sets up variables to store HIV prevalence.

Prevalence <- array(NA,dim=c(trials, sgroup, generations))

Prevalence\_y <- array(NA,dim=c(trials, sgroup, generations/12))

Prevalence\_total <- array(0,dim=c(trials, 1, generations))

Prevalence\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

Target\_prev <- matrix(NA,nrow=trials, ncol=generations/12)

# This command sets up variables to store the number of new infection.

NI <- array(NA,dim=c(trials, sgroup, generations))

NI\_y <- array(NA,dim=c(trials, sgroup, generations/12))

NI\_total <- array(0,dim=c(trials, 1, generations))

NI\_total\_y <- matrix(NA,nrow=trials, ncol=generations/12)

# This command sets up variables to store total population for high-risk urban women.

N\_hr\_f<- array(NA,dim=c(trials, sgroup, generations))

# This command sets up variables to store total population for low-risk urban women.

N\_lr\_u\_f<- array(NA,dim=c(trials, sgroup, generations))

# This command sets up variables to store total population for low-risk rural women.

N\_lr\_r\_f<- array(NA,dim=c(trials, sgroup, generations))

# This command sets up variables to store total population for urban men.

N\_u\_m<- array(NA,dim=c(trials, sgroup, generations))

# This command sets up variables to store total population for rural men.

N\_r\_m<- array(NA,dim=c(trials, sgroup, generations))

# This command sets up variables to store the value of model parameters.

# Rate at which people enter the susceptible population compartment.

rho <- array(NA,dim=c(trials, sgroup, generations))

rho\_2 <- array(NA,dim=c(trials, sgroup, generations))

# Rate of natural history HIV disease pregression.

delta\_1 <- array(NA,dim=c(trials, sgroup, generations))

delta\_2 <- array(NA,dim=c(trials, sgroup, generations))

delta\_3 <- array(NA,dim=c(trials, sgroup, generations))

delta\_4 <- array(NA,dim=c(trials, sgroup, generations))

delta\_5 <- array(NA,dim=c(trials, sgroup, generations))

delta\_6 <- array(NA,dim=c(trials, sgroup, generations))

m\_delta <- array(NA,dim=c(trials, sgroup, generations))

# Rate of getting diagnosed with HIV.

alpha\_1 <- array(NA,dim=c(trials, sgroup, generations))

alpha\_1\_a <- array(NA,dim=c(trials, sgroup, generations))

alpha\_1\_b <- array(NA,dim=c(trials, sgroup, generations))

alpha\_2 <- array(NA,dim=c(trials, sgroup, generations))

#alpha\_2\_a <- array(NA,dim=c(trials, sgroup, generations))

#alpha\_2\_b <- array(NA,dim=c(trials, sgroup, generations))

alpha\_3 <- array(NA,dim=c(trials, sgroup, generations))

#alpha\_3\_a <- array(NA,dim=c(trials, sgroup, generations))

#alpha\_3\_b <- array(NA,dim=c(trials, sgroup, generations))

alpha\_4 <- array(NA,dim=c(trials, sgroup, generations))

#alpha\_4\_a <- array(NA,dim=c(trials, sgroup, generations))

#alpha\_4\_b <- array(NA,dim=c(trials, sgroup, generations))

alpha\_m <- array(NA,dim=c(trials, sgroup, generations))

alpha\_m\_mu <- array(NA,dim=c(trials, sgroup, generations))

alpha\_m\_sd <- array(NA,dim=c(trials, sgroup, generations))

# Rate of being on ART and virally suppressed.

sigma\_1 <- array(NA,dim=c(trials, sgroup, generations))

sigma\_2 <- array(NA,dim=c(trials, sgroup, generations))

sigma\_3 <- array(NA,dim=c(trials, sgroup, generations))

sigma\_4 <- array(NA,dim=c(trials, sgroup, generations))

# Rate of getting lost from care.

gamma\_1 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_2 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_3 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_4 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_5 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_6 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_7 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_8 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_9 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_10 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_11 <- array(NA,dim=c(trials, sgroup, generations))

gamma\_12 <- array(NA,dim=c(trials, sgroup, generations))

# Rate of failure to maintain viral suppression.

psi\_1 <- array(NA,dim=c(trials, sgroup, generations))

psi\_2 <- array(NA,dim=c(trials, sgroup, generations))

psi\_3 <- array(NA,dim=c(trials, sgroup, generations))

psi\_4 <- array(NA,dim=c(trials, sgroup, generations))

# Rate of being on ART and suppressed.

theta\_1 <- array(NA,dim=c(trials, sgroup, generations))

theta\_1\_lb <- array(NA,dim=c(trials, sgroup, generations))

theta\_1\_ub <- array(NA,dim=c(trials, sgroup, generations))

theta\_2 <- array(NA,dim=c(trials, sgroup, generations))

theta\_2\_lb <- array(NA,dim=c(trials, sgroup, generations))

theta\_2\_ub <- array(NA,dim=c(trials, sgroup, generations))

theta\_3 <- array(NA,dim=c(trials, sgroup, generations))

theta\_3\_lb <- array(NA,dim=c(trials, sgroup, generations))

theta\_3\_ub <- array(NA,dim=c(trials, sgroup, generations))

theta\_4 <- array(NA,dim=c(trials, sgroup, generations))

theta\_4\_lb <- array(NA,dim=c(trials, sgroup, generations))

theta\_4\_ub <- array(NA,dim=c(trials, sgroup, generations))

# Rate of return to ART and being suppressed.

tau\_1 <- array(NA,dim=c(trials, sgroup, generations))

# Mortality rate.

mu\_1 <- array(NA,dim=c(trials, sgroup, generations))

mu\_2 <- array(NA,dim=c(trials, sgroup, generations))

mu\_3 <- array(NA,dim=c(trials, sgroup, generations))

mu\_4 <- array(NA,dim=c(trials, sgroup, generations))

mu\_5 <- array(NA,dim=c(trials, sgroup, generations))

mu\_6 <- array(NA,dim=c(trials, sgroup, generations))

mu\_7 <- array(NA,dim=c(trials, sgroup, generations))

mu\_8 <- array(NA,dim=c(trials, sgroup, generations))

mu\_9 <- array(NA,dim=c(trials, sgroup, generations))

mu\_10 <- array(NA,dim=c(trials, sgroup, generations))

mu\_11 <- array(NA,dim=c(trials, sgroup, generations))

mu\_12 <- array(NA,dim=c(trials, sgroup, generations))

mu\_13 <- array(NA,dim=c(trials, sgroup, generations))

mu\_14 <- array(NA,dim=c(trials, sgroup, generations))

mu\_15 <- array(NA,dim=c(trials, sgroup, generations))

mu\_16 <- array(NA,dim=c(trials, sgroup, generations))

m\_mu <- array(NA,dim=c(trials, sgroup, generations))

# Percentage of individuas consistently using a condom in sub-group r.

c\_r <- array(NA,dim=c(trials, sgroup, generations))

# Reduction in probability of HIV transmission when using a condom.

epsilon <- array(NA,dim=c(trials, sgroup, generations))

# Weight appleid for condom use in sub-group r.

omega\_r <- array(NA,dim=c(trials, sgroup, generations))

# Reduction in probability of HIV transmission when on ART and viral suppressed.

upsilon <- array(NA,dim=c(trials, sgroup, generations))

# Weight applied for viral suppreessed individuals

kappa <- array(NA,dim=c(trials, sgroup, generations))

# probability of HIV transmission per unprotected sex contact when not virally suppressed.

beta <- array(NA,dim=c(trials, sgroup, generations))

# Average number of sexual acts per time period in sub-group r

n\_r <- array(NA,dim=c(trials, sgroup, generations))

# Model calibration goodness of fit parameters

sls\_prev\_total <- numeric(trials)

sls\_incidence\_total <- numeric(trials)

sls\_onARTN\_total <- numeric(trials)

sls\_onARTP\_total <- numeric(trials)

sls\_vs\_total <- numeric(trials)

sls\_vsonART\_total <- numeric(trials)

sls\_total <- numeric(trials)

sls\_prev <- matrix(NA,nrow=trials, ncol=36)

sls\_incidence <- matrix(NA,nrow=trials, ncol=36)

sls\_onARTN <- matrix(NA,nrow=trials, ncol=36)

sls\_onARTP <- matrix(NA,nrow=trials, ncol=36)

sls\_vs <- matrix(NA,nrow=trials, ncol=36)

sls\_vsonART <- matrix(NA,nrow=trials, ncol=36)

pd\_prev\_total <- numeric(trials)

pd\_incidence\_total <- numeric(trials)

pd\_onARTN\_total <- numeric(trials)

pd\_onARTP\_total <- numeric(trials)

pd\_vs\_total <- numeric(trials)

pd\_vsonART\_total <- numeric(trials)

pd\_total <- numeric(trials)

pd\_prev <- matrix(NA,nrow=trials, ncol=36)

pd\_incidence <- matrix(NA,nrow=trials, ncol=36)

pd\_onARTN <- matrix(NA,nrow=trials, ncol=36)

pd\_onARTP <- matrix(NA,nrow=trials, ncol=36)

pd\_vs <- matrix(NA,nrow=trials, ncol=36)

pd\_vsonART <- matrix(NA,nrow=trials, ncol=36)

# Read the file that stores model parameter inputs  
files <- list.files(path="T:/Health Behavior and Policy/Faculty/Kimmel/Common/Personnel work/Deo/IeDEA/HIV Transmission Model-Rwanda/R Code/R/Subgroup", pattern = "csv")  
setwd("T:/Health Behavior and Policy/Faculty/Kimmel/Common/Personnel work/Deo/IeDEA/HIV Transmission Model-Rwanda/R Code/R/Subgroup")  
input <- lapply(files, read.csv)  
  
#Initalize parameters for each sub-group

for (t in 1:(trials)) {

for (s in 1:(sgroup)) {

# specify the line starting to read data

l <- 1

# Read initial susceptible population (15-64 years) in 2004 in Rwanda for each group

S[t,s,1] <- input[[s]][l,"S\_t"]

# Read initial population distribution of HIV infected individuals in Rwanda in 2004. We assume all infected individuals are undiagnosed and not on ART. The distribution will vary based on the sub-group considered.

X1[t,s,1] <- input[[s]][l,"X1\_t"]

X2[t,s,1] <- input[[s]][l,"X2\_t"]

X3[t,s,1] <- input[[s]][l,"X3\_t"]

X4[t,s,1] <- input[[s]][l,"X4\_t"]

X5[t,s,1] <- input[[s]][l,"X5\_t"]

X6[t,s,1] <- input[[s]][l,"X6\_t"]

X7[t,s,1] <- input[[s]][l,"X7\_t"]

X8[t,s,1] <- input[[s]][l,"X8\_t"]

X9[t,s,1] <- input[[s]][l,"X9\_t"]

X10[t,s,1] <- input[[s]][l,"X10\_t"]

X11[t,s,1] <- input[[s]][l,"X11\_t"]

X12[t,s,1] <- input[[s]][l,"X12\_t"]

X13[t,s,1] <- input[[s]][l,"X13\_t"]

X14[t,s,1] <- input[[s]][l,"X14\_t"]

X15[t,s,1] <- input[[s]][l,"X15\_t"]

X16[t,s,1] <- input[[s]][l,"X16\_t"]

X17[t,s,1] <- input[[s]][l,"X17\_t"]

X18[t,s,1] <- input[[s]][l,"X18\_t"]

X19[t,s,1] <- input[[s]][l,"X19\_t"]

X20[t,s,1] <- input[[s]][l,"X20\_t"]

X21[t,s,1] <- input[[s]][l,"X21\_t"]

X22[t,s,1] <- input[[s]][l,"X22\_t"]

X23[t,s,1] <- input[[s]][l,"X23\_t"]

X24[t,s,1] <- input[[s]][l,"X24\_t"]

# READ PARAMETER INPUTS

for (g in 1:(generations))

{

if (g<input[[s]][l+1,"generation"])

{

# Constant monthly rate of increase in the susceptible population, calculated using data from World Bank to reflect population growth over time.

rho[t,s,g] <- input[[s]][l,"rho\_t",g]

#Multiplier for monthly probability of getting diagnosed when CD4<200; Import the parameters for the distribution

#alpha\_m[t,s,g] <- input[[s]][l,"alpha\_m",g]

alpha\_m\_mu[t,s,g] <- input[[s]][l,"alpha\_m\_mu\_t",g]

alpha\_m\_sd[t,s,g] <- input[[s]][l,"alpha\_m\_sd\_t",g]

# Monthly probability of getting diagnosed from compartment X1 — X4; Import the prameters for the distribution;

#alpha\_1[t,s,g] <- input[[s]][l,"alpha\_1\_t",g]

alpha\_1\_a[t,s,g] <- input[[s]][l,"alpha\_1\_a\_t",g]

alpha\_1\_b[t,s,g] <- input[[s]][l,"alpha\_1\_b\_t",g]

if (input[[s]][l,"year",g]==2004 & g==1)

{

# generate random multiplier drawn from the log-normal distribution

mediate <- rlnorm(1,alpha\_m\_mu[t,s,g],alpha\_m\_sd[t,s,g])

# generate random parameter estimates from beta distribution

mediate\_a1 <- rbeta(1,alpha\_1\_a[t,s,g],alpha\_1\_b[t,s,g])

}

alpha\_m[t,s,g] <- mediate

alpha\_1[t,s,g] <- mediate\_a1

alpha\_2[t,s,g] <- mediate\_a1

alpha\_3[t,s,g] <- mediate\_a1

alpha\_4[t,s,g] <- mediate\_a1\*mediate

# Multiplier applied to linkage to care compartments (X9 — X12) to reduce the probability disease progression.

m\_delta[t,s,g] <- input[[s]][l,"m\_delta\_t",g]

# Monthly probability of natural history disease progression. We assume individuals who are undiagnosed (X1 — X4), diagnosed and not linked to care (X5 — X8), and those lost from care (X13 — X16), experience similar natural history disease progression in respective compartments.

delta\_1[t,s,g] <- input[[s]][l,"delta\_1\_t",g]

delta\_2[t,s,g] <- input[[s]][l,"delta\_2\_t",g]

delta\_3[t,s,g] <- input[[s]][l,"delta\_3\_t",g]

delta\_4[t,s,g] <- delta\_1[t,s,g]\*m\_delta[t,s,g]

delta\_5[t,s,g] <- delta\_2[t,s,g]\*m\_delta[t,s,g]

delta\_6[t,s,g] <- delta\_3[t,s,g]\*m\_delta[t,s,g]

# Monthly probability of linkage to care from compartments X5 — X8

sigma\_1[t,s,g] <- input[[s]][l,"sigma\_1\_t",g]

sigma\_2[t,s,g] <- input[[s]][l,"sigma\_2\_t",g]

sigma\_3[t,s,g] <- input[[s]][l,"sigma\_3\_t",g]

sigma\_4[t,s,g] <- input[[s]][l,"sigma\_4\_t",g]

# Probability of getting lost from care from compartments X9 — X12, X17 — X24

gamma\_1[t,s,g] <- input[[s]][l,"gamma\_1\_t",g]

gamma\_2[t,s,g] <- input[[s]][l,"gamma\_2\_t",g]

gamma\_3[t,s,g] <- input[[s]][l,"gamma\_3\_t",g]

gamma\_4[t,s,g] <- input[[s]][l,"gamma\_4\_t",g]

gamma\_5[t,s,g] <- input[[s]][l,"gamma\_5\_t",g]

gamma\_6[t,s,g] <- input[[s]][l,"gamma\_6\_t",g]

gamma\_7[t,s,g] <- input[[s]][l,"gamma\_7\_t",g]

gamma\_8[t,s,g] <- input[[s]][l,"gamma\_8\_t",g]

gamma\_9[t,s,g] <- input[[s]][l,"gamma\_9\_t",g]

gamma\_10[t,s,g] <- input[[s]][l,"gamma\_10\_t",g]

gamma\_11[t,s,g] <- input[[s]][l,"gamma\_11\_t",g]

gamma\_12[t,s,g] <- input[[s]][l,"gamma\_12\_t",g]

# Monthly probability of individuals returning to ART from compartment X16

tau\_1[t,s,g] <- input[[s]][l,"tau\_1\_t",g]

# Monthly probability of being on ART and viral suprressed from compartments X9 — X12

theta\_1\_lb[t,s,g] <- input[[s]][l,"theta\_1\_lb",g]

theta\_1\_ub[t,s,g] <- input[[s]][l,"theta\_1\_ub",g]

theta\_2\_lb[t,s,g] <- input[[s]][l,"theta\_2\_lb",g]

theta\_2\_ub[t,s,g] <- input[[s]][l,"theta\_2\_ub",g]

theta\_3\_lb[t,s,g] <- input[[s]][l,"theta\_3\_lb",g]

theta\_3\_ub[t,s,g] <- input[[s]][l,"theta\_3\_ub",g]

theta\_4\_lb[t,s,g] <- input[[s]][l,"theta\_4\_lb",g]

theta\_4\_ub[t,s,g] <- input[[s]][l,"theta\_4\_ub",g]

if (g==input[[s]][l,"generation"] & (input[[s]][l,"year"]==2004))

{

# generate random parameter estimates from beta distribution in year 2010 and 2015, which will apply for later years

mediate\_theta1 <- runif(1,theta\_1\_lb[t,s,g],theta\_1\_ub[t,s,g])

mediate\_theta2 <- runif(1,theta\_2\_lb[t,s,g],theta\_2\_ub[t,s,g])

mediate\_theta3 <- runif(1,theta\_3\_lb[t,s,g],theta\_3\_ub[t,s,g])

mediate\_theta4 <- runif(1,theta\_4\_lb[t,s,g],theta\_4\_ub[t,s,g])

}

theta\_1[t,s,g] <- mediate\_theta1

theta\_2[t,s,g] <- mediate\_theta2

theta\_3[t,s,g] <- mediate\_theta3

theta\_4[t,s,g] <- mediate\_theta4

# Monthly probability of being on ART and not viral suppressed from compartments X17 — X20

psi\_1[t,s,g] <- input[[s]][l,"psi\_1\_t",g]

psi\_2[t,s,g] <- input[[s]][l,"psi\_2\_t",g]

psi\_3[t,s,g] <- input[[s]][l,"psi\_3\_t",g]

psi\_4[t,s,g] <- input[[s]][l,"psi\_4\_t",g]

# Multiplier applied to compartments for mortality rate reduction

m\_mu[t,s,g] <- input[[s]][l,"m\_mu\_t",g]

# Monthly probability of death from compartments X1 — X24. We assume individuals who are undiagnosed (X1 — X4), diagnosed and not linked to care (X5 — X8), and those lost from care (X13 — X16), experience similar mortality rate in respective compartments.

mu\_1[t,s,g] <- input[[s]][l,"mu\_1\_t",g]

mu\_2[t,s,g] <- input[[s]][l,"mu\_2\_t",g]

mu\_3[t,s,g] <- input[[s]][l,"mu\_3\_t",g]

mu\_4[t,s,g] <- input[[s]][l,"mu\_4\_t",g]

mu\_5[t,s,g] <- mu\_1[t,s,g]\*m\_mu[t,s,g]

mu\_6[t,s,g] <- mu\_2[t,s,g]\*m\_mu[t,s,g]

mu\_7[t,s,g] <- mu\_3[t,s,g]\*m\_mu[t,s,g]

mu\_8[t,s,g] <- mu\_4[t,s,g]\*m\_mu[t,s,g]

mu\_9[t,s,g] <- input[[s]][l,"mu\_9\_t",g]

mu\_10[t,s,g] <- input[[s]][l,"mu\_10\_t",g]

mu\_11[t,s,g] <- input[[s]][l,"mu\_11\_t",g]

mu\_12[t,s,g] <- input[[s]][l,"mu\_12\_t",g]

mu\_13[t,s,g] <- input[[s]][l,"mu\_13\_t",g]

mu\_14[t,s,g] <- input[[s]][l,"mu\_14\_t",g]

mu\_15[t,s,g] <- input[[s]][l,"mu\_15\_t",g]

mu\_16[t,s,g] <- input[[s]][l,"mu\_16\_t",g]

# Weight applied to the force of infection expression to account for condom use and effectiveness in reducing HIV transmission. The weight varies based on the subgroup.

c\_r[t,s,g] <- input[[s]][l,"c\_r\_t",g]

epsilon[t,s,g] <- input[[s]][l,"epsilon\_t",g]

omega\_r[t,s,g] <-(1-(epsilon[t,s,g]\*c\_r[t,s,g]))

# Weight applied to the force of infection expression to account for the benefits of being on ART and viral suppressed.

upsilon[t,s,g] <- input[[s]][l,"upsilon\_t",g]

kappa[t,s,g] <- (1-upsilon[t,s,g])

# Risk of HIV transmission. per sexual act. We assume the risk of HIV transmission is the same from male to female and from female to male.

beta[t,s,g] <- input[[s]][l,"beta\_t",g]

# Average number of sexual contacts per month. This number will vary based on the subgroup considered.

n\_r[t,s,g] <- input[[s]][l,"n\_r\_t",g]

} else {

l=l+1

# Constant monthly rate of increase in the susceptible population, calculated using data from World Bank to reflect population growth over time.

rho[t,s,g] <- input[[s]][l,"rho\_t",g]

#Mulitplier for monthly probability of getting diagnosed when CD4<200

alpha\_m\_mu[t,s,g] <- input[[s]][l,"alpha\_m\_mu\_t",g]

alpha\_m\_sd[t,s,g] <- input[[s]][l,"alpha\_m\_sd\_t",g]

# Monthly probability of getting diagnosed from compartment X1 — X4

#alpha\_1[t,s,g] <- input[[s]][l,"alpha\_1\_t",g]

alpha\_1\_a[t,s,g] <- input[[s]][l,"alpha\_1\_a\_t",g]

alpha\_1\_b[t,s,g] <- input[[s]][l,"alpha\_1\_b\_t",g]

if (g==input[[s]][l,"generation"] & (input[[s]][l,"year"]==2010 | input[[s]][l,"year"]==2015))

{

# generate random parameter estimates from beta distribution in year 2010 and 2015, which will apply for later years

mediate\_a1 <- rbeta(1,alpha\_1\_a[t,s,g],alpha\_1\_b[t,s,g])

}

alpha\_m[t,s,g] <- mediate

alpha\_1[t,s,g] <- mediate\_a1

alpha\_2[t,s,g] <- mediate\_a1

alpha\_3[t,s,g] <- mediate\_a1

alpha\_4[t,s,g] <- mediate\_a1\*mediate

# Multiplier applied to linkage to care compartments (X9 — X12) to reduce the probability disease progression.

m\_delta[t,s,g] <- input[[s]][l,"m\_delta\_t",g]

# Monthly probability of natural history disease progression. We assume individuals who are undiagnosed (X1 — X4), diagnosed and not linked to care (X5 — X8), and those lost from care (X13 — X16), experience similar natural history disease progression in respective compartments.

delta\_1[t,s,g] <- input[[s]][l,"delta\_1\_t",g]

delta\_2[t,s,g] <- input[[s]][l,"delta\_2\_t",g]

delta\_3[t,s,g] <- input[[s]][l,"delta\_3\_t",g]

delta\_4[t,s,g] <- delta\_1[t,s,g]\*m\_delta[t,s,g]

delta\_5[t,s,g] <- delta\_2[t,s,g]\*m\_delta[t,s,g]

delta\_6[t,s,g] <- delta\_3[t,s,g]\*m\_delta[t,s,g]

# Monthly probability of linkage to care from compartments X5 — X8

sigma\_1[t,s,g] <- input[[s]][l,"sigma\_1\_t",g]

sigma\_2[t,s,g] <- input[[s]][l,"sigma\_2\_t",g]

sigma\_3[t,s,g] <- input[[s]][l,"sigma\_3\_t",g]

sigma\_4[t,s,g] <- input[[s]][l,"sigma\_4\_t",g]

# Probability of getting lost from care from compartments X9 — X12, X17 — X24

gamma\_1[t,s,g] <- input[[s]][l,"gamma\_1\_t",g]

gamma\_2[t,s,g] <- input[[s]][l,"gamma\_2\_t",g]

gamma\_3[t,s,g] <- input[[s]][l,"gamma\_3\_t",g]

gamma\_4[t,s,g] <- input[[s]][l,"gamma\_4\_t",g]

gamma\_5[t,s,g] <- input[[s]][l,"gamma\_5\_t",g]

gamma\_6[t,s,g] <- input[[s]][l,"gamma\_6\_t",g]

gamma\_7[t,s,g] <- input[[s]][l,"gamma\_7\_t",g]

gamma\_8[t,s,g] <- input[[s]][l,"gamma\_8\_t",g]

gamma\_9[t,s,g] <- input[[s]][l,"gamma\_9\_t",g]

gamma\_10[t,s,g] <- input[[s]][l,"gamma\_10\_t",g]

gamma\_11[t,s,g] <- input[[s]][l,"gamma\_11\_t",g]

gamma\_12[t,s,g] <- input[[s]][l,"gamma\_12\_t",g]

# Monthly probability of individuals returning to ART from compartment X16

tau\_1[t,s,g] <- input[[s]][l,"tau\_1\_t",g]

# Monthly probability of being on ART and viral suprressed from compartments X9 — X12

theta\_1\_lb[t,s,g] <- input[[s]][l,"theta\_1\_lb",g]

theta\_1\_ub[t,s,g] <- input[[s]][l,"theta\_1\_ub",g]

theta\_2\_lb[t,s,g] <- input[[s]][l,"theta\_2\_lb",g]

theta\_2\_ub[t,s,g] <- input[[s]][l,"theta\_2\_ub",g]

theta\_3\_lb[t,s,g] <- input[[s]][l,"theta\_3\_lb",g]

theta\_3\_ub[t,s,g] <- input[[s]][l,"theta\_3\_ub",g]

theta\_4\_lb[t,s,g] <- input[[s]][l,"theta\_4\_lb",g]

theta\_4\_ub[t,s,g] <- input[[s]][l,"theta\_4\_ub",g]

if (g==input[[s]][l,"generation"] & (input[[s]][l,"year"]==2008 | input[[s]][l,"year"]==2014 | input[[s]][l,"year"]==2016))

{

# generate random parameter estimates from beta distribution in year 2010 and 2015, which will apply for later years

mediate\_theta1 <- runif(1,theta\_1\_lb[t,s,g],theta\_1\_ub[t,s,g])

mediate\_theta2 <- runif(1,theta\_2\_lb[t,s,g],theta\_2\_ub[t,s,g])

mediate\_theta3 <- runif(1,theta\_3\_lb[t,s,g],theta\_3\_ub[t,s,g])

mediate\_theta4 <- runif(1,theta\_4\_lb[t,s,g],theta\_4\_ub[t,s,g])

}

theta\_1[t,s,g] <- mediate\_theta1

theta\_2[t,s,g] <- mediate\_theta2

theta\_3[t,s,g] <- mediate\_theta3

theta\_4[t,s,g] <- mediate\_theta4

# Monthly probability of being on ART and not viral suppressed from compartments X17 — X20

psi\_1[t,s,g] <- input[[s]][l,"psi\_1\_t",g]

psi\_2[t,s,g] <- input[[s]][l,"psi\_2\_t",g]

psi\_3[t,s,g] <- input[[s]][l,"psi\_3\_t",g]

psi\_4[t,s,g] <- input[[s]][l,"psi\_4\_t",g]

# Multiplier applied to compartments for linkage to care for mortality rate reduction

m\_mu[t,s,g] <- input[[s]][l,"m\_mu\_t",g]

# Monthly probability of death from compartments X1 — X24. We assume individuals who are undiagnosed (X1 — X4), diagnosed and not linked to care (X5 — X8), and those lost from care (X13 — X16), experience similar mortality rate in respective compartments.

mu\_1[t,s,g] <- input[[s]][l,"mu\_1\_t",g]

mu\_2[t,s,g] <- input[[s]][l,"mu\_2\_t",g]

mu\_3[t,s,g] <- input[[s]][l,"mu\_3\_t",g]

mu\_4[t,s,g] <- input[[s]][l,"mu\_4\_t",g]

mu\_5[t,s,g] <- mu\_1[t,s,g]\*m\_mu[t,s,g]

mu\_6[t,s,g] <- mu\_2[t,s,g]\*m\_mu[t,s,g]

mu\_7[t,s,g] <- mu\_3[t,s,g]\*m\_mu[t,s,g]

mu\_8[t,s,g] <- mu\_4[t,s,g]\*m\_mu[t,s,g]

mu\_9[t,s,g] <- input[[s]][l,"mu\_9\_t",g]

mu\_10[t,s,g] <- input[[s]][l,"mu\_10\_t",g]

mu\_11[t,s,g] <- input[[s]][l,"mu\_11\_t",g]

mu\_12[t,s,g] <- input[[s]][l,"mu\_12\_t",g]

mu\_13[t,s,g] <- input[[s]][l,"mu\_13\_t",g]

mu\_14[t,s,g] <- input[[s]][l,"mu\_14\_t",g]

mu\_15[t,s,g] <- input[[s]][l,"mu\_15\_t",g]

mu\_16[t,s,g] <- input[[s]][l,"mu\_16\_t",g]

# Weight applied to the force of infection expression to account for condom use and effectiveness in reducing HIV transmission. The weight varies based on the subgroup.

c\_r[t,s,g] <- input[[s]][l,"c\_r\_t",g]

epsilon[t,s,g] <- input[[s]][l,"epsilon\_t",g]

omega\_r[t,s,g] <-(1-(epsilon[t,s,g]\*c\_r[t,s,g]))

# Weight applied to the force of infection expression to account for the benefits of being on ART and viral suppressed.

upsilon[t,s,g] <- input[[s]][l,"upsilon\_t",g]

kappa[t,s,g] <- (1-upsilon[t,s,g])

# Risk of HIV transmission. per sexual act. We assume the risk of HIV transmission is the same from male to female and from female to male.

beta[t,s,g] <- input[[s]][l,"beta\_t",g]

# Average number of sexual contacts per month. This number will vary based on the subgroup considered.

n\_r[t,s,g] <- input[[s]][l,"n\_r\_t",g]

}

}

}

}

# Initalize baseline population characteristics

for (t in 1:(trials)) {

#Initalize parameters for each sub-group

for (s in 1:(sgroup)) {

# Calculate the initial HIV prevalence

# Total HIV infected individuals

N[t,s,1] <- X1[t,s,1]+X2[t,s,1]+X3[t,s,1]+X4[t,s,1]+X5[t,s,1]+X6[t,s,1]+X7[t,s,1]+X8[t,s,1]+X9[t,s,1]+X10[t,s,1]+X11[t,s,1]+X12[t,s,1]+X13[t,s,1]+X14[t,s,1]+X15[t,s,1]+X16[t,s,1]+X17[t,s,1]+X18[t,s,1]+X19[t,s,1]+X20[t,s,1]+X21[t,s,1]+X22[t,s,1]+X23[t,s,1]+X24[t,s,1]

# Total population of individuals who are not viral suppressed

Z[t,s,1] <- X1[t,s,1]+X2[t,s,1]+X3[t,s,1]+X4[t,s,1]+X5[t,s,1]+X6[t,s,1]+X7[t,s,1]+X8[t,s,1]+X9[t,s,1]+X10[t,s,1]+X11[t,s,1]+X12[t,s,1]+X13[t,s,1]+X14[t,s,1]+X15[t,s,1]+X16[t,s,1]+X21[t,s,1]+X22[t,s,1]+X23[t,s,1]+X24[t,s,1]

# Total population of individuals who are viral suppressed

V[t,s,1] <- X17[t,s,1]+X18[t,s,1]+X19[t,s,1]+X20[t,s,1]

# Total population of individuals infected but undiagnosed

# Undiagnosed[t,s,1] <- X1[t,s,1]+X2[t,s,1]+X3[t,s,1]+X4[t,s,1]

# Total population of individuals diagnosed with HIV

Diagnosed[t,s,1] <- X5[t,s,1]+X6[t,s,1]+X7[t,s,1]+X8[t,s,1]+X9[t,s,1]+X10[t,s,1]+X11[t,s,1]+X12[t,s,1]+X13[t,s,1]+X14[t,s,1]+X15[t,s,1]+X16[t,s,1]+X17[t,s,1]+X18[t,s,1]+X19[t,s,1]+X20[t,s,1]+X21[t,s,1]+X22[t,s,1]+X23[t,s,1]+X24[t,s,1]

# Total population of individuals newly diagnosed

New\_Diagnosed[t,s,1] <- X1[t,s,1]\*alpha\_1[t,s,g]+X2[t,s,1]\*alpha\_2[t,s,g]+X3[t,s,1]\*alpha\_3[t,s,g]+X4[t,s,1]\*alpha\_4[t,s,g]

# Total population of inidivudlas linked to care

# Linked[t,s,1] <- X9[t,s,1]+X10[t,s,1]+X11[t,s,1]+X12[t,s,1]

# Total population of individuals lost

# Lost[t,s,1] <- X13[t,s,1]+X14[t,s,1]+X15[t,s,1]+X16[t,s,1]

# Total population of individuals on ART and suppressed

Suppressed[t,s,1] <- X17[t,s,1]+X18[t,s,1]+X19[t,s,1]+X20[t,s,1]

# Total population of individuals on ART and not suppressed

# Not\_Suppressed[t,s,1] <- X21[t,s,1]+X22[t,s,1]+X23[t,s,1]+X24[t,s,1]

# Total population of individuals on ART

On\_ART[t,s,1] <- X17[t,s,1]+X18[t,s,1]+X19[t,s,1]+X20[t,s,1]+X21[t,s,1]+X22[t,s,1]+X23[t,s,1]+X24[t,s,1]

if (N[t,s,1]==0) {

p\_r[t,s,1] <- 0

p\_r\_v[t,s,1] <- 0

} else {

# percentage of individuals who are not viral supprressed among people living with HIV

p\_r[t,s,1] <- Z[t,s,1]/N[t,s,1]

# Prevalence of individuals who are viral supprressed among people living with HIV

p\_r\_v[t,s,1] <- V[t,s,1]/N[t,s,1]

}

# initial value for force of infection and new infection (set to null)

lambda\_t\_r[t,s,1] <- NA

NI[t,s,1] <- NA

N\_total[t,1,1]

}

}

# End of initializing the parameters for sub-group  
 # Start the loop by generations  
 # The loop of generations goes before loops of sub-groups since operations in time period t+1 depends on the outcomes from all sub-groups in time period t.  
 # The generations starts from time 2 so that the equations in R are in line with the difference equation.   
 for (g in 2:(generations))  
 {  
 for (s in 1:(sgroup))  
 {  
 # Based on the number of sub-group, calculate the force of infection with mixing  
 # group 1-5: low risk urban women  
 # group 6-10: high risk urban women  
 # group 11-15: low risk rural women  
 # group 16-20: urban men  
 # group 21-25: rural men  
   
 # total population of low risk urban women from last generation, who will mix with urban men   
 N\_lr\_u\_f[t,s,g-1] <- S[t,1,g-1]+S[t,2,g-1]+S[t,3,g-1]+S[t,4,g-1]+S[t,5,g-1]+N[t,1,g-1]+N[t,2,g-1]+N[t,3,g-1]+N[t,4,g-1]+N[t,5,g-1]  
   
 # total population of high risk urban women from last generation, who will mix with urban and rural men   
 N\_hr\_f[t,s,g-1] <- S[t,6,g-1]+S[t,7,g-1]+S[t,8,g-1]+S[t,9,g-1]+S[t,10,g-1]+N[t,6,g-1]+N[t,7,g-1]+N[t,8,g-1]+N[t,9,g-1]+N[t,10,g-1]  
   
 # total population of low risk rural women from last generation, who will mix with rural men   
 N\_lr\_r\_f[t,s,g-1] <- S[t,11,g-1]+S[t,12,g-1]+S[t,13,g-1]+S[t,14,g-1]+S[t,15,g-1]+N[t,11,g-1]+N[t,12,g-1]+N[t,13,g-1]+N[t,14,g-1]+N[t,15,g-1]  
   
 # total population of urban men from last generation, who will mix with high risk women and low risk urban women  
 N\_u\_m[t,s,g-1] <- S[t,16,g-1]+S[t,17,g-1]+S[t,18,g-1]+S[t,19,g-1]+S[t,20,g-1]+N[t,16,g-1]+N[t,17,g-1]+N[t,18,g-1]+N[t,19,g-1]+N[t,20,g-1]  
   
 # total population of rural men from last generation, who will mix with high risk women and low risk rural women  
 N\_r\_m[t,s,g-1] <- S[t,21,g-1]+S[t,22,g-1]+S[t,23,g-1]+S[t,24,g-1]+S[t,25,g-1]+N[t,21,g-1]+N[t,22,g-1]+N[t,23,g-1]+N[t,24,g-1]+N[t,25,g-1]  
   
 # low risk urban women mixed with urban men  
 if (s>=1&s<=5)   
 {  
 for (j in 16:20)   
 {  
 # force of infection low risk urban women  
 lambda\_t\_r[t,s,g] <- lambda\_t\_r[t,s,g]+((N[t,j,g-1]+S[t,j,g-1])/N\_u\_m[t,s,g-1])\*(N[t,j,g-1]/(N[t,j,g-1]+S[t,j,g-1]))\*(1-(1-(omega\_r[t,s]\*beta[t,s]\*(kappa[t,s]\*p\_r\_v[t,j,g-1]+p\_r[t,j,g-1])))^n\_r[t,s])  
 }   
 }  
   
 # high risk women mixed with urban and rural men  
 if (s>=6&s<=10)   
 {  
 for (j in 16:25)   
 {  
 # force of infection (high risk urban women)  
 lambda\_t\_r[t,s,g] <- lambda\_t\_r[t,s,g]+((N[t,j,g-1]+S[t,j,g-1])/N\_u\_m[t,s,g-1])\*(N[t,j,g-1]/(N[t,j,g-1]+S[t,j,g-1]))\*(1-(1-(omega\_r[t,s]\*beta[t,s]\*(kappa[t,s]\*p\_r\_v[t,j,g-1]+p\_r[t,j,g-1])))^n\_r[t,s])  
 }  
 }  
   
 # low risk rural women mixed with rural men  
 if (s>=11&s<=15)   
 {  
 for (j in 21:25)   
 {  
 # force of infection (low risk rural women)  
 lambda\_t\_r[t,s,g] <- lambda\_t\_r[t,s,g]+((N[t,j,g-1]+S[t,j,g-1])/N\_u\_m[t,s,g-1])\*(N[t,j,g-1]/(N[t,j,g-1]+S[t,j,g-1]))\*(1-(1-(omega\_r[t,s]\*beta[t,s]\*(kappa[t,s]\*p\_r\_v[t,j,g-1]+p\_r[t,j,g-1])))^n\_r[t,s])  
 }  
 }  
   
 # urban men mixed with low risk urban women and high risk women  
 if (s>=16&s<=20)   
 {  
 for (j in 1:10)   
 {  
 # force of infection (urban men)  
 lambda\_t\_r[t,s,g] <- lambda\_t\_r[t,s,g]+((N[t,j,g-1]+S[t,j,g-1])/N\_u\_m[t,s,g-1])\*(N[t,j,g-1]/(N[t,j,g-1]+S[t,j,g-1]))\*(1-(1-(omega\_r[t,s]\*beta[t,s]\*(kappa[t,s]\*p\_r\_v[t,j,g-1]+p\_r[t,j,g-1])))^n\_r[t,s])  
 }  
 }  
   
 # rural men mixed with low risk rural women and high risk women  
 if (s>=21&s<=25)   
 {  
 for (j in 6:15)   
 {  
 # force of infection (rural men)  
 lambda\_t\_r[t,s,g] <- lambda\_t\_r[t,s,g]+((N[t,j,g-1]+S[t,j,g-1])/N\_u\_m[t,s,g-1])\*(N[t,j,g-1]/(N[t,j,g-1]+S[t,j,g-1]))\*(1-(1-(omega\_r[t,s]\*beta[t,s]\*(kappa[t,s]\*p\_r\_v[t,j,g-1]+p\_r[t,j,g-1])))^n\_r[t,s])  
 }  
 }  
   
 # Calculate new infections  
 NI[t,s,g] <- lambda\_t\_r[t,s,g]\*S[t,s,g-1]  
   
 # Difference equation for each compartments  
 source("T:/Health Behavior and Policy/Faculty/Kimmel/Common/Personnel work/Deo/IeDEA/HIV Transmission Model-Rwanda/R Code/Shared code/diff\_equation\_0224\_subgroup.R")   
   
 # Total population of HIV infected individuals  
 N[t,s,g] <- X1[t,s,g]+X2[t,s,g]+X3[t,s,g]+X4[t,s,g]+X5[t,s,g]+X6[t,s,g]+X7[t,s,g]+X8[t,s,g]+X9[t,s,g]+X10[t,s,g]+X11[t,s,g]+X12[t,s,g]+X13[t,s,g]+X14[t,s,g]+X15[t,s,g]+X16[t,s,g]+X17[t,s,g]+X18[t,s,g]+X19[t,s,g]+X20[t,s,g]+X21[t,s,g]+X22[t,s,g]+X23[t,s,g]+X24[t,s,g]  
   
 # Total population of individuals who are not viral suppressed  
 Z[t,s,g] <- X1[t,s,g]+X2[t,s,g]+X3[t,s,g]+X4[t,s,g]+X5[t,s,g]+X6[t,s,g]+X7[t,s,g]+X8[t,s,g]+X9[t,s,g]+X10[t,s,g]+X11[t,s,g]+X12[t,s,g]+X13[t,s,g]+X14[t,s,g]+X15[t,s,g]+X16[t,s,g]+X21[t,s,g]+X22[t,s,g]+X23[t,s,g]+X24[t,s,g]  
   
 # Total population of individuals who are viral suppressed  
 V[t,s,g] <- X17[t,s,g]+X18[t,s,g]+X19[t,s,g]+X20[t,s,g]  
   
 if (N[t,s,g]==0) {  
   
 p\_r[t,s,g] <- 0  
 p\_r\_v[t,s,g] <- 0  
   
 } else {  
   
 # Prevalence of HIV in the population based on individuals who are not viral supprressed  
 p\_r[t,s,g] <- Z[t,s,g]/N[t,s,g]  
   
 # Prevalence of HIV in the population based on individuals who are viral supprressed  
 p\_r\_v[t,s,g] <- V[t,s,g]/N[t,s,g]   
 }  
 }  
 }  
   
 # EPIDMIC CURVES   
  
for (t in (1:trials)){   
 for (g in (1:generations)) {  
 for (s in (1:sgroup)){  
 NI\_total[t,1,g]=NI\_total[t,1,g]+NI[t,s,g]  
 N\_total[t,1,g]=N\_total[t,1,g]+N[t,s,g]  
 Z\_total[t,1,g]=Z\_total[t,1,g]+Z[t,s,g]  
 V\_total[t,1,g]=V\_total[t,1,g]+V[t,s,g]  
 S\_total[t,1,g]=S\_total[t,1,g]+S[t,s,g]  
 }  
 }  
}   
   
for (t in (1:trials)) {  
 # Assign a vector to store the year variables  
 year <- seq(from=2004, to=2039, by=1)  
  
 # Annual new infection  
 library(zoo)  
 NI\_total\_y[t,] <- rollapply(NI\_total[t,1,],1,sum,by=12)  
 plot(year, NI\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "New HIV Infections", main = "Annual individuals newly infected - Overall", ylim=c(0,max(NI\_total\_y[t,],na.rm=TRUE)))  
   
 # People living with HIV  
 N\_total\_y[t,] <- N\_total[t,1,][seq(1, length(N\_total[t,1,]), 12)]  
 plot(year, N\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Individuals living with HIV", main = "Annual individuals living with HIV - Overall", ylim=c(0,max(N\_total\_y[t,],na.rm=TRUE)))  
 #  
   
 # HIV prevalence  
 Prevalence\_total[t,1,] <- (Z\_total[t,1,]+V\_total[t,1,])/(N\_total[t,1,]+S\_total[t,1,])  
 Prevalence\_total\_y[t,] <- Prevalence\_total[t,1,][seq(1,length(Prevalence\_total[t,1,]),12)]  
 plot(year, 100\*Prevalence\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Prevalence (%)", main = "Annual HIV prevalence - Overall", ylim=c(0,max(100\*Prevalence\_total\_y[t,],na.rm=TRUE)))  
  
}   
   
 # Subgroups  
 # # For now, we only report new HIV infections, people living with HIV, and HIV prevalence for each sub-group.   
 for (s in (1:sgroup)){  
 # New infections   
 NI\_y[t,s,] <- rollapply(NI[t,s,],1,sum,by=12)  
   
 # People living with HIV  
 N\_y[t,s,] <- N[t,s,][seq(1, length(N[t,s,]), 12)]  
  
 # HIV prevalence  
 Prevalence[t,s,] <- (Z[t,s,]+V[t,s,])/(N[t,s,]+S[t,s,])  
 Prevalence\_y[t,s,] <- Prevalence[t,s,][seq(1,length(Prevalence[t,s,]),12)]  
 }  
   
 for (s in (1:sgroup)){  
 plot(year, NI\_y[t,s,], type = "l", col="red", xlab = "Year", ylab = "New HIV Infections", main = "Annual individuals newly infected", ylim=c(0,max(NI\_y[t,,],na.rm=TRUE)))  
 plot(year, N\_y[t,s,], type = "l", col="red", xlab = "Year", ylab = "Individuals living with HIV", main = "Annual individuals living with HIV", ylim=c(0,max(N\_y[t,,],na.rm=TRUE)))  
 plot(year, 100\*Prevalence\_y[t,s,], type = "l", col="red", xlab = "Year", ylab = "Prevalence (%)", main = "Annual HIV prevalence", ylim=c(0,max(100\*Prevalence\_y[t,,],na.rm=TRUE)))  
 }  
   
}

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

#End of the loop

Difference equation: Difference equation is prepared as a separate R code so that it is re-usable and easy to modify. The codes were presented as follows.

# Difference equation for susceptable population  
 S[t,s,g] <- S[t,s,g-1]+rho[t,s]\*S[t,s,g-1]-lambda\_t\_r[t,s,g]\*S[t,s,g-1]  
   
 # Difference equations for compartments X1 - X4 (undiagnosed individuals)  
   
 X1[t,s,g] <- X1[t,s,g-1]+lambda\_t\_r[t,s,g]\*S[t,s,g-1]-delta\_1[t,s]\*X1[t,s,g-1]-alpha\_1[t,s]\*X1[t,s,g-1]-mu\_1[t,s]\*X1[t,s,g-1]  
   
 X2[t,s,g] <- X2[t,s,g-1]+delta\_1[t,s]\*X1[t,s,g-1]-delta\_2[t,s]\*X2[t,s,g-1]-alpha\_2[t,s]\*X2[t,s,g-1]-mu\_2[t,s]\*X2[t,s,g-1]  
   
 X3[t,s,g] <- X3[t,s,g-1]+delta\_2[t,s]\*X2[t,s,g-1]-delta\_3[t,s]\*X3[t,s,g-1]-alpha\_3[t,s]\*X3[t,s,g-1]-mu\_3[t,s]\*X3[t,s,g-1]  
   
 X4[t,s,g] <- X4[t,s,g-1]+delta\_3[t,s]\*X3[t,s,g-1]-alpha\_4[t,s]\*X4[t,s,g-1]-mu\_4[t,s]\*X4[t,s,g-1]  
   
 # Difference equations for compartments X5 - X8 (diagnosed individuals)  
   
 X5[t,s,g] <- X5[t,s,g-1]+alpha\_1[t,s]\*X1[t,s,g-1]-delta\_1[t,s]\*X5[t,s,g-1]-sigma\_1[t,s]\*X5[t,s,g-1]-mu\_1[t,s]\*X5[t,s,g-1]  
   
 X6[t,s,g] <- X6[t,s,g-1]+alpha\_2[t,s]\*X2[t,s,g-1]+delta\_1[t,s]\*X5[t,s,g-1]-delta\_2[t,s]\*X6[t,s,g-1]-sigma\_2[t,s]\*X6[t,s,g-1]-mu\_2[t,s]\*X6[t,s,g-1]  
   
 X7[t,s,g] <- X7[t,s,g-1]+alpha\_3[t,s]\*X3[t,s,g-1]+delta\_2[t,s]\*X6[t,s,g-1]-delta\_3[t,s]\*X7[t,s,g-1]-sigma\_3[t,s]\*X7[t,s,g-1]-mu\_3[t,s]\*X7[t,s,g-1]  
   
 X8[t,s,g] <- X8[t,s,g-1]+alpha\_4[t,s]\*X4[t,s,g-1]+delta\_3[t,s]\*X7[t,s,g-1]-sigma\_4[t,s]\*X8[t,s,g-1]-mu\_4[t,s]\*X8[t,s,g-1]  
   
 # Difference equations for compartments X9 - X12 (individuals linked to care)  
   
 X9[t,s,g] <- X9[t,s,g-1]+sigma\_1[t,s]\*X5[t,s,g-1]-delta\_4[t,s]\*X9[t,s,g-1]-theta\_1[t,s]\*X9[t,s,g-1]-gamma\_1[t,s]\*X9[t,s,g-1]-mu\_5[t,s]\*X9[t,s,g-1]  
   
 X10[t,s,g] <- X10[t,s,g-1]+sigma\_2[t,s]\*X6[t,s,g-1]+delta\_4[t,s]\*X9[t,s,g-1]-delta\_5[t,s]\*X10[t,s,g-1]-theta\_2[t,s]\*X10[t,s,g-1]-gamma\_2[t,s]\*X10[t,s,g-1]-mu\_6[t,s]\*X10[t,s,g-1]  
   
 X11[t,s,g] <- X11[t,s,g-1]+sigma\_3[t,s]\*X7[t,s,g-1]+delta\_5[t,s]\*X10[t,s,g-1]-delta\_6[t,s]\*X11[t,s,g-1]-theta\_3[t,s]\*X11[t,s,g-1]-gamma\_3[t,s]\*X11[t,s,g-1]-mu\_7[t,s]\*X11[t,s,g-1]  
   
 X12[t,s,g] <- X12[t,s,g-1]+sigma\_4[t,s]\*X8[t,s,g-1]+delta\_6[t,s]\*X11[t,s,g-1]-theta\_4[t,s]\*X12[t,s,g-1]-gamma\_4[t,s]\*X12[t,s,g-1]-mu\_8[t,s]\*X12[t,s,g-1]  
   
 # Difference equations for compartments X13 - X16 (individuals lost from care)  
   
 X13[t,s,g] <- X13[t,s,g-1]+gamma\_1[t,s]\*X9[t,s,g-1]+gamma\_5[t,s]\*X17[t,s,g-1]+gamma\_9[t,s]\*X21[t,s,g-1]-delta\_1[t,s]\*X13[t,s,g-1]-mu\_1[t,s]\*X13[t,s,g-1]  
   
 X14[t,s,g] <- X14[t,s,g-1]+gamma\_2[t,s]\*X10[t,s,g-1]+gamma\_6[t,s]\*X18[t,s,g-1]+gamma\_10[t,s]\*X22[t,s,g-1]+delta\_1[t,s]\*X13[t,s,g-1]-delta\_2[t,s]\*X14[t,s,g-1]-mu\_2[t,s]\*X14[t,s,g-1]  
   
 X15[t,s,g] <- X15[t,s,g-1]+gamma\_3[t,s]\*X11[t,s,g-1]+gamma\_7[t,s]\*X19[t,s,g-1]+gamma\_11[t,s]\*X23[t,s,g-1]+delta\_2[t,s]\*X14[t,s,g-1]-delta\_3[t,s]\*X15[t,s,g-1]-mu\_3[t,s]\*X15[t,s,g-1]  
   
 X16[t,s,g] <- X16[t,s,g-1]+gamma\_4[t,s]\*X12[t,s,g-1]+gamma\_8[t,s]\*X20[t,s,g-1]+gamma\_12[t,s]\*X24[t,s,g-1]+delta\_3[t,s]\*X15[t,s,g-1]-tau\_1[t,s]\*X16[t,s,g-1]-mu\_4[t,s]\*X16[t,s,g-1]  
   
 # Difference equations for compartments X17 - X20 (individuals who are viral suppressed)  
   
 X17[t,s,g] <- X17[t,s,g-1]+theta\_1[t,s]\*X9[t,s,g-1]-gamma\_5[t,s]\*X17[t,s,g-1]-psi\_1[t,s]\*X17[t,s,g-1]-mu\_9[t,s]\*X17[t,s,g-1]  
   
 X18[t,s,g] <- X18[t,s,g-1]+theta\_2[t,s]\*X10[t,s,g-1]-gamma\_6[t,s]\*X18[t,s,g-1]-psi\_2[t,s]\*X18[t,s,g-1]-mu\_10[t,s]\*X18[t,s,g-1]  
   
 X19[t,s,g] <- X19[t,s,g-1]+theta\_3[t,s]\*X11[t,s,g-1]-gamma\_7[t,s]\*X19[t,s,g-1]-psi\_3[t,s]\*X19[t,s,g-1]-mu\_11[t,s]\*X19[t,s,g-1]  
   
 X20[t,s,g] <- X20[t,s,g-1]+theta\_4[t,s]\*X12[t,s,g-1]+tau\_1[t,s]\*X16[t,s,g-1]-gamma\_8[t,s]\*X20[t,s,g-1]-psi\_4[t,s]\*X20[t,s,g-1]-mu\_12[t,s]\*X20[t,s,g-1]  
   
 # Difference equations compartments X21 - X24 (individuals who are not viral suppressed)  
   
 X21[t,s,g] <- X21[t,s,g-1]+psi\_1[t,s]\*X17[t,s,g-1]-gamma\_9[t,s]\*X21[t,s,g-1]-mu\_13[t,s]\*X21[t,s,g-1]  
   
 X22[t,s,g] <- X22[t,s,g-1]+psi\_2[t,s]\*X18[t,s,g-1]-gamma\_10[t,s]\*X22[t,s,g-1]-mu\_14[t,s]\*X22[t,s,g-1]  
   
 X23[t,s,g] <- X23[t,s,g-1]+psi\_3[t,s]\*X19[t,s,g-1]-gamma\_11[t,s]\*X23[t,s,g-1]-mu\_15[t,s]\*X23[t,s,g-1]  
   
 X24[t,s,g] <- X24[t,s,g-1]+psi\_4[t,s]\*X20[t,s,g-1]-gamma\_12[t,s]\*X24[t,s,g-1]-mu\_16[t,s]\*X24[t,s,g-1]  
   
 # Difference equations for compartment D (individuals who have died)  
   
 D[t,s,g] <- D[t,s,g-1]+mu\_1[t,s]\*(X1[t,s,g-1]+X5[t,s,g-1]+X13[t,s,g-1])+mu\_2[t,s]\*(X2[t,s,g-1]+X6[t,s,g-1]+X14[t,s,g-1])+mu\_3[t,s]\*(X3[t,s,g-1]+X7[t,s,g-1]+X15[t,s,g-1])+ mu\_4[t,s]\*(X4[t,s,g-1]+X8[t,s,g-1]+X16[t,s,g-1])+mu\_5[t,s]\*X9[t,s,g-1]+mu\_6[t,s]\*X10[t,s,g-1]+mu\_7[t,s]\*X11[t,s,g-1]+mu\_8[t,s]\*X12[t,s,g-1]+mu\_9[t,s]\*X17[t,s,g-1]+mu\_10[t,s]\*X18[t,s,g-1]+mu\_11[t,s]\*X19[t,s,g-1]+mu\_12[t,s]\*X20[t,s,g-1]+mu\_13[t,s]\*X21[t,s,g-1]+mu\_14[t,s]\*X22[t,s,g-1]+mu\_15[t,s]\*X23[t,s,g-1]+mu\_16[t,s]\*X24[t,s,g-1]

# Calibration targets #

# Import calibration targets #

files <- list.files(path="~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Calibration/", pattern = "csv")

setwd("~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Calibration")

Target\_all <- read.csv("Targets\_00.csv")

# EPIDMIC CURVES

# Generate predicted outcomes for overall population

for (t in (1:trials)){

for (g in (1:generations)) {

NI\_total[t,1,g] <- 0

N\_total[t,1,g] <- 0

Z\_total[t,1,g] <- 0

V\_total[t,1,g] <- 0

S\_total[t,1,g] <- 0

Diagnosed\_total[t,1,g] <- 0

New\_Diagnosed\_total[t,1,g] <- 0

# Linked\_total[t,s,g] <- 0

# Lost\_total[t,s,g] <- 0

Suppressed\_total[t,1,g] <- 0

# Not\_Suppressed\_total[t,s,g] <- 0

On\_ART\_total[t,1,g] <- 0

for (s in (1:sgroup)){

# Initialize parameter values NI\_total[t,1,g]=NI\_total[t,1,g]+NI[t,s,g]

N\_total[t,1,g]=N\_total[t,1,g]+N[t,s,g]

Z\_total[t,1,g]=Z\_total[t,1,g]+Z[t,s,g]

V\_total[t,1,g]=V\_total[t,1,g]+V[t,s,g]

S\_total[t,1,g]=S\_total[t,1,g]+S[t,s,g]

# Undiagnosed\_total[t,1,g]=Undiagnosed\_total[t,1,g]+Undiagnosed[t,s,g]

Diagnosed\_total[t,1,g]=Diagnosed\_total[t,1,g]+Diagnosed[t,s,g]

New\_Diagnosed\_total[t,1,g]=New\_Diagnosed\_total[t,1,g]+New\_Diagnosed[t,s,g]

# Linked\_total[t,1,g]=Linked\_total[t,1,g]+Linked[t,s,g]

# Lost\_total[t,1,g]=Lost\_total[t,1,g]+Lost[t,s,g]

Suppressed\_total[t,1,g]=Suppressed\_total[t,1,g]+Suppressed[t,s,g]

# Not\_Suppressed\_total[t,1,g]=Not\_Suppressed\_total[t,1,g]+Not\_Suppressed[t,s,g]

On\_ART\_total[t,1,g]=On\_ART\_total[t,1,g]+On\_ART[t,s,g]

}

}

}

# Generate annual outcomes for overall population

library(zoo)

for (t in (1:trials)) {

# Assign a vector to store the year variables

year <- seq(from=2004, to=2039, by=1)

# Annual prevalence (end of the year)

Prevalence\_total[t,1,] <- (Z\_total[t,1,]+V\_total[t,1,])/(N\_total[t,1,]+S\_total[t,1,])

Prevalence\_total\_y[t,] <- Prevalence\_total[t,1,][seq(12,length(Prevalence\_total[t,1,]),12)]

# Annual incidence (end of the year)

New\_Diagnosed\_total\_y[t,] <- rollapply(New\_Diagnosed\_total[t,1,],12,sum,by=12)

S\_total\_y[t,] <- S\_total[t,1,][seq(12, length(S\_total[t,1,]), 12)]

Incidence\_total\_y[t,] <- New\_Diagnosed\_total\_y[t,]/S\_total\_y[t,]

# Annual number of people on ART

On\_ART\_total\_ny[t,] <- (On\_ART\_total[t,1,])[seq(12,length(On\_ART\_total[t,1,]),12)]

# Annual proportion on ART

On\_ART\_total\_y[t,] <- (On\_ART\_total[t,1,]/Diagnosed\_total[t,1,])[seq(12,length(On\_ART\_total[t,1,]),12)]

# Annual proportion viral suppressed

Suppressed\_total\_y[t,] <- (Suppressed\_total[t,1,]/N\_total[t,1,])[seq(12,length(Suppressed\_total[t,1,]),12)]

# Annual proportion viral suppressed when on ART

Suppressed\_onART\_total\_y[t,] <- (Suppressed\_total[t,1,]/On\_ART\_total[t,1,])[seq(12,length(Suppressed\_total[t,1,]),12)]

}

for (t in (1:trials)) {

# Initialize calibration storage parameters

sls\_prev\_total[t] <- 0

sls\_incidence\_total[t] <- 0

sls\_onARTN\_total[t] <- 0

sls\_onARTP\_total[t] <- 0

sls\_vs\_total[t] <- 0

sls\_vsonART\_total[t] <- 0

sls\_total[t] <- 0

pd\_prev\_total[t] <- 0

pd\_incidence\_total[t] <- 0

pd\_onARTN\_total[t] <- 0

pd\_onARTP\_total[t] <- 0

pd\_vs\_total[t] <- 0

pd\_vsonART\_total[t] <- 0

pd\_total[t] <- 0

# Calculate sum of least squares for each trial

for (y in (1:length(Prevalence\_total\_y[t,]))) {

# Prevalence

Target\_prev[t,y]<- Target\_all[y,"Prevalence"]

if (is.na(Target\_prev[t,y]))

{

sls\_prev[t,y] <- NA

sls\_prev\_total[t] <- sls\_prev\_total[t]

pd\_prev[t,y] <- NA

pd\_prev\_total[t] <- pd\_prev\_total[t]

} else {

sls\_prev\_total[t] <- sls\_prev\_total[t]+(Prevalence\_total\_y[t,y]-Target\_prev[t,y]/100)^2

sls\_prev[t,y] <- (Prevalence\_total\_y[t,y]-Target\_prev[t,y]/100)^2

pd\_prev\_total[t] <- pd\_prev\_total[t]+abs(Prevalence\_total\_y[t,y]-Target\_prev[t,y]/100)/(Target\_prev[t,y]/100)

pd\_prev[t,y] <- abs(Prevalence\_total\_y[t,y]-Target\_prev[t,y]/100)/(Target\_prev[t,y]/100)

}

}

for (y in (1:length(Incidence\_total\_y[t,]))) {

#Incidence

Target\_incidence[t,y]<- Target\_all[y,"Incidence"]

if (is.na(Target\_incidence[t,y]))

{

sls\_incidence[t,y] <- NA

sls\_incidence[t] <- sls\_incidence[t]

pd\_incidence[t,y] <- NA

pd\_incidence[t] <- pd\_incidence[t]

} else {

sls\_incidence\_total[t] <- sls\_incidence\_total[t]+(Incidence\_total\_y[t,y]-Target\_incidence[t,y]/100)^2

sls\_incidence[t,y] <- (Incidence\_total\_y[t,y]-Target\_incidence[t,y]/100)^2

pd\_incidence\_total[t] <- pd\_incidence\_total[t]+abs(Incidence\_total\_y[t,y]-Target\_incidence[t,y]/100)/ (Target\_incidence[t,y]/100)

pd\_incidence[t,y] <- abs(Incidence\_total\_y[t,y]-Target\_incidence[t,y]/100)/ (Target\_incidence[t,y]/100)

}

}

for (y in (1:length(On\_ART\_total\_ny[t,]))) {

# Annual number of people on ART

Target\_onART\_N[t,y]<- Target\_all[y,"N\_onART"]

if (is.na(Target\_onART\_N[t,y]))

{

sls\_onARTN[t,y] <- NA

sls\_onARTN\_total[t] <- sls\_onARTN\_total[t]

pd\_onARTN[t,y] <- NA

pd\_onARTN\_total[t] <- pd\_onARTN\_total[t]

} else {

sls\_onARTN\_total[t] <- sls\_onARTN\_total[t]+(On\_ART\_total\_ny[t,y]-Target\_onART\_N[t,y])^2

sls\_onARTN[t,y] <- (On\_ART\_total\_ny[t,y]-Target\_onART\_N[t,y])^2

pd\_onARTN\_total[t] <- pd\_onARTN\_total[t]+abs(On\_ART\_total\_ny[t,y]-Target\_onART\_N[t,y])/Target\_onART\_N[t,y]

pd\_onARTN[t,y] <- abs(On\_ART\_total\_ny[t,y]-Target\_onART\_N[t,y])/Target\_onART\_N[t,y]

}

}

for (y in (1:length(On\_ART\_total\_y[t,]))) {

# Annual proportion on ART

Target\_onART\_P[t,y]<- Target\_all[y,"OnART"]

if (is.na(Target\_onART\_P[t,y]))

{

sls\_onARTP[t,y] <- NA

sls\_onARTP\_total[t] <- sls\_onARTP\_total[t]

pd\_onARTP[t,y] <- NA

pd\_onARTP\_total[t] <- pd\_onARTP\_total[t]

} else {

sls\_onARTP\_total[t] <- sls\_onARTP\_total[t]+(On\_ART\_total\_y[t,y]-Target\_onART\_P[t,y]/100)^2

sls\_onARTP[t,y] <- (On\_ART\_total\_y[t,y]-Target\_onART\_P[t,y]/100)^2

pd\_onARTP\_total[t] <- pd\_onARTP\_total[t]+abs(On\_ART\_total\_y[t,y]-Target\_onART\_P[t,y]/100)/(Target\_onART\_P[t,y]/100)

pd\_onARTP[t,y] <- abs(On\_ART\_total\_y[t,y]-Target\_onART\_P[t,y]/100)/(Target\_onART\_P[t,y]/100)

}

}

for (y in (1:length(Suppressed\_total\_y[t,]))) {

# Annual proportion viral suppressed

Target\_VS[t,y]<- Target\_all[y,"Suppressed"]

if (is.na(Target\_VS[t,y]))

{

sls\_vs[t,y] <- NA

sls\_vs\_total[t] <- sls\_vs\_total[t]

pd\_vs[t,y] <- NA

pd\_vs\_total[t] <- pd\_vs\_total[t]

} else {

sls\_vs\_total[t] <- sls\_vs\_total[t]+(Suppressed\_total\_y[t,y]-Target\_VS[t,y]/100)^2

sls\_vs[t,y] <- (Suppressed\_total\_y[t,y]-Target\_VS[t,y]/100)^2

pd\_vs\_total[t] <- pd\_vs\_total[t]+abs(Suppressed\_total\_y[t,y]-Target\_VS[t,y]/100)/(Target\_VS[t,y]/100)

pd\_vs[t,y] <- abs(Suppressed\_total\_y[t,y]-Target\_VS[t,y]/100)/(Target\_VS[t,y]/100)

}

}

for (y in (1:length(Suppressed\_onART\_total\_y[t,]))) {

# Annual proportion viral suppressed when on ART

Target\_VSonART[t,y]<- Target\_all[y,"Suppresed\_onART"]

if (is.na(Target\_VSonART[t,y]))

{

sls\_vsonART[t,y] <- NA

sls\_vsonART\_total[t] <- sls\_vsonART\_total[t]

pd\_vsonART[t,y] <- NA

pd\_vsonART\_total[t] <- pd\_vsonART\_total[t]

} else {

sls\_vsonART\_total[t] <- sls\_vsonART\_total[t]+(Suppressed\_onART\_total\_y[t,y]-Target\_VSonART[t,y]/100)^2

sls\_vsonART[t,y] <- (Suppressed\_onART\_total\_y[t,y]-Target\_VSonART[t,y]/100)^2

pd\_vsonART\_total[t] <- pd\_vsonART\_total[t]+abs(Suppressed\_onART\_total\_y[t,y]-Target\_VSonART[t,y]/100)/(Target\_VSonART[t,y]/100)

pd\_vsonART[t,y] <- abs(Suppressed\_onART\_total\_y[t,y]-Target\_VSonART[t,y]/100)/(Target\_VSonART[t,y]/100)

}

}

sls\_total[t]=sls\_prev\_total[t] + sls\_incidence\_total[t] + sls\_onARTN\_total[t] + sls\_onARTP\_total[t]+sls\_vs\_total[t]+sls\_vsonART\_total[t]

pd\_total[t]=pd\_prev\_total[t] + pd\_incidence\_total[t] + pd\_onARTN\_total[t] + pd\_onARTP\_total[t]+pd\_vs\_total[t]+pd\_vsonART\_total[t]

}

df\_total <- data.frame(Trial=seq(from=1,to=trials,by=1), SLS\_prev\_total=sls\_prev\_total, SLS\_incidence\_total=sls\_incidence\_total,SLS\_N\_onART\_total=sls\_onARTN\_total, SLS\_P\_onART\_total=sls\_onARTP\_total, SLS\_P\_VS\_total=sls\_vs\_total, SLS\_P\_VSonART\_total=sls\_vsonART\_total, SLS\_total=sls\_total)

write.csv(df\_total,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_Total.csv")

write.csv(sls\_prev,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_Prev.csv")

write.csv(sls\_incidence,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_incidence.csv")

write.csv(sls\_onARTN,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_N\_ART.csv")

write.csv(sls\_onARTP,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_P\_ART.csv")

write.csv(sls\_vs,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_VS.csv")

write.csv(sls\_vsonART,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_VSonART.csv")

df\_total <- data.frame(Trial=seq(from=1,to=trials,by=1), pd\_prev\_total=pd\_prev\_total, pd\_incidence\_total=pd\_incidence\_total,pd\_N\_onART\_total=pd\_onARTN\_total, pd\_P\_onART\_total=pd\_onARTP\_total, pd\_P\_VS\_total=pd\_vs\_total, pd\_P\_VSonART\_total=pd\_vsonART\_total, pd\_total=pd\_total)

write.csv(df\_total,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_Total\_PD.csv")

write.csv(pd\_prev,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_Prev\_PD.csv")

write.csv(pd\_incidence,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_incidence\_PD.csv")

write.csv(pd\_onARTN,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_N\_ART\_PD.csv")

write.csv(pd\_onARTP,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_P\_ART\_PD.csv")

write.csv(pd\_vs,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_VS\_PD.csv")

write.csv(pd\_vsonART,"~/Dropbox/VCU\_PhD\_Year 3&4/GRA/Rwanda-Model/Full model/calibration outcomes\_VSonART\_PD.csv")

#Plot calibration figures

# library(zoo)

#

# for (t in (1:trials)) {

# setEPS()

#

# # # Annual Diagnosed

# # Diagnosed\_total\_y[t,] <- Diagnosed\_total[t,1,][seq(1, length(Diagnosed\_total[t,1,]), 12)]

# # plot(year,Diagnosed\_total\_y[t,],type="l",col="red",xlab="Year", ylab="Number Diagnosed", main="Annual number of diagosed PLWH", ylim=c(0,max(Diagnosed\_total\_y,na.rm=TRUE)))

# #

# # # Annual percent diagnosed

# # Diagnosed\_total\_py[t,] <- (Diagnosed\_total[t,1,]/N\_total[t,1,])[seq(1, length(N\_total[t,1,]), 12)]

# # plot(year,Diagnosed\_total\_py[t,],type="l",col="red",xlab="Year", ylab="Percent Diagnosed", main="Annual percent of diagosed PLWH", ylim=c(0,max(Diagnosed\_total\_py,na.rm=TRUE)))

# #

# # # Annual new infection

# # NI\_total\_y[t,] <- rollapply(unlist(NI\_total[t,1,]),12,sum,by=12)

# # plot(year, NI\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "New HIV Infections", main = "Annual individuals newly infected - Overall", ylim=c(0,max(NI\_total\_y[t,],na.rm=TRUE)))

# #

# # # People living with HIV

# # N\_total\_y[t,] <- N\_total[t,1,][seq(1, length(N\_total[t,1,]), 12)]

# # plot(year, N\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Individuals living with HIV", main = "Annual individuals living with HIV - Overall", ylim=c(0,max(N\_total\_y[t,],na.rm=TRUE)))

# # #

# #

# # HIV prevalence

# postscript("HIV\_prevalence.eps")

# plot(year, 100\*Prevalence\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Prevalence (%)", main = "Annual HIV prevalence - Overall", ylim=c(0,max(100\*Prevalence\_total\_y[t,],Target\_all$Prevalence\_UB,na.rm=TRUE)))

# points(Target\_all$Year, Target\_all$Prevalence)

# arrows (x0=Target\_all$Year, y0=Target\_all$Prevalence\_LB, x1=Target\_all$Year, y1=Target\_all$Prevalence\_UB, code=3, angle=90, length=0.15, col="blue")

# dev.off()

#

# # HIV incidence

# postscript("HIV\_incidence.eps")

# plot(year,100\*Incidence\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Incidence (%)", main = "Annual HIV incidence - Overall", ylim=c(0,max(100\*Incidence\_total\_y[t,],Target\_all$Incidence,Target\_all$Incidence\_UP,na.rm=TRUE)))

# points(Target\_all$Year, Target\_all$Incidence)

# arrows (x0=Target\_all$Year, y0=Target\_all$Incidence\_LB, x1=Target\_all$Year, y1=Target\_all$Incidence\_UP, code=3, angle=90, length=0.15, col="blue")

# dev.off()

#

# # Annual number of individuals on ART

# postscript("HIV\_number on ART.eps")

# plot(year,On\_ART\_total\_ny[t,], type = "l", col="red", xlab = "Year", ylab = "Number of individuals on ART", main = "Number of people on ART - Overall", ylim=c(0,max(On\_ART\_total\_ny[t,],Target\_all$N\_onART,na.rm=TRUE)))

# points(Target\_all$Year, Target\_all$N\_onART)

# # arrows (x0=Target\_all$Year, y0=Target\_all$OnART\_LB, x1=Target\_all$Year, y1=Target\_all$OnART\_UB, code=3, angle=90, length=0.15, col="blue")

# dev.off()

#

# # Annual proportion of individuals on ART

# postscript("HIV\_percent on ART.eps")

# plot(year,100\*On\_ART\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Proportion of individuals on ART (%)", main = "Annual percent on ART - Overall", ylim=c(0,max(100\*On\_ART\_total\_y[t,],Target\_all$OnART,Target\_all$OnART\_UB,na.rm=TRUE)))

# points(Target\_all$Year, Target\_all$OnART)

# arrows (x0=Target\_all$Year, y0=Target\_all$OnART\_LB, x1=Target\_all$Year, y1=Target\_all$OnART\_UB, code=3, angle=90, length=0.15, col="blue")

# dev.off()

#

# # Annual proportion of individuals virally suppressed regardless of known HIV status or on treatment.

# postscript("HIV\_percent suppressed.eps")

# plot(year,100\*Suppressed\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Proportion of individuals viral suppressed (%)", main = "Annual percent viral suppressed - Overall", ylim=c(0,max(100\*Suppressed\_total\_y[t,], Target\_all$Suppressed\_UP,Target\_all$Suppressed,na.rm=TRUE)))

# points(Target\_all$Year, Target\_all$Suppressed)

# arrows (x0=Target\_all$Year, y0=Target\_all$Suppressed\_LB, x1=Target\_all$Year, y1=Target\_all$Suppressed\_UP, code=3, angle=90, length=0.15, col="blue")

# dev.off()

#

# # Annual proportion of individuals virally suppressed conditional on individuals on ART.

# postscript("HIV\_percent suppresed when on ART.eps")

# plot(year,100\*Suppressed\_onART\_total\_y[t,], type = "l", col="red", xlab = "Year", ylab = "Proportion of individuals viral suppressed (%)", main = "Annual percent on ART and viral suppressed - Overall", ylim=c(0,max(100\*Suppressed\_onART\_total\_y[t,], Target\_all$Suppresed\_onART\_UB,Target\_all$Suppresed\_onART,na.rm=TRUE)))

# points(Target\_all$Year, Target\_all$Suppresed\_onART)

# arrows (x0=Target\_all$Year, y0=Target\_all$Suppresed\_onART\_LB, x1=Target\_all$Year, y1=Target\_all$Suppresed\_onART\_UB, code=3, angle=90, length=0.15, col="blue")

# dev.off()

# }

# References

1. World Health Organization. *Antiretroviral Therapy for HIV Infection in Adults and Adolescents: Recommendations for a Public Health Approach*.; 2006.

2. World Health Organization. *Antiretroviral Therapy for HIV Infection in Adults and Adolescents: Recommendations for a Public Health Approach*.; 2010.

3. World Health Organization. *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach*.; 2013. https://apps.who.int/iris/bitstream/handle/10665/85321/9789241505727\_eng.pdf?sequence=1

4. World Health Organization. *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach*.; 2016. doi:10.1016/j.jped.2014.04.007

5. Nsanzimana S, Kanters S, Remera E, et al. HIV care continuum in Rwanda: A cross-sectional analysis of the national programme. *Lancet HIV*. 2015;2(5):e208-e215. doi:10.1016/S2352-3018(15)00024-7

6. Nsanzimana S, Remera E, Kanters S, et al. Life expectancy among HIV-positive patients in Rwanda: A retrospective observational cohort study. *Lancet Glob Heal*. 2015;3(3):e169-e177. doi:10.1016/S2214-109X(14)70364-X

7. Ministry of Health (MOH) [Rwanda]. Circular of key changes in HIV prevention and management guidelines.

8. Suthar AB, Granich RM, Kato M, Nsanzimana S, Montaner JSG, Williams BG. Programmatic implications of acute and early HIV infection. *J Infect Dis*. 2015;212(9):1351-1360. doi:10.1093/infdis/jiv430

9. Consortium WTS. Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. *Lancet*. 2009;373(9672):1352-1363.

10. Ross J, Sinayobye J d’Amour, Yotebieng M, et al. Early outcomes after implementation of *treat all* in Rwanda: an interrupted time series study. *J Int AIDS Soc*. 2019;22(4):e25279. doi:10.1002/jia2.25279

11. Nsanzimana S, Remera E, Kanters S, et al. Effect of baseline cd4 cell count at linkage to hiv care and at initiation of antiretroviral therapy on mortality in hivpositive adult patients in Rwanda: A nationwide cohort study. *Lancet HIV*. 2015;2(9):e376-e384. doi:10.1016/S2352-3018(15)00112-5

12. Sharma M, Ying R, Tarr G, Barnabas R, Division ID, Hutchinson F. A systematic review and meta-analysis of community and facility-based approaches to address gaps in HIV testing and linkage in sub-Saharan Africa. *Nature*. 2015;528(7580):S77-S85. doi:10.1038/nature16044.A

13. Lancaster KE, Cernigliaro D, Zulliger R, et al. HIV care and treatment experiences among female sex workers living with HIV in sub Saharan Africa: A systematic review. 2017;15(4):377-386. doi:10.2989/16085906.2016.1255652.HIV

14. Bendavid E, Stauffer D, Remera E, Nsanzimana S, Kanters S, Mills EJ. Mortality along the continuum of HIV care in Rwanda: A model-based analysis. *BMC Infect Dis*. 2016;16(1):1-9. doi:10.1186/s12879-016-2052-7

15. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection undetectable equals untransmittable. *JAMA - J Am Med Assoc*. Published online 2019. doi:10.1001/jama.2018.21167

16. Kitahata MM, Gange SJ, Abraham AG, et al. Effect of early versus deferred antiretroviral therapy for HIV on survival. *N Engl J Med*. 2009;360(18):1815-1826.

17. Mutagoma M, Kayitesi C, Gwiza A, et al. Estimation of the size of the female sex worker population in Rwanda using three different methods. *Int J STD AIDS*. 2015;26(11):810-814.

18. National Institute of Statistics of Rwanda (NISR) [Rwanda] Ministry of Health (MOH) [Rwanda] and ICF International. *Rwanda Demographic and Health Survey 2014-15*.; 2015. doi:March, 2016

19. Population-based HIV Impact Assessment. Rwanda population-based HIV impact assessment, RPHIA 2018–2019. (February 2020):2-8.

20. National Institute of Statistics of Rwanda. *Demographic and Health Survey Report 2010*. Vol 2003.; 2010. doi:10.1017/S0266673100000246

21. UNAIDS. *Prevention Gap Report 2016*. Vol 83.; 2016. doi:10.1371/journal.pone.0154893.

22. National Institute of Statistics of Rwanda (NISR) [Rwanda] Ministry of Health (MOH) [Rwanda] and ICF International. Demographic and Health Survey, 2005. Rwanda Mininstry of Health.

23. Ross J, Ribakare M, Remera E, et al. High levels of viral load monitoring and viral suppression under Treat All in Rwanda – a cross-sectional study. *J Int AIDS Soc*. Published online 2020. doi:10.1002/jia2.25543

24. Ramjee G, Daniels B, Ramjee G, et al. Women and HIV in Sub-Saharan Africa. *AIDS Res Ther*. 2013;10(1):30. doi:10.1186/1742-6405-10-30

25. Test FS, Mehta SD, Handler A, Mutimura E, Bamukunde AM, Cohen M. Gender inequities in sexual risks among youth with HIV in Kigali, Rwanda. *Int J STD AIDS*. 2012;23(6):394-399.

26. Mutagoma M, Samuel MS, Kayitesi C, et al. High HIV prevalence and associated risk factors among female sex workers in Rwanda. *Int J STD AIDS*. 2017;28(11):095646241668813. doi:10.1177/0956462416688137

27. Mutagoma M, Nyirazinyoye L, Riedel DJ, Ntaganira J. Sexual risk behaviors and practices of female sex workers in Rwanda in over a decade , 2006 – 2015. *Int J STD AIDS*. 2018;1(0):1-8. doi:10.1177/0956462418785297

28. PEPFAR. *Rwanda Country Operational Plan (COP/ROP) 2018 Strategic Direction Summary*.; 2018.

29. National Institute of Statistics of Rwanda (NISR) [Rwanda] Ministry of Health (MOH) [Rwanda] and ICF International. Demographic and Health Survey, 2015. Rwanda Mininstry of Health.

30. Rwanda Biomedical Center. *Republic of Rwanda Ministry of Health NATIONAL HIV / AIDS TARGETS 2018-2020-2030*.; 2015.

31. Ingabire R, Parker R, Nyombayire J, et al. Female sex workers in Kigali, Rwanda: a key population at risk of HIV, sexually transmitted infections, and unplanned pregnancy. *Int J STD AIDS*. Published online 2019. doi:10.1177/0956462418817050

32. Mountain E, Pickles M, Mishra S, Vickerman P, Alary M, Boily MC. The HIV care cascade and antiretroviral therapy in female sex workers: Implications for HIV prevention. *Expert Rev Anti Infect Ther*. Published online 2014. doi:10.1586/14787210.2014.948422

33. Nsanzimana S, Remera E, Kanters S, et al. Household survey of HIV incidence in Rwanda: a national observational cohort study. *Lancet HIV*. 2017;4(10):e457-e464.

34. Kerr CC, Stuart RM, Gray RT, et al. Optima: a model for HIV epidemic analysis, program prioritization, and resource optimization. *JAIDS J Acquir Immune Defic Syndr*. 2015;69(3):365-376.

35. Bernard CL, Brandeau ML, Humphreys K, et al. Cost-effectiveness of HIV preexposure prophylaxis for people who inject drugs in the United States. *Ann Intern Med*. 2016;165(1):10-19. doi:10.7326/M15-2634

36. Rwanda Biomedical Center. *National HIV Annual Report, 2013-2014*.; 2014.

37. Boily M-C, Baggaley RF, Wang L, et al. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infect Dis*. 2009;9(2):118-129.

38. Powers KA, Poole C, Pettifor AE, Cohen MS. Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis. *Lancet Infect Dis*. Published online 2008. doi:10.1016/S1473-3099(08)70156-7

39. Weller SC, Davis‐Beaty K. Condom effectiveness in reducing heterosexual HIV transmission. *Cochrane database Syst Rev*. 2002;(1).

40. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. *J Infect Dis*. Published online 2012. doi:10.1093/infdis/jir747

41. Alsallaq RA, Buttolph J, Cleland CM, et al. The potential impact and cost of focusing HIV prevention on young women and men: A modeling analysis in western Kenya. *PLoS One*. 2017;12(4):e0175447.

42. Sorensen SW, Sansom SL, Brooks JT, et al. A mathematical model of comprehensive test-and-treat services and HIV incidence among men who have sex with men in the United States. *PLoS One*. 2012;7(2):e29098.

43. National Institute of Statistics of Rwanda (NISR) [Rwanda] Ministry of Health (MOH) [Rwanda] and ICF International. Demographic and Health Survey, 2010. Rwanda Mininstry of Health.

44. Rwanda Ministry of Health. *Behavioral and Biological Surveillance Survey among Female Sex Workers, Rwanda – 2010*.; 2010. doi:10.1016/B978-0-12-420118-7.00008-1.Dopamine

45. Supervie V, Viard JP, Costagliola D, Breban R. Heterosexual risk of HIV transmission per sexual act under combined antiretroviral therapy: Systematic review and bayesian modeling. *Clin Infect Dis*. Published online 2014. doi:10.1093/cid/ciu223

46. Braunstein SL, Ingabire CM, Geubbels E, et al. High burden of prevalent and recently acquired HIV among female sex workers and female HIV voluntary testing center clients in Kigali, Rwanda. *PLoS One*. Published online 2011. doi:10.1371/journal.pone.0024321

47. Braunstein SL, Umulisa MM, Veldhuijzen NJ, et al. HIV diagnosis, linkage to HIV Care, and HIV risk behaviors among newly diagnosed HIV-positive female sex workers in Kigali, Rwanda. *J Acquir Immune Defic Syndr*. 2011;57(4):70-76. doi:10.1097/QAI.0b013e3182170fd3

48. World Bank. Population Estimates And Projections. Published 2020. Accessed June 8, 2020. https://datacatalog.worldbank.org/dataset/population-estimates-and-projections.

49. Joint United Nations Programme on HIV/AIDS. Country factsheets: Rwanda 2019. Published 2021. Accessed May 1, 2021. https://www.unaids.org/en/regionscountries/countries/rwanda

50. Joint United Nations Programme on HIV/AIDS. Sex workers: Population size estimate. Published 2015. http://data.un.org/Data.aspx?d=UNAIDS&f=inID%3A111

51. National Institute of Allergy and Infectious Diseases. International Epidemiological Databases to Evaluate AIDS. https://www.iedea.org/

52. Institut National de la Statistique du Rwanda (INSR) and ORC Macro. *Rwanda Demographic and Health Survey 2007-2008*.; 2007.

53. Braunstein SL, Umulisa M-M, Veldhuijzen NJ, et al. HIV Diagnosis, Linkage to HIV Care, and HIV Risk Behaviors Among Newly Diagnosed HIV-Positive Female Sex Workers in Kigali, Rwanda. *JAIDS J Acquir Immune Defic Syndr*. 2011;57(4):e70-e76. doi:10.1097/QAI.0b013e3182170fd3

54. Rachlis B, Ochieng D, Geng E, et al. Evaluating outcomes of patients lost to follow-up in a large comprehensive care treatment program in Western Kenya. *J Acquir Immune Defic Syndr*. 2015;68(4):e46-e55. doi:10.1097/QAI.0000000000000492

55. Geng EH, Bwana MB, Muyindike W, et al. Failure to initiate antiretroviral therapy, loss to follow-up and mortality among HIV-infected patients during the pre-ART period in Uganda. *J Acquir Immune Defic Syndr*. 2013;63(2):64-71. doi:10.1097/QAI.0b013e31828af5a6

56. Haas AD, Zaniewski E, Anderegg N, et al. Retention and mortality on antiretroviral therapy in sub-Saharan Africa: Collaborative analyses of HIV treatment programmes: Collaborative. *J Int AIDS Soc*. 2018;21(2):1-7. doi:10.1002/jia2.25084

57. Okal DO, Oyaro B, Zeh C, et al. Effect of point-of-care CD4 cell count results on linkage to care and antiretroviral initiation during a home-based HIV testing campaign: a non-blinded, cluster-randomised trial. *Lancet HIV*. 2017;4(9):e393-e401. doi:10.1016/s2352-3018(17)30091-7

58. Smith JA, Sharma M, Levin C, et al. Cost-effectiveness of community-based strategies to strengthen the continuum of HIV care in rural South Africa: A health economic modelling analysis. *Lancet HIV*. 2015;2(4):e159-e168. doi:10.1016/S2352-3018(15)00016-8

59. Rwanda Ministry of Health. *Rwanda HIV and AIDS National Strategic Plan 2013–2018: Extension: 2018–2020*.; 2018.

60. Stalter R, Chen M, Uwizeye G, et al. Association of sexual risk behaviour with previous HIV testing among voluntary HIV counselling and testing clients in Kigali, Rwanda. *Int J STD AIDS*. 2016;27(14):1317-1325. doi:10.1177/0956462415617590

61. Nsanzimana S, Mills EJ, Harari O, et al. Prevalence and incidence of HIV among female sex workers and their clients: modelling the potential effects of intervention in Rwanda. *BMJ Glob Heal*. 2020;5(8):e002300.

62. Langford SE, Ananworanich J, Cooper DA. Predictors of disease progression in HIV infection: A review. *AIDS Res Ther*. Published online 2007. doi:10.1186/1742-6405-4-11

63. Brinkhof MWG, Dabis F, Myer L, et al. Early loss of HIV-infected patients on potent antiretroviral therapy programmes in lower-income countries. *Bull World Health Organ*. 2008;86(7):559-567. doi:10.2471/BLT.07.044248

64. Nuwagaba-Biribonwoha H, Kiragga AN, Yiannoutsos CT, et al. Adolescent pregnancy at antiretroviral therapy (ART) initiation: a critical barrier to retention on ART. *J Int AIDS Soc*. 2018;21(9):1-9. doi:10.1002/jia2.25178

65. Grimsrud A, Cornell M, Schomaker M, Fox MP. CD4 count at antiretroviral therapy initiation and the risk of loss to follow-up: results from a multicentre cohort study. *J Epidemiol Community Health*. 2016;70(6):549-555. doi:10.1136/jech-2015-206629.CD4

66. Johnson LF, Anderegg N, Zaniewski E, et al. Global variations in mortality in adults after initiating antiretroviral treatment: An updated analysis of the International epidemiology Databases to Evaluate AIDS cohort collaboration. *Aids*. 2019;33(July):S283-S294. doi:10.1097/QAD.0000000000002358

67. Rwanda Ministry of Health. *National Guidelines for Prevention and Management of HIV and STIs. Edition 201 6*.; 2016.

68. Rwanda Ministry of Health. *National Guidelines for Prevention and Managment of HIV - Edition 2018*.; 2018.

69. Drummond MF, O’Brien B, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes, Second Edition. *Am J Prev Med*. Published online 1998. doi:10.1016/S0749-3797(97)00069-X

70. Rwanda Biomedical Center. *National HIV Annual Report 2014-2015*.; 2015.

71. Ministry of Health (MOH) [Rwanda]. National HIV Annual Report 2013-2014. Published online 2014.

72. Kong CY, McMahon PM, Gazelle GS. Calibration of disease simulation model using an engineering approach. *Value Heal*. Published online 2009. doi:10.1111/j.1524-4733.2008.00484.x

73. Enns EA, Cipriano LE, Simons CT, Kong CY. Identifying best-fitting inputs in health-economic model calibration: A pareto frontier approach. *Med Decis Mak*. Published online 2015. doi:10.1177/0272989X14528382

# Appendix A

## Datasets

1. The population estimation and projections data (1960-2050). Downloaded on Feb 15, 2019. Accessible at: <https://datacatalog.worldbank.org/dataset/population-estimates-and-projections>

2. Demographic and Health Survey data (2005). Downloaded on Nov 26, 2019. Accessible at: <https://dhsprogram.com/what-we-do/survey/survey-display-252.cfm>

3. Demographic and Health Survey data (2010). Downloaded on Dec 02, 2019. Accessible at: <https://dhsprogram.com/what-we-do/survey/survey-display-364.cfm>

4. Demographic and Health Survey data (2015). Downloaded on Dec 02, 2019. Accessible at: <https://dhsprogram.com/what-we-do/survey/survey-display-468.cfm>

## Codes and documents to derive parameter inputs

1. The code to derive probability of consistent condom use. Accessible at: T:\Health Behavior and Policy\Faculty\Kimmel\Common\Personnel work\Deo\IeDEA\HIV Transmission Model-Rwanda\Parameter Inputs\Consistent condom use\consistent condom use.dta
2. The excel to estimate the sub-population growth rate. Accessible at: T:\Health Behavior and Policy\Faculty\Kimmel\Common\Personnel work\Deo\IeDEA\HIV Transmission Model-Rwanda\Susceptible Population\Popn Projection Method\Growth rate sub-group, 2003-2045.xlsx
3. The code to derive probability of HIV diagnosis. Accessible at: T:\Health Behavior and Policy\Faculty\Kimmel\Common\Personnel work\Deo\IeDEA\HIV Transmission Model-Rwanda\Parameter Inputs\Probability of diagnosis\Data\RW\_DHS05\_14.dta
4. The code to derive HIV prevalence as calibration targets. Accessible at: Pending
5. The codes using IeDEA data: Pending.

## System of Differential Equations

*Susceptible Population*

*Infected, Undiagnosed*

*Infected, Diagnosed*

*Infected, Diagnosed and Linked to care*

*Infected, Lost from Care*

*Infected, On ART and Virally Suppressed*

*Infected, On ART and Not Virally Suppressed*

*Dead*

## Methods to estimate CD4 strata and time-at-risk patient living with HIV

Estimation of stratum-specific residence time will be used to inform a variety of analysis including estimated for natural history, LTFU, On ART and viral suppressed, On ART and not viral suppressed, and death. It will be performed sequentially in the steps below:

* Estimate a trendline for CD4 cell count. A linear trendline will be estimated based on a patient’s Pre-ART CD4 cell count measurements starting from time of enrollment in care (i.e., first clinic date recorded) until the censor date of ART initiation or, given no ART initiation, the date of death, transfer, or database closure.
* Estimate the CD4 cell count at the date when event or competing risks happens. We will use the simple linear regression expression to estimate the CD4 when the event or competing risk happens based on the time interval ( between enrollment date and date of event or competing risks. The event and competing risks are defined depending on the analysis it informs:
  + For disease progression process:
    - Censor: Database closer, transfer, LTFU
    - Event: Disease progression
    - Competing risk: Death, ART initiation
  + For LTFU:
    - Censor: Database closer, transfer, ART initiation for pre-ART
    - Event: LTFU
    - Competing risk: Death
  + For ART and viral suppressed:
    - Censor: Database closer, transfer, ART initiation
    - Event: Viral suppression
    - Competing risk: Death, LTFU
  + For ART and not viral suppressed
    - Censor: Database closer, transfer
    - Event: Virological failure
    - Competing risk: Death, LTFU
  + For Death:
    - Censor: Database closer, transfer, LTFU, ART initiation for pre-ART
    - Event: Death
* Estimate time interval to reach CD4 stratum threshold. We will use the fitted regression line in step 1 to estimate the time interval (for example, in days) each patient takes to reach each of the CD4 cell count stratum thresholds (500, 350, and 200) before Event date.
* Calculate CD4 stratum-specific residence time. We calculate CD4 stratum-specific residence time, which represents the time at risk for LTFU within a given CD4 cell count stratum, by taking the difference in time to reach successive CD4 stratum thresholds before Event date.
* Step 5: Assign Event value: We will assign the events a censor (), event (), and competing risk (). Death is the only exception where no competing risk is assigned.

## Average monthly population growth rate, by sub-populationa

|  |  |  |
| --- | --- | --- |
| **Sub-population** | **Monthly growth rate** | **95% Confidence Interval** |
| Low-risk urban women (15-24) | 0.20% | (0.17%, 0.23%) |
| Low-risk urban women (25-34) | 0.30% | (0.26%, 0.34%) |
| Low-risk urban women (35-44) | 0.35% | (0.30%, 0.39%) |
| Low-risk urban women (45-54) | 0.36% | (0.31%, 0.41%) |
| Low-risk urban women (55-64) | 0.45% | (0.40%, 0.50%) |
| Low-risk rural women (15-24) | 0.08% | (0.05%, 0.11%) |
| Low-risk rural women (25-34) | 0.18% | (0.14%, 0.23%) |
| Low-risk rural women (35-44) | 0.23% | (0.17%, 0.28%) |
| Low-risk rural women (45-54) | 0.24% | (0.19%, 0.30%) |
| Low-risk rural women (55-64) | 0.33% | (0.29%, 0.37%) |
| Urban men (15-24) | 0.21% | (0.18%, 0.24%) |
| Urban men (25-34) | 0.32% | (0.28%, 0.35%) |
| Urban men (35-44) | 0.33% | (0.27%, 0.39%) |
| Urban men (45-54) | 0.36% | (0.30%, 0.42%) |
| Urban men (55-64) | 0.48% | (0.42%, 0.53%) |
| Rural men (15-24) | 0.09% | (0.06%, 0.12%) |
| Rural men (25-34) | 0.20% | (0.15%, 0.25%) |
| Rural men (35-44) | 0.21% | (0.16%, 0.27%) |
| Rural men (45-54) | 0.24% | (0.18%, 0.30%) |
| Rural men (55-64) | 0.36% | (0.30%, 0.41%) |

1. The monthly population growth rate is an average monthly growth between 2003 and 2045.

## Comparing different for measuring the goodness of fit in the calibration process.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **Advantage** | **Disadvantages** | **Discussion** | **Source** |
| **Visual – Calibration plot**  This is a qualitative approach of measuring the goodness-of-fit by visually comparing the model output with the calibration targets. This goodness of fit is determined by the analyst and no theoretical justification or consensuses on how that is determined.  Calibration targets with/without confidence interval are the only data required to perform the method. | The method is simple and does not require complex statistical methods. | This approach involves subjective judgment and there’s no clear methodology for determining the goodness of fit.  The approach can be difficult to justify. | This method is not selected because it is not reliable and has not methodological approach to justify it. However, the visual method can be used to identify the parameter inputs included in the calibration process.  This method can also be inappropriate when multiple calibration targets are used. | 2 |
| **Sum of least-squares**  This method minimizes the sum of square errors. The sum of square errors is calculated by taking the difference between the model output values and the calibration target values. Parameter values that minimize the sum of errors are selected.  Point estimates of calibration targets are required to perform the method. The sum of the squares of the differences between the observed data and the model predictions should be at the global minimum.  **Assumptions:**  Statistically, the least squares method assumes the random error in each data point is normally distributed and the constant variance (meaning the data is homoscedastic) across all data points. | The sum of least squares method is intuitive and easier to implement. | The sum of least squares method does take into account the precision (uncertainty in the data) of the observed data used for calibration. That means, the method assumes all the data points have the same variance which is not always true.  The least squares method may generate biased estimates when the assumptions are violated. | The sum of least squares method is one of the commonly used methods for measuring the goodness of fit. Although the assumption of normality is usually not satisfied when calibrating to external targets, the least squares method generates acceptable model fits. | 2,3 |
| **Percentage deviation**  The percentage deviation method minimizes the deviation from the observed value. The percentage deviation is calculated as the absolute value of the difference between projected and observed value divided by the observed value.  Point estimates of calibration targets are required to perform the method. | The percentage deviation is intuitive and is comparable between calibration targets with different units | The percentage deviation is not a common method used in the calibration paper. | The percentage deviation generates acceptable goodness of fit scores especially when multiple calibration targets co-exist. |  |
| **Weighted Least-squares**  The weighted least squares (WLS) method is a continuation of the sum of least-squares method. The difference is that the WLS takes into account the variations in the calibration targets.  WLS is a special case of a generalized least squares methods. WLS works by allocating a weight to each data point. The size of the weight shows the precision of information in the associated observation. Optimizing the weighted model fitting criterion to find parameter estimates allows the weights to determine the contribution of each data point to the final parameter estimates.  One way of allocating weights when the standard deviation is not constant across all data points is to use the inverse of the variance for each data point.  Point estimates of the calibration targets and the variances are required to perform the analysis. | The weighted least squares method considers the variation in data points which generates unbiased parameter estimates.  WLS method can produce reasonably good parameter estimates with small data sets | The weighted least squares method depends on the assumption that the weights are known. When weights are not known, and the estimated weights are not close to the true weight value, the parameter estimates will not be accurate.  Weighted least squares method is also sensitive to the effects of outliers. If potential outliers are not investigated and dealt with appropriately, they will likely have a negative impact on the parameter estimation. | The weighted least square is not applicable since we do not always have the variance to calculate weights that applied to the data points (e.g. the number of patients on ART and the percent of viral suppression condition on ART). | 4 |
| **Chi-square**  The chi-square method uses the similar approach as the sum of least-squares method discussed above. The difference is that the chi-square divides the sum of least squares by the standard deviation to take into account the variation across data points.  Point estimates of the calibration targets and the variances are required to perform the analysis.  **Assumption**  Statistically, the chi-square method assumes that the data points are independent of each other and normally distributed. | The chi-square method divides the sum of square errors by the standard deviation, which gives more weight to calibration targets with lower standard deviations. | The chi-square method is highly sensitive to the sample size. This method requires a sufficient sample size to have a valid chi-square approximation.  The chi-square method does not test whether the assumption of normality has been fulfilled but it assumes normality in the data. If the assumption is significantly violated, then the test will not be valid. | The chi-square method is somewhat similar to the WSL method as it accounts for variation in calibration points by diving the sum of squares by the standard deviation. The method requires the standard deviation of the data point to be known.  The chi-square method may not be applicable since the standard deviation is not available in some calibration targets. | 2 |
| **Likelihood method**  Maximum likelihood is a statistical method used for estimating unknown parameters of a probability model by maximizing the likelihood function to generate a set of parameter estimates that fit the empirical data.  Individual-level data associated with calibration targets will be needed to derive the likelihood function. | The likelihood method considers the level of uncertainty in the observed data, which generates reliable parameter estimates.  Maximum likelihood is a consistent measure of parameter estimation cross a variety of estimation situations and hence it can be applicable to multiple problems.  Most statistical softwares support maximum likelihood estimation which minimizes computation errors. | The likelihood method requires more empirical data to estimate parameter values. The likelihood method is sensitive to small samples in empirical data and can produce biased estimates.  The likelihood method is sensitive to the starting point. This means that the estimated values can be change depending on the starting point. Multiple starting points are recommended. | The maximum likelihood method might not be selected since we have no individual-level data for our calibration targets. | 2,5 |
| **Multiple goodness of fit (GOF)estimates**  This method uses multiple calibration targets to obtain a combined measure of goodness of fit across all calibration targets. In this approach, individual calibration target is treated as independent target and then sum the GOF measures across the different targets.  Two methods can be used perform this task are the global criterion and lexicographic method. The global criterion method is the sum of GOF of each weighted calibration target. The weights are allocated based on importance by the analyst.  For the lexicographic approach, the calibration targets are ranked in order of importance, and the process of finding the optimal parameter values is carried out step by step, starting with the most important calibration target and proceeding according to the order of importance. | The multiple goodness of fit method is helpful in circumstances where the calibration targets need to be calibrated independently. | For the lexicographic approach, the calibration targets are ranked in order of importance but not specific criteria is given to rank the calibration targets which leave a lot to personal judgment.  The global criterion method uses weights allocated based on importance by the analyst. This requires the weights to be known and criteria for allocating importance on the weights is subjective. | The multiple goodness of fit method could potentially be considered as an option for calibrating the model since we have calibration targets on HIV prevalence, HIV incidence, number of patients on ART, etc. | 2 |

## Methods for searching for parameters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **Advantages** | **Disadvantages** | **Decision** | **Source** |
| **Manual search**  This method is based on intuition, experience and guessing to generate parameter values that the analysts is satisfied with. | This approach does not follow any specific set of guidelines and can be easy to implement in models that are not complex. | The manual search process can be time consuming and hard to justify especially for models with multiple parameters. | The manual search method is not applicable alone since we have hundreds of parameter inputs and the search is not systematic. |  |
| **Grid Search Method**  In the grid search method, the parameter search takes place across the different possible combinations of parameter values (parameter space).  The method involves setting up a suitable grid in the parameter space, evaluating the GOF estimate at all the grid points and finding the grid point that best minimizes the GOF. For each additional parameter, the number of dimensions required to represent the space also increases accordingly and, in most practical problems, the grid search method requires large numbers of model evaluations. | The grid search method does not require complex computation to implement and offers a systematic way of identifying parameters. | The grid search method requires multiple evaluations for models with many parameters. This can be time consuming. | The grid search method is a potential method to use in our model despite that we have a lot of parameter inputs to estimate and the evaluation will be complex and time consuming. | 2,3 |
| **Random Search Method**  The random search method is the most commonly used method for parameter search. In this method, distributions are assigned to each parameter in the model and multiple sets of parameter values are sampled using a random number generator. Each set is then used in the model and GOF calculated. The set that results in the optimum GOF is selected.  The method uses sampling to obtain parameter values from the parameter distribution. The methods of sampling used to obtain the parameter values vary in efficiency. Random sampling can be used but it is less efficient. One of the popular sampling methods used is the Latin hypercube. In the Latin hypercube sampling, for each parameter, a probability density function is defined and divided into n intervals with the same probability. A parameter value is picked randomly from every interval and this procedure is performed for every parameter. | The random search method is intuitive and relatively easy to program compared to other methods. | The random search method is not efficient in covering the whole parameter space. Although increasing the number of searches improves the chance that the global min or max has been identified, we cannot be certain that the identified is global min or max and not local.  In more complex models with many parameters and larger parameter space, the random search method has limitations in the processing time required to search for the global extremum.  The random search is likely to generate parameter estimates with high variance. | The random search method is a potential method to use. | 2,3 |
| **Bayesian calibration methods**  A calibration approach based on the Bayesian theorem. Bayesian approaches use simple probability rules to combine three sources of information: 1) evidence about the distribution of model parameters, 2) evidence about the distribution of modelled outcomes, and 3) model structural assumptions that relate parameters and modelled outcomes. The prior distribution is assigned to model parameter inputs and the outputs will be projected using the parameters under prior distribution. The likelihood of projected outputs and observed outputs will be calculated. The posterior distribution of parameter inputs conditional on outputs will be derived by combining prior distribution and likelihood function. At each iteration, the distribution of parameter inputs will be updated compared to random search method. | The Bayesian calibration methods is similar to random sampling approach in parameter value selecting that assigned distribution of each parameter inputs.  The process is intuitive and easy to program.  The Bayesian calibration method and its adaptations are commonly used in epidemiological models for HIV, Malaria, and TB. | With increased number of parameters, the method has limitations in the processing time required to search for the global extremum since a significant number of iterations will be needed to estimate posterior distribution of parameters  Bayesian calibration methods requires information on the distribution of calibration targets. For calibration targets that has unknown distribution, the estimation can be biased if the assumption on distribution fails to fit the actual data. | Bayesian calibration methods might be usable for our model. However, the calibration targets we collected have limited data on the distribution of calibration targets. For certain calibration targets, we have only a single point estimate without any information on ranges or confidence interval. This might prevent us from assigning an appropriate distribution to calibration targets. | 6 |
| **Generalized Reduced Gradient Method.**  An extension of reduced gradient method that allows non-linear constraints and arbitrary bounds for variables. The method assumes that the objective function is differentiable.  This approach is used in Microsoft Excel Solver as an optimization tool, which assumes the model to be non-linear. Optimization methods that use the gradient vector can be expected to find the minimum point faster. In the spreadsheet format, the optimization problem can be solved using Microsoft Excel Solver.  Solver extracts the information (problem to optimize) from the spreadsheet cells and internally builds a representation of the model that is suitable for the generalized reduced gradient method. Solver assumes the model to be nonlinear as default. The path and scaling factors used by the generalized reduced gradient method depend on the starting point. It is recommended that different starting points are tried. If the software reaches roughly the same final point, then that this is a global extremum. | The generalized reduced gradient method can be implemented in Microsoft Excel using Solver and is faster in obtaining the minimum points. Microsoft Excel is free and available to most people. | The generalized reduced gradient method when using the Microsoft Excel Solver assumes the model to be nonlinear as default. This can lead to biased estimates incase the model is linear.  Since it is an optimization algorithm, the method generates only one solution for a given starting point. The final outcome depends on the starting point and multiple points have be tried. | This is an optimization method that generates one set of the parameter estimate.  The method finds local extremes instead of global optimization. Multiple initial simplex will be tried.  However, it is applied in Excel and no existing package is found in R to perform the method. | 2,3 |
| **Downhill Simplex Method (Nelder-Mead).**  Nelder-Mead method is an optimization method that establishes a nonlinear simplex (downhill simplex) to identify global extreme in objective function. Unlike reduced gradient method, the Nelder Mead method does not require differentiable objective function.  Downhill simplex is a geometrical figure consisting, N dimensions, of N + 1 points (or vertices) and all their interconnecting line segments. In two dimensions the simplex is a triangle. In three dimensions it is a tetrahedron. The number of dimensions is determined by the number of input parameters varied in the optimization process.  The downhill simplex method must be initialized with N+ 1 points in order to constitute an initial simplex. The model’s GOF indicates poorly fitting parameter sets at surface with peaks and valleys indicate better fitting parameter sets. This method takes a series of steps (reflection, expansion, contraction, and reduction) to generate the parameter set. The method is slow and generates only one best-fit parameter set at the end of the process. In order to gain more confidence that the best-fit parameter set does not represent a local extremum, the algorithm is usually run a few times from different starting points (different simplexes). | This method makes no assumption about the function being minimized. | This method is not as fast compared to other methods such as the generalized reduced gradient method. In addition, the convergence becomes increasingly difficult with more than 10 parameter inputs.  This method only generates one best-fit parameter after the process. | The number of parameter inputs to calibrate is large so that the method becomes complex in computing. The existing R package for this method limits the number of parameter inputs calibrated to 10-20 with bounds or non-linear constraints. | 2 |
| **Simulated Annealing Method**  ﻿Simulated annealing is a more complex parameter search method and an efficient alternative for largescale optimization problems, particularly those where a desired global extremum is hidden among many poorer local extrema.  Unlike the Nelder Mead method that searches the entire space described by the simplex, simulated annealing is based on the thermodynamics of the crystallization of metal, where parameter searching involves the introduction of an artificial parameter that determines the probability of accepting a set of random parameter values.  At initial high value, the probability of accepting a new set of parameter values is higher, which means that the algorithm is allowed to widely explore the parameter space. Like in the downhill simplex, by conceptualizing the model’s GOF as a surface with peaks (poorly fitting parameter sets) and valleys (better fitting parameter sets), it is apparent that bigger ‘jumps’ avoid the algorithm falling into a local minimal GOF. Slowly decreasing the values of the artificial parameter allows the algorithm to find the parameter set with the lowest GOF.  Only one parameter set emerges at the end of the process. However, simulated annealing is efficient, and it can also be used in problems of combinatorial optimization. In the case of disease models, this would allow us to consider sets of possible model structures in the calibration process. | The method is fast and efficient and does not make assumptions on the objective functions.  The method would provide a global optimum compared to reduced gradient method that generates local extreme.  Applicable to largescale optimization problems | Not applicable to small optimization problems.  The method only generates one best-fit parameter after the process. | The method requires a large sample size to estimate the model and process might be complicated. The R package for simulated annealing method applied only for estimating parameter inputs under maximum likelihood methods. | 2 |
| **Mixed Approaches**  Mixed approaches have been proposed where methods such as random search or grid search can be used to predict the region of the parameter space in which the global extremum is placed.  Once this region is located, more efficient guided techniques can be used to find the precise location of the global extremum. In general, it is encouraged if time allow for analysts to consider the application of more than one method or combinations of methods in a comparative way. | Mixed approaches are helpful to generate the best parameter estimates using multiple techniques. | The approach may not be applicable in case of data limitations | The mixed method approach could potentially be used in the model calibration for our study. | 2 |

## Probabilistic distribution assigned to model parameters

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Description** | **Distribution** | **Justification** | **Parameterization** | | | **Data availability** |
|  | Population growth rate for group r | Normal | The distribution of population growth is unknown. Population growth rate is not bounded and we assume the same rate overtime suggesting that it is independent to time. Thus, normal distribution of population growth rate might be suitable | The model will adopt the mean and variance estimated from the data. | | | Both mean and variances can be estimated from data. |
|  | Probability of diagnosed with HIV in compartment i for group r | Beta | The parameter is a probability bounded within the interval [0,1]. Beta distribution is suitable for probabilities and proportions not only because is defined on interval [0,1] and commonly used to calibrate probabilities but also because beta distribution is a standard result in statistics | The shape parameters can be estimated from the mean and variances of the data . The distribution will be scaled based on the range.  Since  Thus, | | | For the probabilities estimated from the survey data, we can easily derive the parameters.  For the probabilities reported in the literature the mean and sample size are reported. |
|  | Probability of loss to follow-up in compartment i for group r | The shape parameters can be estimated via mean and sample size . | | |
|  | Probability of on ART and suppressed in compartment i for group r |
|  | Probability of on ART and not suppressed in compartment i for group r |
|  | Probability of return to ART and suppressed in compartment i for group r |
|  | Probability of HIV disease progression in compartment i for group r |
|  | Probability of death in compartment i for group r |
|  | Probability of link to care in compartment i for group r |
| **In force of infection equation** | | | | | | | |
|  | Probability of HIV transmission per sex act for group r | Log-normal distribution | Despite that it is a probability, the beta distribution is hard to apply since the original data comes from meta-analysis and estimates the pooled probabilities using random effect model based on inverse-variance method. According to the methods used, the underlying assumption to derive the pooled probability estimate and confidence interval is that the logarithm of probabilities is normally distributed. | The model will require the mean and variance of the natural log of the probabilities of HIV transmission per sex act. | | | The means and variances can be estimated from the point estimates and 95% confidence interval reported in the literature. |
|  | Proportion of consistent condom use for group r | Beta | The parameter is a probability bounded within the interval [0,1]. Beta distribution is suitable for probabilities and proportions not only because is defined on interval [0,1] and commonly used to calibrate probabilities but also because beta distribution is a standard result in statistics | The shape parameters can be estimated via mean and sample size . | | | For the probabilities estimated from the survey data, we can easily derive the parameters.  For the probabilities reported in the literature the mean and sample size are reported. |
|  | Average number of sexual acts for group r | Uniform distribution | The data is reported using median and IQR instead of mean and variances. Despite that gamma distribution could fit the distribution of data better, it is unable to estimate the parameters in gamma distribution using the information reported in the literature, | The bounds will be used. | | | The bounds come from the literature. |
|  | Effectiveness of condom | Uniform | Despite that it is a probability, the data does not report the sample size or variance. The confidence interval reported is not symmetric suggesting a non-normal distribution. Thus, uniform distribution is used in this analysis. | The bounds will be used. | | | The bounds come from the literature. |
|  | Effectiveness of viral suppression |
|  | **Initial distribution** | | | |  |
|  | Proportion of population diagnosed for group r | Beta | The parameter is a probability bounded within the interval [0,1]. Beta distribution is suitable for probabilities and proportions since it is defined on interval [0,1] and is commonly used to calibrate probabilities. | The shape parameters can be estimated via mean and sample size . | | | The original data does not report the variances but reported the sample size to derive the estimates, which might be used. |
|  | Proportion of population on ART for group r |
|  | Proportion of population on ART but virally not suppressed for group r |
|  | Proportion of population loss to follow-up for group r |
|  | Proportion of population linked for group r |
| Distribution of CD4 strata | Proportions of population in each CD4 stratum for group r | Grid search | I am proposing grid search for the distribution of CD4 strata since the four probabilities are associated with each other and follows certain patterns (i.e. most people will be in earlier stage of HIV disease) |  | | |  |