Technical description of a mathematical model of the impact of antiretroviral therapy on HIV incidence in South Africa

Jeffrey W Eaton, Geoffrey P Garnett, Timothy B Hallett

March 29, 2012

This document describes version of a mathematical model of the impact of antiretroviral therapy (ART) on HIV incidence in South Africa used for a systematic model comparison exercise [1]. This model is identified as the Eaton model in that exercise.

The model uses ordinary differential equations to simulate heterosexual HIV transmission in a two-sex population. The sexually active population is divided into three sexual risk groups which mix semi-assortatively. Individuals can move between risk groups and sexual behaviour can change over the course of the epidemic. The model includes progression through five stages of HIV infection according to CD4 cell count, and allows for initiation of ART at any level of CD4 count. Parameter values for HIV transmission and progression are based on the best available data from discordant couple studies and longitudinal HIV cohorts in sub-Saharan Africa. The model is calibrated to national HIV prevalence data from South Africa by adjusting sexual behaviour and sexual mixing parameters in a Bayesian framework.

Section 1 gives a non-technical description of the model, including diagrams of key model processes. Section 2 gives a mathematical description of the model. Section 3 describes the calibration of the model to the South African epidemic context.

1 Model description

1.1 Population structure

The model simulates a two-sex adult population age 15 and older (see diagram in Figure 1). The population is divided into the age groups 15 to 49, presumed to be sexually active age group, and age 50 and older, who are assumed not to form new sexual contacts. Individuals move from the younger to older age group at a rate $\nu = 1/35$ per year and die from the age 50 plus population at an annual rate $\mu = 1/11.45$ per year, calibrated to match the relative sizes of the 15 to 49 year-old population and the age 50 plus population in the year 1990 [2]. Individuals enter the age 15–49 population at a rate $\alpha = 0.0226$ per year, calibrated such that the population growth over the period 1990 to 2010 approximately matches the population growth estimates over that period published by Statistics South Africa [2]. All new sexual contacts, and hence HIV transmission, occur in the 15–49 age group, but the older age group is included in the model to assess the total ART need in the intervention scenarios. The sexually active population is divided into three sexual risk groups (termed ‘low’, ‘medium’, and ‘high’).

As a crude means of simulating natural variability in individuals’ sexual risk behaviour, individuals move from high risk to medium risk, high to low risk, and medium risk to low risk at a rate $\psi$. This rate is varied in the model calibration (see Section 3), and is the same for both genders.

The proportion of the population in each risk group before the introduction of HIV into the population is allowed to be different for each gender and is estimated in the model calibration. The proportion of the new entrants into the population that enter each risk group is calculated such that, in the absence of HIV, this proportion of the population in each risk group would remain constant. The proportion entering each risk group remains fixed over the duration of the simulation (meaning that the proportional size of risk groups may change as a result of the differential burden of HIV in each risk group).
1.2 Sexual mixing

As described in the previous section, the population is divided into three sexual risk groups (see Figure 1). The sexual contact rate in each of the risk groups is determined by the overall population average sexual contact rate, the relative rate of new sexual contacts between the three risk groups, and the size of the risk group. The underlying population average rate of new unprotected sexual contacts, $\bar{c}(t)$, is allowed to vary over the course of the epidemic in order to model potential behaviour change in response to the epidemic such as increased condom usage [3] or reductions in the number of new sexual partners [4]. The functional form for the reduction is a logistic logistic function parameterized by the initial contact rate, $c_0$, the percentage reduction in the contact rate that will occur, $\Delta c$, the start year of the behaviour change, $t_c$, and the number duration (in years) over which the behaviour change occurs, $d_c$. Together, the overall average contact rate at time $t$ is given by

$$\bar{c}(t) = \bar{c}_0 \cdot (1 - \Delta c) + \bar{c}_0 \cdot \Delta c \cdot \frac{1}{1 + \exp \left( \frac{t - (t_c + d_c)/2}{d_c/10} \right)}$$ (1)

Each of these parameters (the initial contact rate, the percentage reduction in the contact rate, the timing of the start of the reduction, and the duration of the change) are estimated in the model calibration. The relative contact rates between high and low risk females and medium and low risk females are also estimated. The relative contact rates for males are calculated based on the relative contact rates for females and the sizes of each of the risk groups for males and females so that the total number of contacts offered by males and females in the same risk group is the same.

The number of sexual contacts formed between members of each risk group is determined by the sexual mixing parameter $\varepsilon$, as proposed by [5]. A proportion $\varepsilon$ of sexual contacts are reserved exclusively to be formed with other members of the same risk group, while the remaining $(1 - \varepsilon)$ proportion of partnerships are formed at random. The value of $\varepsilon$ is estimated in the model calibration. Thus if $\varepsilon = 0$ sexual mixing is completely random, while if $\varepsilon = 1$, mixing is fully assortative. As HIV mortality differentially affects
males and females of each risk group, the total number of contacts offered by males may not balance with the number of partnerships offered by females. In this case the desired number of contacts desired by males and females are geometrically weighted by the parameter $\theta_G$ ranging between zero and one; $\theta_G = 0$ indicates that the females’ preferences determines the number of contacts while $\theta_G = 1$ indicates that the males’ desired number of contacts dominates. For this exercise, the value is fixed at $\theta_G = 0.5$.

1.3 Natural history of HIV infection

![Diagram depicting average duration of and relative HIV transmission rate during each stage of HIV infection.](image)

HIV infection is divided into five stages according to the HIV infection is divided into stages according to CD4 cell count progression associated with the duration of infection (Figure 2). The stages are:

1. Primary infection,
2. CD4 count greater than 350 cells/µl,
3. CD4 count between 200 and 350 cells/µl,
4. CD4 count between 100 and 200 cells/µl, and
5. CD4 count below 100 cells/µl.

HIV infected individuals progress from one stage of HIV infection to the next at a rate which is the reciprocal of the average duration in the stage. The average duration of primary infection is an 2.9 months, as estimated by Hollingsworth et al. [6]. The rates of progression to subsequent CD4 cell stages and the overall duration from HIV infection to death is based on estimates of the duration until CD4 cell count thresholds in sub-Saharan African cohort from the eART-linc collaboration [7], which estimated mean durations of 4.8 years and 9.4 years to reaching CD4 cell counts of below 350 and 200 cells/µl, respectively. The estimate of an average of 4.17 years between CD4 count $\leq 200$ cells/µl and CD4 $\leq 100$ cells/µl is based on extrapolating the rate of decline in square-root transformed CD4 count between CD4 $\leq 350$ cells/µl to CD4 $\leq 200$. The overall average duration from infection to HIV death is 14.6 years, from [7], and the resulting median duration from infection to HIV death is 13.2 years.

The HIV transmission rate of an infected individual varies according to these stages of infection. The overall baseline weighted average transmission rate over the period from the end of primary HIV infection...
to 1.6 years before death is set to be 0.106 per year, as estimated by Hollingsworth, Anderson, and Fraser [6] using data from discordant couples in Rakai, Uganda [8]. The relative rates of transmission during the CD4 stages CD4 $> 350$, $350 > CD4 > 200$, and $CD4 < 200$ are based on the relative rates of transmission observed for these stages by Donnell et al [12], although the rate of transmission during the CD4 $< 100$ stages has been reduced in accordance with the estimate from Hollingsworth, Anderson and Fraser that no transmission occurs during the final 9 months of infection [6], presumably because individuals are sick and not very sexually active during this period. The transmission rate during primary HIV infection is set at 2.76 per year, as estimated by Hollingsworth, Anderson, and Fraser [6].

1.4 ART model

![Figure 3: Diagram of stages of antiretroviral therapy](image)

Individuals may initiate ART from any of the above stages of HIV infection. Antiretroviral therapy is divided into a multistage process (Figure 3). Upon treatment initiation, all individuals first enter a ‘virally suppressing’ stage during which they are on ART but their viral load is not yet fully suppressed. This stage lasts for an average of 3 months and transmission is assumed to be reduced by half compared to the CD4 stage from which they initiated treatment.

After this stage, most patients enter a long period of ‘effective ART’, while a proportion of patients for whom treatment is not successful go directly to the final stage of the ART model of being ‘very sick’, which lasts for an average of 6.2 months before death. The probability of immediately failing treatment depends on the CD4 count stage from which treatment was initiated to allow for high early mortality when starting treatment at lower CD4 cell counts, but then relatively similar long-term survival if treatment effectively suppresses viral load and symptoms are controlled [9, 10, 11]. The proportion of patients who fail ART is calibrated to such that mortality in the first year after initiating ART matches the crude first-year mortality rate observed for each CD4 count stratum in a collaborative analysis of ART cohorts from sub-Saharan Africa [10].
For the majority of patients for whom ART is effective, they are virally suppressed and transmission is reduced by 92% [12] compared to the HIV transmission rate in the CD4 count between 200 to 350 cells/µl stage. The period of ‘effective ART’ is divided into two stages—first a period of ‘early effective ART’ lasting an average of 1.75 years, and then a long period of sustained viral suppression. The reduction in transmission is assumed to be the same in both of these stages, but this is implemented as separate ART stages so that the dropout rate from treatment can be varied according to the duration on ART, if, for example, we may assume that there is high dropout in the years following ART initiation, but patients who remain on treatment for two years are likely to have accommodated treatment and have high retention thereafter. In addition to the previously described higher probability for immediate treatment failure and death for those starting ART at low CD4 cell counts, the failure rate for long-term effective ART is assumed to be slightly lower for those who start treatment at high CD4 cell counts (see Figure 3), in line with observations that mortality is modestly higher even several years after treatment initiation for those who start at low CD4 cell counts [11] and to ensure that there is no ‘survival benefit’ in the model from delaying treatment initiation.

After patients fail long-term effective ART, they enter a stage of ‘treatment-failing’ in which they are viraemic and are assumed to have the same infectiousness as individuals in the CD4 cell count category 200 to 350 cells/µl. The average duration of this stage is 2.3 years. Finally individuals enter a stage of being ‘very sick’ just before death, which lasts for an average duration of 6.2 months. During this period of being ‘very sick’, transmission is reduced and assumed to be at the same level as during the CD4 < 100 cells/µl stage.

### 1.4.1 Dropping out from ART

Individuals may dropout from any of the first three stages of ART: ‘virally suppressing’, ‘early effective ART’, and ‘effective ART’. The model is designed and implemented to permit the rate of dropout from treatment to vary according to duration on ART and the CD4 count category from which treatment was initiated, in line with data suggesting that those starting treatment at higher CD4 cell counts may have poorer retention in treatment programmes [13], perhaps because they did not experience AIDS-related symptoms before initiating ART. But, in accordance with the specification of intervention scenarios for the model comparison exercise [1], the application of the model for this exercise assumes that dropout from treatment is constant across all of these strata.

<table>
<thead>
<tr>
<th>CD4 stage at ART initiation</th>
<th>Treatment stage at dropout</th>
<th>Percentage of dropouts returning to CD4 category</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4 &gt; 350</td>
<td>Virally suppressing</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Early effect. ART</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Effective ART</td>
<td>100%</td>
</tr>
<tr>
<td>CD4: 200 - 350</td>
<td>Virally suppressing</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Early effect. ART</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Effective ART</td>
<td>100%</td>
</tr>
<tr>
<td>CD4: 100 - 200</td>
<td>Virally suppressing</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>Early effect. ART</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Effective ART</td>
<td>100%</td>
</tr>
<tr>
<td>CD4 &lt; 100</td>
<td>Virally suppressing</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>Early effect. ART</td>
<td>100%</td>
</tr>
</tbody>
</table>

After dropping out from treatment, the untreated CD4 stage category that individuals enter depends on their pre-treatment CD4 category and the duration on treatment to simulate rebounding CD4 cell count associated with ART. Individuals who dropout from the ‘virally suppressing’ stage all return to the same CD4 stage from which they initiated treatment. For those who dropout during the ‘early effective ART’ stage, half move one CD4 stage higher, while half increase two CD4 stages. Those who dropout from the ‘effective ART’ stage all increase two CD4 stages. This is summarized in Table 1. However, individuals who have dropped out of treatment progress through subsequent CD4 stages twice as fast as treatment
naïve individuals (rates described in Figure 2).

After individuals have dropped out of treatment, they are eligible to restart treatment again once. Upon restarting treatment individuals progress through the same stages of ART as when first initiating ART. The dropout rate may be different, but in the implementation for the model comparison they were assumed to be the same as dropout from first ART initiation. Because individuals may restart treatment, but only once, the model separately tracks treated and untreated people according to the number of times which they have initiated ART. To summarize this, Table 2 lists all of the stages of antiretroviral treatment through which infected individuals can progress, and the subscript identifying each stages in the technical model description that follows.

### Table 2: Stages of antiretroviral therapy.

<table>
<thead>
<tr>
<th>Subscript</th>
<th>Stage</th>
<th>Duration</th>
<th>Infectiousness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Untreated, no access to treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Untreated, access to ART</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Virally suppressing</td>
<td>3 months</td>
<td>50% lower than prev. stage</td>
</tr>
<tr>
<td>3</td>
<td>Early effective ART, virally suppressed</td>
<td>1.75 years</td>
<td>92% lower than CD4 &lt; 350</td>
</tr>
<tr>
<td>4</td>
<td>Effective ART, virally suppressed</td>
<td>(Figure 3)</td>
<td>92% lower than CD4 &lt; 350</td>
</tr>
<tr>
<td>5</td>
<td>Treatment failing, viraemic</td>
<td>2.3 years</td>
<td>Same as CD4 &gt; 100</td>
</tr>
<tr>
<td>6</td>
<td>Very sick</td>
<td>6.2 months</td>
<td>Same as CD4 &lt; 100</td>
</tr>
<tr>
<td>7</td>
<td>Untreated, dropped out after first initiation, eligible to restart ART</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Re-initiated ART, virally suppressing</td>
<td>3 months</td>
<td>50% lower than prev. stage</td>
</tr>
<tr>
<td>9</td>
<td>Re-initiated ART, early effective ART</td>
<td>1.75 years</td>
<td>92% lower than CD4 &lt; 350</td>
</tr>
<tr>
<td>10</td>
<td>Re-initiated ART, effective ART</td>
<td>(Figure 3)</td>
<td>92% lower than CD4 &lt; 350</td>
</tr>
<tr>
<td>11</td>
<td>Re-initiated ART, treatment failing, viraemic</td>
<td>2.3 years</td>
<td>Same as CD4 &gt; 100</td>
</tr>
<tr>
<td>12</td>
<td>Re-initiated ART, very sick</td>
<td>6.2 months</td>
<td>Same as CD4 &lt; 100</td>
</tr>
<tr>
<td>13</td>
<td>Untreated, dropped out after second initiation, not eligible to restart</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 1.5 HIV transmission

The probability of transmission in a contact between a susceptible with an infected individual depends on the annual transmission rate $\beta_{m,u}$ in HIV stage $m$, the “intensity” of a contact $\kappa_{rM,rF}$ between a male in risk group $r_M$ and female in risk group $r_F$, and a factor $\eta_G$ that is the relative transmission rate for males to females compared to females to males. The intensity parameter accounts for factors that affect the probability of transmission in different types of partnerships such as the partnership duration, coital frequency, and condom usage. The per-contact female-to-male transmission probability based on these parameters is $1 - e^{-\beta_{m,u}\kappa_{rM,rF}\eta_G}$ and the male-to-female transmission probability is $1 - e^{-\beta_{m,u}\kappa_{rM,rF}\eta_G}$ where $m$ is the HIV stage of the infected partner.
2 Model equations

We divide the population into four categories according to their infection and treatment status:

- $S^{g,r}$: HIV uninfected and sexually active (susceptible) individuals of gender $g$ in risk group $r$.

- $I^{g,r}_{m,u}$: HIV infected and sexually active individuals of gender $g$ in risk group $r$ and HIV infection stage $m$. The subscript $u = 0$ indicates untreated individuals who do not have access to treatment, $u = 1$ indicates untreated individuals who do have access to treatment, $u = 7$ indicates individuals who have dropped out of treatment and are eligible to restart, and $u = 13$ indicates those who have dropped out from treatment a second time and are not eligible to restart.

- $T^{g,r}_{m,u}$: HIV infected and sexually active individuals on ART (treated) who began treatment in HIV infection stage $m$ and are in treatment stage $u$ (see Table 2).

- $R^{g}_{m,u}$: Individuals removed from the sexually active population of gender $g$, in HIV infection stage $m$, and treatment stage $u$. (Uninfected individuals are indicated by $m = 0$ and untreated individuals by $u \in \{0, 1, 7, 13\}$).

In the above and throughout the following mathematical description, the superscript $g \in \{M, F\}$ corresponds to gender; superscript $r \in \{H, M, L\}$ corresponds to sexual risk group for the sexually active population; subscript $m \in \{0, \ldots, 5\}$ corresponds to HIV infection stage with 0 representing uninfected and 1 to 5 corresponding to the stages of infection from primary infection to CD4 $< 100$; and subscript $u \in \{0, \ldots, 13\}$ corresponds to ART status with 0 indicating untreated individuals without access to treatment, 1 indicating untreated individuals with access to treatment, and the remaining stages indicating different stages of ART as indicated in Table 2.
The following differential equations define the dynamics of the groups:

\[
\begin{align*}
\frac{dS_{r}}{dt} &= \frac{\alpha + \nu}{2} \pi^g_{r} \left( S'_{r} + I_{r} + T_{r} \right) + \sum_{r'} \psi_{r',r} \cdot S^g_{r'} - \left( f^g_{r}(t) + \nu + \sum_{r'} \psi^g_{r,r'} \right) S^g_{r'} \\
\frac{dI_{0,1}}{dt} &= (1 - \chi) f^g_{0,1}(t) S^g_{0,1} + \sum_{r'} \psi_{0,r'} \cdot I^g_{0,1} - \left( \sigma_1 + \nu + \sum_{r'} \psi^g_{r,r'} \right) I^g_{0,1} \\
\frac{dI_{1,1}}{dt} &= \chi f^g_{1,1}(t) S^g_{1,1} + \sum_{r'} \psi_{1,r'} \cdot I^g_{1,1} - \left( \sigma_1 + \lambda_1 + \nu + \sum_{r'} \psi^g_{r,r'} \right) I^g_{1,1} \\
\frac{dI_{m,0}}{dt} &= \sigma_{m-1} I^g_{m-1,0} + \sum_{r'} \psi_{m,r'} \cdot I^g_{m,0} - \left( \sigma_m + \nu + \sum_{r'} \psi^g_{r,r'} \right) I^g_{m,0} \quad \text{for } m \geq 2 \\
\frac{dI_{m,1}}{dt} &= \sigma_{m-1} I^g_{m-1,1} + \sum_{r'} \psi_{m,r'} \cdot I^g_{m,1} - \left( \sigma_m + \lambda_m + \nu + \sum_{r'} \psi^g_{r,r'} \right) I^g_{m,1} \quad \text{for } m \geq 2 \\
\frac{dT_{m,2}}{dt} &= \lambda^m_{m,2} I^g_{m,2} + \sum_{r'} \psi_{m,r'} \cdot T^g_{m,2} - \left( \phi^g_{m,2} + \eta_{m,2} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,2} \\
\frac{dT_{m,3}}{dt} &= (1 - \xi_m) \phi^g_{m,3} T^g_{m,3} + \sum_{r'} \psi_{m,r'} \cdot T^g_{m,3} - \left( \phi^g_{m,3} + \eta_{m,3} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,3} \\
\frac{dT_{m,u}}{dt} &= \phi^g_{m,u} T^g_{m,u-1} + \sum_{r'} \psi_{m,r'} \cdot T^g_{m,u} - \left( \phi^g_{m,u} + \eta_{m,u} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,u} \quad \text{for } u \in \{4, 5\} \\
\frac{dT_{m,6}}{dt} &= \xi \phi^g_{m,6} T^g_{m,6} + \phi^g_{m,5} T^g_{m,6} + \sum_{r'} \psi_{m,r'} \cdot T^g_{m,6} - \left( \phi^g_{m,6} + \eta_{m,6} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,6} \\
\frac{dT_{m,7}}{dt} &= \sum_{m' u'} \phi^m_{m',u'} \eta_{m',u'} T^g_{m',u'} + \sigma_{m-1} I^g_{m-1,7} + \\
&\quad \sum_{r'} \psi_{r',r} \cdot I^g_{m,7} - \left( \sigma_m + \lambda_m + \nu + \sum_{r'} \psi^g_{r,r'} \right) I^g_{m,7} \quad \text{for } m \geq 1 \\
\frac{dT_{m,8}}{dt} &= \lambda^m_{m,8} I^g_{m,8} + \sum_{r'} \psi_{r',r} \cdot T^g_{m,8} - \left( \phi^g_{m,8} + \eta_{m,8} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,8} \\
\frac{dT_{m,9}}{dt} &= (1 - \xi_m) \phi^g_{m,9} T^g_{m,9} + \sum_{r'} \psi_{r',r} \cdot T^g_{m,9} - \left( \phi^g_{m,9} + \eta_{m,9} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,9} \quad \text{for } u \in \{10, 11\} \\
\frac{dT_{m,12}}{dt} &= \xi \phi^g_{m,12} T^g_{m,12} + \phi^g_{m,11} T^g_{m,11} + \sum_{r'} \psi_{r',r} \cdot T^g_{m,12} - \left( \phi^g_{m,12} + \eta_{m,12} + \nu + \sum_{r'} \psi^g_{r,r'} \right) T^g_{m,12} \\
\frac{dT_{m,13}}{dt} &= \sum_{m' u'} \phi^m_{m',u'} \eta_{m',u'} T^g_{m',u'} + \sigma_{m-1} I^g_{m-1,13} + \\
&\quad \sum_{r'} \psi_{r',r} \cdot I^g_{m,13} - \left( \sigma_m + \lambda_m + \nu + \sum_{r'} \psi^g_{r,r'} \right) I^g_{m,13} \quad \text{for } m \geq 1
\end{align*}
\]
\[
\begin{align*}
\frac{dR_{0,0}^g}{dt} &= \nu \sum_r S_{0,r}^g - \mu R_{0,0}^g \\
\frac{dR_{m,0}^g}{dt} &= \sigma_{m-1} R_{m-1,0}^g + \nu \sum_r I_{m,r}^g - (\sigma_m + \mu) R_{m,0}^g \\
\frac{dR_{m,1}^g}{dt} &= \sigma_{m-1} R_{m-1,1}^g + \nu \sum_r I_{m,r}^g - (\sigma_m + \lambda_m^g + \mu) R_{m,1}^g \\
\frac{dR_{m,2}^g}{dt} &= \lambda_m^g R_{m,0}^g + \nu \sum_r T_{m,2}^g - (\phi_{m,2}^g + \eta_{m,2} + \mu) R_{m,2}^g \\
\frac{dR_{m,3}^g}{dt} &= (1 - \xi_m) \phi_{m,2}^g R_{m,2}^g + \nu \sum_r T_{m,3}^g - (\phi_{m,3}^g + \eta_{m,3} + \mu) R_{m,3}^g \\
\frac{dR_{m,4}^g}{dt} &= \phi_{m,u-1}^g R_{m,u-1}^g + \nu \sum_r R_{m,u-1}^g - (\phi_{m,u}^g + \eta_{m,u} + \mu) R_{m,u}^g \\
\frac{dR_{m,5}^g}{dt} &= \xi_m^g R_{m,5}^g + \phi_{m,5}^g R_{m,5}^g + \nu \sum_r T_{m,6}^g - (\phi_{m,6}^g + \eta_{m,6} + \mu) R_{m,6}^g \\
\frac{dR_{m,7}^g}{dt} &= \sum_{m'=2}^{5} \sum_{u'=2}^{m'} \rho_{m',u'}^{m,u} \eta_{m',u'} R_{m',u'}^g + \hat{\sigma}_{m-1} R_{m-1,7}^g + \nu \sum_r I_{m,7}^g - \left(\hat{\sigma}_m + \lambda_m^g + \mu\right) I_{m,7}^g \\
\frac{dR_{m,8}^g}{dt} &= \hat{\lambda}_m^g R_{m,7}^g + \nu \sum_r T_{m,8}^g - (\phi_{m,8}^g + \eta_{m,8} + \mu) R_{m,8}^g \\
\frac{dR_{m,9}^g}{dt} &= (1 - \xi_m) \phi_{m,9}^g R_{m,9}^g + \nu \sum_r T_{m,9}^g - (\phi_{m,9}^g + \eta_{m,9} + \mu) R_{m,9}^g \\
\frac{dR_{m,10}^g}{dt} &= \phi_{m,u-1}^g R_{m,u-1}^g + \nu \sum_r R_{m,u-1}^g - (\phi_{m,u}^g + \eta_{m,u} + \mu) R_{m,u}^g \\
\frac{dR_{m,11}^g}{dt} &= \xi_m^g R_{m,11}^g + \phi_{m,11}^g R_{m,11}^g + \nu \sum_r T_{m,12}^g - (\phi_{m,12}^g + \eta_{m,12} + \mu) R_{m,12}^g \\
\frac{dR_{m,13}^g}{dt} &= \sum_{m'=8}^{12} \sum_{u'=8}^{m'} \rho_{m',u'}^{m,u} \eta_{m',u'} R_{m',u'}^g + \hat{\sigma}_{m-1} R_{m-1,13}^g + \nu \sum_r I_{m,13}^g - \left(\hat{\sigma}_m + \lambda_m^g + \mu\right) R_{m,13}^g
\end{align*}
\]

In the above equations, the parameter \( \tilde{\pi}^{g,r} \) is the proportion of new susceptible individuals of gender \( g \) that should enter risk group \( r \) in order to maintain a constant proportion \( \pi^{g,r} \) in risk group \( r \) in the absence of HIV. Solving the above equations with \( f^{g,r}(t) = 0 \) gives that

\[
\tilde{\pi}^{g,r} = \frac{\sum_r \psi_{r,r}^{g} \pi^{g,r} - \sum_r \psi_{r,r}^{g} \pi^{g,r'}}{\alpha + \nu}.
\]

The function \( f^{g,r}(t) \) is the force of infection for the group \( S^{g,r} \) which depends on the contact rate \( c^{g,r}(t) \) at time \( t \) in that group, the probability that a contact is formed with an infectious partner, and the probability of transmission in that contact.

The contact rate for a risk group depends on the average underlying contact rate \( \tau(t) \) which changes over time according to equation 1, the size of the risk groups at the beginning of the epidemic \( \pi^{g,r} \) and the relative contact rates of the high and medium risk groups to the low risk group \( \omega^{g,r} \) (where \( \omega^{g,L} := 1 \)). The weighted average of the relative contact rate by the size of the initial risk group yields the annual contact rate for each risk group at time \( t \):

\[
c^{g,r}(t) = \frac{\tau(t) \cdot \omega^{g,r}}{\sum_r \pi^{g,r} \omega^{g,r}}.
\]
The total number of contacts desired to be formed by members of risk group \( r \) of gender \( g \) is thus

\[
J_{r,g}^g(t) = c_{r,g}^g(t) \left( S_{r,g}^g + \sum_m I_{m,r}^g + \sum_u \sum_{r'} T_{m,u}^g r_{r',r} \right)
\]  

(3)

The number of these contacts \( J_{r,g}^g \) desired to be formed with each risk group \( r' \) of the opposite gender \( g' \) depends on the value of the assortativity parameter \( \varepsilon \). A proportion \( \varepsilon \) of the partnerships are desired only to be formed with the members of the same risk group \( r = r' \), while the remaining \( 1 - \varepsilon \) proportion of the partnerships are formed with each risk group of the opposite gender proportionally to the number of partnerships \( J_{r,g}^g r_{r',r} \) offered by those risk groups. Formally, the proportion of contacts desired to be formed by gender \( g \) and risk group \( r \) that are formed with the risk group \( r' \) of the opposite gender is defined as

\[
Q_{r,r'}^g = \varepsilon \delta_{r,r'} + (1 - \varepsilon) \sum_{r''} J_{r'',r'}^g r_{r',r''}
\]

(4)

where \( \delta_{r,r'} \) is the Kronecker delta defined as \( \delta_{r,r'} = 1 \) if \( r = r' \) and 0 otherwise.

In the case that males in risk group \( r_M \) and females in group \( r_F \) do not agree on the number of partnerships to be formed between the risk groups, i.e. \( Q_{r_M,r_F}^M \neq Q_{r_M,r_F}^F \), the discrepancy is balanced according to the parameter \( \theta_G \) as

\[
\tilde{Q}_{r_M,r_F}^M = Q_{r_M,r_F}^M \left( \frac{Q_{r_M,r_F}^M \delta_{r_M,r_F}}{Q_{r_M,r_F}^M \delta_{r_M,r_F} + Q_{r_M,r_F}^F \delta_{r_M,r_F}} \right)^{\theta_G} - 1
\]

\[
\tilde{Q}_{r_M,r_F}^F = Q_{r_M,r_F}^F \left( \frac{Q_{r_M,r_F}^M \delta_{r_M,r_F}}{Q_{r_M,r_F}^M \delta_{r_M,r_F} + Q_{r_M,r_F}^F \delta_{r_M,r_F}} \right)^{\theta_G}
\]

(5)

The probability that transmission occurs in a contact between a susceptible and an infected individual depends on the stage of infection and treatment status of the infection according to the transmission rate parameter \( \beta_{m,u} \) and on the value of the partnership intensity multiplier \( \kappa_{r_M,r_F} \) for a partnership between the male in risk group \( r_M \) and the female in risk group \( r_F \). The force of infection is then calculated by summing over the probability that each contact is with an infectious individual and the probability that transmission occurs according to the infection stage and treatment status of the partner:

\[
f_{r,g}^g(t) = c_{r,g}^g(t) \sum_{r'} \left[ \tilde{Q}_{r,r'}^g \frac{\sum_{r'} \sum_{u} T_{m,u}^g r_{r',r} p_{m,u}^g r_{r',r}}{\sum_{m} \sum_{u} I_{m,u}^g + \sum_{m} \sum_{u} T_{m,u}^g r_{r',r}} \right]
\]

(6)

where \( p_{m,u}^g r_{r',r} \) is the probability of transmission per contact by an infected individual of gender \( g \) in risk group \( r \), HIV stage \( m \) and treatment status \( u \) to a susceptible individual of the opposite gender in risk group \( r' \), defined as

\[
p_{m,u}^g r_{r',r} = 1 - \exp \left\{ -\beta_{m,u} \cdot \kappa_{r_M,r_F} \cdot \eta_{G}(\delta_{r,r'}) \right\}
\]

(7)

### 3 Model calibration

The model is calibrated to nationally representative HIV prevalence data and ART scale-up data from South Africa. The general strategy for model calibration is that parameters related to the natural history of HIV infection, the effect of antiretroviral therapy on individual infection, and patterns of access and retention in the existing ART program are fixed and informed from the literature as described in Section 1. Parameters relating to sexual behaviour and mixing, the start time of the epidemic, the timing and magnitude of sexual behaviour change, and the timing and rate of existing ART scale-up are estimated using a Bayesian approach. This yields a joint distribution of parameter combinations representing different sexual mixing patterns consistent with the observed epidemic.
Table 3: Model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Population growth rate (in absence of HIV)</td>
<td>0.023 per year</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Rate of progression from 15-49 age group to 50+</td>
<td>1/35 per year</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mortality rate out of the 50+ age group</td>
<td>1/11.45 per year</td>
</tr>
<tr>
<td>$\pi_{g,r}$</td>
<td>Proportion of the age 15-49 population of gender $g$ in risk group $r$ in absence of HIV</td>
<td>derived from $\pi$ and $\psi$</td>
</tr>
<tr>
<td>$\tilde{\pi}_{g,r}$</td>
<td>Proportion of new entrants of gender $g$ entering risk group $r$</td>
<td>estimated</td>
</tr>
<tr>
<td>$\psi_{r,r'}^{g}$</td>
<td>Rate of moving from risk group $r$ to $r'$ for gender $g$</td>
<td>estimated</td>
</tr>
<tr>
<td>$\tau(t)$</td>
<td>Population mean contact rate per year at time $t$</td>
<td>estimated</td>
</tr>
<tr>
<td>$\omega_{g,r}^{t}$</td>
<td>Relative contact rate between risk group $r$ and low risk group for gender $g$</td>
<td>estimated</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Degree of assortative mixing</td>
<td>0.5</td>
</tr>
<tr>
<td>$\theta_G$</td>
<td>Balance between male and female partner preference</td>
<td>estimated</td>
</tr>
<tr>
<td>$r_{rM,rF}$</td>
<td>Intensity of partnership between male in risk group $rM$ and female in $rF$</td>
<td>(see Figure 2, 3)</td>
</tr>
<tr>
<td>$\beta_{m,u}$</td>
<td>Annual HIV transmission rate in stage $m$ and ART status $u$</td>
<td>(Figure 3)</td>
</tr>
<tr>
<td>$\sigma_m$</td>
<td>Rate of progression from HIV stage $m$ to stage $m+1$</td>
<td>0.24, 4.56, 4.53, 4.28, 0.94, $2\sigma$</td>
</tr>
<tr>
<td>$\sigma_m$</td>
<td>Rate of progression from HIV stage $m$ to stage $m+1$ after treatment dropout</td>
<td></td>
</tr>
<tr>
<td>$\chi$</td>
<td>Proportion of HIV infected population with access to ART</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{m}^{g}$</td>
<td>Rate of ART initiation for gender $g$ in stage $m$</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{m}^{a}$</td>
<td>Rate of re-initiating ART after dropout for gender $g$ in stage $m$</td>
<td></td>
</tr>
<tr>
<td>$\phi_{m,u}^{g}$</td>
<td>Rate of progression from ART stage $u$ to $u+1$ when initiating ART in HIV stage $m$ for gender $g$</td>
<td>(Table 1)</td>
</tr>
<tr>
<td>$\xi_m$</td>
<td>Probability of immediate treatment failure if initiating ART in stage $m$</td>
<td>0, 0, 0.025, 0.067, 0.189</td>
</tr>
<tr>
<td>$\eta_{m,u}$</td>
<td>Rate of dropping out of ART if initiated in stage $m$ and currently in stage $u$</td>
<td></td>
</tr>
<tr>
<td>$\bar{\eta}_{m,u}$</td>
<td>Rate of dropping out of ART after re-initiating if re-initiated in stage $m$ and currently in stage $u$</td>
<td></td>
</tr>
<tr>
<td>$\rho_{m',u'}^{m,u}$</td>
<td>Probability of entering CD4 stage $m$ after dropping out of stage $(m',u')$</td>
<td>estimated</td>
</tr>
<tr>
<td>$t_0$</td>
<td>Date at which HIV epidemic is seeded into population</td>
<td></td>
</tr>
<tr>
<td>$\delta_{g,r}$</td>
<td>Seed HIV prevalence for gender $g$ and risk group $r$</td>
<td></td>
</tr>
</tbody>
</table>


3.1 Data

The model is calibrated using HIV prevalence data from two sources. The first is HIV prevalence amongst 15–49 year-old pregnant women from the annual national antenatal prevalence surveys from 1990 to 2008 [14]. The second is national HIV prevalence amongst 15–49 year-old males and females from the three nationally representative household surveys conducted by the Human Sciences Research Council in 2002, 2005, and 2008 [15, 16, 17]. Figure 4 shows the mean and 95% confidence intervals for each of these estimates. The discrepancy between the level of HIV prevalence in the antenatal surveillance and the household survey based prevalence is reconciled by incorporating a bias parameter in the antenatal prevalence compared to prevalence amongst the general 15–49 year old population from the household surveys. This is described in detail below in Section 3.3.

![South Africa HIV prevalence data](image)

Figure 4: Age 15–49 HIV prevalence from three nationally representative HSRC household surveys and HIV prevalence amongst antenatal care attendees.

The model is also calibrated to the percentage of the adult population on antiretroviral therapy. This is calculated by dividing the total number of adults reported to be on ART by the South Africa Department of Health in June of each year from 2005 to 2010 [18] by the annual mid-year population size estimate of those over 15 years old from Statistics South Africa [2]. The resulting estimates for the proportion of the adult population on ART for 2005 to 2010 to which the model is calibrated are shown in Figure 5.

3.2 Estimated model parameters

A vector $\theta$ of 20 parameter values are estimated, and are either used directly or to derive a number of the model inputs in the equations described in Section 2 and Table 3. The mathematical model parameters which are estimated from fitting to HIV prevalence data from South Africa are given in Table 4 along with the prior distribution from which they are estimated. This collection of model parameters estimated from the data will be referenced together as the parameter vector $\theta$. 
Figure 5: Percentage of adult population (age 15+) on ART at midpoint of each year (South Africa Department of Health)

Table 4: Estimated model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_0$</td>
<td>Start date of the epidemic</td>
</tr>
<tr>
<td>$1 - \pi^{M,L}$</td>
<td>Proportion of males not in the low risk group</td>
</tr>
<tr>
<td>$\frac{1}{1 - \pi^{M,L}}$</td>
<td>Proportion of males in high risk group of those not in low risk group</td>
</tr>
<tr>
<td>$1 - \pi^{F,L}$</td>
<td>Proportion of females not in the low risk group</td>
</tr>
<tr>
<td>$\frac{1}{1 - \pi^{F,L}}$</td>
<td>Proportion of females in high risk group of those not in low risk group</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Rate of movement from higher to lower risk groups</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>Mean annual contact rate at start of the epidemic</td>
</tr>
<tr>
<td>$\Delta_\tau$</td>
<td>Proportion reduction in average contact rate</td>
</tr>
<tr>
<td>$t_\tau$</td>
<td>Year behaviour change start</td>
</tr>
<tr>
<td>$t_\tau + d_\tau$</td>
<td>Year behaviour change ends</td>
</tr>
<tr>
<td>$\omega^{F,M}$</td>
<td>Relative contact rate between medium risk and low risk females</td>
</tr>
<tr>
<td>$\omega^{F,H} - \omega^{F,M}$</td>
<td>Additional relative contact rate for high risk women</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Assortativity of sexual mixing</td>
</tr>
<tr>
<td>$\kappa_H$</td>
<td>Partnership intensity for partnership involving a high risk partner</td>
</tr>
<tr>
<td>$\kappa_M$</td>
<td>Partnership intensity for partnership between medium and low risk</td>
</tr>
<tr>
<td>$\kappa_L$</td>
<td>Partnership intensity for partnership between low risk partners</td>
</tr>
<tr>
<td>$t_{A_0}$</td>
<td>Date of ART scale-up</td>
</tr>
<tr>
<td>$\lambda_{A_0}$</td>
<td>ART initiation rate during initial ART scale-up</td>
</tr>
<tr>
<td>$t_{A_1} - t_{A_0}$</td>
<td>Time between initial ART scale-up and intensified ART scale-up</td>
</tr>
</tbody>
</table>

† These parameters have a joint uniform prior distribution such that $0 < \kappa_H < \kappa_M < \kappa_H < 1$. 
3.3 Statistical methods

We use a Bayesian analysis to estimate probability distributions for the unknown parameters given the model and available HIV prevalence data, $W$. Let $\theta$ denote the set of parameters to be estimated from section 3.2 and $M$ denote the mathematical model and the associated fixed parameter values. For a given set of parameter values, the model outputs a set of predicted HIV prevalence values $\zeta = M(\theta)$ corresponding to the data. Then we can produce a likelihood function $p(W|M(\theta))$ for the probability of the data given the model and parameter values. If we let $p(\theta)$ denote a prior distribution on the unknown parameters, using Bayes theorem the posterior distribution of $\theta$ given the model and data is given by

$$p(\theta|W, M) \propto p(\theta)p(W|M(\theta)).$$

3.3.1 Likelihood function

We now derive the combined likelihood for the national seroprevalence survey data and antenatal clinic data. First we define the likelihood for an individual datum.

For the national survey estimates, let $W_{F,t}$ and $W_{M,t}$ be the HIV prevalence in the age group 15-49 reported by a survey in year $t$. As above, define $\zeta_{g,t}$ to be the predicted population HIV prevalence for gender $g$ at time $t$ by the model $M(\theta)$. We assume that national survey prevalence provides an unbiased estimate of the true population prevalence, but logit-transform the data to stabilise the error variance, so that

$$\log \left( \frac{W_{g,t}}{1 - W_{g,t}} \right) = \log \left( \frac{\zeta_{g,t}}{1 - \zeta_{g,t}} \right) + \epsilon_{g,t}$$

where $\epsilon_{g,t} \sim \text{Normal}(0, \sigma_{g,t}^2)$ and the errors are conditionally independent given $\zeta$. Thus the likelihood function for an estimate from a national prevalence survey is

$$p(W_{g,t}|M(\theta)) = \frac{1}{\sqrt{2\pi\sigma_{g,t}^2}} \exp \left\{ -\frac{1}{2\sigma_{g,t}^2} (\logit(W_{g,t}) - \logit(\zeta_{g,t}))^2 \right\}$$

In calculating the likelihood, the value of $\sigma_{g,t}^2$ is replaced by an estimate $\hat{\sigma}_{g,t}^2$ based on the confidence intervals reported by the HSRC which account for the complex sampling strategy. The confidence intervals in the HSRC reports are on the inverse-logit scale, so the error variances are estimated by

$$\hat{\sigma}_{g,t}^2 = \left( \frac{\logit(CI_{g,t}^{\text{max}}) - \logit(CI_{g,t}^{\text{min}})}{2 \cdot \Phi^{-1}(.975)} \right)^2$$

where $\Phi^{-1}$ is the inverse of standard normal cumulative distribution function.

For the antenatal clinic data, similar to [19, 3], we assume that HIV prevalence $W_{C,t}$ amongst antenatal clinic attendees at time $t$ is linearly related to prevalence in the general female age 15 to 49 population on the logit scale and that this affect is fixed over time. To model this, we introduce an additional parameter $\gamma$ such that

$$\log \left( \frac{W_{C,t}}{1 - W_{C,t}} \right) = \log \left( \frac{\zeta_{F,t}}{1 - \zeta_{F,t}} \right) + \gamma + \epsilon_{C,t},$$

where $\epsilon_{C,t} \sim \text{Normal}(0, \sigma_{C,t}^2)$ and recalling that $\zeta_{F,t}$ is the HIV prevalence amongst females aged 15 to 49 predicted by the model at time $t$. Thus

$$p(W_{C,t}|M(\theta)) = \frac{1}{\sqrt{2\pi\sigma_{C,t}^2}} \exp \left\{ -\frac{1}{2\sigma_{C,t}^2} (\logit(W_{C,t}) - (\logit(\zeta_{F,t}) + \gamma))^2 \right\}$$

Once again, when evaluating the likelihood, we replace $\sigma_{C,t}^2$ with an estimate $\hat{\sigma}_{C,t}^2$ computed from the confidence intervals published by the South African Department of Health. In the case of the antenatal clinic data, the reported confidence intervals are symmetric, suggesting that they have been estimated on the untransformed scale rather than logit-transformed prevalence. Deriving an estimate of the sampling
error variance on the logit scale involves two steps: first estimating the untransformed error variance \( \hat{\sigma}^2_{C,t} \), and then using the delta method to approximate the variance \( \tilde{\sigma}^2_{C,t} \) of the logit-transformed distribution, which depends on the estimated antenatal clinic prevalence at time \( t \), \( W_{C,t} \). The equations for this are

\[
\hat{\sigma}^2_{C,t} = \left( \frac{\text{logit}(\text{CI}^{\text{max}}_{t}) - \text{logit}(\text{CI}^{\text{min}}_{t})}{2 \cdot \Phi^{-1}(0.975)} \right)^2, \quad \text{and}
\]

\[
\tilde{\sigma}^2_{C,t} = \frac{\hat{\sigma}^2_{C,t}}{W_{C,t}^2 (1 - W_{C,t})^2}
\]

To arrive at the likelihood for the full data \( W \), we assume that the data points are conditionally independent given the predicted prevalences \( \xi = M(\theta) \) and the antenatal clinic bias parameter \( \gamma \). Defining \( T_N \) to be the set of years when national survey prevalence estimate are available and \( T_C \) the set of years where antenatal clinic estimates are available, the full likelihood is

\[
p(W|M(\theta), \gamma) = \prod_{t \in T_N} \prod_{g \in \{M,F\}} p(W_{g,t}|M(\theta)) \cdot \prod_{t \in T_C} p(W_{C,t}|M(\theta), \gamma)
\]

\[
= \prod_{t \in T_N} \prod_{g \in \{M,F\}} \frac{1}{\sqrt{2\pi \sigma^2_{g,t}}} \exp \left\{ -\frac{1}{2\sigma^2_{g,t}} \left( \text{logit}(W_{g,t}) - \text{logit}(\xi_{g,t}) \right)^2 \right\}
\]

\[
= \prod_{t \in T_C} \frac{1}{\sqrt{2\pi \sigma^2_{C,t}}} \exp \left\{ -\frac{1}{2\sigma^2_{C,t}} \left( \text{logit}(W_{C,t}) - (\text{logit}(\xi_{C,t}) + \gamma) \right)^2 \right\}
\]

### 3.3.2 Priors

The prior distributions for the estimated model parameters \( \theta \) are given in Table 4. The antenatal bias parameter \( \gamma \) is assumed to have Uniform(0, 1) prior distribution. This amounts to assuming that the odds ratio of ANC prevalence to adult females prevalence is between 1 and 2.7. Restricting the prior value of \( \gamma \) to be positive ensures that the model does not mistakenly interpret the ANC prevalence as the true population prevalence when fitting the model.

### 3.3.3 Estimating the posterior distribution

Multiplying the likelihood and prior distribution yields the joint posterior distribution of the parameters \( \theta \) and \( \gamma \) given the model and data up to a scaling constant:

\[
p(\theta, \gamma|W, M) \propto p(W|\theta, \gamma, M)p(\theta, \gamma)
\]

We are principally interested in the values of the model parameters \( \theta \). The posterior distribution for \( \theta|W, M \) can be calculated by integrating out the parameter \( \gamma \). Observe that

\[
p(\theta|W, M) \propto \int_{\Omega_\gamma} p(W|\theta, \gamma, M)p(\theta, \gamma) \, d\gamma
\]

\[
= \int_{\Omega_\gamma} p(\theta)p(\gamma) \prod_{t \in T_N} \prod_{g \in \{M,F\}} p(W_{g,t}|M(\theta)) \cdot \prod_{t \in T_C} p(W_{C,t}|M(\theta), \gamma) \, d\gamma
\]

\[
= \prod_{t \in T_N} \prod_{g \in \{M,F\}} p(W_{g,t}|M(\theta))p(\theta) \cdot \int_{\Omega_\gamma} p(\gamma) \prod_{t \in T_C} p(W_{C,t}|M(\theta), \gamma) \, d\gamma,
\]

so if we can efficiently evaluate \( \int_{\Omega_\gamma} p(\gamma) \prod_{t \in T_C} p(W_{C,t}|M(\theta), \gamma) \, d\gamma \), then we can efficiently ... Before
attacking this, let us define three useful quantities:

\[ S^2 = \left( \sum_{t \in T_C} \frac{1}{\sigma_{C,t}^2} \right)^{-1} \]

\[ D = S^2 \cdot \sum_{t \in T_C} \frac{W_{C,t} - \zeta_{F,t}}{\sigma_{C,t}^2} \]

\[ D^2 = S^2 \cdot \sum_{t \in T_C} \frac{(W_{C,t} - \zeta_{F,t})^2}{\sigma_{C,t}^2} \]

The first can be thought of as the pooled variance of the ANC prevalence estimates, the second as the precision-weighted mean difference between ANC data prevalence and the model predicted female prevalence, and the third as the precision-weight mean squared distance between the ANC data and the predicted female prevalence.

Now, again consider our integral. We will do a bit of rearranging to show that the integral can be evaluated as a normal cumulative distribution function. For brevity, denote \( \dot{W}_t = \logit(W_{C,t}) \) and \( \dot{\zeta}_t = \logit(\zeta_{F,t}) \), and all sums and products are over the set \( T_C \).

\[
\int_{\Omega} \frac{1}{\prod_t 2\pi \sigma_{C,t}^2} \exp \left\{ -\frac{1}{2\sigma_{C,t}^2} \left( \dot{W}_t - (\dot{\zeta}_t + \gamma) \right)^2 \right\} \, d\gamma
= \int_0^1 \prod_t \frac{1}{2\pi \sigma_{C,t}^2} \exp \left\{ -\frac{1}{2\sigma_{C,t}^2} \left( \gamma - (\hat{W}_t - \dot{\zeta}_t) \right)^2 \right\} \, d\gamma
= K \cdot \int_0^1 \exp \left\{ -\frac{1}{2 \sigma_{C,t}^2} \left( \frac{\gamma^2}{S^2} - 2\gamma D/S^2 + D^2/S^2 \right) \right\} \, d\gamma
= K \cdot \int_0^1 \exp \left\{ -\frac{1}{2 S^2} \left( \gamma - \frac{D}{S^2} \right)^2 + \frac{1}{2 S^2} \left( \frac{D^2}{S^2} - \frac{D^2}{S^2} \right) \right\} \, d\gamma
= K \cdot e^{\frac{(D^2 - D^2)/(2S^2)}{2\pi S^2}} \int_0^1 \frac{1}{\sqrt{2\pi S^2}} e^{-\frac{(\gamma - D)^2}{2S^2}} \, d\gamma
= \sqrt{\frac{2\pi S^2}{\prod_t 2\pi \sigma_{C,t}^2}} \cdot e^{\frac{(\bar{D}^2 - \bar{D}^2)/(2S^2)}{2\pi S^2}} \left[ \Phi \left( \frac{1 - \bar{D}}{\sqrt{S^2}} \right) - \Phi \left( \frac{0 - \bar{D}}{\sqrt{S^2}} \right) \right]
\]

Using this expression, we can efficiently evaluate the posterior density function \( p(\theta | \mathbf{W}, M) \) up to a constant. We use the incremental mixture importance sampling (IMIS) algorithm to approximate and sample from the posterior distribution [20].
References


