An uncertainty analysis and sensitivity analysis of the ASSA2002 AIDS and Demographic model

Centre for Actuarial Research working paper
March 2006

By Leigh Johnson1, Rob Dorrington1, Alan Matthews2

1 Centre for Actuarial Research, University of Cape Town
2 School of Physics, University of KwaZulu-Natal

Available: www.commerce.uct.ac.za/care

Correspondence to:
Leigh Johnson
Centre for Actuarial Research
10 University Avenue
University of Cape Town
Private Bag
Rondebosch 7701
South Africa

Tel: +27 21 650 5761
Fax: +27 21 650 5475
E-mail: ljohnson@commerce.uct.ac.za
Synopsis

Background

HIV/AIDS statistics and models are a source of much controversy in South Africa. Until recently, only estimates of the expected impact of the epidemic were produced by most HIV/AIDS models, with no indication being given of the uncertainty around these projections. The estimates of different agencies have differed substantially, and the lack of uncertainty analysis has made the estimates of the different agencies difficult to reconcile, which in turn has undermined the credibility of HIV/AIDS models. The lack of uncertainty analysis has also made it difficult for policy makers to anticipate the range of possible HIV/AIDS scenarios that might be expected in future.

The objective of this paper is to describe an uncertainty analysis and sensitivity analysis of the ASSA2002 AIDS and Demographic model, a model of the HIV/AIDS epidemic in South Africa. The uncertainty analysis aims to quantify the extent of the uncertainty in the model projections, taking into account uncertainty regarding the epidemiological parameters in the model as well as uncertainty regarding the extent to which the HIV prevalence data and mortality data in South Africa may be biased. The sensitivity analysis aims to assess how sensitive the model outputs are to changes in the epidemiological parameters. This will lead to a deeper understanding of the principal factors driving the epidemic at different stages in its development.

Method

The scope of both the uncertainty analysis and sensitivity analysis is restricted to 24 epidemiological parameters that significantly affect the correspondence between modelled and actual HIV prevalence and mortality levels, up to 2005. The uncertainty analysis was conducted using a Bayesian approach similar to Generalized Likelihood Uncertainty Estimation (GLUE) and Sampling/Importance Resampling (SIR). Prior distributions for the 24 parameters were set based on a review of the literature relating to these parameters. 100 000 parameter combinations were then sampled from these prior distributions. For each parameter combination, a likelihood function was calculated, representing the extent to which model estimates produced by the parameter combination were consistent with reported mortality and HIV prevalence levels. Four data sets were used in defining this likelihood function:

- the HIV prevalence rates observed in the annual antenatal clinic surveys (1991-2004),
- the reported deaths from the Department of Home Affairs (1997-2002),
- the HIV prevalence rates observed in the Nelson Mandela/HSRC household surveys (2002 and 2005) and
- the HIV prevalence rates observed in the RHRU/loveLife youth survey (2003).

After the likelihood values were calculated for all 100 000 parameter combinations, a sample of 500 parameter combinations was drawn (with replacement) from the 100 000,
using the likelihood values as weights. These 500 parameter combinations therefore represent a sample from the posterior distribution. Average model outputs and 95% prediction intervals were determined using these 500 parameter combinations.

**Results of uncertainty analysis**

In general, levels of HIV prevalence and mortality, averaged across the 500 parameter combinations, were consistent with the actual data. Notable exceptions were antenatal prevalence rates in the 25 to 29 age band (recent survey results tend to lie above the average levels predicted by the model) and male and female mortality in the 35 to 44 age band (recent reported deaths tend to be below the average levels predicted by the model). The failure of the model to reproduce the empirical data in these age bands may be due to a number of factors, such as the model assumption that effectiveness of HIV prevention and treatment programmes is the same at all adult ages. Other differences between actual reported deaths and model estimates of reported deaths suggest that some of the demographic assumptions in the ASSA2002 model may need to be revised.

![Graphs showing annual new HIV infections, total HIV infections, total AIDS deaths, and total AIDS sick](image)

The above figure shows average model outputs (solid lines) and 2.5 and 97.5 percentiles of the model output distributions (dashed lines). The results of the uncertainty analysis indicate that HIV incidence rates in South Africa have been declining since about 1998. The average estimated number of new HIV infections occurring in South Africa between mid-2004 and mid-2005 is 490 000, with a 95% prediction interval of 370 000 to 590 000. The average estimated number of HIV infections in mid-2005 is 5.1 million (95% interval: 4.2 – 6.0 million), which is equivalent to an HIV prevalence rate of 11.1% (95% interval: 9.1 – 13.1%). AIDS mortality has increased dramatically over the last decade. The average estimated number of AIDS deaths occurring between mid-2004 and mid-2005 is 320 000, with a 95% prediction interval of 270 000 to 380 000. Total AIDS
cases have followed a similar trend, and the average estimated number of AIDS cases in mid-2005 is 590 000 (95% interval: 500 000 – 680 000).

The prior and posterior distributions were compared, for each of the 24 parameters, in order to determine whether the HIV prevalence data and mortality data suggested epidemiological parameters significantly different from those assumed. The average HIV survival time, for an individual infected at age 29, has a posterior distribution with a mean of 11.5 years, substantially greater than the mean of 8 to 9 years that is usually assumed when modelling the HIV/AIDS epidemic in Africa. The posterior distributions also suggest that the effect of HIV on fertility in South Africa is smaller than that suggested by other African studies. The HIV prevalence data and mortality data also support a greater variance of HIV survival times than that originally assumed, and a less significant effect of HIV viral load on HIV infectiousness than originally assumed.

**Sensitivity analysis**

The sensitivity analysis was performed using the same 500 parameter combinations previously sampled, after removing those parameter combinations which had been resampled. Six model outputs were calculated for each parameter combination, reflecting HIV transmission dynamics at different stages in the epidemic. Partial rank correlation coefficients (PRCCs) were calculated for each combination of model outputs and model parameters. Most parameters were found to be significantly correlated with the model outputs, though the extent of the correlation changed significantly over the course of the epidemic in most cases. For example, the relative frequency of sex in symptomatic HIV infection and the effect of HIV viral load on HIV infectiousness are both negatively related to the incidence of HIV the early stages of the epidemic, but are positively related to the incidence of HIV in the later stages of the epidemic. This is due to the increasing proportion of HIV-positive individuals in later disease stages as the epidemic matures.

**Comparison with UNAIDS estimates**

To date, UNAIDS is the only other agency to have produced 95% intervals around its HIV/AIDS model estimates. Its estimates of total HIV infections in South Africa are slightly greater than those obtained using the ASSA2002 approach, though the two approaches produce 95% intervals which are similar in width. The UNAIDS estimates of AIDS deaths in South Africa, however, are substantially greater than those produced using the ASSA2002 model, and its 95% intervals around these AIDS mortality estimates are almost three times the width of the ASSA2002 intervals. The UNAIDS model is calibrated mainly against HIV prevalence data from the antenatal clinic surveys, and the fact that it does not take mortality data into account explains the greater range of uncertainty around the UNAIDS mortality estimates. The higher levels of AIDS mortality in the UNAIDS model are largely the result of the shorter HIV survival times assumed in the UNAIDS model. They are also partly the result of the higher HIV prevalence in the UNAIDS model, which is due to the antenatal data not being adjusted for various sources of bias. These include bias towards urban antenatal clinics in the early antenatal surveys, bias due to the exclusion of women seeking antenatal care in the private health sector,
and bias due to false positive reactions on HIV antibody tests. The comparison of prior
and posterior distributions has shown that HIV has little effect on fertility in South
Africa, and antenatal data are therefore less likely to understate HIV prevalence in
women than has been previously assumed. These unusual features of the South African
epidemic need to be considered in future models of the HIV/AIDS epidemic.

**Strengths and limitations**

A strength of the approach developed here is that it is able to integrate data from a
number of sources, both mortality data sources and HIV prevalence data sources. The use
of Bayesian techniques is particularly appropriate in this context, as there is much prior
knowledge regarding HIV/AIDS epidemiology which needs to be incorporated into the
statistical analysis. A limitation of the approach, however, is that a degree of subjectivity
is unavoidable in the specification of the likelihood function.

This analysis has not considered uncertainty with respect to the effects of antiretroviral
treatment and HIV prevention programmes. Antiretroviral treatment and HIV vaccines
could both have a significant impact on the future evolution of the South African
HIV/AIDS epidemic, and this uncertainty analysis and sensitivity analysis has therefore
been based only on model outputs up to 2005. Future work will explore the uncertainty
introduced by these vaccine and antiretroviral parameters, and the sensitivity of the
model projections to changes in these parameters. This analysis has also not considered
uncertainty relating to paediatric HIV transmission and paediatric HIV survival, which
will also be considered in future work. Further work is also required to assess uncertainty
with respect to demographic parameters.
Table of contents

Synopsis .......................................................................................................................... 2
Table of contents ............................................................................................................. 6
List of abbreviations ....................................................................................................... 7

1. Introduction ................................................................................................................. 8
2. Review of uncertainty and sensitivity analysis techniques ........................................ 10
   2.1 Uncertainty analysis for calibrated models ......................................................... 10
   2.2 Uncertainty analysis for models that are not calibrated ...................................... 14
   2.3 Sensitivity analysis techniques ........................................................................... 16
3. Method ...................................................................................................................... 20
   3.1 Overview of the method ...................................................................................... 20
   3.2 Description of the ASSA2002 model ................................................................ 21
   3.3 Specification of prior assumptions ...................................................................... 23
   3.4 Specification of the likelihood function ............................................................... 41
   3.5 Programming ....................................................................................................... 46

4. Uncertainty analysis results ...................................................................................... 48
   4.1 Comparison of observations and model outputs ................................................. 48
   4.2 Model projections up to 2005 ............................................................................. 52
   4.3 Comparison of prior and posterior distributions ................................................. 53

5. Sensitivity analysis results ........................................................................................ 57
   5.1 Bias parameters ................................................................................................... 58
   5.2 HIV transmission parameters ............................................................................ 58
   5.3 HIV survival parameters ..................................................................................... 59
   5.4 Sexual behaviour parameters .............................................................................. 60

6. Discussion ................................................................................................................. 62
   6.1 Comparison with other uncertainty and sensitivity analyses ................................ 62
   6.2 Strengths and limitations ..................................................................................... 63
   6.3 Key findings ......................................................................................................... 64
   6.4 Future work ......................................................................................................... 66

Acknowledgements ....................................................................................................... 69
References ..................................................................................................................... 70
Appendix A: Changes made to the ASSA2002 model ................................................. 81
Appendix B: Review of ELISA sensitivities and specificities in African populations 83
Appendix C: Calculation of male rates of partner change and proportions of male partners in different risk groups ................................................................. 85
Appendix D: Approach to setting Nu parameters .......................................................... 88
Appendix E: Antenatal clinic prevalence data ............................................................... 96
Appendix F: Nelson Mandela/HSRC household prevalence survey data ................... 98
**List of abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS</td>
<td>Acquired immune deficiency syndrome</td>
</tr>
<tr>
<td>ANC</td>
<td>Antenatal clinic</td>
</tr>
<tr>
<td>ASSA</td>
<td>Actuarial Society of South Africa</td>
</tr>
<tr>
<td>DHA</td>
<td>Department of Home Affairs</td>
</tr>
<tr>
<td>DOH</td>
<td>Department of Health</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunoassay</td>
</tr>
<tr>
<td>GGB</td>
<td>General growth balance</td>
</tr>
<tr>
<td>GLUE</td>
<td>Generalized likelihood uncertainty estimation</td>
</tr>
<tr>
<td>GSA</td>
<td>Global sensitivity analysis</td>
</tr>
<tr>
<td>HAART</td>
<td>Highly active antiretroviral treatment</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HSRC</td>
<td>Human Sciences Research Council</td>
</tr>
<tr>
<td>HSV-2</td>
<td>Herpes simplex virus type 2</td>
</tr>
<tr>
<td>LHS</td>
<td>Latin hypercube sampling</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov chain Monte Carlo</td>
</tr>
<tr>
<td>MCS</td>
<td>Monte Carlo simulation</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal components analysis</td>
</tr>
<tr>
<td>PCC</td>
<td>Partial correlation coefficient</td>
</tr>
<tr>
<td>PRCC</td>
<td>Partial rank correlation coefficient</td>
</tr>
<tr>
<td>RHRU</td>
<td>Reproduction Health Research Unit</td>
</tr>
<tr>
<td>SIR</td>
<td>Sampling/importance resampling</td>
</tr>
<tr>
<td>RSA</td>
<td>Regional sensitivity analysis</td>
</tr>
<tr>
<td>SEG</td>
<td>Synthetic extinct generations</td>
</tr>
<tr>
<td>STD</td>
<td>Sexually transmitted disease</td>
</tr>
<tr>
<td>VBA</td>
<td>Visual Basic for Applications</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
1. Introduction

HIV/AIDS statistics and models are a source of much controversy in South Africa. HIV prevalence data and vital registration data, by themselves, provide an incomplete and potentially misleading picture of how the epidemic is progressing, and it is therefore necessary to use mathematical models to extrapolate from these data. However, the models that have been applied in South Africa produce very different results, with recent estimates of the number of HIV infections in South Africa ranging between 4.5 million (Statistics South Africa 2005) and 6.3 million (Department of Health 2005). The much-publicized debate around these differences has added to public confusion and undermined the credibility of HIV/AIDS models and modelling in general.

Until recently, most deterministic HIV/AIDS models have been used to project only the expected impact of HIV/AIDS, and published estimates from these models have not included any indication of the range of uncertainty around these projections. There is a need for modellers to indicate the extent of this uncertainty, not only to determine the significance of the differences between estimates of different agencies, but also to enable those planning for the potential impact of HIV/AIDS to assess the range of possible future scenarios.

The objective of this paper is to describe a Bayesian approach to assessing uncertainty in HIV/AIDS models. This approach acknowledges uncertainty with respect to major epidemiological parameters, and integrates data from four South African data sets into a single statistical framework. The statistical framework is developed using a C++ version of the ASSA2002 AIDS and Demographic model, a model of the HIV/AIDS epidemic in South Africa. This will make it possible to derive 95% prediction intervals around the model outputs and to obtain credibility intervals for the key epidemiological parameters in the model. Although the approach is described in relation to the ASSA2002 model, it could also be applicable to other HIV/AIDS models.

This paper also aims to assess the sensitivity of the model estimates to the major parameters in the model. Such a sensitivity analysis will assist in developing a deeper understanding of the key factors driving the epidemic at different stages in its development. It can also be used to identify those model parameters which require further empirical investigation, for the purpose of increasing confidence in the output of the model.

The focus of this analysis is restricted to a selection of the ASSA2002 parameters which determine sexual behaviour, HIV transmission, HIV survival and bias in HIV prevalence data and mortality data. While not accounting for all of the uncertainty in the model, these parameters are believed to account for almost all of the uncertainty when assessing the epidemiological impact of the HIV/AIDS epidemic up to the current time.

The paper begins with a review of uncertainty and sensitivity analysis techniques (section 2). This is followed by a description of the approach taken in this analysis.
(section 3), much of which is devoted to reviewing the empirical evidence on which the model parameters are based. In section 4, the results of the uncertainty analysis are presented, including a detailed comparison of the model estimates with the data to which the model is calibrated. The sensitivity of model outputs to different parameters is examined in section 5. Finally, the key findings are discussed and the limitations of both the uncertainty analysis technique and the ASSA2002 model are considered (section 6).
2 Review of uncertainty and sensitivity analysis techniques

Mathematical models of complex systems often fail to reproduce observed experience, for a number of reasons. Firstly, reality is governed by random processes, but deterministic models represent the experience one might expect to observe on average, in the absence of random variation. Secondly, model parameters are often not known precisely, and to the extent that the true parameters differ from those assumed, the output of the model can be expected to differ from the actual experience. Thirdly, the structure of the model is almost invariably a simplification of reality. Inadequate allowance for complexity and inaccurate specification of relationships between variables can result in a model that has poor predictive validity.

The first source of error can be quantified by stochastic models, which allow explicitly for random processes. The second source of error is the source of error with which modellers are often most concerned. Many techniques have been developed to assess variability in model outputs associated with uncertainty regarding input parameters (uncertainty analysis) and to assess the relative significance of different parameters in influencing model outputs (sensitivity analysis). These techniques are discussed in this chapter. The third source of error is the most difficult to quantify, and no standard techniques have been developed to assess the extent of potential error due to model misspecification.

In many models, the range of parameter values that can reasonably be entered into the model is to some extent constrained by the need to maintain consistency between model outputs and empirical data. In these ‘calibrated models’, uncertainty analysis is more complex because of this requirement for consistency. Techniques for uncertainty analysis are therefore discussed separately for calibrated models and models that are not calibrated. Sensitivity analysis techniques, however, are equally applicable to both types of model.

2.1 Uncertainty analysis for calibrated models

Approaches to uncertainty analysis for calibrated models can be classified as falling within either the frequentist or Bayesian paradigms. In the frequentist paradigm, inferences about the parameters in the model and the uncertainty relating to these parameters are based only on empirical data. In the Bayesian paradigm, the same empirical data are used to draw inferences, but in addition the modeller’s a priori beliefs are incorporated. Because of the estimation of parameters at different ‘levels’ (a priori and empirical), the Bayesian approach is sometimes also referred to as the ‘multilevel modelling’ approach (Greenland 2000).

An example of the use of the frequentist approach to uncertainty analysis is the approach adopted by UNAIDS in deriving ‘plausibility bounds’ for HIV/AIDS projections in countries with generalized HIV/AIDS epidemics. A model with four variable parameters
is fitted to antenatal clinic prevalence data using a least squares criterion. Following this, a bootstrap approach is followed; this involves ‘resampling’ the antenatal data 1000 times, on the assumption that the originally fitted parameters determine the distributions from which the subsequent samples are drawn. The four parameters in the model are then refitted for each simulated sample, according to the same least squares criterion (Grassly et al, 2004).

Another frequentist approach, used by Salomon and Murray (2001), is to construct a distribution around each of the parameters estimated, using the standard error of the maximum likelihood estimate produced by the statistical software. Salomon and Murray use this approach in fitting an eight-parameter adaptation of Epimodel to antenatal clinic HIV prevalence data from four African countries. 20 000 parameter combinations are generated by randomly sampling from the distributions constructed, and 95% confidence intervals for model outputs are obtained by rejecting those parameter combinations that give a significantly poorer fit to the observed data than the parameter combination which maximizes the likelihood, using a likelihood ratio test.

The Bayesian approach can be explained mathematically as follows. Suppose that \( \vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_d) \) is a vector of \( d \) model parameters which the modeller wishes to explore in an uncertainty analysis. The modeller’s a priori beliefs about the likely values of these parameters are specified by the joint distribution \( p(\vec{\theta}) \), referred to as the prior distribution. Further assume that \( \vec{x}_i \) is a vector of observations to which the model is calibrated, and that the likelihood function \( L(\vec{x}_i | \vec{\theta}) \) represents the probability of observing the outcome \( \vec{x}_i \) given that \( \vec{\theta} \) represents the parameters governing the process. Then, by Bayes Theorem, the probability that \( \vec{\theta} \) is the true parameter set, given the observed outcomes \( \vec{x}_i \), is

\[
p(\vec{\theta} | \vec{x}_i) = \frac{p(\vec{\theta})L(\vec{x}_i | \vec{\theta})}{\int \int \ldots \int p(\vec{\theta})L(\vec{x}_i | \vec{\theta})d\theta_1 \cdots d\theta_d}.
\]  

Since the denominator is constant with respect to \( \vec{\theta} \), equation (1) can be rewritten as

\[
p(\vec{\theta} | \vec{x}_i) \propto p(\vec{\theta})L(\vec{x}_i | \vec{\theta}).
\]  

\( p(\vec{\theta} | \vec{x}_i) \) is referred to as the posterior distribution, and represents the modeller’s beliefs about the parameters \( \vec{\theta} \), updated to take into account the empirical data \( \vec{x}_i \).

A number of Bayesian techniques have been developed for the purpose of uncertainty analysis. The remainder of this section focuses on two important Bayesian approaches in uncertainty analysis: firstly, the Generalized Likelihood Uncertainty Estimation (GLUE) and related Sampling/Importance Resampling (SIR) approaches, and secondly, the Markov Chain Monte Carlo (MCMC) approach. Both make use of Monte Carlo
simulation techniques to approximate the posterior distribution, as it is impossible to calculate the posterior distribution analytically, except in the simplest models.

In the GLUE and SIR approaches, prior distributions are usually specified independently for each parameter. The key difference between the two approaches is that the GLUE approach requires a less formal specification of the likelihood than the SIR approach. Under the GLUE approach, the ‘likelihood’ function is at the discretion of the modeller, and does not have to conform to the statistical formulation of the likelihood used in maximum likelihood theory (Beven and Binley 1992). Ratto et al (2001), for example, use the likelihood function

\[
\left( \frac{1}{2n} \sum_{i=1}^{n} (E_i - O_i)^2 \right)^{-N}
\]

where \( n \) is the number of observations to which the model is calibrated, \( O_i \) is the \( i^{\text{th}} \) observation (so that \( \mathbf{x}_i \) in equations (1) and (2) corresponds to the vector of \( O_i \) values), \( E_i \) is the model prediction corresponding to the \( i^{\text{th}} \) observation, and \( N \) is an arbitrary constant chosen by the modeller. The higher the value of \( N \), the greater the relative weight given to those simulations that best fit the empirical data.

Once the prior and likelihood functions have been specified, a large number of combinations of parameters, \( \theta_j \) (\( j = 1, 2, \ldots, m \)), are randomly drawn from the prior distributions. For each parameter combination, the likelihood is calculated. ‘Non-behavioral’ parameter combinations (i.e. those parameter combinations which have likelihood values below a predetermined threshold, or which in some way violate the constraints of the model) are excluded from further analysis. The posterior distribution can then be simulated using weighted bootstrapping (also known as Sampling/Importance Resampling, or SIR) applied to the behavioural parameter combinations. This involves weighting each behavioural parameter combination according to the likelihood value calculated for it, and then sampling from the behavioural parameter combinations in proportion to their likelihood values (Smith and Gelfand 1992).

Under the MCMC approach, different parameter combinations are sampled sequentially to approximate the posterior distribution. The key difference between the GLUE/SIR approach and the MCMC approach is that parameter combinations are sampled independently of one another under the GLUE/SIR approach, but the parameter combination sampled under the MCMC approach is dependent on the previous parameter combinations sampled. If parameter combinations \( \theta_1, \theta_2, \ldots, \theta_t \) have been generated already, the parameter combination \( \theta_{t+1} \) sampled will depend on \( \theta_t \), but will be independent of \( \theta_1, \theta_2, \ldots, \theta_{t-1} \) if \( \theta_t \) is known. The sequence \( \{ \theta_1, \theta_2, \ldots \} \) thus constitutes a Markov chain – hence the term Markov Chain Monte Carlo simulation.

The Gibbs sampler is the most basic application of the MCMC algorithm, which can only be used when the conditional posterior distribution, \( p(\theta_j | \theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_d, \mathbf{x}_i) \), has a
standard form which is invertible and thus can be sampled from. An arbitrary starting parameter combination \( \theta_i \) is chosen, and parameters are then updated one at a time, so that the next value sampled for parameter \( \theta_j \) is dependent on the most recently sampled values of \( \theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_d \) and on \( x_i \). In the common situation in which it is not possible to sample from the conditional distribution directly, it is necessary to use the Metropolis algorithm or Metropolis-Hastings algorithm instead.

Using the Metropolis algorithm, the initial parameter combination \( \theta_i \) is again chosen arbitrarily. For the purpose of sampling \( \theta_i \), based on \( \theta_{i-1} \), it is necessary to specify a jumping distribution (also known as a proposal distribution or a candidate-generating distribution), \( J(\theta^* | \theta_{i-1}) \). The posterior distribution is calculated for both \( \theta_{i-1} \) and the proposed new parameter combination \( \theta^* \). If \( p(\theta^* | x_i) > p(\theta_{i-1} | x_i) \), then \( \theta_i \) is set to be \( \theta^* \), otherwise \( \theta_i \) is randomly set to be either \( \theta^* \) or \( \theta_{i-1} \) (see Gelman et al (2004) for a more detailed description). This means that the proposed parameter set \( \theta^* \) will automatically be accepted as the parameter combination \( \theta_i \) if its posterior likelihood is greater than that of \( \theta_{i-1} \), otherwise \( \theta_i \) is set to \( \theta_{i-1} \) or \( \theta^* \).

Using the Metropolis algorithm, parameter combinations are updated one at a time, while with the Gibbs sampler, individual parameters are updated one at a time. It is not necessary to specify a jumping distribution when using the Gibbs sampler, since the conditional posterior distribution can be sampled directly. The jumping distribution used in the Metropolis algorithm must be symmetric, i.e. of such a form that \( J(\theta_a | \theta_b) = J(\theta_b | \theta_a) \). This requirement is removed when using the Metropolis-Hastings algorithm, which uses a different procedure for determining whether \( \theta_i \) is set to \( \theta^* \) or \( \theta_{i-1} \).

All MCMC techniques involve generating sequences of parameter combinations \{\( \theta_1, \theta_2, \ldots \}\), and then sampling from these sequences to simulate the posterior distribution (see Gelman et al (2004) for a more detailed explanation). Uncertainty analysis then proceeds by running the model using these different parameter combinations and examining the distribution of the results produced.

Complex models with large numbers of parameters are often non-invertible, i.e. a number of different parameter combinations can give the same or very similar correspondence to a given set of output targets. This is often referred to as ‘model equifinality’ or over-parameterization. The GLUE approach is often recommended as the appropriate uncertainty analysis tool for these non-invertible models (Beven and Binley 1992; Ratto et al, 2001; Romanowicz and Young 2003). MCMC techniques may prove difficult to apply with such models. O'Neill et al (2000), for example, note that a large number of parameters in a model can make convergence more difficult to achieve when using MCMC algorithms, i.e. the Markov chain may tend towards different points on the
posterior likelihood surface depending on the arbitrarily chosen initial parameter combination $\tilde{\theta}_1$.

The Bayesian techniques described above can easily be updated to incorporate new sample information. Using the GLUE approach, for example, Beven and Binley (1992) illustrate how a new vector of observations $\bar{x}_2$ can be incorporated, essentially by treating $p(\tilde{\theta} | \bar{x}_1)$ as the prior distribution and $p(\tilde{\theta} | \bar{x}_1, \bar{x}_2)$ as the posterior distribution. Similarly, the Static Updating Theorem can be used to update estimates of parameters for models in which outputs are linearly related to inputs (Catlin 1989). The same recursive procedures can be used as further data ($\bar{x}_3, \bar{x}_4, \ldots$) become available.

### 2.2 Uncertainty analysis for models that are not calibrated

Five uncertainty analysis techniques are considered here:

- Monte Carlo simulation
- Latin Hypercube Sampling
- Factorial or fractional factorial sampling
- Rosenblueth’s point estimation method
- Harr’s point estimation method

Monte Carlo simulation (MCS) can be considered a simplification of the GLUE technique described in the previous section. Again, the modeller’s *a priori* beliefs about the likely values of the $d$ parameters under consideration are specified by the joint distribution $p(\tilde{\theta})$. However, because the model is not calibrated to empirical data, it is not necessary to specify a likelihood function. The modeller simply samples parameter combinations from $p(\tilde{\theta})$, enters these into the model, and then examines the range of model outputs generated. Often the *a priori* distributions for each of parameters $\theta_1, \theta_2, \ldots, \theta_d$ are specified to be independent of one another, but it is also possible to allow for correlation between parameters.

Latin Hypercube Sampling (LHS) is similar to Monte Carlo simulation, but involves sampling *without* replacement rather than sampling with replacement. Suppose one wishes to generate $m$ parameter combinations. The distribution specified for each parameter is then divided into $m$ intervals, the spacing of the intervals being such that the integral of the distribution over each interval is equal to $1/m$ (i.e. each interval is equally likely to be randomly sampled). One then samples the intervals without replacement, and for each interval sampled, one randomly samples a parameter value within the interval. In this way, $m$ parameter combinations are generated (Iman and Helton 1988; Blower and Dowlatabadi 1994). As with the MCS approach, the distributions for each of the $d$ parameters are usually specified to be independent of one another, though techniques have been developed to allow for correlation between parameters (Iman and Conover 1982).
The chief advantage of the LHS method, relative to MCS, is that it gives a better approximation to the distribution of the model output, for a given number of parameter combinations (Iman and Conover 1982; Yu et al, 2001). If the run times of the model being considered are very long, it is therefore a more efficient method for simulating the model outputs than Monte Carlo simulation, i.e. the LHS method will converge on the ‘true’ distribution of the model output more rapidly than MCS as the size of \( m \) increases. However, Iman & Conover (1982) show that although the approximation to the first and second moments of the output distribution is better under LHS than MCS, LHS does not appear to perform any better than MCS in approximating the higher order moments of the distribution.

The remaining three uncertainty techniques differ from the first two techniques in two respects. Firstly, the \textit{a priori} assumptions about the model parameters are discretized, so that instead of specifying a distribution for each parameter, one specifies only ‘high’ and ‘low’ values. Secondly, there is no ‘randomness’ in the way in which the parameter combinations are generated. In the factorial design, all possible combinations of ‘high’ and ‘low’ values are run, while in the ‘fractional factorial’ design a subset of all possible combinations are selected (Iman and Helton 1988). The factorial design can also be applied when there are more than two levels that the modeller wishes to examine for certain parameters. Boily \textit{et al} (2004), for example, use this approach when assessing uncertainty around the effect of antiretroviral treatment on the incidence of sexually transmitted diseases (STDs) other than HIV, setting five to nine levels for each of the key antiretroviral parameters.

Rosenblueth’s point estimation method is an extension of the factorial design, which takes into account \textit{a priori} assumptions about correlations between parameters. All possible parameter combinations are generated, but each combination is weighted by a probability function that takes into account the extent to which the parameter combination is consistent with the assumed correlations (Yu \textit{et al}, 2001). The factorial design can be regarded as a special case of Rosenblueth’s point estimation method, in which all correlations between parameters are set to zero.

A problem with the factorial design and Rosenblueth’s point estimation method is that the techniques become unwieldy if there are a large number of parameters under consideration; for each additional parameter, there is at least a doubling of the number of parameter combinations to be considered. Harr’s point estimation method is an adaptation of Rosenblueth’s point estimation method, which attempts to overcome this problem. If there are \( p \) parameters under consideration, and two levels are being considered for each parameter, a total of \( 2^p \) simulations would be required when using Rosenblueth’s point estimation method. Using Harr’s point estimation method, this number is reduced considerably to \( 2p \). This technique is mathematically complex, and the approach cannot be presented in simple terms here. For a description of the technique, the reader is referred to Yu \textit{et al} (2001).
2.3 Sensitivity analysis techniques

The primary objective of a sensitivity analysis is to assess the relative significance of the parameters that influence the model outputs. A secondary objective may be to assess how much of this influence is due to the independent effect of the parameter being considered, and how much is due to its interaction with other parameters in the model. A model output may be strongly correlated with a particular parameter, but this does not necessarily imply that the variation in that parameter is a significant cause of variation in the output – the parameter may merely be correlated with other parameters that significantly affect the model output. The modeller may therefore be interested in a measure of sensitivity that is independent of interactions with other parameters, and may wish to explore the interactions between parameters.

These two objectives are met to varying degrees by the five broad sensitivity analysis techniques that are considered here:

- Regional sensitivity analysis (RSA)
- Scatter plots
- Global sensitivity analysis (GSA)
- Principal components analysis (PCA)
- Correlation coefficients between model parameters and outputs

Sensitivity analysis and uncertainty analysis often go hand in hand, and the techniques described here build upon the uncertainty analysis techniques described in the previous sections.

Regional sensitivity analysis (RSA) techniques are visualization techniques that were developed primarily to meet the first objective. These techniques are used with calibrated models, for which some sort of likelihood measure is calculated for each of the parameter combinations considered in the uncertainty analysis. One approach is to separate the ‘prior distribution’ for a given parameter into behavioural and non-behavioural components, i.e. plot separately the distribution of behavioural parameter combinations and the distribution of non-behavioural parameter combinations, with respect to the parameter of interest. The extent to which the two distributions diverge provides some measure of sensitivity (Spear et al., 1994). A related visualization technique is the construction of scatter plots; for each parameter combination, the likelihood measure is plotted against the value of the parameter of interest (Ratto et al., 2001). The greater the variation in the likelihood over the range of the parameter concerned, the more sensitive the ‘fit’ of the model to the parameter concerned.

Global sensitivity analysis (GSA) techniques were developed to assess both overall sensitivity and the independent effect of a parameter on model outputs (Archer et al., 1997). Like the RSA techniques, they are applicable mainly to calibrated models. Suppose that the model output of interest is \( Y \) and the modeller wishes to consider its sensitivity to parameter \( \theta_i \). The main effect or first order sensitivity index \( S_i \) is defined as
This represents the degree of variation in $Y$ that is attributable to parameter $\theta_i$, independent of any interaction between $\theta_i$ and other parameters. The total sensitivity index is defined as

$$S_{\theta_i} = \frac{\text{Var}[E(Y \mid \theta_i)]}{\text{Var}[Y]}.$$ (4)

The total sensitivity index represents the overall effect of $\theta_i$ on output $Y$ through interactions with other factors (Ratto et al, 2001). The $S_i$ and $S_{\theta_i}$ statistics are considered in combination. For example, a high value of $S_{\theta_i}$ and a low value of $S_i$ suggests that output $Y$ is strongly correlated with parameter $\theta_i$, but mostly as a result on interactions with other parameters. This in turn suggests that there is a degree of model equifinality or over-parameterization.

Principal components analysis (PCA) is used to examine interactions between parameters in calibrated models. If there are $d$ parameters, the objective of a PCA is to identify the $d \times 1$ vector which captures the greatest variation in the $d$-dimensional parameter space (the first principal component), the vector orthogonal to this which captures the most remaining variation (second principal component), the vector orthogonal to the first and second principle components which captures the most residual variation (third principal component), and so on up to the $d$th principal component. Finding the vector which captures the greatest variation in $d$-dimensional space is equivalent to finding the vector for which the sum of squared differences between each sampled parameter combination and the vector is minimized.

The principal components can be used to assess which parameters are most variable and how the parameters interact with each other (see Ratto et al (2001) for an example). One can also think of the first principal component as the vector along which the greatest degree of equifinality is preserved, i.e. for a given behavioural parameter combination, likelihood values are likely to be similar for other parameter combinations that are expressed as scalar multiples of the first principal component plus the given behavioural parameter combination (Ratto et al, 2001).

Correlation coefficients are generally used with non-calibrated models. The Pearson correlation coefficient between output $Y$ and parameter $\theta_i$ is calculated as

$$S_{\theta_i} = \frac{\text{E}[\text{Var}(Y \mid \theta_i, \theta_{i+1}, \theta_{i+2}, \ldots, \theta_d)]}{\text{Var}[Y]}.$$ (5)
\[
\frac{\sum_{j=1}^{m} (y_j - \mu_y) (\theta_i - \mu_i)}{\sqrt{\sum_{j=1}^{m} (y_j - \mu_y)^2 \sum_{j=1}^{m} (\theta_i - \mu_i)^2}},
\]

where \( \theta_i \) is the \( j \)th value of parameter \( \theta_i \) sampled, \( y_j \) is the model output associated with the \( j \)th parameter combination, \( m \) is the number of parameter combinations sampled, and \( \mu_y \) and \( \mu_i \) are the average values of \( Y \) and \( \theta_i \) respectively, from the sample of size \( m \).

This correlation coefficient provides information on the amount of correlation between \( Y \) and \( \theta_i \), but cannot be used to assess the independent contribution of \( \theta_i \) to the variation in \( Y \). The partial correlation coefficient (PCC) is used to estimate this independent contribution. First define the \((d+1) \times (d+1)\) matrix \( C \) as follows:

\[
c_{ij} = \frac{\sum_{k=1}^{m} (\theta_{ik} - \mu_i) (\theta_{jk} - \mu_j)}{\sqrt{\sum_{k=1}^{m} (\theta_{ik} - \mu_i)^2 \sum_{k=1}^{m} (\theta_{jk} - \mu_j)^2}} \quad i, j = 1, 2, \ldots, d
\]

where \( d \) is the number of parameters in the model. For \( i = d + 1 \), \( \theta_{ik} \) is replaced in equation (7) by \( y_k \), and \( \mu_i \) is replaced by \( \mu_y \). Similarly, for \( j = d + 1 \), \( \theta_{jk} \) is replaced by \( y_k \), and \( \mu_j \) is replaced by \( \mu_y \). Matrix \( C \) is then inverted; suppose \( B \) is the inverse. The PCC between parameter \( \theta_i \) and output \( Y \) is then calculated as

\[
\frac{-b_{i,d+1}}{\sqrt{b_i b_{d+1,d+1}}}
\]

The PCC then represents the strength of the linear relationship between \( Y \) and \( \theta_i \), after removing the linear effect on \( Y \) of all the other model parameters (Iman and Helton 1988; Blower and Dowlatabadi 1994).

A shortcoming of the PCC measure is that it does not provide a reliable measure of sensitivity if the relationship between the parameter and model output of interest is non-linear. This problem can be overcome by using rank statistics, i.e. by replacing \( \theta_{ik} \) by \( r_{ik} \), where \( r_{ik} \) is the rank given to the \( k \)th sample of parameter \( \theta_i \) (after ordering \( \theta_{i1}, \theta_{i2}, \ldots, \theta_{im} \) according to size). The formula for \( c_{ij} \) (equation 7) is modified by replacing \( \theta_{ik} \) with \( r_{ik} \), and by replacing all \( \mu_i \) with \((m + 1)/2\). The resulting correlation measure is known as the partial rank correlation coefficient (PRCC) (Iman and Helton 1988; Blower and Dowlatabadi 1994). The Spearman rank correlation coefficient is similarly defined as the ‘rank equivalent’ of the Pearson correlation coefficient (Saltelli et al., 1993).
A number of other sensitivity analysis techniques for non-calibrated models have been described by Saltelli \textit{et al} (1993) and Cacuci and Ionescu-Bujor (2004), but are not described here.
3. Method

3.1 Overview of the method

The uncertainty analysis of the ASSA2002 model was conducted using an approach based on the generalized likelihood uncertainty estimation (GLUE) and sampling/importance resampling (SIR) techniques, described in section 2.1. The approach is Bayesian, and can therefore incorporate both prior knowledge regarding HIV/AIDS epidemiology and empirical data indicating the magnitude of the HIV/AIDS epidemic. It also has the advantage that it is relatively simple to incorporate data from different sources, and to update the estimates as new data become available.

Prior distributions were set for 24 of the parameters in the ASSA2002 model, based on a review of the literature. 100,000 parameter combinations were generated by sampling from these prior distributions. For each of these 100,000 parameter combinations, a likelihood function was evaluated, representing the degree of correspondence between empirical data and the model outputs produced by the parameter combination sampled. Four sources of empirical data were used to define the likelihood function: the public antenatal clinic surveys (1991 to 2004), the reported deaths from the Department of Home Affairs (1997 to 2002), the Nelson Mandela/HSRC household surveys (2002 and 2005) and the RHRU/loveLife youth survey (2003).

After calculating likelihood values for all 100,000 parameter combinations, a sample of 500 parameter combinations was drawn (with replacement) from the initial set of 100,000 parameter combinations. The likelihood values were used as sample weights, so that the most weight was given to those parameter combinations which produced model results most consistent with the empirical data. The sample of 500 parameter combination therefore represents a sample from the posterior distribution. More detailed model outputs were generated for these 500 parameter combinations, so that average values and 95% prediction intervals could be calculated for selected model outputs. The 500 parameter combinations sampled were used to determine the posterior distributions for each parameter.

The sensitivity analysis of the ASSA2002 model was performed on the 500 parameter combinations drawn, after removing parameter combinations that had been resampled. Six model outputs were considered for the purpose of the uncertainty analysis: HIV incidence, HIV prevalence and increase in the crude mortality rate due to AIDS, in both 1995 and 2005. Partial rank correlation coefficients (PRCCs) were calculated for each combination of the 24 model parameters and the six model outputs. This measure of association indicates the significance of each parameter independently of interactions with other parameters, and is also more appropriate in cases where model outputs are not linearly related to input parameters.
Aspects of the methodology are described in more detail in the sections that follow. Section 3.2 contains a brief description of the ASSA2002 model and an explanation of the parameters considered in the uncertainty analysis. The prior distributions are described in section 3.3, and the literature on which these prior distributions are based is discussed. In section 3.4, the method used to define the likelihood function is explained and the four data sources used in the calculation of the likelihood function are described in more detail. Lastly, the programming of the uncertainty analysis is described in section 3.5.

3.2 Description of the ASSA2002 model

The ASSA2002 AIDS and Demographic model is a combined cohort component projection and HIV/AIDS model, developed by the Actuarial Society of South Africa (ASSA) to model the HIV/AIDS epidemic in South Africa. Two modes of HIV transmission are modelled: heterosexual transmission and mother-to-child transmission. The model contains a large number of parameters, which can be broadly grouped into the following classes:

- Demographic parameters
- Sexual behaviour parameters
- Heterosexual HIV transmission parameters
- Adult HIV survival parameters
- Paediatric HIV transmission and survival parameters
- HIV prevention and treatment parameters
- Calibration parameters

The non-demographic parameters are described in detail elsewhere (Johnson and Dorrington 2005). The uncertainty analysis that follows is restricted to a subset of the sexual behaviour parameters, heterosexual HIV transmission parameters, adult HIV survival parameters and calibration parameters. The remaining parameters are not considered, either because they are believed to have little effect on the correspondence between model outputs and observed levels of HIV prevalence and mortality, or because the empirical evidence used to determine the parameters is sufficiently reliable to warrant treating the parameters as known.

Sexual behaviour is modelled by dividing the sexually active population, between the ages of 14 and 59, into four risk groups. The ‘PRO’ group consists of sex workers and their regular clients, while the ‘STD’ group consists of other individuals who are regularly infected with sexually transmitted diseases (STDs). Although not regularly infected with STDs, individuals in the ‘RSK’ group are assumed to be at risk of HIV infection. Individuals in the ‘NOT’ group are assumed not to be at risk of HIV infection, either because they are abstaining from sexual intercourse or because they are in monogamous long-term relationships with partners who remain faithful to them. Individuals in the STD group can interact sexually with individuals in the PRO and RSK groups, but no other interactions between risk groups are assumed to occur. A two-parameter function is used to determine the ‘sex activity curve’ for women, which is a measure of the relative frequency of sex and the relative rate of partner change at
different ages. Most sexual behaviour parameters are specified for females, and male parameters are automatically calculated to be consistent with these.

Adults infected with HIV are assumed to progress through four stages of disease before dying from AIDS, in the absence of antiretroviral treatment. These four stages correspond to the four stages of the WHO Clinical Staging System, with the first two stages representing asymptomatic infection, the third phase representing minor HIV-related symptoms, and the fourth stage representing AIDS. The time spent in each stage of disease is assumed to follow a Weibull distribution, and the means and shape parameters of these Weibull distributions are assumed to be linearly related, with a ‘gradient parameter’ determining the slope of this linear relationship. For a specified mean survival time, the gradient parameter determines the variance of the survival time; an increase in the gradient parameter implies a reduction in the variance. The mean HIV survival time is assumed to depend on the age at which individuals become infected with HIV, with individuals infected at older ages progressing to death more rapidly than those infected at young ages.

The probability of heterosexual transmission of HIV, per act of sex with an infected partner, is assumed to depend on a number of factors. The most important of these are the sex of the susceptible partner and the risk groups of the susceptible and infected partners. It is also assumed that the disease stage of the infected partner affects the probability of HIV transmission. The levels of infectiousness in the different stages are estimated by assuming average levels of HIV viral load in each disease stage, as well as a factor by which the probability of HIV transmission increases per log increase in viral load. It is also assumed that in young women, age affects HIV susceptibility, with susceptibility increasing by a particular factor for each year below the age of 25.

Four HIV prevention programmes are allowed for in the model: social marketing, improved treatment for STDs, voluntary counselling and testing (VCT) and prevention of mother-to-child transmission (PMTCT). These interventions are assumed to be introduced at rates consistent with South Africa’s experience. Social marketing programmes are assumed to lead to increases in condom usage and reductions in average annual numbers of partners.

Highly active antiretroviral treatment (HAART) is also modelled. The effects of HAART are modelled by adding two stages to the basic four-stage model of adult HIV survival: one representing people receiving HAART and another representing people who have discontinued treatment. Treatment is assumed to be started at the time of the first AIDS-defining illness, and assumptions about proportions of people starting HAART have been based on historical trends in numbers of people on HAART. The frequency of sex in the different stages of disease is assumed to depend on both the proportion of people in the different disease stages who know their HIV status and on the severity of symptoms in the different disease stages. Symptoms are assumed to be most severe in people who have untreated AIDS and people who have discontinued HAART.
For the purpose of calibrating the model to HIV prevalence data and vital registration data, it is necessary to take into account various sources of bias associated with these data. In the case of the vital registration data, comparison of the modelled and actual deaths is only valid if the modelled deaths are adjusted to make some allowance for incompleteness of the reporting of deaths. In the case of HIV prevalence data, it is necessary to allow for possible inaccuracies in the HIV tests that are used; these tests sometimes produce false positive reactions. In the case of the antenatal clinic survey data, it is also necessary to allow for two sources of sampling bias: bias towards urban antenatal clinics in the early years of the survey, and exclusion of women seeking antenatal care in private health facilities. Surveys of HIV prevalence in pregnant women are also biased because HIV is known to reduce fertility, and the model therefore assumes that fertility in HIV-positive women reduces exponentially, relative to that in HIV-negative women, the longer they have been infected with HIV.

The AIDS epidemic in South Africa is assumed to start in 1985, a few years before the first heterosexual AIDS cases were reported. The results of the model are presented here up to 2005. Longer term projections are not included in this paper, as the antiretroviral parameters have not been included as a source of uncertainty in this analysis. These parameters are a significant source of uncertainty when the model is projected beyond 2005, but have little impact on the results up to the current time, due to the very limited access to HAART in South Africa historically.

The ASSA2002 model is publicly available (Actuarial Society of South Africa 2004), and is programmed in Excel and VBA. A C++ version of the ASSA2002 ‘lite’ model has also been developed for the purpose of this paper. Although the Excel and C++ models are very similar, a number of changes were made to the C++ model for the purpose of this uncertainty analysis. The differences between the Excel and C++ versions are summarized in Appendix A.

3.3 Specification of prior assumptions

As outlined in section 3.2, the 24 parameters included in the uncertainty analysis can be divided into four groups: calibration parameters, heterosexual HIV transmission parameters, adult HIV survival parameters and sexual behaviour parameters. The prior distributions for each group of parameters are specified and motivated in the sections that follow. In general, beta priors are used for those parameters restricted to the range [0, 1] and gamma priors are used for those parameters restricted to the range [0, ∞). Exponential priors are used for the false positive rates, as the exponential distribution matches the strongly skewed distribution of false positive rates more closely than the beta distribution. For a summary of the prior distributions assumed, see Table 4.2 in section 4.3.
3.3.1 Calibration parameters

3.3.1.1 Completeness of vital registration

For the purpose of calibrating the ASSA2002 model to vital registration data, it is necessary to make assumptions about the proportions of deaths that are recorded. Using estimates of numbers of deaths from the population register, and data from the 1996 and 2001 censuses, Dorrington et al (2004) estimated that between 10 October 1996 and 10 October 2001, 83.4% of adult male deaths and 84.5% of adult female deaths were reported. They followed the approach suggested by Hill and Choi (2004), namely to use the General Growth Balance (GGB) to estimate the completeness of one census relative to the other, adjust the less complete census, and then apply the Synthetic Extinct Generation (SEG) method to estimate the completeness of death reporting (United Nations Population Division 2002). The estimates of completeness thus derived are estimates of completeness relative to the census populations, and therefore underestimation of the population size by the censuses would mean that vital registration was lower than estimated by Dorrington et al – and vice versa if the censuses overestimated the population size.

The GGB and SEG methods require a number of assumptions. Hill and Choi (2004) examined the extent to which the estimated levels of completeness can be biased if these assumptions are not valid. The combination of the GGB and SEG methods was found to produce mortality estimates that were relatively insensitive to age misreporting in death data and census data (less than 5% absolute error in most cases). However, the mortality estimates were sensitive to violations of the assumptions that census coverage and completeness of death are constant with respect to age (absolute errors exceeded 10% in most cases).

In South Africa, the completeness of vital registration has improved substantially since the early 1990s (Timæus et al, 2001), and this needs to be allowed for when calibrating the model to vital registration data. Dorrington et al (2004) estimated the change in completeness in individual years between the censuses by assuming that completeness changed linearly and that the trend in the mortality rate of those over the age of 65, between the censuses, was flat. Under these assumptions, completeness was estimated to have improved by 1.7% per annum for men and 2.1% per annum for women. If mortality over the age of 65 in fact increased between the censuses, the extent of this improvement would be over-estimated, and vice versa if the mortality rates over the age of 65 decreased. However, comparison of estimates of mortality rates for the period 1984-86 with those for the period 1996-2001 suggests that the mortality of those aged 65 and older has remained relatively unchanged over the period.

The parameter explored in this analysis is the rate of completeness of the adult death data in the 1999 calendar year. The prior chosen for this variable is beta with parameters 179.8 and 34.2 (mean 84% and standard deviation 2.5%). Although the standard deviation is small relative to the potential errors estimated by Hill and Choi (2004), the quality of data in South Africa is believed to be better than that in other developing countries, and a
narrower range of uncertainty is therefore reasonable. The rate of completeness is assumed to increase by 1.9% for each year after 1999, up to 2002 (the most recent year for which reasonably complete vital registration data are available). Completeness is also assumed to decrease by 1.9% for each year prior to 1999. Thus, the average simulated completeness of adult death data rises from 80.2% in 1997 to 89.7% in 2002.

3.3.1.2 Bias towards urban areas in early antenatal surveys

In the early antenatal surveys, there was a tendency to over-sample the clinics in urban areas, and rural clinics were under-represented (Webb 1994). This is thought to have distorted significantly HIV prevalence estimates, as the prevalence of HIV in urban areas was higher than that in rural areas. The extent of this bias is illustrated in Table 3.1 below, which shows the ratio of urban to rural prevalence estimated in various surveys.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Population</th>
<th>Year</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1992</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1993</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1994</td>
<td>4.5</td>
</tr>
<tr>
<td>HSRC (2002)</td>
<td>SA households</td>
<td>2002</td>
<td>1.6</td>
</tr>
<tr>
<td>Shisana et al (2005b)</td>
<td>SA households</td>
<td>2005</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 3.1: Ratio of HIV prevalence in urban areas to HIV prevalence in rural areas

Definitions of ‘urban’ and ‘rural’ differ between surveys, and this may partially explain differences in urban to rural prevalence ratios. Differences in age groups sampled may also account for some variation.

The table suggests that there has been a decline in the extent to which prevalence in urban areas exceeds that in rural areas; the ratio of urban to rural prevalence has declined from around 5 in the early 1990s to close to 1 more recently. Over the same period, the samples have become more representative of rural clinics, particularly following the phasing in of a new survey protocol between 1997 and 2000 (Department of Health 2001). Recent antenatal prevalence estimates are therefore likely to be a more accurate measure of true antenatal prevalence than the earlier surveys were.

The exact extent of the bias towards urban clinics in the early antenatal surveys is unknown. However, if rural clinics were 50% less likely to be sampled than urban clinics in the 1990 survey, and 50% of the population were living in urban areas at this time (a proportion slightly lower than that observed in the 1996 census), then the ratio of the ‘true’ prevalence of HIV in antenatal clinics to that observed in the survey might be estimated as
where $v$ is the prevalence of HIV in rural antenatal clinics, assumed to be a fifth of the prevalence in urban antenatal clinics. This expression is equal to 0.82. The prior distribution chosen for the ratio is beta with parameters 19.2 and 4.8 (mean 0.80 and standard deviation of 0.08). It is further assumed that any bias would have reduced to 0 by 1999 (by which time the new sampling protocol was almost completely implemented), and that the ratio declines linearly between 1990 and 1999. Although this is crude, the lack of data on the extent of the bias prevents a more sophisticated approach.

### 3.3.1.3 Ratio of odds of infection in the private health sector to that in the public health sector

A further source of bias affecting antenatal clinic data is socioeconomic bias. Women attending the public antenatal clinics that are annually surveyed are typically poorer and at a higher risk of HIV infection than pregnant women seeking antenatal care in the private health sector. Race is significantly associated with health seeking behaviour and is also significantly associated with HIV risk, even after controlling for socioeconomic differences between race groups (Johnson and Budlender 2002). Any assessment of public antenatal bias therefore needs to take into account both the difference in socioeconomic profile and the difference in racial profile between users of public and private health facilities.

Suppose that survey data of HIV prevalence in pregnant women are available, and a logistic regression model of the following form is fitted to these data:

\[
\log\left(\frac{\pi_{irx}}{1-\pi_{irx}}\right) = \alpha + \beta_i + \gamma_r + \phi_x, \tag{10}
\]

where $\pi_{irx}$ is the probability of HIV infection in pregnant women using sector $i$ ($0 =$ private, $1 =$ public), of race $r$ and age group $x$. When $\pi_{irx}$ is small, it can be approximated as $\exp(\alpha + \beta_i + \gamma_r + \phi_x)$. Further suppose that $\omega_{irx}$ is the proportion of pregnant women in age group $x$ who are of race $r$ and are receiving antenatal care in sector $i$, so that $\sum_i \sum_r \omega_{irx} = 1$. In order to estimate the extent of antenatal bias, it is necessary to estimate the ratio of HIV prevalence in women attending public antenatal clinics, in age group $x$, to that in all pregnant women in age group $x$. This ratio is

\[
\left(\frac{\sum_i \omega_{irx} \pi_{irx}}{\sum_r \omega_{0rx} \pi_{0rx} + \sum_r \omega_{irx} \pi_{irx}}\right)
\]
Although the approximation in the second line is technically only valid when all $\pi_{irx}$ values are small, it is also valid when the same approximation is used in both the numerator and denominator, and the ratio is close to 1 (as is the case here). Values of $\omega_{irx}$ are estimated from the 1998 Demographic and Health Survey (Department of Health 1999). The parameter $\exp(\gamma_r)$ is set at 1.00 for African women, 0.23 for Indian women, 0.17 for coloured women and 0.13 for white women, based on a logistic regression of HIV prevalence data in women attending public antenatal clinics in 1998 and 1999 (Johnson and Budlender 2002).

The parameter $\exp(\beta_0 - \beta_1)$ represents the ratio of the odds of HIV infection in women using private facilities to that in women using public facilities, after controlling for age and race. A study of pregnant women attending private practices in KwaZulu-Natal in 1998 estimated a prevalence of 30.4% in African women (Wilkinson 1999), which was marginally less than the prevalence in public antenatal clinic attenders in KwaZulu-Natal in the same year (32.5%), almost all of whom were African (Department of Health 1999). Surveys in workforce populations in recent years have yielded estimates of HIV prevalence in African employees of between 9% and 21% (Shisana et al, 2004; Shisana et al, 2005a), which tend to be lower than the prevalence levels of 18 to 20% observed in national household surveys, in Africans between the ages of 15 and 49 (Human Sciences Research Council 2002; Shisana et al, 2005b). Although estimates of HIV prevalence in other race groups are available, these estimates are low and are typically based on small samples, which makes it difficult to reliably estimate relative levels of HIV prevalence in private and public health facility users for these race groups.

Taking into account the empirical data and the uncertainty regarding racial minorities, it is believed that the factor $\exp(\beta_0 - \beta_1)$ is unlikely to be below 0.4 and unlikely to be above 1.0. The prior chosen for this factor is therefore a beta distribution with parameters 5.83 and 2.50 (mean 0.7 and standard deviation 0.15). The factor is used to determine the amount by which the model estimate of HIV prevalence in pregnant women must be scaled up in order to obtain the estimate of HIV prevalence in public antenatal clinics. The scaling up is conducted separately for each age group, using equation (11).
3.3.1.4 Specificity of the ELISA

The ELISA is the antibody test used to detect HIV infection in the antenatal clinic surveys and in other prevalence surveys. This test sometimes produces false positive results (i.e. individuals who are HIV-negative may test positive), and failure to take this into account may therefore result in some over-estimation of HIV prevalence in a population. Sensitivity is defined as the proportion of infected individuals who test positive, while specificity is defined as the proportion of uninfected individuals who test negative. Although both measures of ELISA accuracy are discussed here, specificity is the only variable included in the uncertainty analysis, since sensitivity is usually higher and is less significant as a source of uncertainty in the model.

Most frequently, ELISAs are applied to blood specimens. Table B.1 in Appendix B summarizes estimates of the sensitivity and specificity of the ELISA when applied to blood specimens. For all studies, sensitivity and specificity were established relative to Western blot or combinations of the Western blot and other tests. Only studies of African populations are included in this table, as sera from African populations yield higher rates of false positives on ELISAs than sera from developed countries (Van Kerckhoven et al., 1991), and these studies are therefore more appropriate to South Africa.

The distributions of ELISA sensivities and specificities are strongly negatively skewed, with most estimates above 98%, but several estimates below 90%. The unusually low sensitivity and specificity estimates are mostly from studies conducted in the 1980s, when HIV tests were less accurate. Low specificity estimates have also been observed in studies of hospital patients (Tswana et al., 1993; Louis et al., 1999), but these may be due to a high prevalence of diseases that can cause ELISA cross-reactions in a hospital setting. These studies and the less recent studies should therefore be given less weight when setting prior assumptions. The prior distribution for the false positive rate (1 – the specificity) is thus chosen to be exponential with parameter 50, so that the mean specificity assumed is 0.98 and the standard deviation is 0.02. This is roughly consistent with the distribution of specificity rates in Table B.1 after studies from the 1980s and studies of hospital patients are excluded.

In some South African prevalence surveys, notably the 2002 Nelson Mandela/HSRC household prevalence survey (Human Sciences Research Council 2002) and the RHRU/loveLife youth prevalence survey (Reproductive Health Research Unit 2004), ELISA tests were applied to saliva or oral mucosal transudate specimens. Few African studies have examined the accuracy of the ELISA when applied to these specimens, but those which have been conducted do not suggest them to be any less specific or sensitive than those applied to blood (Behets et al., 1991; Luo et al., 1995; Matee et al., 1996; Ettiegne-Traore et al., 1998). In this analysis, two parameters are considered: the average false positive rate on the Abbott Axsym test used in the antenatal surveys, and the average false positive rate on the Orasure test used in the HSRC and RHRU surveys. For both parameters, the prior distribution is exponential with parameter 50.
Confirmatory testing substantially reduces the rate of false positives, and adjustments for specificity are therefore made only for those surveys in which no confirmatory HIV testing has been done. In line with the WHO guidelines (Sato et al, 1994), the Department of Health has conducted confirmatory testing up to 1996, when antenatal HIV prevalence was 14%, and thereafter ceased to do confirmatory testing in all provinces other than the Western Cape (Department of Health 1997; Department of Health 1998). No confirmatory testing was done in either the 2002 Nelson Mandela/HSRC household prevalence survey or the RHRU/loveLife youth prevalence survey, but confirmatory testing was part of the protocol for the 2005 HSRC household survey.

3.3.1.5 Reduction in fertility per year of HIV infection

Because women with HIV are less likely to fall pregnant than women who are uninfected, antenatal surveys may be biased towards women with a lower HIV prevalence (Zaba and Gregson 1998). If antenatal surveys are a major source of prevalence data in a country, then the assumptions about the extent to which HIV reduces fertility become important in extrapolating from the antenatal clinic data to the general female population.

It would appear that the extent to which HIV reduces fertility increases the greater the degree of HIV disease progression (Ryder et al, 1991; Gray et al, 1998; du Loû et al, 1999). However, few studies have evaluated the reduction in fertility per year of increase in the duration of infection, the parameter of interest from the point of view of the ASSA2002 model. Two studies provide some insight into this parameter. Firstly, Heuveline (2003) fitted a maximum likelihood model to demographic and HIV data from eastern and southern Africa, and estimated that the ratio of fertility in HIV-positive women to that in HIV-negative women was 0.85 (95% CI: 0.80-0.91) for women infected for around 5 years, 0.36 (95% CI: 0.28-0.45) for women infected for around 10 years, and 0.29 (95% CI: 0.08-0.61) for women infected for more than 10 years. Secondly, Lee et al (2000), in a study of an indigent US population, estimated the ratio of fertility in women with an AIDS diagnosis to that in uninfected women to be 0.85 (95% CI: 0.71-1.03) 7.5 to 10 years before AIDS diagnosis, 0.74 (95% CI: 0.63-0.86) 5 to 7.5 years before AIDS diagnosis, 0.55 (95% CI 0.47-0.64) 2.5 to 5 years before AIDS diagnosis, and 0.45 (95% CI: 0.38-0.55) less than 2.5 years before AIDS diagnosis.

Neither the Heuveline (2003) study nor the Lee et al (2000) study can be used to estimate accurately the reduction in fertility per year of infection. However, if it were assumed that the difference in the average duration of infection between successive duration categories was 5 years in the Heuveline study, the factors by which the odds of childbirth are reduced per year of infection might be estimated at 0.84 between 5 and 10 years after infection, and at 0.96 more than 10 years after infection. Similarly, if it were assumed that the difference in the average duration of infection between successive duration categories was 2.5 years in the Lee et al study, the annual reduction in the odds of childbirth might be estimated at 0.95 between 6 and 9 years before AIDS diagnosis, 0.89 between 4 and 6 years before AIDS and 0.92 between 1 and 4 years before AIDS. There
is some evidence to suggest that the effect of HIV on fertility is less significant in women using contraceptives (Gray et al, 1998; Terceira et al, 2003; Lewis et al, 2004), and the factors estimated here might therefore underestimate those that would apply in South Africa, where rates of contraceptive usage are relatively high. The prior distribution for the factor by which fertility is reduced per year of infection is therefore chosen to be beta with parameters of 49.2 and 2.6 (mean 0.95 and standard deviation 0.03).

### 3.3.2 Heterosexual HIV transmission parameters

#### 3.3.2.1 The effect of sex and risk group on HIV transmission probabilities

Four key HIV transmission parameters are considered here:

- The average probability that an HIV-positive male in the RSK group transmits the virus to an HIV-negative female in the RSK group, in a single act of unprotected sex;
- The average probability that an HIV-positive female in the RSK group transmits the virus to an HIV-negative male in the RSK group, in a single act of unprotected sex;
- The average probability that an HIV-positive male in the STD group transmits the virus to an HIV-negative female in the STD group, in a single act of unprotected sex; and
- The average probability that an HIV-positive female in the STD group transmits the virus to an HIV-negative male in the STD group, in a single act of unprotected sex.

The probability of HIV transmission that applies when one partner is in the STD group and the other in the RSK group is assumed to be the average of the probability that applies when both partners are in the RSK group and the probability that applies when both partners are in the STD group. Sex acts involving serodiscordant partners in the PRO group are assumed to be as likely to result in HIV transmission as sex acts between serodiscordant partners in the STD group.

To estimate the transmission probabilities that apply when both partners are in the RSK group, results from studies of stable serodiscordant couples are considered. Estimates of transmission probabilities from these studies are summarized in Table 3.2 below.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Region</th>
<th>Type of analysis</th>
<th>Probability per-contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray et al (2001), Gray et al (2003)</td>
<td>Uganda</td>
<td>Prospective</td>
<td>0.0009 0.0013 0.0011</td>
</tr>
<tr>
<td>Leynaert et al (1998)</td>
<td>Europe</td>
<td>Retrospective</td>
<td>0.0006 - 0.0008</td>
</tr>
<tr>
<td>Duerre et al (1996)</td>
<td>Thailand</td>
<td></td>
<td>0.0016 0.0003</td>
</tr>
<tr>
<td>Downs &amp; de Vincenzi (1996)</td>
<td>Europe</td>
<td>Retrospective</td>
<td>0.0005 0.0009</td>
</tr>
<tr>
<td>De Vincenzi et al (1994)</td>
<td>Europe</td>
<td>Prospective</td>
<td>0.0015 0.0009 0.001</td>
</tr>
<tr>
<td>Peterman et al (1988), Wiley et al (1989)</td>
<td>USA</td>
<td>Retrospective</td>
<td>0.0013 0.0005 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 0.0014  - 0.0011</td>
</tr>
</tbody>
</table>

Table 3.2: Empirical estimates of per-contact HIV transmission probabilities in serodiscordant heterosexual couples

These studies estimate the probability of male-to-female HIV transmission per coital act to be between 0.0005 and 0.0016, and the probability of female-to-male transmission to be between 0.0003 and 0.0013. However, it is likely that these studies under-estimate the true likelihood of HIV transmission. There is substantial heterogeneity between individuals, both in terms of susceptibility to HIV infection and HIV infectiousness. Prospective studies of serodiscordant couples are subject to significant selection bias because, by sampling couples in which transmission has not yet occurred, the studies introduce a bias towards couples in which the seronegative partner has reduced susceptibility and/or the seropositive partner has reduced infectiousness (O'Farrell 2001). Retrospective studies of partnerships are also subject to selection bias because they are less likely to include short-term or irregular partnerships. These prospective and retrospective studies usually estimate per-contact transmission probabilities on the assumption of a constant probability per sexual contact, which is at odds with the finding that the per-contact transmission probability decreases as the cumulative number of sexual contacts with the infected partner increases (Downs and De Vincenzi 1996).

A further factor explaining heterogeneity in the findings in Table 3.2 is variation between populations in rates of male circumcision. Circumcision has been found to reduce significantly the risk of HIV infection in men, particularly in ‘high risk’ groups such as truck drivers and men with STDs (Weiss et al, 2000; Auvert et al, 2005). Approximately 70% of men in the US are circumcised (Moses et al, 1998), and this may explain the relatively large differences in male-to-female and female-to-male transmission probabilities observed in the US. Rates of male circumcision are around 15% in Uganda (Wawer et al, 1999), and this may explain why female-to-male HIV transmission probabilities are relatively high in this setting. In South Africa, 35% of men over the age of 15 are circumcised (Human Sciences Research Council 2002). Female-to-male and
male-to-female transmission probabilities in South Africa might therefore not be as different as those observed in the US.

The prior distribution chosen for the probability of male-to-female transmission in the RSK group is beta with parameters 24.95 and 12,449 (mean 0.002 and standard deviation 0.0004), while that chosen for the probability of female-to-male transmission is beta with parameters 24.97 and 24,949 (mean 0.001 and standard deviation 0.0002).

Estimates of HIV transmission probabilities in casual partnerships and in partnerships involving sex workers are rare. In a study of Kenyan men with a single sexual exposure to a sex worker, the cumulative proportion of men seroconverting was 13% (Cameron et al., 1989). HIV prevalence in female sex workers was estimated to be 85%, and it is therefore unlikely that the true HIV transmission probability per sex act would have been substantially greater than 0.13. However, the fact that the men were recruited from an STD treatment service implies some overestimation of the transmission efficiency that might have been expected in men who did not have STD symptoms. In addition, it is possible that some of the men acquired HIV from other sources.

Another study, conducted among Thai military conscripts, estimated that the probability of HIV transmission from an HIV-positive sex worker to an uninfected male, in a single act of sex, was between 0.031 and 0.056 (Mastro et al., 1994). This assumed, however, that there was no under-reporting of past contacts with sex workers, and that contacts with sex workers were the only source of infection. This study may therefore also overestimate the true HIV transmission probability. A study conducted among Kenyan truck drivers (Baeten et al., 2005), who reported having sex only with casual partners and sex workers, estimated a female-to-male HIV transmission probability of 0.0105 per unprotected sex act.

In the default version of the ASSA2002 model, it is assumed that probability of HIV transmission per sex act between serodiscordant couples in the STD and/or PRO groups is 0.007 when the infected partner is male and 0.005 when the infected partner is female. Since the empirical data for these parameters are not reliable, the priors chosen for these parameters have been set to be centred on the ASSA2002 default parameters, with a high variance. Thus the prior chosen for the male-to-female transmission probability in the STD/PRO group is beta with parameters 24.82 and 3,521 (mean of 0.007 and standard deviation of 0.0014), and that chosen for the female-to-male transmission probability is beta with parameters 24.87 and 4,949 (mean 0.005 and standard deviation 0.001).

### 3.3.2.2 Increase in transmission probability per log increase in viral load

After the initial drop in HIV viral load following acute infection, viral load levels typically increase steadily over the course of chronic infection. The assumed increase in transmission probability per log increase in viral load therefore determines the extent to which individuals in the later HIV disease stages contribute to HIV transmission, and thus affects the change in HIV incidence over the course of the epidemic, as more infected individuals enter the later stages of disease.
In a study of serodiscordant couples in Uganda, it was estimated that the rate of HIV transmission was increased by a factor of 2.45 (95% CI: 1.85-3.26) for each log increase in the viral load of the index partner (Quinn et al, 2000). A similar study of serodiscordant couples, conducted in Zambia, estimated that female-to-male and male-to-female transmission rates increased by factors of 2.5 (95% CI: 1.5-4.0) and 1.8 (95% CI: 1.2-2.8) respectively, per log increase in viral load of the index partner (Fideli et al, 2001). Both studies estimated the effect of viral load on the incidence rate in the HIV-negative partners, not the effect on transmission probabilities per sex act. Taking into account likely reductions in the frequency of sex as viral load increases (Gray et al, 2001), one might expect the factor by which transmission probability increases, per log rise in viral load, to be greater than that by which HIV incidence increases.

In another study, a mathematical model was fitted to data from a variety of surveys to estimate the effect of HIV RNA count in semen on the risk of male-to-female transmission (Chakraborty et al, 2001). When the viral load data were transformed to fit on the interval [0, 1], it was estimated that a 10-fold increase in the transformed RNA count was associated with a 6-fold increase (10^{0.778}) in the risk of HIV transmission per coital act. However, this estimate is likely to over-estimate the increase in transmissibility per log increase in blood viral load, as RNA concentration in blood is not as strongly correlated with infectiousness as that in semen. In addition, the study of Chakraborty et al considered only non-synctium-inducing RNA, which is known to be more transmissible than synctium-inducing RNA, and which occurs less frequently in more advanced disease. This too would lead to some over-estimation of the effect on infectiousness of an increase in total HIV RNA.

Because of the biases associated with the estimate of Chakraborty et al, it would be reasonable to give more weight to the estimates of Quinn et al and Fideli et al in setting the prior distribution. The prior distribution for the percentage increase in infectiousness per sex act, per log increase in viral load, is chosen to be gamma with parameters 6.25 and 3.57 (mean 175% and standard deviation 70%).

### 3.3.2.3 Increase in susceptibility to HIV in young women

There is evidence to suggest that the probability of male-to-female transmission of HIV is higher in young women than in older women. In a study of HIV incidence rates in spouses of HIV-positive individuals in Uganda, Carpenter et al (1999) found that incidence rates in women under the age of 25 were roughly double those in women over the age of 25. In another Ugandan study of discordant couples, Gray et al (2001) found that HIV transmission probabilities below the age of 30 were roughly double those over the age of 30, though these age-specific transmission frequencies were not reported separately for males and females. It is thought that these higher transmission probabilities in younger women may be the result of the high prevalence of cervical ectopy in young women, as cervical ectopy has been found to increase susceptibility to HIV (Moss et al, 1991; Plourde et al, 1994). Alternatively, the lower susceptibility in older women may reflect a degree of acquired resistance to HIV infection.
In the ASSA2002 model, it is assumed that the male-to-female transmission probability per coital act increases by a certain factor for each year of age below 25. If this factor is set to 1.12 (i.e. a 12% increase in female susceptibility to HIV infection for each year below age 25), the average male-to-female HIV transmission probability per coital act between the ages of 14 and 24 is roughly double that in women over the age of 25. The prior distribution for the percentage increase in female susceptibility to HIV, for each year below 25, is chosen to be gamma with parameters 16 and 133.33 (mean 12% and standard deviation 3%).

3.3.3 HIV survival parameters in adults in the absence of HAART

The median time from HIV infection to death is generally believed to be shorter in developing countries than in developed countries (UNAIDS Reference Group on Estimates Modelling and Projections 2002). However, there are relatively few studies from developing countries in which HIV survival has been monitored from the point of HIV infection. The median time from HIV infection to death in adults has been estimated at 9.8 years in a Ugandan study (Morgan et al, 2002), at approximately 9 years in a Thai study of sex workers (Kilmarx et al, 2000), at 7.4 years in a small Haitian cohort (Deschamps et al, 2000), and at just over 6 years in a study of Filipino sex workers (Manaloto et al, 1994). Other studies have not accumulated sufficient follow-up data to estimate median survival reliably, but have nevertheless estimated the median time to AIDS, which is slightly shorter than the median time to death. This has been estimated at 7.4 years in a study of Thai military recruits (Rangsin et al, 2004) and at 4.4 years in a study of Kenyan sex workers (Anzala et al, 1995).

These empirical estimates are problematic for a number of reasons. Firstly, many of the studies have been conducted in sex workers. Sex workers may be at a higher risk of HIV dual infection, and therefore might experience more rapid HIV disease progression (Gottlieb et al, 2004). Right-censoring of individuals lost to follow-up could also bias results, as the implicit assumption underlying this censoring is that the survival prospects are the same for those remaining in the cohort and those leaving the cohort (Porter and Zaba 2004). A further limitation of the studies listed is that most do not report the proportion of mortality in HIV-positive individuals that is due to AIDS. For the purpose of estimating AIDS mortality rates, the effect of non-AIDS mortality needs to be removed. This would increase slightly the medians quoted in the previous paragraph.

A further problem is that none of these studies have been conducted in settings in which HIV-1 subtype C is the predominant HIV strain present. Subtype C is the subtype that is most common in southern Africa, and studies from non-subtype C settings may not be representative. Although there is little consistent evidence to suggest that subtypes differ in terms of their pathogenesis (Hu et al, 1999), it has been noted that individuals infected with subtype C develop the synctium-inducing HIV phenotype less frequently than individuals infected with other subtypes, and this phenotype is associated with more rapid disease progression (Morris et al, 2000; Peeters et al, 2003). In addition, comparison of non-synctium-inducing isolates suggests that subtype C is less fit than subtype B, which
has fitness comparable to subtypes A, D and E (Ball et al., 2003). Although this may suggest longer survival in subtype C-infected individuals, there are no data to confirm this.

Limited data from South African cohorts suggest that rates of disease progression are similar in this setting to those in the developed world (Maartens et al., 1997; Badri et al., 2004; Glynn et al., 2005). The median survival of 10.9 years, observed in the developed world in individuals infected between the ages of 25 and 34 (Collaborative Group on AIDS Incubation and HIV Survival 2000), may therefore serve as a reasonable approximation to the mean survival (ignoring non-AIDS mortality) in South Africa in individuals infected at the age of 29. The prior chosen for this parameter is gamma with parameters 110.25 and 10.50 (mean of 10.5 years and standard deviation 1 year). Using the default ASSA2002 parameters, median HIV survival times are roughly half a year shorter than mean survival times.

In adults, the median time to death appears to depend on the age at infection. A linear regression of HIV survival data from the developed world in the pre-HAART era (Collaborative Group on AIDS Incubation and HIV Survival 2000) suggests that the median time to death decreases by 1.6 years for every 10 year increase in the age at HIV infection. When a Weibull model was applied to earlier data from homosexual and haemophiliac cohorts, the estimated median times to AIDS were 10.4 years for individuals infected at the age of 20 and 7.4 years for individuals infected at the age of 40 (Rosenberg et al., 1994), implying a reduction of roughly 1.5 years in median survival per 10-year increase in age at seroconversion. The prior chosen for the decrease in mean survival per 10-year increase in age at infection is chosen to be gamma with parameters 25 and 16.67 (mean of 1.5 years and standard deviation of 0.3).

The variability of the time from HIV infection to death is another important parameter to consider when modelling HIV survival. The Weibull distribution, which is most commonly used to model HIV survival, has a variance which is determined principally by the ‘shape’ parameter. Gregson et al. (1998) found that a shape parameter of 2.5 provided a reasonable fit to early HIV survival data from Uganda. When fitting Weibull distributions to data from developed countries, Williams, Kochi and Dye estimated a Weibull shape parameter of 2.28 and noted that this parameter did not vary significantly between the ages of 15 and 55 (manuscript in preparation). This shape parameter may well be an under-estimate of the ‘true’ shape parameter, as the data on which this analysis is based are drawn from a number of different countries, and the variability in survival times estimated may be greater than that which might be expected within a single country such as South Africa. In addition, neither Gregson et al. nor Williams et al. were able to remove the effect of non-AIDS mortality. If this effect is not removed, it might result in over-estimation of variability in HIV survival times, independent of non-AIDS mortality, and hence under-estimation of the shape parameter of the HIV survivor function.

In the ASSA2002 model, it is assumed that the time spent in each of the four adult disease stages follows a Weibull distribution, and a ‘gradient parameter’ is used to determine the shape parameters of each of these Weibull distributions. The explanation of
this gradient parameter is given elsewhere (Johnson and Dorrington 2005). A gradient parameter of 0.1 produces a distribution of HIV survival times that is roughly consistent, in shape, with that obtained when a single Weibull distribution with shape parameter 2.28 is used to model the time from HIV infection to death. Similarly, a gradient parameter of 0.5 produces a distribution of HIV survival times similar to a single Weibull distribution with a shape parameter of 3. The prior distribution chosen for the gradient parameter is gamma with parameters 4 and 13.33 (mean 0.3 and standard deviation 0.15).

### 3.3.4 Sexual behaviour parameters

#### 3.3.4.1 Relative frequencies of sex in different disease stages

Individuals experiencing HIV symptoms may be less likely to acquire new partners and less likely to engage in sex than individuals not experiencing HIV-related symptoms. As with the viral load parameter described in section 3.3.2.2, this is important in determining the extent to which individuals in the later stages of disease contribute to overall HIV transmission. This in turn influences the trends in HIV incidence as the epidemic matures and as more individuals enter the late stages of disease.

The most detailed data available to date are those collected in a longitudinal study of HIV-positive and HIV-negative women in Uganda (Ross et al., 2004). HIV-positive women were classified according to the WHO clinical staging system; Table 3.3 below shows the results. The average frequency of sex in a four-week period, for women in relationships, is calculated on the assumption that the average frequency for women reporting sex 1 to 5 times is 3, the average frequency of sex for women reporting sex 6 to 10 times is 8, and the average frequency of sex for women reporting sex more than 10 times is 13. The results show a trend towards decreasing probability of being in a partnership and decreasing frequencies of sex within partnerships with increasing disease progression. Age-adjusted results show a greater effect of disease progression than results not adjusted for age.

<table>
<thead>
<tr>
<th>Measure of activity</th>
<th>HIV-</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>% in partnerships</td>
<td>63%</td>
<td>68%</td>
<td>48%</td>
<td>49%</td>
<td>38%</td>
</tr>
<tr>
<td>OR for being in partnership (95% CI)†</td>
<td>1</td>
<td>1.62</td>
<td>0.70</td>
<td>0.45</td>
<td>0.14</td>
</tr>
<tr>
<td>Average frequency of sex (women in partnerships)</td>
<td>5.6</td>
<td>5.3</td>
<td>5.1</td>
<td>4.7</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Table 3.3: Probabilities of being in partnerships and frequencies of sex within partnerships, by WHO clinical stage (Uganda)

† Age-adjusted. * Last four weeks

A similar study was conducted among HIV-positive and HIV-negative women in Zimbabwe (Terceira et al., 2003). Table 3.4 shows the results, in a similar format to the previous table, but with HIV-positive women classified only according to whether they
are symptomatic or asymptomatic. The average frequency of sex for women reporting sex 1 to 5 times is again assumed to be 3, and the average frequency for women reporting sex more than 6 times in the last two weeks is assumed to be 8. Although the average frequencies of sex, in women in partnerships, are roughly consistent with the Ugandan data, there is not as marked a difference in sexual behaviour patterns between women who are asymptomatic and women who are symptomatic.

### Measure of activity

<table>
<thead>
<tr>
<th></th>
<th>HIV-negative</th>
<th>Asymptomatic</th>
<th>Symptomatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>% in partnerships</td>
<td>69%</td>
<td>60%</td>
<td>55%</td>
</tr>
<tr>
<td>OR for not being in partnership (95% CI)†</td>
<td>1 (1.42-1.90)</td>
<td>1.64 (1.33-2.83)</td>
<td></td>
</tr>
<tr>
<td>Average frequency of sex (women in partnerships)*</td>
<td>2.9</td>
<td>2.4</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Table 3.4: Probabilities of being in partnerships and frequencies of sex within partnerships, by symptom status (Zimbabwe)

† Age-adjusted. * Last two weeks

Studies of North American women have also examined the effect of declining CD4+ count on the frequency of sex. Hankins *et al* (1998) found that the percentage of women who reported sexual activity in the last six months was strongly associated with CD4+ count. The percentage ranged from 77% in women reporting a CD4+ count greater than or equal to 500 to 59% in women with CD4+ counts less than 200. Greenblatt *et al* (1999) obtained similar results in their cohort: the proportion of women reporting sex in the last six months declined from about two thirds in women with a CD4+ count above 200, to 55% in women with CD4+ counts below 200. In the same cohort, Wilson *et al* (2004) found that the odds of being sexually active (relative to women with CD4+ count > 500) was 0.75 in women with CD4 counts from 201 to 500 and 0.56 in women with CD4+ counts of 200 or lower.

The studies reviewed here show that disease progression – whether assessed clinically or immunologically – is associated with reduced probability of being in a partnership and reduced coital frequency, even for those women who are in partnerships. In the ASSA2002 model, it is assumed that there is negligible reduction in coital frequency during WHO clinical stages 1 and 2, as these stages of infection are largely asymptomatic. The prior distribution for the ratio of coital frequency in WHO stage 3 to that in the asymptomatic stages is chosen to be beta with parameters 17.45 and 7.48 (mean 0.7 and standard deviation 0.09), and the prior distribution for the ratio of coital frequency in WHO stage 4 to that in the asymptomatic stages is chosen to be beta with parameters 3.99 and 8.10 (mean 0.33 and standard deviation 0.13). (Both assumptions ignore the independent effect of counselling on sexual risk behaviour.) Individuals on antiretroviral treatment would be largely restored to health, although some reduction in coital frequency (relative to asymptomatic individuals) might still be expected due to drug side effects, immune reconstitution disease and opportunistic infections. The prior distribution for the ratio of coital frequency in HAART recipients to that in asymptomatic
individuals (again ignoring the independent effect of counselling) is therefore chosen to be beta with parameters 46.02 and 9.42 (mean 0.83 and standard deviation 0.05). The symptom profile for individuals who have discontinued HAART is assumed to be similar to that of untreated individuals in WHO stage 4, so the same prior distribution for the relative coital frequency (beta with parameters 3.99 and 8.10) is used for those who have discontinued HAART.

Suppose that the frequency of sex in stage $s$ of disease, as a multiple of that in the asymptomatic stages (stages 1 and 2), is $K_s$. Ignoring the effects of counselling on the frequency of sex in the different disease stages, the following constraints are set:

$$K_4 = K_6 < K_3 < K_5 < 1,$$

(12)

where stage 5 represents individuals on HAART and stage 6 represents individuals who have started HAART but subsequently discontinued treatment. To ensure that these inequality constraints are met, the same uniform $(0, 1)$ variate, $K$, is used to sample from the inverted cumulative distribution functions of variables $K_3, K_4, K_5$ and $K_6$. It is the variable $K$ which is examined in the uncertainty and sensitivity analyses, rather than the individual $K_s$ parameters.

### 3.3.4.2 Change in sexual behaviour in response to the HIV/AIDS epidemic

Several studies conducted in other African countries have found evidence of change in sexual behaviour over time, in response to the HIV/AIDS epidemic. For example:

- Asiimwe-Okiror et al (1997) found that in urban Uganda there had been significant increases in condom usage, increases in age at sexual debut and marriage, and reductions in proportions of men reporting sex involving exchange for money. However, there was no significant change in the reporting of non-regular partners or extramarital sex. Similar findings in rural Uganda were reported by Kamali et al (2000); this study also found increased condom use and delayed sexual debut and marriage, but no reduction in numbers of partners.

- In a study of male factory workers in Tanzania, Ng'weshemi et al (1996) found significant reductions in numbers of partners and contacts with casual partners, but condom use only increased with casual partners. In another study in rural Tanzania, however, no significant changes were observed over time, either in age at sexual debut, numbers of partners per annum, or condom use with casual partners (Mwaluko et al, 2003).

- Nationally representative household surveys conducted in Zambia suggest that there has been an overall increase in the age at sexual debut and a reduction in unprotected sex with non-cohabiting partners and multiple partners (Slaymaker and Buckner 2004). However, another Zambian study found that behaviour change (in terms of increased condom usage and reduction in numbers of partners) was limited to urban areas (Fylkesnes et al, 2001).

The evidence suggests that changes in sexual behaviour vary between settings. Studies tend to show that changes in sexual behaviour are greatest in urban areas, and this is
supported by HIV prevalence data, which show greater declines in HIV prevalence in urban areas than in rural areas (Asiimwe-Okiror et al, 1997; Fylkesnes et al, 2001). Fylkesnes et al (2001) also found that reductions in risk behaviours were greatest amongst youth and amongst the more educated. In analyzing other Zambian data, however, Slaymaker & Buckner (2004) found significant anomalies suggestive of social desirability bias and associated exaggeration of changes in sexual behaviour. It is also possible that reductions in average numbers of partners and HIV prevalence may be the result of individuals with high rates of partner change being removed from the population through AIDS mortality, rather than spontaneous change in behaviour. Caution therefore needs to be applied when interpreting data on trends in sexual behaviour.

In South Africa, there is evidence of consistent increases in condom usage since the late 1990s, but no evidence of significant changes in age at sexual debut or proportions of births to women under the age of 20 (Bradshaw et al, 2004). Data on numbers of partners have been collected (Department of Health 1999; Human Sciences Research Council 2002; Reproductive Health Research Unit 2004), but differences in survey designs and inconsistencies in survey results make it difficult to assess whether or not a decline in numbers of partners has occurred. The percentage reduction in average annual numbers of partners is therefore included as a parameter in the uncertainty and sensitivity analyses. This percentage reduction is applied only to women in the PRO, STD and RSK risk groups, and corresponding changes in male sexual behaviour are determined to be consistent with this. The reduction in number of partners is assumed to be the result of social marketing programmes, and is therefore assumed to be phased in at the same rate as the increase in the rate of condom usage, which is already allowed for in the ASSA2002 model, and which is also assumed to result from social marketing. The prior distribution chosen for the reduction in the number of partners, after social marketing programmes are fully introduced, is beta with parameters 0.8 and 7.2 (mean 10% and standard deviation 10%).

3.3.4.3 Other sexual behaviour parameters

There are few data to support many of the sexual behaviour parameters. The approach taken in incorporating these parameters into the uncertainty analysis is therefore to set the priors for each parameter such that the mean is the current ASSA2002 default parameter value and the standard deviation is 20% of the default parameter value. The priors and the corresponding means and standard deviations are specified for each parameter in Table 3.5 below.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior</th>
<th>Mean</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of population in STD risk group*</td>
<td>Beta (19.8, 79.2)</td>
<td>20%</td>
<td>4%</td>
</tr>
<tr>
<td>% of population in RSK risk group*</td>
<td>Beta (19.0, 63.7)</td>
<td>23%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Average annual # partners in PRO group†</td>
<td>Gamma (25, 0.1)</td>
<td>250</td>
<td>50</td>
</tr>
<tr>
<td>Average annual # partners in STD group†</td>
<td>Gamma (25, 2.08)</td>
<td>12</td>
<td>2.4</td>
</tr>
<tr>
<td>Average annual # partners in RSK group†</td>
<td>Gamma (25, 25)</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>Sexual mixing parameter</td>
<td>Beta (12, 12)</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>Shape factor of female sex activity curve</td>
<td>Beta (24.9, 4949)</td>
<td>0.005</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 3.5: Prior distributions for other sexual behaviour parameters

* The same parameter values are used for males, females, the initial population and subsequent immigrants.
† The parameter is specified for females. Male parameters are calculated to be consistent with the female assumptions.

In order to allow for uncertainty with respect to mixing patterns between risk groups, the ASSA2002 model has been altered so that the assumptions about proportions of partners in different risk groups are determined by a single ‘sexual mixing parameter’. Suppose \( w^*_j \) represents the proportion of female partners who are in risk group \( j \), for a man who is in risk group \( i \). Garnett and Anderson (1996) use an approach similar to that described below to define \( w^*_j \):

\[
 w^*_j = (1 - \varepsilon)\delta^*_j + \varepsilon \left( \frac{\Psi_jP_j}{\sum_u \Psi_uP_u} \right) 
\]  

where
\( \delta^*_j = 1 \) when \( i = j \) and \( 0 \) when \( i \neq j \),
\( \Psi_j \) is the proportion of sexually experienced females (aged 14-59) who are in risk group \( j \) at the start of the epidemic, and
\( P_j \) is the average annual number of partnerships for a sexually experienced female in risk group \( j \) at the start of the epidemic.

\( \varepsilon \) is a parameter which can take on any value between 0 and 1, and is the sexual mixing parameter referred to above. This parameter determines the degree of sexual mixing; the higher the value of \( \varepsilon \), the more mixing there is between the risk groups. The default parameters in ASSA2002 are roughly consistent with an \( \varepsilon \) value of 0.5, and the prior distribution for this parameter has therefore been chosen to be centred on this value. Once the value of \( \varepsilon \) has been specified and the parameters \( w^*_j \) have been calculated, the average annual numbers of partnerships for males and the proportion of male partners in each risk group can be calculated. The procedure for calculating these parameters is explained in more detail in Appendix C.
The shape factor of the sexual activity curve determines the extent to which sexual activity is concentrated at young ages. The higher the value of this parameter is, the more sexual activity occurs at young ages, relative to older ages.

### 3.4 Specification of the likelihood function

Four sources of data are used in specifying the likelihood function: the antenatal clinic prevalence data, reported deaths from the Department of Home Affairs, the Nelson Mandela/HSRC household prevalence surveys and the RHRU/loveLife youth prevalence survey. Table 3.6 summarizes the four data sources in terms of the years to which they relate, the age range considered for the purpose of the uncertainty analysis, and the total number of records used for the purpose of the uncertainty analysis. All of these data sources can be considered nationally representative, and are therefore important in the calibration of the model.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Years</th>
<th>Age range</th>
<th># records</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV prevalence from antenatal clinic (ANC) surveys</td>
<td>1991-2004</td>
<td>15-39</td>
<td>205 865</td>
</tr>
<tr>
<td>Reported deaths from the Department of Home Affairs (DHA)</td>
<td>1997-2002</td>
<td>20-54</td>
<td>1 099 010</td>
</tr>
<tr>
<td>HIV prevalence from Nelson Mandela/HSRC household surveys</td>
<td>2002, 2005</td>
<td>15+</td>
<td>18 112</td>
</tr>
<tr>
<td>HIV prevalence from RHRU/loveLife youth survey</td>
<td>2003</td>
<td>15-24</td>
<td>11 904</td>
</tr>
</tbody>
</table>

Table 3.6: Characteristics of the four data sets

The likelihood function is defined as the product of the likelihood measures for each of the four data sources, i.e.

\[
L = L_{\text{ANC}}^{N_{\text{ANC}}} \times L_{\text{DHA}}^{N_{\text{DHA}}} \times L_{\text{HSRC}}^{N_{\text{HSRC}}} \times L_{\text{RHRU}}^{N_{\text{RHRU}}},
\]

where \( L \) is the aggregate likelihood measure, \( L_u \) is the likelihood measure for data source \( u \), and \( N_u \) is a weighting factor applied to data source \( u \).

In theory, it should not be necessary to include the \( N_u \) factors in the above equation. However, in situations in which it is not possible to define the likelihood in respect of a data source precisely, it would be inappropriate to weight the different data sources equally. In addition, if these \( N_u \) adjustments were not included, it is likely that the confidence intervals produced would be too narrow. Heuveline (2003) observes that when attempting to maximize the likelihood in a complex model, confidence intervals are likely to be unrealistically narrow if certain sources of variation are not incorporated into the model, or if certain parameters are treated as fixed when there is in fact uncertainty as to their true value. The same would apply in this context, where the demographic
parameters and many of the epidemiological parameters are being treated as fixed, and the complexity of the HIV transmission process is not being fully captured. Beven and Binley (1992) suggest that when using the GLUE approach, \( N_u \) factors, similar to those in equation (14), can be included. They further suggest that these \( N_u \) factors should be adjusted, through an iterative process, until the 95% prediction intervals around the model outputs are wide enough to include a reasonable proportion of the observations. The term ‘95% prediction interval’ is used here to refer to the range between the 2.5 and 97.5 percentiles of the outputs sampled from the posterior distribution.

Following the approach suggested by Beven and Binley, the \( N_u \) factors were set to achieve prediction intervals of appropriate width. For each data source \( u \), the value of \( N_u \) was set such that roughly 95% of the observations from the data source lay within the hypothetical model 95% prediction intervals. These prediction intervals are hypothetical in the sense that the prediction intervals for each data source are constructed independently of the other data sources, i.e. by setting the other \( N_u \) values to zero. The prediction intervals are also hypothetical in the sense that they represent the 95% prediction interval around the observation rather than around the mean, i.e. they have been adjusted to reflect random variation. Values of 0.01 and 0.0002 were chosen for \( N_{ANC} \) and \( N_{DHA} \) respectively, while \( N_{HSRC} \) and \( N_{RHRU} \) were set to 0.001. A more detailed description of the method used to obtain these prediction intervals and the \( N_u \) values is given in Appendix D. The prediction intervals presented in section 4 represent the confidence in the model prediction of the mean after taking into account all data sources, and are not adjusted to reflect random variation. The prediction intervals in section 4.1 are therefore narrower than the corresponding prediction intervals shown in Appendix D for the chosen \( N_u \) values.

In the following sections, the procedure for setting the likelihood is described, for each of the four data sources. As in the GLUE approach, the likelihood function should be interpreted as a ‘goodness of fit’ measure rather than as a likelihood in the probabilistic sense. The inclusion of the \( N_u \) parameters precludes a formal statistical interpretation of the likelihood.

### 3.4.1 Antenatal clinic prevalence data

Surveys of HIV prevalence in pregnant women attending public antenatal clinics have been conducted in South Africa every year since 1990. The sample of clinics is intended to be nationally representative, although there was a bias towards urban clinics in the early surveys, as explained previously. HIV prevalence rates for each five-year age band (together with 95% confidence intervals) are reported annually, but reporting of the sample sizes is less regular. These prevalence estimates, confidence intervals and sample sizes are included in Appendix E.

Suppose that in year \( t \), \( n_t(a,t) \) women with age in the interval \([a, a + 5)\) are sampled in the national antenatal survey, and a proportion \( x_t(a,t) \) is estimated to be HIV-positive. Suppose that \( M_t(a,t) \) is the proportion of pregnant women of this age group, attending public antenatal clinics, who one would expect to test positive (as estimated by the
The sampled number of women who test HIV positive is then binomially distributed with parameters \( n_1(a,t) \) and \( M_1(a,t) \), and the contribution to the likelihood from the observation \( x_1(a,t) \) is thus

\[
\frac{n_1(a,t)!}{(n_1(a,t)x_1(a,t))!(n_1(a,t)(1-x_1(a,t)))!} M_1(a,t)^{n_1(a,t)x_1(a,t)}(1-M_1(a,t))^{n_1(a,t)(1-x_1(a,t))}.
\] (15)

The method used to calculate \( x_1(a,t) \) in the survey is to weight the province-specific estimates of prevalence by estimated numbers of births to women aged \([a, a+5]\) in each province (Department of Health 1997). The binomial assumption is thus not strictly correct. In addition, it is not feasible to calculate the factorial function when \( n_1(a,t) \) is large. It may therefore be appropriate to use the normal approximation to the binomial, so that the contribution to the likelihood from observation \( x_1(a,t) \) becomes

\[
\frac{\exp\left(-\frac{n_1(a,t)(x_1(a,t)-M_1(a,t))^2}{2M_1(a,t)(1-M_1(a,t))}\right)}{\sqrt{2\pi M_1(a,t)(1-M_1(a,t))/n_1(a,t)}}.
\] (16)

This approximation is usually only valid if the number testing positive and the number testing negative are both greater than five (Rice 1995). This requirement is not met in the first year of the survey (Department of National Health and Population Development 1991), and also is not met in many of the later surveys for women over the age of 40 (sample sizes are small in these age categories). Data for 1990 and data for women over the age of 40 are therefore excluded from the likelihood measure.

Assuming the prevalence rates in different years and age bands are independent, the likelihood for the antenatal clinic data is then calculated as

\[
L_{\text{ANC}} = \prod_{a} \prod_{t=1991}^{2003} \exp\left(-\frac{n_1(a,t)(x_1(a,t)-M_1(a,t))^2}{2M_1(a,t)(1-M_1(a,t))}\right) \frac{1}{\sqrt{2\pi M_1(a,t)(1-M_1(a,t))/n_1(a,t)}}.
\] (17)

Other likelihood functions, with suitable adjustments to eliminate over- or under-dispersion, could also have been used, and there is thus a degree of arbitrariness in the choice of likelihood function.

### 3.4.2 Vital registration data

Statistics South Africa (2005) recently published an analysis of death notification forms submitted to the Department of Home Affairs (DHA) over the 1997 to 2003 period. Numbers of reported deaths were tabulated for each calendar year, by sex and by five-year age band. The detailed numbers were not presented for 2003, however, since these
were incomplete due to late reporting of deaths. In the analysis that follows, only the reported deaths in the 1997 to 2002 period are considered. A small proportion of the reported deaths had no age recorded on the death notification form. These deaths were treated as being unreported for the purposes of estimating the extent of completeness and thus can be ignored.

Suppose that the model estimates that in year $t$ there are $M_2(s,a,t)$ deaths among individuals aged $a$ to $a+4$, of sex $s$, and that the corresponding number of recorded deaths, based on the DHA data, is $x_2(s,a,t)$. Further suppose that $c_i$ is the estimated completeness of death registration in year $t$. The quantity $x_2(s,a,t)$ should then be compared with $M_2(s,a,t) \times c_i$. The exact number of individuals in the population, aged $a$ to $a+4$, of sex $s$, at the start of year $t$, is not known, but if the model estimate of this quantity is $n_2(s,a,t)$, then it might be assumed that $x_2(s,a,t)$ is a realization of a binomially distributed random variable with parameters $n_2(s,a,t)$ and $M_2(s,a,t) c_i$. Alternatively, it might be assumed that $x_2(s,a,t)$ is a realization of a Poisson distributed random variable with parameter $M_2(s,a,t) c_i$. The normal approximations to these two distributions are similar, with the same mean and a slightly greater variance for the Poisson distributed variable. Since the variance of the distribution is effectively rescaled by parameter $N_{DHA}$ anyway, it makes little difference whether the binomial or Poisson assumption is used. Since the exact population size is unknown and since the Poisson assumption is simpler, the normal approximation to the Poisson is used here. The contribution of $x_2(s,a,t)$ to the likelihood is thus

$$
\exp \left( - \frac{(x_2(s,a,t) - M_2(s,a,t) c_i)^2}{2M_2(s,a,t) c_i} \right) \over \sqrt{2\pi M_2(s,a,t) c_i}.
$$

(18)

The vast majority of the deaths under the age of 20 and over the age of 55 are likely to be non-AIDS-related. Since the focus of this analysis is on the epidemiological parameters rather than the demographic parameters, these data can be ignored. The likelihood for the death data is thus

$$
L_{DHA} = \prod_{s} \prod_{a} \prod_{t=1997}^{2003} \exp \left( - \frac{(x_2(s,a,t) - M_2(s,a,t) c_i)^2}{2M_2(s,a,t) c_i} \right) \over \sqrt{2\pi M_2(s,a,t) c_i},
$$

(19)

where the product is across all age groups from 20-24 to 50-54.

As before, there is some arbitrariness in this choice of likelihood function, and other parametric forms with suitably chosen means and variances may be equally appropriate.
3.4.3 Nelson Mandela/HSRC household prevalence survey data

The Nelson Mandela/HSRC survey is a nationally representative household survey that was conducted in 2002 and repeated in 2005. In this analysis, only HIV prevalence rates over the age of 15 are considered, as the paediatric HIV parameters are not included in the uncertainty analysis.

HIV prevalence rates and 95% confidence intervals for the 2002 survey are presented in the full report (Human Sciences Research Council 2002), for each quinquennial age band from 15-19 to 50-54, and for all individuals over the age of 54, separately for males and females. Sample sizes in each of these age and sex categories were provided by Thomas Rehle (Director, Social Aspects of HIV/AIDS and Health Programme, HSRC). Similar information is provided in the full report for the 2005 survey (Shisana et al, 2005b). The reported prevalence rates, confidence intervals and sample sizes are given in Appendix F.

The method used to define the contribution of the HSRC data to the likelihood is similar to that used for the antenatal clinic data. Suppose that \( n_{s,a,t} \) is the number of individuals sampled in the HSRC survey in year \( t \), of sex \( s \) and age \( a \), and \( x_{s,a,t} \) is the proportion estimated to be HIV-positive. If \( M_{s,a,t} \) is the corresponding model estimate of the proportion that would test positive in year \( t \), and the normal approximation to the binomial distribution is again used, then the contribution of the HSRC data to the likelihood is

\[
L_{\text{HSRC}} = \prod_{s} \prod_{a} \prod_{t} \frac{\exp\left(-\frac{n_{s,a,t} (x_{s,a,t} - M_{s,a,t})^2}{2M_{s,a,t} (1 - M_{s,a,t})}\right)}{\sqrt{2\pi M_{s,a,t} (1 - M_{s,a,t}) / n_{s,a,t}}},
\]

where the product is across both survey years, all age groups 15 and older, and both sexes. As with the antenatal clinic data, the binomial assumption is not strictly correct, as the \( x_{s,a,t} \) values are obtained by weighting the prevalence rates in each race, province and locality type stratum by the estimated population size in the stratum (Shisana et al, 2005b).

3.4.4 RHRU/loveLife prevalence survey data

The Reproductive Health Research Unit, together with loveLife, conducted a survey of HIV prevalence in youth in 2003. The survey was nationally representative, and sampled only individuals in the 15 to 24 age category. HIV prevalence rates, confidence intervals and sample sizes are given in the full report (Reproductive Health Research Unit 2004), and are reproduced in Table 3.7 below.
Table 3.7: RHRU/loveLife prevalence survey results
Source: Reproductive Health Research Unit (2004)

The approach followed in calculating \( L_{\text{RHRU}} \) is similar to that followed for the HSRC data. If \( n_s(s,a) \) is the number of individuals sampled, of sex \( s \) and age \( a \), \( x_s(s,a) \) is the proportion estimated to be HIV-positive, and \( M_s(s,a) \) is the corresponding model estimate of the proportion that would test positive, then the contribution of the RHRU data to the likelihood is

\[
L_{\text{RHRU}} = \prod_s \prod_a \frac{\exp\left( -\frac{n_s(s,a)x_s(s,a) - M_s(s,a))^2}{2M_s(s,a)[1 - M_s(s,a)]} \right)}{\sqrt{2\pi M_s(s,a)[1 - M_s(s,a)]/n_s(s,a)}},
\]

where the product is across both sexes and both the 15-19 and 20-24 age groups. As with the HSRC survey data, prevalence estimates in Table 3.7 are obtained by weighting prevalence measured in each race, province and locality type stratum by the corresponding estimates of population size. The binomial assumption implicit in the definition of the likelihood is therefore not strictly correct.

A potential criticism of the approach followed with the RHRU data, as well as with the HSRC and ANC prevalence data, is that the likelihood functions do not take into account clustering in the samples drawn. As a result, the variance of the binomial distribution is under-estimated. However, since the effect of the \( N_{\text{RHRU}} \) parameter is to rescale the variance to an acceptable degree, the initial under-estimation of the variance is of little consequence.

3.5 Programming

The ASSA2000 and ASSA2002 models that are currently publicly available are spreadsheet models, programmed in Excel/VBA. The ASSA2000 model was also coded in Visual C++ Version 6.0, by Alan Matthews (University of KwaZulu-Natal, School of Physics). In the interests of computational speed, the C++ code was chosen as the model on which the uncertainty analysis would be based. This code has been updated and modified to incorporate the new features of the ASSA2002 spreadsheet model, using Visual C++ .NET. A reconciliation of the C++ and Excel models was performed, and
after correcting a bug in the ASSA2002 Excel model, it was found that the C++ and Excel models produced identical results.

Two libraries were copied into the ASSA2002 C++ code for the purpose of the uncertainty analysis. The DCDFLIB library (downloaded on 4 March 2005 from http://www.csit.fsu.edu/~burkardt/cpp_src/dcdflib/dcdflib.html) is used for its statistical functions, notably the cumulative beta and gamma distributions. The ‘randomc’ library (downloaded from http://www.agner.org/random/randomc.htm on 3 May 2005) is used to generate random numbers from the uniform (0,1) distribution. The ‘Mersenne Twister’ random number generator is used for this purpose.

Four functions were added to the ASSA2002 C++ code for the purpose of the uncertainty analysis:

- A function which randomly generates 100,000 parameter combinations and calculates the four components of the log likelihood for each.
- A function which rescales each of the four log likelihood components, calculates the aggregate likelihood, and then samples (with replacement) the 100,000 parameter combinations, using the aggregate likelihood as the sample weights. A sample of 500 is drawn.
- A function which runs the ASSA2002 model for each of the 500 parameter combinations sampled, generating more detailed model outputs.
- A function which performs both the first and second functions, but without storing the intermediate results for the 100,000 parameter combinations in an output file.

The user can specify which of the functions is to be run, and can alter the number of parameter combinations initially generated as well as the size of the sample that is drawn.
4. Uncertainty analysis results

4.1 Comparison of observations and model outputs

Figure 4.1 below compares the levels of HIV prevalence observed in the antenatal clinic surveys, \( x_i(a,t) \), with the levels of HIV prevalence predicted by the model for pregnant women attending public clinics, \( M_i(a,t) \). Almost all of the observed prevalence points are contained in the 95% prediction intervals for \( M_i(a,t) \). The model tends to underestimate recent HIV prevalence rates in the 25 to 29 age band, although there is reasonable consistency between modelled and observed prevalence levels for other age bands. The model outputs, \( M_i(a,t) \), represent the prevalence one would expect to measure, taking into account false positive reactions and other sources of bias.

Figure 4.1: Levels of HIV prevalence observed and modelled in pregnant women
Observed values, \( x_i(a,t) \), represented by crosses (×). Mean of 500 simulated prevalence levels represented by closed squares (■). 2.5 and 97.5 percentiles of simulated prevalence levels represented by dashed lines.
Figure 4.2: Numbers of reported male deaths, predicted and actual

Actual reported numbers, \( x(s,a,t) \), represented by crosses (×). Mean of 500 simulated mortality levels represented by closed squares (■). 2.5 and 97.5 percentiles of simulated mortality levels represented by dashed lines.

Figure 4.2 compares actual numbers of reported male deaths, \( x(s,a,t) \), with the numbers of reported male deaths predicted by the model, \( M(s,a,t) \). In recent years, model estimates of reported male deaths in the 35 to 39 and 40 to 44 age categories tend to exceed the actual reported deaths. In the earlier years, there are significant
discrepancies between predicted and actual reported deaths in the 20 to 24 and 45 to 54 age ranges.

Figure 4.3: Numbers of reported female deaths, predicted and actual
Actual reported numbers, \( x_i(s,a,t) \), represented by crosses (×). Mean of 500 simulated mortality levels represented by closed squares (■). 2.5 and 97.5 percentiles of simulated mortality levels represented by dashed lines.

Figure 4.3 compares actual numbers of female deaths reported with those predicted by the model. The predicted trend in reported female deaths is roughly consistent with the trend in numbers actually reported, except in the 20 to 24, 35 to 44 and 50 to 54 age ranges.
categories. In the 20 to 24 and 50 to 54 age groups, the rate of increase in the actual reported deaths tends to exceed the predicted rate of increase. However, in the 35 to 39 and 40 to 44 age groups, the increase in the actual reported deaths tends to be lower than that predicted by the model.

In Figure 4.4, the levels of HIV prevalence observed in the Nelson Mandela/HSRC household surveys, $x_3(s,a,t)$, are compared with the levels of prevalence predicted by the model, $M_3(s,a,t)$. The model outputs represent the prevalence one would expect to measure, taking into account false positive reactions in the 2002 survey. Modelled levels of HIV prevalence in women are generally consistent with those levels measured in the survey. Modelled levels of prevalence in men, however, tend to exceed the measured levels between the ages of 30 and 54. The model prediction intervals are generally narrower in 2005 than in 2002, as the introduction of confirmatory testing in the 2005 survey reduces uncertainty related to false positive reactions.

![Figure 4.4: Levels of HIV prevalence observed and modelled in the HSRC surveys](image)

Observed values, $x_3(s,a,t)$, represented by crosses ($\times$), with 95% confidence intervals. Mean of 500 simulated prevalence levels represented by closed squares (■). 2.5 and 97.5 percentiles of simulated prevalence levels represented by dashed lines.

Table 4.1 compares modelled levels of HIV prevalence in youth in 2003, $M_4(s,a)$, with levels of HIV prevalence observed in the RHRU/loveLife survey, $x_4(s,a)$. Again, model estimates of prevalence include allowance for false positive reactions. The mean prevalence levels simulated are close to those observed in all cases.
### Table 4.1: Levels of HIV prevalence observed and modelled in 2003 RHRU survey

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age</th>
<th>Observed prevalence</th>
<th>Modelled prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>15-19</td>
<td>2.5% (1.8-3.4%)</td>
<td>2.4% (0.4-8.0%)</td>
</tr>
<tr>
<td></td>
<td>20-24</td>
<td>7.6% (6.3-9.3%)</td>
<td>8.7% (5.5-13.3%)</td>
</tr>
<tr>
<td>Female</td>
<td>15-19</td>
<td>7.3% (5.9-9.0%)</td>
<td>8.5% (5.6-13.1%)</td>
</tr>
<tr>
<td></td>
<td>20-24</td>
<td>24.5% (20.3-29.3%)</td>
<td>24.8% (20.3-30.1%)</td>
</tr>
</tbody>
</table>

Model estimates are means of 500 simulated prevalence. 95% confidence intervals/prediction intervals indicated in brackets.

### 4.2 Model projections up to 2005

The results of the model are presented in Figure 4.5. Trends in total HIV infections and HIV prevalence are shown in Figures 4.5(a) and 4.5(b) respectively. In mid-2005, the average estimated number of HIV infections in South Africa is 5.1 million, with a 95% prediction interval of 4.2 million to 6.0 million. This is equivalent to an HIV prevalence of 11.1% of the total population (9.1-13.1%). Total HIV infections are still rising in most of the 500 scenarios.

Figures 4.5(c) and 4.5(d) show trends in annual numbers of new HIV infections and HIV incidence rates respectively. In most scenarios, HIV incidence rates peaked during the 1997 to 1999 period, and have been declining since then. The average estimated number of new HIV infections that occurred between mid-2004 and mid-2005 is 490 000 (370 000-590 000). The corresponding average HIV incidence rate over the same period is 1.2% (0.9-1.5%).

The rise in AIDS mortality lags the rise in total HIV infections, as is evident from a comparison of Figures 4.5(e) and 4.5(a). There has been a rapid rise in AIDS mortality over the last decade, though increases in recent years have not been as substantial as in previous years. Between mid-2004 and mid-2005, the average estimated number of AIDS deaths is 320 000 (270 000-380 000). This is equivalent to an increase of 7.0 per 1000 (5.8-8.5/1000) in the crude mortality rate, as shown in Figure 4.5(f). These estimates do not reflect the uncertainty with respect to the effects of antiretroviral treatment. Associated with the rise in AIDS mortality is a rise in AIDS morbidity, shown in Figures 4.5(g) and 4.5(h). The average estimated number of AIDS cases in South Africa in mid-2005 is 590 000 (500 000-680 000). The average estimated percentage of the population sick with AIDS is 1.3% (1.1-1.5%).
Figure 4.5: Trends in HIV prevalence, HIV incidence, AIDS mortality and AIDS morbidity
Mean of 500 estimates represented by closed squares (■). 2.5 and 97.5 percentiles of estimates represented by dashed lines.

4.3 Comparison of prior and posterior distributions

The prior distributions for each of the 24 parameters considered in the uncertainty analysis are compared with the posterior distributions in Table 4.2. In most cases, the prior and posterior distributions are similar. The difference between the posterior and
prior means, expressed as a multiple of the standard deviation of the prior distribution, is
greatest in the case of the mean time from HIV infection to death (0.96), the fertility
reduction factor (0.67), the gradient parameter (-0.36) and the average annual number of
partners in the RSK group (0.35). Other parameters for which the posterior and prior
means differ by more than 0.25 prior standard deviations include the increase in HIV
infectiousness per log increase in viral load (-0.32), the mixing parameter (0.31), the
average annual number of partners in the STD group (-0.30) and the proportion of the
sexually active population in the RSK group (0.27). The prior and posterior distributions
of these parameters are compared graphically in Figure 4.6.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior dbn</th>
<th>Prior (mean, 95% interval)</th>
<th>Posterior (mean, 95% interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completeness of reporting</td>
<td>Beta (179.8, 34.2)</td>
<td>0.84 (0.79-0.89)</td>
<td>0.84 (0.79-0.88)</td>
</tr>
<tr>
<td>Ratio national sample: antenatal</td>
<td>Beta (19.2, 4.8)</td>
<td>0.80 (0.62-0.93)</td>
<td>0.81 (0.66-0.94)</td>
</tr>
<tr>
<td>Ratio private: public</td>
<td>Beta (5.83, 2.50)</td>
<td>0.70 (0.37-0.94)</td>
<td>0.68 (0.26-0.93)</td>
</tr>
<tr>
<td>Axsym false positive rate</td>
<td>Exponential (50)</td>
<td>2.0% (0.0-7.4%)</td>
<td>2.3% (0.0-6.5%)</td>
</tr>
<tr>
<td>Orasure false positive rate</td>
<td>Exponential (50)</td>
<td>2.0% (0.0-7.4%)</td>
<td>2.1% (0.0-7.8%)</td>
</tr>
<tr>
<td>Fertility reduction factor</td>
<td>Beta (49.19, 2.59)</td>
<td>0.95 (0.88-0.99)</td>
<td>0.97 (0.93-1.00)</td>
</tr>
<tr>
<td>Transmission parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male to female (RSK-RSK)</td>
<td>Beta (24.95, 12449)</td>
<td>2.0 (1.3-2.9) per 1000</td>
<td>2.0 (1.4-2.9) per 1000</td>
</tr>
<tr>
<td>Female to male (RSK-RSK)</td>
<td>Beta (24.97, 24949)</td>
<td>1.0 (0.6-1.4) per 1000</td>
<td>1.0 (0.6-1.4) per 1000</td>
</tr>
<tr>
<td>Male to female (STD-STD)</td>
<td>Beta (24.82, 3521)</td>
<td>7.0 (4.5-10.0) per 1000</td>
<td>7.0 (4.9-9.3) per 1000</td>
</tr>
<tr>
<td>Female to male (STD-STD)</td>
<td>Beta (24.87, 4949)</td>
<td>5.0 (3.2-7.1) per 1000</td>
<td>5.0 (3.4-6.9) per 1000</td>
</tr>
<tr>
<td>Infectiousness increase/log VL</td>
<td>Gamma (6.25, 0.28)</td>
<td>1.75 (0.66-3.37)</td>
<td>1.52 (0.76-2.70)</td>
</tr>
<tr>
<td>Susceptibility increase/yr of age</td>
<td>Gamma (16, 0.0075)</td>
<td>12.0% (6.9-18.6%)</td>
<td>12.3% (8.2-17.9%)</td>
</tr>
<tr>
<td>Mortality parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean HIV survival (age 29)</td>
<td>Gamma (110, 0.095)</td>
<td>10.5 (8.6-12.5)</td>
<td>11.5 (10.0-12.9)</td>
</tr>
<tr>
<td>Decrease in mean per 10 years</td>
<td>Gamma (25, 0.06)</td>
<td>1.50 (0.97-2.14)</td>
<td>1.50 (1.00-2.00)</td>
</tr>
<tr>
<td>Gradient parameter</td>
<td>Gamma (4, 0.075)</td>
<td>0.30 (0.08-0.66)</td>
<td>0.25 (0.08-0.52)</td>
</tr>
<tr>
<td>Sexual behaviour</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coital reduction factor (K)</td>
<td>Uniform (0, 1)</td>
<td>0.50 (0.025-0.975)</td>
<td>0.48 (0.02-0.97)</td>
</tr>
<tr>
<td>Reduction in annual # partners</td>
<td>Beta (0.8, 7.2)</td>
<td>10.0% (0.1-36.7%)</td>
<td>8.2% (0.2-33.0%)</td>
</tr>
<tr>
<td>% in STD risk group</td>
<td>Beta (19.8, 79.2)</td>
<td>20% (13-28%)</td>
<td>21% (14-28%)</td>
</tr>
<tr>
<td>% in RSK risk group</td>
<td>Beta (19.0, 63.7)</td>
<td>23% (15-33%)</td>
<td>24% (17-33%)</td>
</tr>
<tr>
<td>Annual # partners (PRO)</td>
<td>Gamma (25, 10)</td>
<td>250 (162-357)</td>
<td>261 (179-347)</td>
</tr>
<tr>
<td>Annual # partners (STD)</td>
<td>Gamma (25, 0.48)</td>
<td>12.0 (7.8-17.1)</td>
<td>11.3 (8.1-15.5)</td>
</tr>
<tr>
<td>Annual # partners (RSK)</td>
<td>Gamma (25, 0.04)</td>
<td>1.00 (0.65-1.43)</td>
<td>1.07 (0.70-1.48)</td>
</tr>
<tr>
<td>Sexual mixing parameter</td>
<td>Beta (12, 12)</td>
<td>0.50 (0.31-0.69)</td>
<td>0.53 (0.34-0.70)</td>
</tr>
<tr>
<td>Shape factor (sex activity)</td>
<td>Beta (24.9, 4949)</td>
<td>0.0050 (0.0032-0.0071)</td>
<td>0.0047 (0.0036-0.0062)</td>
</tr>
</tbody>
</table>

Table 4.2: Comparison of prior and posterior distributions
Figure 4.6 demonstrates that the empirical data strongly suggest a longer average time from HIV infection to death than was previously assumed, as well as a higher variance of the distribution of HIV survival times. The data also suggest that in the South African context, HIV does not substantially reduce fertility, and it would also appear from the empirical data that the increase in HIV infectiousness, per log increase in viral load, is not as substantial as previously assumed. A better fit to the data is usually obtained when
one assumes a higher average annual number of partners in the RSK group, a higher proportion of the population in the RSK group, and a higher mixing parameter. All three of these parameter changes tend to be associated with significant increases in HIV incidence in the later stages of the HIV/AIDS epidemic (see Table 5.1).

In some cases, particular relationships between parameters are of interest. Table 4.3 compares the posterior and prior distributions for a number of quantities which can be expressed as functions of two or more of the parameters considered in the uncertainty analysis. The prior distributions for these quantities cannot be determined analytically, and therefore the prior distributions have been approximated numerically, using 500 simulations.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Prior (mean, 95% interval)</th>
<th>Posterior (mean, 95% interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% in NOT group</td>
<td>56% (43-68%)</td>
<td>54% (44-63%)</td>
</tr>
<tr>
<td>Ratio of male-to-female transmission to female-to-male transmission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in RSK group</td>
<td>2.07 (1.16-3.34)</td>
<td>2.14 (1.11-3.79)</td>
</tr>
<tr>
<td>in STD group</td>
<td>1.45 (0.78-2.41)</td>
<td>1.45 (0.80-2.33)</td>
</tr>
<tr>
<td>For individuals infected at age 29, with no access to HAART</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median survival (years)</td>
<td>9.9 (8.1-11.9)</td>
<td>10.8 (9.4-12.0)</td>
</tr>
<tr>
<td>% surviving after 5 years</td>
<td>91% (82-97%)</td>
<td>93% (89-96%)</td>
</tr>
<tr>
<td>% surviving after 10 years</td>
<td>49% (31-66%)</td>
<td>57% (45-67%)</td>
</tr>
<tr>
<td>% surviving after 15 years</td>
<td>15% (6-28%)</td>
<td>21% (11-31%)</td>
</tr>
<tr>
<td>% sick with AIDS after 5 years</td>
<td>10% (5-16%)</td>
<td>9% (6-12%)</td>
</tr>
<tr>
<td>% sick with AIDS after 10 years</td>
<td>35% (27-45%)</td>
<td>31% (25-38%)</td>
</tr>
<tr>
<td>% sick with AIDS after 15 years</td>
<td>55% (42-73%)</td>
<td>49% (39-63%)</td>
</tr>
</tbody>
</table>

Table 4.3: Comparison of prior and posterior distributions for other quantities

The posterior average proportion of the sexually active population in the NOT group is marginally lower than the prior average, reflecting the difference in prior and posterior averages for the STD and RSK proportions. The ratio of the male-to-female transmission probability to the female-to-male transmission probability has similar prior and posterior averages, both in the case of transmission within the RSK group and in the case of transmission within the STD group. As might be expected from Table 4.2, the rates of HIV survival implied by the posterior distribution are higher than those suggested by the prior distribution, at all durations of infection. The longer survival times also imply lower proportions of HIV-positive survivors sick with AIDS (i.e. in WHO clinical stage 4) at each duration of infection.
5. Sensitivity analysis results

Of the 500 parameter combinations sampled, 362 remained after resampled parameter combinations were excluded from the analysis. The PRCCs calculated using these 362 parameter combinations and associated model outputs are given in Table 5.1. Model outputs are arranged such that the HIV transmission dynamics in the early stages of the epidemic most affect the outputs listed first, and the last outputs listed reflect the more recent HIV transmission dynamics.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completeness of reporting</td>
<td>-0.14*</td>
<td>-0.19**</td>
<td>-0.17**</td>
<td>-0.16**</td>
<td>-0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>Ratio national sample: antenatal</td>
<td>0.15**</td>
<td>0.21**</td>
<td>0.15**</td>
<td>0.11*</td>
<td>0.07</td>
<td>-0.05</td>
</tr>
<tr>
<td>Ratio private: public</td>
<td>0.09</td>
<td>0.12*</td>
<td>0.12*</td>
<td>0.15*</td>
<td>0.24**</td>
<td>0.14*</td>
</tr>
<tr>
<td>Axysym false positive rate</td>
<td>-0.28**</td>
<td>-0.35**</td>
<td>-0.37**</td>
<td>-0.39**</td>
<td>-0.35**</td>
<td>-0.17**</td>
</tr>
<tr>
<td>Orasure false positive rate</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.05</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>Fertility reduction factor</td>
<td>-0.11*</td>
<td>-0.20**</td>
<td>-0.25**</td>
<td>-0.16**</td>
<td>-0.23**</td>
<td>-0.02</td>
</tr>
<tr>
<td>Transmission parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male to female (RSK-RSK)</td>
<td>-0.14*</td>
<td>-0.17**</td>
<td>-0.10</td>
<td>0.01</td>
<td>0.20**</td>
<td>0.35**</td>
</tr>
<tr>
<td>Female to male (RSK-RSK)</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.04</td>
<td>0.03</td>
<td>0.14*</td>
<td>0.15**</td>
</tr>
<tr>
<td>Male to female (STD-STD)</td>
<td>0.28**</td>
<td>0.22**</td>
<td>0.14*</td>
<td>0.13*</td>
<td>0.02</td>
<td>-0.05</td>
</tr>
<tr>
<td>Female to male (STD-STD)</td>
<td>0.36**</td>
<td>0.40**</td>
<td>0.42**</td>
<td>0.43**</td>
<td>0.34**</td>
<td>0.16**</td>
</tr>
<tr>
<td>Infectiousness increase/log VL</td>
<td>-0.38**</td>
<td>-0.36**</td>
<td>-0.28**</td>
<td>-0.22**</td>
<td>-0.02</td>
<td>0.14*</td>
</tr>
<tr>
<td>Susceptibility increase/yr of age</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>Mortality parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean HIV survival (age 29)</td>
<td>-0.60**</td>
<td>0.35**</td>
<td>0.24**</td>
<td>-0.70**</td>
<td>0.53**</td>
<td>-0.06</td>
</tr>
<tr>
<td>Decrease in mean per 10 years</td>
<td>0.21**</td>
<td>0.08</td>
<td>0.05</td>
<td>0.04</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>Gradient parameter</td>
<td>-0.63**</td>
<td>0.14*</td>
<td>0.08</td>
<td>0.09</td>
<td>0.23**</td>
<td>-0.01</td>
</tr>
<tr>
<td>Sexual behaviour</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coital reduction factor (K)</td>
<td>-0.38**</td>
<td>-0.37**</td>
<td>-0.32**</td>
<td>-0.29**</td>
<td>-0.06</td>
<td>0.13*</td>
</tr>
<tr>
<td>Reduction in annual # partners</td>
<td>0.11*</td>
<td>0.15**</td>
<td>0.11*</td>
<td>-0.16**</td>
<td>-0.46**</td>
<td>-0.61**</td>
</tr>
<tr>
<td>% in STD risk group</td>
<td>-0.38**</td>
<td>-0.28**</td>
<td>0.16**</td>
<td>0.48**</td>
<td>0.66**</td>
<td>0.65**</td>
</tr>
<tr>
<td>% in RSK risk group</td>
<td>-0.11*</td>
<td>-0.07</td>
<td>-0.04</td>
<td>0.28**</td>
<td>0.58**</td>
<td>0.76**</td>
</tr>
<tr>
<td>Annual # partners (PRO)</td>
<td>0.66**</td>
<td>0.47**</td>
<td>-0.05</td>
<td>-0.04</td>
<td>-0.22**</td>
<td>-0.14*</td>
</tr>
<tr>
<td>Annual # partners (STD)</td>
<td>-0.15**</td>
<td>0.17**</td>
<td>0.59**</td>
<td>0.49**</td>
<td>0.34**</td>
<td>-0.05</td>
</tr>
<tr>
<td>Annual # partners (RSK)</td>
<td>-0.14*</td>
<td>-0.20**</td>
<td>-0.16**</td>
<td>0.27**</td>
<td>0.53**</td>
<td>0.65**</td>
</tr>
<tr>
<td>Sexual mixing parameter</td>
<td>0.18**</td>
<td>0.12*</td>
<td>-0.06</td>
<td>0.20**</td>
<td>0.44**</td>
<td>0.52**</td>
</tr>
<tr>
<td>Shape factor (sex activity)</td>
<td>0.06</td>
<td>0.07</td>
<td>-0.02</td>
<td>-0.18**</td>
<td>-0.24**</td>
<td>-0.27**</td>
</tr>
</tbody>
</table>

Table 5.1: Partial rank correlation coefficients (PRCCs)

Associations significant at the 5% level are indicated with a * and those significant at the 0.5% level are indicated with a **. ‘HIV prevalence’ refers to the percentage of the population infected with HIV at the middle of the relevant year. ‘HIV incidence’ is the proportion of the HIV negative population becoming infected with HIV between the middle of the year and the middle of the next year, and ‘extra mortality’ is the proportion of the population dying from AIDS over the same period.
5.1 Bias parameters

The completeness of vital registration is strongly negatively related to the model outputs, but is not significantly correlated with HIV incidence and prevalence in 2005. High levels of predicted AIDS mortality in the early years of the epidemic can only be consistent with vital registration data if it is assumed that a high proportion of deaths are not reported, and this appears to be the reason for the negative associations. Similarly, high levels of predicted HIV prevalence are more likely to be consistent with antenatal prevalence data if it is assumed that HIV-positive women are significantly less fertile, and this would explain the negative association between the fertility reduction factor and the model outputs. The fertility reduction factor ceases to be strongly negatively associated with the model outputs in 2005, perhaps because a high fertility reduction factor implies more children born to HIV-positive women, and hence more paediatric HIV, which would counteract the previously described negative association.

The false positive rate on the Abbott Axsym ELISA is strongly negatively associated with all of the model outputs. The explanation is that a high false positive rate implies a true antenatal prevalence much lower than that observed, and hence a smaller epidemic. The false positive rate on the Orasure ELISA, however, is not significantly correlated with the model outputs. This may be due to the HSRC and RHRU surveys having less of an influence on the likelihood function, relative to the antenatal survey and vital registration data. Any bias in the HSRC and RHRU survey results, due to false positive reactions on the Orasure ELISA, is therefore likely to have little effect on the likelihood values calculated.

The two sources of antenatal sampling bias appear to have a significant effect on the outputs of the model. The ratio of the initial antenatal sample prevalence to the prevalence in all pregnant women using the public sector is significantly positively associated with the model outputs in 1995, when the extent of urban bias in the antenatal surveys was significant, but it is not strongly associated with the more recent HIV prevalence and incidence. The odds ratio for HIV infection in women attending private health facilities, relative to women attending public facilities, is strongly positively correlated with most of the model outputs, which indicates that the prevalence of HIV in users of private health facilities significantly influences the magnitude of the epidemic at a national level.

5.2 HIV transmission parameters

As might be expected, both female-to-male and male-to-female HIV transmission probabilities in the STD group strongly influenced the model outputs for 1995. However, the male-to-female transmission probabilities became steadily less significant as determinants of the more recent HIV transmission dynamics, while the female-to-male transmission probabilities remained highly significant. A possible explanation is that
there are relatively more male susceptibles than female susceptibles in the later stages of the epidemic, which implies that HIV incidence in the later stages of the epidemic would be expected to depend more on female-to-male transmission probabilities than on male-to-female. HIV infection in the RSK group occurs mostly in the later stages of the epidemic, and it is therefore only the HIV incidence and prevalence rates in 2005 that are strongly positively associated with the transmission probabilities in the RSK group. There is some negative correlation between RSK transmission probabilities and the model outputs in 1995, but the explanation for this is not clear.

The increase in infectiousness per log increase in viral load is strongly negatively associated with the early model outputs, but the association becomes less strong the more recent the transmission dynamics influencing the model output, and ultimately switches to being positive. A high increase in infectiousness per log increase in viral load implies a low probability of transmitting the virus when the infected individual is in the early stages of disease, relative to the likelihood that applies when the infected individual is in the late stages of disease. A high value of this parameter therefore favours a lower level of HIV incidence early in the epidemic, when most infected individuals are in the early stages of disease. The association becomes weaker and then switches to being positive as the proportion of infected people in the later stages of disease increases.

The increase in HIV susceptibility in young females is not significantly associated with any of the model outputs.

5.3 HIV survival parameters

Not surprisingly, the mean term from HIV infection to death is strongly negatively associated with the increase in mortality due to AIDS, both in 1995 and 2005. However, the mean term from HIV infection to death is also strongly positively associated with the HIV prevalence and incidence in 1995. A possible explanation is that when high mean survival terms are sampled, the remaining parameters sampled will favour a high HIV prevalence and incidence in 1995, to compensate for the low rates of early AIDS mortality implied by the high mean survival terms (this compensation is necessary in order to ensure that the modelled trends in mortality are consistent with those observed). There is also a strong positive correlation between the mean survival term and the HIV prevalence in 2005, as greater survival of HIV-positive individuals implies more HIV-positive individuals alive in the later stages of the epidemic.

The gradient parameter determines the variance of the time to death. A high gradient parameter implies a low variance, which in turn implies few deaths in the very early stages of the epidemic – hence the strong negative correlation between the gradient parameter and the extra mortality in 1995. As before, it appears that the other parameters sampled favour a high prevalence of HIV in 1995, to compensate for the low mortality implied by the high gradient parameter. This may explain the strong positive association between the gradient parameter and the prevalence rates in 1995 and 2005.
The decrease in mean survival time per 10-year increase in age is strongly positively associated with the extra mortality due to AIDS in 1995, but is not significantly related to the other model outputs. A possible explanation is that in the very early stages of the epidemic, the first individuals dying from AIDS are likely to be those infected at older ages, since those infected at older ages have the shortest survival times. The greater the decrease in mean survival per 10-year increase in age, the higher the rate of AIDS mortality in these older individuals will be.

5.4 Sexual behaviour parameters

The relationship between the coital reduction factor and the model outputs is similar to the relationship between the model outputs and the increase in infectiousness per log increase in viral load. The higher the coital reduction factor is, the more infectious HIV-positive individuals are in the late stages of disease, relative to the early stages of disease. A high coital reduction factor is therefore likely to be associated with a relatively low average level of infectiousness in the early stages of the epidemic, when most HIV-infected individuals are in the early stages of disease, and a high average level of infectiousness in the later stages of the epidemic.

The reduction in the annual number of partners is strongly negatively associated with model outputs in 2005, but is positively associated with earlier model outputs. The negative association is to be expected, as the reduction in partner numbers is assumed to be the result of social marketing programmes, which were only phased in to any significant extent in the late 1990s. However, the explanation for the positive correlation coefficients in 1995 is unclear.

The percentage of the sexually active population in the RSK group is strongly positively associated with the model outputs for 2005, but is not positively associated with the outputs in 1995, due to there being relatively little transmission in the RSK group early in the epidemic. The percentage of the sexually active population in the STD group is significantly associated with all the model outputs, though the association is negative in the case of the extra mortality and HIV prevalence in 1995. A possible explanation is that the rate of partner change in males in the PRO group is negatively related to the proportion of the sexually active population in the STD group (see equations (C4) and (C6) in Appendix C). A high percentage of the sexually active population in the STD group therefore implies that the epidemic takes longer to become established in the PRO group, so that the epidemic starts slightly later, with resulting lower levels of AIDS mortality and HIV prevalence in 1995.

The annual numbers of partners for females in the different risk groups are strongly positively correlated with outputs at different stages in the epidemic. The annual number of partners in the PRO group has a significant impact on the very early dynamics of the epidemic, as is evident from the strong positive correlation (0.66) between this parameter and the extra mortality due to AIDS in 1995. The annual number of partners in the STD group is most significantly correlated with the HIV incidence in 1995, and the number of
partners in the RSK group is most significantly associated with HIV incidence in 2005. This reflects the progression of the epidemic over time: starting in the PRO group, spreading into the STD group, and finally reaching the RSK group.

The average annual numbers of partners for females in the PRO, STD and RSK groups also appear to be negatively associated with some of the model outputs. It is clear from equation (C6) in Appendix C that the average annual number of partners in the male STD group is negatively related to the average annual numbers of partners in both the female PRO group and the female STD group. This probably explains why the average annual numbers of female PRO partners and female RSK partners are negatively related to the magnitude of the epidemic in its late and early stages respectively. It is also clear from equations (C4) and (C6) in Appendix C that the average annual number of partners in the male PRO group is negatively related to the average annual number of partners in the female STD group, which may explain why the number of partners in the female STD group is negatively related to the level of AIDS mortality in 1995.

The sexual mixing parameter, which determines the extent of sexual mixing between risk groups, is strongly positively correlated with all model outputs other than the HIV incidence in 1995. The correlation appears to decrease in significance in the early stages of the epidemic, and then increase in significance in the later stages of the epidemic. A possible explanation is that in the early stages of the epidemic, the spread of HIV is heavily dependent on the extent of the interaction between the PRO and STD groups, while in the later stages of the epidemic, the spread of HIV is strongly dependent on the extent of the interaction between the STD and RSK groups. In the intermediate stage of the epidemic, the mixing parameter may be a less significant determinant of HIV incidence.

A high shape factor implies a concentrated distribution of sex activity with respect to age. This in turn implies an epidemic which reaches a ‘saturation point’ at an early stage; in the later stages of the epidemic, relatively few new HIV infections occur at older ages, and almost all new HIV infections occur in youth who have recently become sexually active. A high shape factor therefore favours an epidemic characterized by relatively low incidence rates in the later stages of the epidemic – hence the strong negative correlations with the 2005 model outputs.
6. Discussion

6.1 Comparison with other uncertainty and sensitivity analyses

The only other modelling group which has thus far published HIV/AIDS estimates for South Africa together with prediction intervals is UNAIDS. The approach followed by the UNAIDS group is described in section 2.1 (see Grassly et al (2004) for more detail). Table 6.1 below compares the point estimates and plausibility bounds produced by UNAIDS for four statistics (UNAIDS 2004) with the corresponding estimates and 95% prediction intervals produced using the ASSA2002 model. The estimates of HIV infections in South Africa in 2001 and 2003, obtained using the ASSA2002 approach, are both substantially lower than those obtained using the UNAIDS approach, though there is nevertheless considerable overlap between the ranges of uncertainty. The estimates of AIDS deaths are much lower when the ASSA2002 approach is used, and there is little overlap between the ranges of uncertainty produced using the two approaches. The possible reasons for the differences in the point estimates are discussed in section 6.3 below. While the ranges of uncertainty around the numbers of infections are approximately equal in width, the ranges of uncertainty around the numbers of AIDS deaths, obtained using the UNAIDS approach, are almost three times as wide as those obtained using the ASSA2002 approach. This is to be expected, as the UNAIDS approach does not take into account the South African vital registration data, and the approach therefore cannot achieve the same degree of confidence when used to estimate mortality.

<table>
<thead>
<tr>
<th>Output</th>
<th>UNAIDS</th>
<th>ASSA2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total HIV infections, ages 0-49, end of 2003</td>
<td>5 300 000 (4 500 000 – 6 200 000)</td>
<td>4 600 000 (3 700 000 – 5 300 000)</td>
</tr>
<tr>
<td>Total HIV infections, ages 0-49, end of 2001</td>
<td>5 000 000 (4 200 000 – 5 900 000)</td>
<td>4 200 000 (3 400 000 – 4 800 000)</td>
</tr>
<tr>
<td>Total AIDS deaths, ages 0-49, during 2003</td>
<td>370 000 (270 000 – 520 000)</td>
<td>250 000 (210 000 – 300 000)</td>
</tr>
<tr>
<td>Total AIDS deaths, ages 0-49, during 2001</td>
<td>270 000 (190 000 – 390 000)</td>
<td>180 000 (140 000 – 220 000)</td>
</tr>
</tbody>
</table>

Table 6.1: Comparison of UNAIDS estimates with ASSA2002 estimates
Plausibility bounds/95% prediction intervals indicated in brackets.

The results of the sensitivity analysis are similar to those produced by other modellers, although none of the other sensitivity analyses have been conducted using models of the HIV/AIDS epidemic in South Africa specifically. Other studies have demonstrated that the length of the incubation period is negatively associated with levels of AIDS morbidity (Blower and Dowlatabadi 1994) and AIDS mortality in the early stages of the epidemic (Anderson 1996), but positively associated with HIV prevalence in the later stages of the
epidemic (Anderson 1996). As might be expected, the magnitude of the epidemic has been shown by others to be positively related to rates of partner change (Garnett and Anderson 1993; Blower and Dowlatabadi 1994), transmission probabilities (Blower and Dowlatabadi 1994) and coital frequencies (Auvert et al, 2000). A less obvious finding is that for highly random sexual mixing, early levels of HIV prevalence may be lower than those obtained with more assortative sexual mixing, although levels of HIV prevalence over the longer term are substantially higher when sexual mixing is less assortative (Garnett and Anderson 1993; Anderson 1996). The sensitivity analysis presented in section 5 also suggests that the relationship between this parameter and the model estimates changes significantly over the course of the epidemic.

Garnett and Anderson (1993) also confirm that a more concentrated distribution of HIV incidence (greater heterogeneity in sexual risk behaviour) leads to an epidemic with a more rapid initial spread but a lower endemic prevalence. Relative levels of infectiousness in different stages of disease have also been shown by Anderson (1996) to influence the evolution of an AIDS epidemic, with high levels of infectiousness in early disease favouring a very rapid epidemic growth.

6.2 Strengths and limitations

An advantage of the approach developed in this paper is that it can handle data from multiple sources. Until recently, UNAIDs estimates for countries with generalized HIV/AIDS epidemics were based almost exclusively on data collected from antenatal clinics, and the approach currently used by UNAIDS to determine plausibility bounds does not take into account other empirical data sources (Grassly et al, 2004). Increasingly, however, modellers are having to validate their models against HIV prevalence data from household surveys (Boerma et al, 2003) and mortality statistics (Timæus and Jasseh 2004). It is therefore important that uncertainty analysis techniques capable of handling multiple data sources be further explored. Ideally, this should be done in a Bayesian framework, so that prior knowledge of HIV/AIDS epidemiology can also be incorporated into the model.

A further advantage of the approach developed in this paper is that it allows the user the flexibility to alter the weights given to the different data sources and hence the width of the prediction intervals. To some extent, this is also a disadvantage, as it implies a significant degree of subjectivity. The subjectivity is partially overcome if there is an explicit criterion for determining what the weights should be. In this analysis, the criterion is that the prediction intervals should be wide enough to include approximately 95% of the data points to which the model is calibrated, after adjusting the prediction intervals to make allowance for random variation. Because of the rescaling of the likelihood, the ‘likelihood function’ is not a formal likelihood in the statistical sense, but rather a somewhat arbitrary measure of goodness of fit.

South Africa is a country with more detailed data on HIV/AIDS and its demographic impact than are available in most other developing countries. This analysis makes full use
of the demographic detail, with the likelihood function being evaluated for each sex and each five-year age band. This makes it possible to obtain calibrated projections of the impact of HIV/AIDS at an age- and sex-specific level. As far as we are aware, no other HIV/AIDS model does this. As shown in section 4.1, however, the correspondence between the age-specific data and the model estimates is poor in some age bands, and this suggests that further refinements to the model are necessary.

Uncertainty with respect to demographic parameters has not been taken into account in this uncertainty analysis. The vital registration data may therefore favour a particular AIDS mortality pattern when in fact a wider range of AIDS mortality patterns might be favoured if demographic uncertainty were accounted for. A related limitation is that the non-AIDS mortality assumptions in the ASSA2002 model were not derived independently of the vital registration data. The vital registration data have therefore been used both in setting the model parameters and in determining the model posterior distribution.

6.3 Key findings

One of the most notable findings is that it is generally not possible to obtain a good fit to the South African data unless a relatively long mean survival time is assumed. The posterior mean of the average survival time in an individual infected at age 29 is 11.5 years. This is substantially longer than the mean survival times assumed by many other modellers of the African HIV/AIDS epidemic, such as 4.5 years (Robinson et al., 1997), 8.0 years (Korenromp et al., 2002) and 8.5 years (Bracher et al., 2004). UNAIDS assume a median survival term of 9 years for developing countries (Grassly et al., 2004), and this may partly explain why the UNAIDS estimates of AIDS mortality in South Africa are higher than those obtained when using the ASSA2002 approach. Most models set HIV survival assumptions based on net rates of mortality observed in HIV-positive individuals, though these individuals typically experience a high rate of non-AIDS mortality. This may account for some of the difference between the mean survival time estimated here (in the absence of non-AIDS mortality) and the mean survival time assumed in other models.

A number of possible explanations can be given for the longer HIV survival times observed in South Africa. As discussed in section 3.3.3, it is possible that the natural history of HIV-1 subtype C may differ from that of other HIV-1 group M subtypes. It is also possible that HIV is evolving over time towards a less pathogenic virus, as a recent study has suggested (Ariën et al., 2005). South Africa’s heterosexual HIV/AIDS epidemic began later than the epidemic in other African countries, and it is thus possible that the virus may have evolved to a more attenuated form by the time it entered South Africa’s heterosexual population – and might still be evolving. It is also necessary to investigate further the possibility that better access to prophylaxis against opportunistic infections may explain the better HIV survival prospects in South Africa (Badri et al., 2004). In addition, it is possible that better survival prospects in South Africa may be linked to the high prevalence of GB virus C in South Africa. This virus has been found to be
associated with delayed HIV disease progression in longitudinal studies, and has also been shown to inhibit HIV replication in vitro (Tillmann et al, 2001; Xiang et al, 2001). Since GB virus C RNA has been found in 29-38% of black South African blood donors and pregnant women (Casteling et al, 1998; Sathar et al, 1999; Sathar et al, 2004), the virus could be mitigating AIDS mortality in South Africa significantly.

The posterior distribution for the false positive rate on the Abbott Axsym ELISA has a mean of 2.3% and 95% credibility interval of 0.0% to 6.5%. This is a source of bias that is usually not considered by modellers calibrating their models against HIV prevalence data, although it has been acknowledged as a potential source of bias (Grassly et al, 2004). The sensitivity analysis demonstrates, however, that the model estimates are very sensitive to this parameter. If the true HIV prevalence in a country is 10%, and the false positive rate is 2%, the prevalence one would expect to observe in a national household survey would be 11.8%, in the absence of confirmatory testing. In this instance, the number of infections in the country would be overestimated by some 18% if the false positive rate was not taken into account. This may partially explain why the numbers of HIV infections estimated using the UNAIDS approach are higher than those produced using the ASSA2002 approach. Modellers need to interrogate seroprevalence data more closely and make allowance for potential biases where confirmatory testing has not been done. It is also necessary that studies be conducted to determine the specificity of the Abbott Axsym ELISA used in the South African antenatal clinic surveys, as this would reduce uncertainty regarding the true prevalence of HIV in South African women.

This analysis also suggests that the reduction in fertility due to HIV infection is likely to be small in South Africa. This is consistent with findings in other African countries with high rates of contraception and low rates of fertility (Terceira et al, 2003). The implication of this is that antenatal prevalence data are less likely to understate HIV prevalence in women generally, in a country like South Africa, all other things being the same. This may partially explain why the ASSA2002 estimates of numbers of infections in South Africa are lower than those of the UNAIDS model.

The sensitivity analysis demonstrates that the concentration of HIV incidence, with respect to age, plays an important role in determining the evolution of the epidemic. The more concentrated this distribution, the lower the prevalence of the virus over the longer term. The sensitivity analysis also demonstrates that the evolution of the HIV/AIDS epidemic is strongly influenced by the relative levels of infectiousness and relative levels of sexual activity in the different stages of disease. This has important implications for HIV prevention. As South Africa enters the more mature phase of its epidemic, it is becoming increasingly important to reduce levels of infectiousness and sexual risk behaviour in the later HIV disease stages. This can be achieved by promoting voluntary counselling and testing more intensively, especially to individuals with symptoms suggestive of HIV. The potential for antiretroviral treatment to reduce infectiousness in the later stages of disease also needs to be explored.
6.4 Future work

This uncertainty and sensitivity analysis considers only a subset of the parameters in the ASSA2002 model. It does not include uncertainty relating to the antiretroviral parameters, as the focus of this analysis has been on the uncertainty with respect to the evolution of the epidemic up to the current time, and access to antiretroviral treatment in South Africa has until recently been very limited. It also does not include the parameters determining mother-to-child transmission of HIV and paediatric HIV survival. This will be explored in future work. Work is currently being done to determine the range of uncertainty around projections into the future, taking into account uncertainty with respect to antiretroviral parameters and uncertainty with respect to the potential impact of HIV vaccines.

As noted previously, the uncertainty and sensitivity analysis also ignores uncertainty relating to the demographic parameters in the model (fertility, migration, non-AIDS mortality and initial population profile). The comparison of the model outputs with observed trends in mortality in section 4.1 does, however, suggest that certain modifications to these demographic parameters may be necessary. For example, model estimates of non-AIDS deaths in females appear to be on the high side, particularly in the 50 to 54 age category. It is possible that some of these discrepancies may be the result of problems with the data rather than problems with the model, for example age exaggeration in order to qualify for state old age pensions. It was found that replacing the non-AIDS mortality rates from the ASSA2002 model with those from the ASSA2003 model resulted in significantly improved correspondence with the vital registration data (results not shown). This suggests that some of the non-AIDS mortality rates in the ASSA2002 model may have been mis-specified. Further work is required to assess the extent of demographic uncertainty, which may require alternative statistical techniques (Lutz and Goldstein 2004).

In several cases, the trends and patterns predicted by the model deviate significantly from the data points, and the deviation appears to be non-random. For example, the observed trends in antenatal prevalence in recent years are not consistent with those predicted by the model for the 25 to 29 age band. A possible explanation for the model estimates being lower is that rates of partner change are assumed to be reduced to an equal degree at all ages, in response to social marketing programmes. If it were the case that reductions were mostly in youth (aged 15 to 24), predicted levels of HIV prevalence at older ages would be higher in recent years. Similarly, it is assumed that the effect of improved STD treatment on HIV transmission probabilities is the same at all ages. If curable STDs such as chlamydia and gonorrhoea are relatively more prevalent at younger ages (Buvé et al., 2001), and incurable STDs such as HSV-2 are relatively more prevalent at older ages (Smith and Robinson 2002), improved STD treatment may have less of an impact on HIV incidence at older ages.

It is also clear that the modelled rate of increase in mortality in females aged 50 to 54 is too low, relative to that observed. This may be partly due to the ASSA2002 assumption that HIV survival rates are the same for all individuals infected at age 39 or older.
Empirical evidence suggests that in fact the median survival time decreases roughly linearly with respect to the age at infection (Collaborative Group on AIDS Incubation and HIV Survival 2000). Rates of AIDS mortality in the 50 to 54 age category are therefore likely to be underestimated as a result of assuming no deterioration in survival prospects at older ages of infection.

For the recent years, modelled estimates of male deaths tend to be too high in the 35 to 39 and 40 to 44 age categories, and too low in the 50 to 54 age category. The latter problem is the same as that in females, and is probably also explained, in part, by the survival rates being assumed to be the same for all individuals infected over age 39. The former problem may be due to the ASSA2002 assumption that access to antiretroviral treatment is constant with respect to age. In 2001 and 2002, the years in which the deviation is most apparent, access to antiretroviral treatment was limited mainly to members of medical schemes. A comparison of the 1999 October Household Survey and ASSA2002 estimates of the population profile in 1999 shows that in the black African population, rates of medical scheme membership are highest in men aged 35 to 49. The relatively low rates of mortality observed in men aged 35 to 44 might therefore be partly explained by better access to antiretroviral treatment.

Further work may be necessary to improve the correspondence between the model and the results of the HSRC surveys. Firstly, modelled levels of HIV prevalence in men between the ages of 30 and 55 tend to be higher than the levels of HIV prevalence observed in the HSRC surveys. It is possible that the structure of the ASSA2002 model may be such that it overestimates HIV prevalence in these male age groups. However, it is also possible that the surveys may under-estimate HIV prevalence in this group. In the 2005 survey, only 38% of respondents over the age of 15 who were tested for HIV were male (Shisana et al, 2005b), which indicates the potential for a significant response bias in this group. A further discrepancy between the model and the survey is apparent in the 15 to 19 age group. Levels of HIV prevalence measured in the 15 to 19 age group in the 2005 survey tend to exceed the levels of HIV prevalence predicted by the model, for both males and females, which suggests that the model is underestimating the prevalence of HIV at young ages. The difference is not apparent in the 2002 HSRC survey and 2003 RHU survey comparisons, due to the absence of confirmatory testing in these surveys, which increases the uncertainty around the true prevalence of HIV.

This suggests that it will be necessary to explore whether different approaches to modelling sexual behaviour and HIV transmission result in lower prevalence of HIV in older men and higher prevalence of HIV in the 15 to 19 age group. The preceding discussion also suggests that future versions of the ASSA model may need to be adapted to model HIV survival at older ages more realistically. It will also be necessary to conduct research into whether prevention and treatment programmes are more effective or more widely accessed in some age groups than in others.

Future work will also explore possible correlations between AIDS mortality and non-AIDS mortality. Although HIV/AIDS modellers generally assume that AIDS mortality and non-AIDS mortality rates are independent, it is plausible that HIV risk behaviours
may be correlated with other health risk behaviours. Several African studies have shown smoking and alcohol consumption to be correlated with HIV infection and sexual activity (Tengia-Kessy et al., 1998; Williams et al., 2000; Taylor et al., 2003; Simbayi et al., 2004), and other factors such as socioeconomic status and exposure to violence may also influence an individual’s risk of HIV infection. All of these factors are known to influence mortality rates in the absence of AIDS. If non-AIDS and AIDS mortality rates were positively correlated, one might expect rising AIDS mortality to be associated with declining non-AIDS mortality rates. Allowance for correlation between non-AIDS mortality and AIDS mortality could therefore significantly affect the correspondence between the model estimates of reported deaths and the actual numbers of reported deaths.
Acknowledgements

This research is funded by the South African AIDS Vaccine Initiative (SAAVI), a consortium with primary funding from the Department of Health, the Department of Science and Technology and Eskom. We would like to thank Lize van der Merwe and Carl Lombard for statistical advice, Nathan Geffen and Michael Cohen for advice on C++ programming, and Thomas Rehle for providing unpublished data from the Nelson Mandela/HSRC survey. We would also like to thank Debbie Bradshaw for helpful advice on earlier drafts of this paper.
References


Morgan D., Mahe C., Mayanja B., et al. (2002) HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? AIDS. 16: 597-603


78


Appendix A: Changes made to the ASSA2002 model

The ASSA2002 model that is currently publicly available is described by Johnson and Dorrington (2005). For the purpose of this uncertainty analysis, a number of changes have been made to the model, which are detailed below. For the sake of convenience, the version of the ASSA2002 model used in the uncertainty analysis is referred to as the C++ model, and the version of the ASSA2002 model that is currently publicly available is referred to as the Excel model.

Firstly, a sexual mixing parameter was introduced. This mixing parameter is used to determine the proportions of female partners in each risk group, as described in section 3.3.4.3. These proportions of female partners in each risk group are then used to determine the proportions of male partners in each risk group, as described in Appendix C. In the Excel model, assumptions are made about the proportions of male partners in each risk group, and these are then used to determine the proportions of female partners in each risk group. As a result of this difference in approach, the method used to determine the average annual number of partners in each male risk group in the C++ model also differs from that used in the Excel model. The method used in the C++ model to determine the average annual number of partners in the male risk groups is described in more detail in Appendix C.

The second change made to the C++ model was to allow for a reduction in the average annual number of partners, as a result of social marketing programmes. In the Excel model, it is assumed that the only effect of social marketing programmes is to increase levels of condom usage, while in the C++ model it is assumed that social marketing programmes lead to changes in both condom usage and average annual numbers of partners. The allowance for the latter is described in more detail in section 3.3.4.2.

The Excel and C++ models differ substantially in terms of the procedure used to calibrate the model to vital registration data. In the Excel model, model estimates of total deaths are compared with total deaths, corrected for incompleteness of reporting, estimated by Dorrington et al. (2004) using recorded deaths of individuals on the population register (i.e. individuals with IDs). In the C++ model, model estimates of reported deaths are compared with the reported deaths as estimated by Statistics South Africa, using death notification forms submitted to the Department of Home Affairs. The estimates produced by Statistics South Africa are more appropriate for this analysis, as no assumptions are required about ID coverage, and it is possible to allow explicitly for uncertainty regarding completeness of reporting.

The Excel and C++ models also differ in terms of the method used to adjust model estimates of HIV prevalence in pregnant women, to correspond with the rates of prevalence observed in public antenatal clinics. The Excel model does not make explicit allowance for false positive test results when comparing model estimates of antenatal clinic prevalence with reported antenatal prevalence rates. The adjustments to allow for
bias towards urban clinics in the early antenatal surveys (described in section 3.3.1.2) and bias towards the public health sector (described in section 3.3.1.3) also differ.

Lastly, an error in the Excel model was corrected in the C++ model. The error in the Excel model occurs in the ‘MaleSTD’ sheet, in column AT. All references to ‘MaleRSK’ in this column should be replaced with ‘FemRSK’. Correction of this error does not significantly change the model results.
## Appendix B: Review of ELISA sensitivities and specificities in African populations

<table>
<thead>
<tr>
<th>Reference</th>
<th>Test</th>
<th>Se (%)</th>
<th>Sp (%)</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akanmu et al (2001)</td>
<td>Murex HIV-1 + 2</td>
<td>100</td>
<td>98.9</td>
<td>Haematology clinic outpatients</td>
</tr>
<tr>
<td></td>
<td>Enzygnost anti HIV-1/HIV-2 plus</td>
<td>100</td>
<td>99.8</td>
<td>Pregnant women, sex workers, TB patients</td>
</tr>
<tr>
<td></td>
<td>Murex HIV-1 + 2</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICE HIV-1.0.2</td>
<td>100</td>
<td>99.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vironostika HIV Uni-form II plus O</td>
<td>100</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Innotest HIV-1/HIV-2 Ab s.p.</td>
<td>100</td>
<td>98.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genelavia Mixt</td>
<td>100</td>
<td>96.8</td>
<td></td>
</tr>
<tr>
<td>Nkengasong et al (1999)</td>
<td>Genelavia Mixt</td>
<td>100</td>
<td>93.4</td>
<td>Blood donors, hospital patients, pregnant women</td>
</tr>
<tr>
<td></td>
<td>Enzygnost anti HIV-1/HIV-2 plus</td>
<td>100</td>
<td>99.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Murex HIV-1 + 2</td>
<td>100</td>
<td>96.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vironostika HIV Uni-form II</td>
<td>100</td>
<td>81.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICE HIV-1.0.2</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Meda et al (1999)</td>
<td>Genelavia Mixt</td>
<td>100</td>
<td>95.4</td>
<td>General population</td>
</tr>
<tr>
<td></td>
<td>Enzygnost anti HIV-1/HIV-2 plus</td>
<td>100</td>
<td>99.8</td>
<td>Hospital patients</td>
</tr>
<tr>
<td></td>
<td>Murex HIV-1 + 2</td>
<td>100</td>
<td>96.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vironostika HIV Uni-form II</td>
<td>100</td>
<td>81.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICE HIV-1.0.2</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vironostika</td>
<td>100</td>
<td>84.4</td>
<td>Hospital patients</td>
</tr>
<tr>
<td></td>
<td>ICE HIV-1.0.2</td>
<td>94.4</td>
<td>89.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biotest anti-HIV-1/2</td>
<td>96.3</td>
<td>67.5</td>
<td></td>
</tr>
<tr>
<td>Andersson et al (1997)</td>
<td>Enzygnost anti HIV-1/HIV-2 plus</td>
<td>100</td>
<td>98</td>
<td>People referred for HIV testing, hospital patients, occupational cohort</td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV-1+2</td>
<td>99.2</td>
<td>99.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genelavia Mixt</td>
<td>96.1</td>
<td>96.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vironostika HIV Uni-form II</td>
<td>95.3</td>
<td>99.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV 1+2</td>
<td>96.1</td>
<td>98.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genelavia Mixt</td>
<td>100</td>
<td>99.5</td>
<td>Hospital patients</td>
</tr>
<tr>
<td></td>
<td>Vironostika HIV Uni-form II</td>
<td>100</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV 1+2</td>
<td>100</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genelavia Mixt</td>
<td>100</td>
<td>99.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biotest anti-HIV-1/2</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV 1+2</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV 1+2</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV-1 recombinant</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Tswana et al (1993)</td>
<td>Abbott recombinant HIV-1/HIV-2 (2nd generation)</td>
<td>90.6</td>
<td>94.1</td>
<td>Patients with symptoms</td>
</tr>
<tr>
<td></td>
<td>Wellcozyme</td>
<td>96.8</td>
<td>53.3</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Test Description</td>
<td>Sensitivity (S)</td>
<td>Specificity (Sp)</td>
<td>Class</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Nunn et al (1993)</td>
<td>Biotest</td>
<td>93.2</td>
<td>89.1</td>
<td>suggestive of HIV/AIDS</td>
</tr>
<tr>
<td></td>
<td>Du Pont HIV-1 recombinant</td>
<td>89.1</td>
<td>81.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recombigen HIV-1</td>
<td>99.5</td>
<td>98.3</td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV-1 recombinant</td>
<td>98.9</td>
<td>98.4</td>
<td>population</td>
</tr>
<tr>
<td></td>
<td>Abbott recombinant anti-HIV-1</td>
<td>99.5</td>
<td>99.2</td>
<td>Blood donors,</td>
</tr>
<tr>
<td></td>
<td>Enzygnost anti HIV-1/HIV-2</td>
<td>99.5</td>
<td>97.0</td>
<td>hospital patients</td>
</tr>
<tr>
<td></td>
<td>Vironostika</td>
<td>97.7</td>
<td>98.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellcozyme HIV-1 recombinant</td>
<td>100</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vironostika HIV Micro ELISA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAV</td>
<td>89.1</td>
<td>96.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>90.1</td>
<td>72.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>84.4</td>
<td>99.1</td>
<td></td>
</tr>
<tr>
<td>Spielberg et al (1989)</td>
<td>Blood donors</td>
<td>98.6</td>
<td>98.1</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>97.4</td>
<td>94.8</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>99.5</td>
<td>98.5</td>
<td></td>
</tr>
</tbody>
</table>

Table B.1: Estimates of the sensitivity and specificity of the ELISA in detecting HIV (blood specimens)
Appendix C: Calculation of male rates of partner change and proportions of male partners in different risk groups

The approach taken to calculating the proportion of partners in different risk groups and annual average numbers of partners differs from that in the publicly available ASSA2002 model. This appendix explains the approach used in the uncertainty analysis, and concludes by contrasting this with the approach used in the ASSA2002 model that is publicly available.

Symbols are defined as follows:

\( \Psi_j \) is the proportion of sexually experienced females (aged 14-59) who are in risk group \( j \) at the start of the epidemic

\( \Psi_j^* \) is the proportion of sexually experienced males (aged 14-59) who are in risk group \( j \) at the start of the epidemic

\( P_j \) is the average annual number of partnerships for a sexually experienced female in risk group \( j \) at the start of the epidemic

\( P_j^* \) is the average annual number of partnerships for a sexually experienced male in risk group \( j \) at the start of the epidemic

\( w_{ij} \) is the proportion of male partners who are in risk group \( j \), for a woman who is in risk group \( i \)

\( w_{ij}^* \) is the proportion of female partners who are in risk group \( j \), for a man who is in risk group \( i \).

Parameters \( \Psi_j \), \( \Psi_j^* \) and \( P_j \) are assumed, for each value of \( j \) from 1 to 3, and the \( w_{ij}^* \) parameters can be calculated once \( \epsilon \) has been specified, as per equation (13) in section 3.3.4.3. (The equation is modified to exclude the possibility of interactions between the PRO and RSK groups and interactions between the NOT and other risk groups.)

To calculate the \( P_j^* \) parameters, it is assumed that the following relationships must hold.

\[
P_1 = \left( w_{11}^* \Psi_1^* P_1^* + w_{12}^* \Psi_2^* P_2^* \right) / \Psi_1^* \quad \text{(C1)}
\]

\[
P_2 = \left( w_{21}^* \Psi_1^* P_1^* + w_{22}^* \Psi_2^* P_2^* + w_{23}^* \Psi_3^* P_3^* \right) / \Psi_2^* \quad \text{(C2)}
\]

\[
P_3 = \left( w_{32}^* \Psi_2^* P_2^* + w_{33}^* \Psi_3^* P_3^* \right) / \Psi_3^* \quad \text{(C3)}
\]

Equation (C1), for example, ensures that the total number of partnerships that women in the PRO group have is equal to the total number of partnerships that men have with women in the PRO group (assuming that the number of sexually active men is approximately equal to the number of sexually active women). Similar constraints are imposed for the STD and RSK groups in equations (C2) and (C3) respectively.
From equation (C1), it follows that

\[ P_1^* = \frac{\Psi_1 P_1 - w_{12}^* \Psi_2^* P_2^*}{w_{11}^* \Psi_1^*}, \quad (C4) \]

and from equation (C3) it follows that

\[ P_3^* = \frac{\Psi_1 P_3 - w_{32}^* \Psi_2^* P_2^*}{w_{33}^* \Psi_3^*}. \quad (C5) \]

Substituting equations (C4) and (C5) into equation (C2), it can be shown that

\[ P_2^* = \frac{w_{11}^* \Psi_2^* P_2 - w_{12}^* w_{33}^* \Psi_1 P_1 - w_{11}^* w_{32}^* \Psi_3 P_3}{w_{11}^* \Psi_2^* \Psi_3 - w_{12}^* w_{33}^* - w_{11}^* w_{32}^*}. \quad (C6) \]

This equation is expressed in terms of the values previously calculated, and hence can be used to calculate \( P_2^* \). \( P_1^* \) and \( P_3^* \) can then be calculated using equations (C4) and (C5).

For most of the parameter combinations generated, the following constraints were met:

\[ P_1^* \geq P_2^* \geq P_3^* \geq 0. \quad (C7) \]

However, in rare instances, these constraints were not met, and this led to negative estimates and error messages. Parameter combinations not satisfying these constraints were therefore excluded from the uncertainty analysis, and new parameter combinations were generated in their place. In an initial sample of 10,000 parameter combinations, only two parameter combinations failed to meet the constraints.

Once the average annual numbers of partnerships have been calculated for sexually experienced males, it is possible to calculate the \( w_{ij} \) parameters. To ensure that the total number of partnerships women in the PRO group have with men in the PRO group is equal to the total number of partnerships men in the PRO group have with women in the PRO group, it is necessary to assume that

\[ w_{11}^* \Psi_1 P_1 = w_{11}^* \Psi_1^* P_1^*. \quad (C8) \]

(As before, it is assumed that the number of sexually active men is approximately equal to the number of sexually active women.) Hence

\[ w_{11} = w_{11}^* \Psi_1^* P_1^*/(\Psi_1 P_1), \quad (C9) \]

and similarly
\[ w_{21} = w_{12}^* \Psi_{11}^* P_{11}^* / (\Psi_{22}^* P_{22}^*) \]
\[ w_{22} = w_{22}^* \Psi_{22}^* P_{22}^* / (\Psi_{22}^* P_{22}^*) \]  \hspace{1cm} (C10)
\[ w_{33} = w_{33}^* \Psi_{33}^* P_{33}^* / (\Psi_{33}^* P_{33}^*) . \]

The remaining \( w_{ij} \) parameters are calculated to satisfy the constraint that the proportions must sum to 1:
\[ w_{12} = 1 - w_{11} \]
\[ w_{23} = 1 - w_{22} - w_{21} \]  \hspace{1cm} (C11)
\[ w_{32} = 1 - w_{33} . \]

In the ASSA2002 model and in the approach described here, all calculations are performed using the proportion of the sexually active population initially in each risk group, and the calculations are not updated over time, as the relative sizes of the risk groups change. A modified version of the ASSA2002 model, in which the proportions were allowed to vary over time, produced similar results to the default version of the model.

In the version of the ASSA2002 model that is publicly available, the approach is also to start with assumed values of \( \Psi_{ij}^*, \Psi_{ij}^* \) and \( P_{ij}^* \), for each value of \( j \) from 1 to 3. However, instead of specifying a value of \( \epsilon \) to determine the \( w_{ij}^* \) parameters, assumptions are made about the \( w_{ij} \) parameters. Using equations similar to (C1) to (C3), with the male-specific and female-specific parameters interchanged, the average annual numbers of partners in the male risk groups can then be estimated. The \( w_{ij}^* \) parameters are then determined using equations (C9) and (C10).
Appendix D: Approach to setting $N_u$ parameters

Consider the method used to define the likelihood function for the antenatal clinic prevalence data, as described in section 3.4.1. Suppose that $X_{(a,t)}$ is a random variable representing the prevalence of HIV in a sample of $n_{(a,t)}$ pregnant women, with age in the interval $[a, a + 5)$, in year $t$. As in section 3.4.1, define $M_{(a,t)}$ as the prevalence predicted by the model in the corresponding year and age group. $M_{(a,t)}$ is itself a random variable, since different values of this variable are generated for each of the randomly selected parameter combinations. The distribution of $M_{(a,t)}$ can be approximated by sampling from the posterior distribution of parameter combinations and calculating the $M_{(a,t)}$ values corresponding to the sampled parameter combinations.

If the model reflects an acceptable degree of uncertainty, one would expect that approximately 95% of the observed $\{X_{(a,t)}\}$ values should lie within the 95% prediction intervals for the corresponding $X_{(a,t)}$ values. It is therefore necessary to set the $N_{\text{ANC}}$ parameter at such a level that this criterion is achieved. However, prediction intervals for $X_{(a,t)}$ cannot be calculated precisely, and it is necessary to derive an approximate adjustment to the $M_{(a,t)}$ prediction intervals for the purpose of obtaining these prediction intervals. This adjustment is obtained by noting that
\[
Var[X_{(a,t)}] = Var[E[X_{(a,t)} | M_{(a,t)}]] + E[Var[X_{(a,t)} | M_{(a,t)}]]
\]
\[
= Var[M_{(a,t)}] + E[Var[X_{(a,t)} | M_{(a,t)}]].
\]

The second term in equation (D1) is the expected variation due to the limited sample size, or random variation, which tends to zero the larger $n_{(a,t)}$ is. Although this can be approximated by assuming that the number of women testing positive in the survey is binomially distributed, this would result in an under-estimate of the true variance, since the true variance should be calculated on the assumption of cluster sampling. In the absence of the full data set and information regarding the weights used in deriving the prevalence estimates, the true variance cannot be calculated accurately. However, it can be approximated from the width of the reported confidence intervals, if it is assumed that these confidence intervals were calculated using the simple asymptotic method. Although the simple asymptotic method was not used to calculate the confidence interval widths, the width of the confidence interval calculated using this method is similar to that obtained using other methods (Newcombe 1998), and hence it serves as a reasonable approximation. $E[Var[X_{(a,t)} | M_{(a,t)}]]$ is thus approximated as
\[
\left( \frac{x^u_i(a,t) - x^l_i(a,t)}{2 \times 1.96} \right)^2,
\]

where \( x^u_i(a,t) \) and \( x^l_i(a,t) \) are the upper and lower limits, respectively, of the reported 95% confidence interval around \( x_i(a,t) \).

To obtain the 95% prediction intervals for \( X_i(a,t) \), the 95% prediction intervals for \( M_i(a,t) \) are widened by the difference between the prediction interval widths that would be expected if both were calculated using the simple asymptotic method. Hence the upper limit on the 95% prediction interval for \( X_i(a,t) \) is calculated as

\[
M^u_i(a,t) + 1.96 \times \left[ \sqrt{\text{Var}[X_i(a,t)]} - \sqrt{\text{Var}[M_i(a,t)]} \right],
\]

where \( M^u_i(a,t) \) is the upper limit of the 95% prediction interval for \( M_i(a,t) \), or the 97.5 percentile of the posterior sample of \( M_i(a,t) \) values. A similar formula is used to calculate the lower limit of the prediction interval.

The \( N_{ANC} \) parameter was set for the antenatal clinic data independently of other data sources. The \( N_{HSRC}, N_{HRHU} \) and \( N_{DH} \) parameters were therefore set to zero for the purpose of determining an appropriate value for \( N_{ANC} \), giving zero weight to the other data sources. Using a sample of 500 parameter combinations, drawn from an initial set of 10 000 parameter combinations, it was found that an \( N_{ANC} \) parameter of 0.01 gave prediction intervals of appropriate width. Figure D.1 shows the prediction intervals for this and two other \( N_{ANC} \) values. As shown in Table D.1, 1.4% (1/70) of antenatal measurements lie outside of the prediction intervals for \( X_i(a,t) \), when \( N_{ANC} \) is set to 0.01. Values of \( N_{ANC} \) substantially greater than 0.01 would result in prediction intervals that are too narrow, while values substantially lower than 0.01 would result in prediction intervals that are too wide.
% of observations outside 95% intervals

<table>
<thead>
<tr>
<th>Value of $N_u$</th>
<th>ANC</th>
<th>DHA</th>
<th>HSRC, RHRU</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0%</td>
<td>2.4%</td>
<td>4.8%*</td>
</tr>
<tr>
<td>0.00005</td>
<td>-</td>
<td>4.8%*</td>
<td>-</td>
</tr>
<tr>
<td>0.0002</td>
<td>-</td>
<td>9.5%†</td>
<td>-</td>
</tr>
<tr>
<td>0.001</td>
<td>-</td>
<td>39.3%‡</td>
<td>4.8%‡</td>
</tr>
<tr>
<td>0.002</td>
<td>0.0%*</td>
<td>-</td>
<td>11.9%</td>
</tr>
<tr>
<td>0.005</td>
<td>-</td>
<td>-</td>
<td>14.3%‡</td>
</tr>
<tr>
<td>0.01</td>
<td>1.4%†</td>
<td>-</td>
<td>16.7%</td>
</tr>
<tr>
<td>0.05</td>
<td>18.6%‡</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table D.1: Proportions of observations outside of 95% prediction intervals for selected $N_u$ values

* Represented by light gray lines in figures below. † Represented by dashed lines in figures below. ‡ Represented by solid black lines in figures below. Proportions of observations outside 95% prediction intervals are calculated after 95% prediction intervals are adjusted for sample variation.
Figure D.1: Levels of HIV prevalence in pregnant women, observed and predicted

Observed values are represented by crosses (×). 95% prediction intervals for $X(a,t)$ represented by light gray lines ($N_{ANC} = 0.002$), dashed lines ($N_{ANC} = 0.01$) and solid black lines ($N_{ANC} = 0.05$).

In setting $N_{HSRC}$ and $N_{RHRU}$, the approach followed was identical to that followed in setting $N_{ANC}$. As before, $X_3(s,a,t)$ and $X_4(s,a)$ are defined as the random variables corresponding to observations $x_3(s,a,t)$ and $x_4(s,a)$ respectively. As both surveys were household surveys, and the number of data points in the RHRU survey was small, the two data sets were combined. The $N_{HSRC}$ and $N_{RHRU}$ parameters were therefore set at the same level, chosen to be such that roughly 95% of the HSRC and RHRU prevalence estimates lay within the corresponding 95% prediction intervals for $X_3(s,a,t)$ and $X_4(s,a)$, when $N_{ANC}$ and $N_{DHA}$ were both set to zero. Prediction intervals were found to be of appropriate width when $N_{HSRC}$ and $N_{RHRU}$ were both set to 0.001, with 4.8% (2/42) of observations lying outside of these intervals (see Table D.1). Figure D.2 and Table D.2 show the 95% prediction intervals for this and two other values of $N_{HSRC}$ and $N_{RHRU}$. At these low levels of $N_{HSRC}$ and $N_{RHRU}$, changes to these parameters make little difference to the upper limits of the 95% prediction intervals, but have a more noticeable effect on the lower limits of the 95% prediction intervals.
Figure D.2: Levels of HIV prevalence in HSRC surveys, observed and predicted
Observed values are represented by crosses (×). 95% prediction intervals for $X_{s,a,t}$ represented by light gray lines ($N_{HSRC} = N_{RHRU} = 0$), dashed lines ($N_{HSRC} = N_{RHRU} = 0.001$) and solid black lines ($N_{HSRC} = N_{RHRU} = 0.005$).

<table>
<thead>
<tr>
<th></th>
<th>Observed prevalence</th>
<th>95% interval $N_{RHRU} = 0$</th>
<th>95% interval $N_{RHRU} = 0.001$</th>
<th>95% interval $N_{RHRU} = 0.005$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males 15-19</td>
<td>2.5%</td>
<td>0.2-8.2%</td>
<td>0.3-7.1%</td>
<td>0.5-6.0%</td>
</tr>
<tr>
<td>20-24</td>
<td>7.6%</td>
<td>2.1-16.5%</td>
<td>4.2-16.6%</td>
<td>6.3-16.2%</td>
</tr>
<tr>
<td>Females 15-19</td>
<td>7.3%</td>
<td>2.2-15.0%</td>
<td>4.2-14.7%</td>
<td>5.7-13.5%</td>
</tr>
<tr>
<td>20-24</td>
<td>24.5%</td>
<td>7.1-35.3%</td>
<td>13.7-36.7%</td>
<td>17.5-34.0%</td>
</tr>
</tbody>
</table>

Table D.2: Levels of HIV prevalence in the RHRU/loveLife household survey, observed and predicted

In setting $N_{DHA}$, a slightly different approach was adopted to that used in determining $N_{ANC}$. Replacing the parameters in equation (D1) with the relevant DHA parameters,

$$Var[X_{s,a,t}] = Var[M_{s,a,t}] + E[Var[X_{s,a,t} | M_{s,a,t}]], \quad (D4)$$

where $X_{s,a,t}$ is the random variable corresponding to $x_{s,a,t}$. As before, $Var[M_{s,a,t}]$ is approximated using the sample variance of the 500 sampled values of $M_{s,a,t}$. Using the normal approximation to the Poisson distribution, $Var[X_{s,a,t} | M_{s,a,t}]$ is estimated to be $M_{s,a,t}$, as described in section 3.4.2. The last term in equation (D4) is thus $E[M_{s,a,t}]$, which can be approximated by the sample mean of the 500 sampled values of $M_{s,a,t}$. After obtaining $Var[X_{s,a,t}]$,
the approach to determining the width of the 95% prediction intervals for \( X_2(s,a,t) \) is the same as for the other data sources.

As before, the \( N_{DHA} \) parameter is set independently of the other data sources. \( N_{ANC} \), \( N_{HSRC} \), and \( N_{RHRU} \) are therefore set to zero for the purpose of determining \( N_{DHA} \). Figure D.3 and Figure D.4 show the 95% prediction intervals for \( X_2(s,a,t) \) for three different \( N_{DHA} \) parameter values, and the corresponding proportions of observations lying outside of the 95% prediction intervals are shown in Table D.1. Although a value of 0.00005 might appear to be most appropriate for \( N_{DHA} \), a value of 0.0002 was chosen, as most of the observations lying outside of the 95% prediction intervals appear to be the result of mis-specified non-AIDS mortality rates in the ASSA2002 model (see section 6.4), rather than inappropriate AIDS mortality rates. This results in 9.5% (8/84) of reported death estimates lying outside of the 95% prediction intervals for \( X_2(s,a,t) \).

The approach described in this appendix has two important features. Firstly, the values of \( N_u \) are not dependent on the other sources of data to which the model is being calibrated. Secondly, the prediction intervals are adjusted to allow for random variation, which ensures that the objective of obtaining 95% of observations within the 95% prediction intervals is statistically valid. The adjustments to allow for random variation are based on several approximations, and the values chosen for the \( N_u \) factors are therefore appropriate only in very crude terms. The allowance for random variation makes a significant difference to the width of the prediction intervals around the HSRC prevalence estimates, due to the relatively small samples on which these prevalence estimates are based. It makes little difference, though, to the width of the prediction intervals around the DHA data, as random variation is in this case extremely small in relation to model variation.
Figure D.3: Numbers of reported male deaths, predicted and actual
Observed values are represented by crosses (×). 95% prediction intervals for \( X(s,a,t) \) represented by light gray lines (\( N_{DHA} = 0.00005 \)), dashed lines (\( N_{DHA} = 0.0002 \)) and solid black lines (\( N_{DHA} = 0.001 \)).
Figure D.4: Numbers of reported female deaths, predicted and actual
Observed values are represented by crosses (×). 95% prediction intervals for $X_2(s,a,t)$ represented by light gray lines ($N_{DHA} = 0.00005$), dashed lines ($N_{DHA} = 0.0002$) and solid black lines ($N_{DHA} = 0.001$).
## Appendix E: Antenatal clinic prevalence data

<table>
<thead>
<tr>
<th>Year</th>
<th>&lt;20 (95% CI)</th>
<th>20-24 (95% CI)</th>
<th>25-29 (95% CI)</th>
<th>30-34 (95% CI)</th>
<th>35-39 (95% CI)</th>
<th>40-44 (95% CI)</th>
<th>45-49 (95% CI)</th>
<th>Total (95% CI)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>1.79 (0.83-2.33)</td>
<td>2.15 (1.47-2.47)</td>
<td>1.37 (0.82-1.73)</td>
<td>0.72 (0.24-0.97)</td>
<td>0.39 (0-0.70)</td>
<td>0.95 (0-2.60)</td>
<td>0 (0-6.44)</td>
<td>0.76 (0.57-0.96)</td>
<td>DNHPD (1991)</td>
</tr>
<tr>
<td>1991</td>
<td>2.62 (1.62-3.63)</td>
<td>3.92 (3.07-4.77)</td>
<td>2.11 (1.42-2.79)</td>
<td>2.04 (1.07-3.02)</td>
<td>1.98 (0.65-3.30)</td>
<td>0 (0-1.00)</td>
<td>0 (0-7.39)</td>
<td>1.35 (1.11-1.59)</td>
<td>DNHPD (1992)</td>
</tr>
<tr>
<td>1992</td>
<td>4.57 (3.3-5.84)</td>
<td>6.06 (4.99-7.13)</td>
<td>5.22 (4.1-6.35)</td>
<td>3.05 (1.97-4.13)</td>
<td>1.76 (0.53-2.99)</td>
<td>5.89 (0-5.89)</td>
<td>0 (0-7.39)</td>
<td>2.42 (2.08-2.76)</td>
<td>DNHPD (1993)</td>
</tr>
<tr>
<td>1993</td>
<td>6.47 (5.45-7.49)</td>
<td>8.94 (7.02-9.86)</td>
<td>8.63 (7.25-10.01)</td>
<td>6.37 (5.28-7.45)</td>
<td>5.32 (4.24-5.32)</td>
<td>0 (0-5.89)</td>
<td>0 (0-7.39)</td>
<td>4.25 (3.80-4.71)</td>
<td>DNHPD (1994)</td>
</tr>
<tr>
<td>1994</td>
<td>9.50 (8.15-10.85)</td>
<td>13.12 (11.97-14.27)</td>
<td>11.03 (9.81-12.25)</td>
<td>8.05 (7.31-9.31)</td>
<td>7.37 (6.24-7.89)</td>
<td>0 (0-1.00)</td>
<td>0 (0-1.00)</td>
<td>10.44 (9.86-11.02)</td>
<td>DNHPD (1995)</td>
</tr>
<tr>
<td>1996</td>
<td>12.7 (11.3-14.2)</td>
<td>19.7 (18.4-21.1)</td>
<td>18.2 (16.8-19.6)</td>
<td>14.5 (12.9-16.2)</td>
<td>9.5 (7.7-11.5)</td>
<td>7.5 (4.4-11.8)</td>
<td>8.8 (1.9-23.7)</td>
<td>17.0 (14.5-19.5)</td>
<td>DNHPD (1997)</td>
</tr>
<tr>
<td>1999</td>
<td>16.1 (14.5-17.7)</td>
<td>29.1 (27.4-30.8)</td>
<td>30.6 (28.8-32.4)</td>
<td>23.3 (21.5-25.1)</td>
<td>18.3 (13.9-17.7)</td>
<td>15.8 (8.9-13.3)</td>
<td>10.2 (2.1-24.0)</td>
<td>24.5 (23.4-25.6)</td>
<td>DNHPD (2000)</td>
</tr>
<tr>
<td>2000</td>
<td>15.4 (13.8-16.9)</td>
<td>28.4 (26.5-30.2)</td>
<td>31.4 (29.5-33.3)</td>
<td>25.6 (23.5-27.7)</td>
<td>19.3 (13.9-17.7)</td>
<td>18.3 (8.5-15.6)</td>
<td>10.2 (5.4-15.9)</td>
<td>24.8 (23.4-25.6)</td>
<td>DNHPD (2001)</td>
</tr>
<tr>
<td>2001</td>
<td>14.8 (13.4-16.1)</td>
<td>29.1 (27.5-30.6)</td>
<td>34.5 (32.6-36.4)</td>
<td>29.5 (23.5-27.7)</td>
<td>17.2 (11.2-15.6)</td>
<td>19.3 (6.2-11.9)</td>
<td>9.1 (4.3-13.4)</td>
<td>26.5 (23.6-26.1)</td>
<td>DNHPD (2002)</td>
</tr>
<tr>
<td>2002</td>
<td>15.8 (14.3-17.2)</td>
<td>30.3 (28.8-31.8)</td>
<td>35.4 (33.6-37.2)</td>
<td>30.9 (28.9-32.9)</td>
<td>23.4 (17.5-22.5)</td>
<td>23.4 (17.5-22.5)</td>
<td>23.4 (17.5-22.5)</td>
<td>26.5 (25.5-27.6)</td>
<td>DNHPD (2003)</td>
</tr>
<tr>
<td>2003</td>
<td>16.1 (14.7-17.5)</td>
<td>30.8 (29.3-32.3)</td>
<td>38.5 (36.8-40.3)</td>
<td>34.4 (32.2-36.6)</td>
<td>24.5 (21.9-27.2)</td>
<td>24.5 (21.9-27.2)</td>
<td>24.5 (21.9-27.2)</td>
<td>29.5 (26.8-28.9)</td>
<td>DNHPD (2004)</td>
</tr>
<tr>
<td>2004</td>
<td>16.1 (14.7-17.5)</td>
<td>30.8 (29.3-32.3)</td>
<td>38.5 (36.8-40.3)</td>
<td>34.4 (32.2-36.6)</td>
<td>24.5 (21.9-27.2)</td>
<td>24.5 (21.9-27.2)</td>
<td>24.5 (21.9-27.2)</td>
<td>29.5 (26.8-28.9)</td>
<td>DNHPD (2005)</td>
</tr>
</tbody>
</table>

Table E.1: South African antenatal prevalence statistics by age and year (95% confidence intervals in brackets)
It would appear that before the 1998 survey, confidence intervals were calculated on the assumption of random sampling (Department of Health 1997), and did not take into account clustering. Cluster sampling appears to have been used thereafter (Department of Health 1999), which may partially explain the increase in the width of confidence intervals after 1997.

<table>
<thead>
<tr>
<th>Year</th>
<th>&lt;20</th>
<th>20-24</th>
<th>25-29</th>
<th>30-34</th>
<th>35-39</th>
<th>40-44</th>
<th>45-49</th>
<th>No age</th>
<th>Total</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>2,063</td>
<td>3,631</td>
<td>2,941</td>
<td>1,690</td>
<td>722</td>
<td>160</td>
<td>34</td>
<td>3,135</td>
<td>14,376</td>
<td>DNHPD (1991)</td>
</tr>
<tr>
<td>1991</td>
<td>3,045</td>
<td>4,850</td>
<td>3,799</td>
<td>2,411</td>
<td>1,159</td>
<td>270</td>
<td>45</td>
<td>1,576</td>
<td>17,155</td>
<td>DNHPD (1992)</td>
</tr>
<tr>
<td>1992</td>
<td>3,414</td>
<td>5,592</td>
<td>4,049</td>
<td>2,678</td>
<td>1,203</td>
<td>399</td>
<td>59</td>
<td>1,819</td>
<td>19,213</td>
<td>DNHPD (1993)</td>
</tr>
<tr>
<td>1993</td>
<td>2,736</td>
<td>4,290</td>
<td>3,313</td>
<td>2,209</td>
<td>1,013</td>
<td>274</td>
<td>67</td>
<td>2,284</td>
<td>16,186</td>
<td>DOH (1994)</td>
</tr>
<tr>
<td>1995</td>
<td>2,096</td>
<td>3,397</td>
<td>2,745</td>
<td>1,859</td>
<td>922</td>
<td>276</td>
<td>45</td>
<td>1,386</td>
<td>12,726</td>
<td>DOH (1996)</td>
</tr>
<tr>
<td>1996</td>
<td>2,151</td>
<td>3,491</td>
<td>2,822</td>
<td>1,861</td>
<td>955</td>
<td>241</td>
<td>37</td>
<td>3,486</td>
<td>15,044</td>
<td>DOH (1997)</td>
</tr>
<tr>
<td>1997</td>
<td>2,109</td>
<td>3,591</td>
<td>2,950</td>
<td>1,853</td>
<td>994</td>
<td>226</td>
<td>34</td>
<td>678</td>
<td>12,435</td>
<td>DOH, unpublished</td>
</tr>
<tr>
<td>1998</td>
<td>2,846</td>
<td>4,470</td>
<td>3,731</td>
<td>2,280</td>
<td>1,265</td>
<td>278</td>
<td>38</td>
<td>181</td>
<td>15,089</td>
<td>DOH, unpublished</td>
</tr>
<tr>
<td>1999</td>
<td>2,994</td>
<td>4,929</td>
<td>4,346</td>
<td>2,689</td>
<td>1,377</td>
<td>325</td>
<td>36</td>
<td>154</td>
<td>16,850</td>
<td>DOH, unpublished</td>
</tr>
<tr>
<td>2000</td>
<td>2,898</td>
<td>4,905</td>
<td>4,083</td>
<td>2,749</td>
<td>1,403</td>
<td>386</td>
<td>38</td>
<td>96</td>
<td>16,558</td>
<td>DOH, unpublished</td>
</tr>
</tbody>
</table>

Table E.2: Antenatal survey sample sizes by age and year
Appendix F: Nelson Mandela/HSRC household prevalence survey data

<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th></th>
<th></th>
<th>Females</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n†</td>
<td>HIV %</td>
<td>95% CI</td>
<td>n†</td>
<td>HIV %</td>
<td>95% CI</td>
</tr>
<tr>
<td>15 to 19</td>
<td>554</td>
<td>4.4%</td>
<td>2.0-7.5%*</td>
<td>589</td>
<td>7.3%</td>
<td>4.7-11.3%</td>
</tr>
<tr>
<td>20 to 24</td>
<td>422</td>
<td>8.3%</td>
<td>5.0-13.0%*</td>
<td>534</td>
<td>17.1%</td>
<td>12.9-22.3%</td>
</tr>
<tr>
<td>25 to 29</td>
<td>237</td>
<td>22.0%</td>
<td>14.5-32.0%*</td>
<td>350</td>
<td>32.0%</td>
<td>24.9-40.1%</td>
</tr>
<tr>
<td>30 to 34</td>
<td>235</td>
<td>24.1%</td>
<td>16.5-34.0%*</td>
<td>324</td>
<td>24.1%</td>
<td>17.3-32.5%</td>
</tr>
<tr>
<td>35 to 39</td>
<td>225</td>
<td>18.4%</td>
<td>11.0-27.5%*</td>
<td>327</td>
<td>13.8%</td>
<td>8.7-21.1%</td>
</tr>
<tr>
<td>40 to 44</td>
<td>205</td>
<td>12.4%</td>
<td>7.0-19.5%*</td>
<td>335</td>
<td>19.0%</td>
<td>12.9-27.2%</td>
</tr>
<tr>
<td>45 to 49</td>
<td>200</td>
<td>11.9%</td>
<td>7.0-19.5%*</td>
<td>258</td>
<td>11.2%</td>
<td>6.5-18.7%</td>
</tr>
<tr>
<td>50 to 54</td>
<td>161</td>
<td>5.4%</td>
<td>2.5-10.5%*</td>
<td>197</td>
<td>8.5%</td>
<td>4.5-14.5%*</td>
</tr>
<tr>
<td>55+</td>
<td>554</td>
<td>7.0%</td>
<td>2.5-16.0%*</td>
<td>589</td>
<td>6.6%</td>
<td>4.0-10.5%*</td>
</tr>
</tbody>
</table>

Table F.1: Household prevalence statistics and sample sizes by age and sex, 2002
* Values approximated from graph (Figure 11, p. 52). † Data provided by Thomas Rehle (Director, Social Aspects of HIV/AIDS and Health Programme, HSRC).

<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th></th>
<th></th>
<th>Females</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>HIV %</td>
<td>95% CI</td>
<td>n</td>
<td>HIV %</td>
<td>95% CI</td>
</tr>
<tr>
<td>15 to 19</td>
<td>1001</td>
<td>3.2%</td>
<td>1.4-7.1%</td>
<td>1153</td>
<td>9.4%</td>
<td>7.1-12.4%</td>
</tr>
<tr>
<td>20 to 24</td>
<td>784</td>
<td>6.0%</td>
<td>3.8-9.4%</td>
<td>1182</td>
<td>23.9%</td>
<td>19.8-28.4%</td>
</tr>
<tr>
<td>25 to 29</td>
<td>383</td>
<td>12.1%</td>
<td>8.0-17.9%</td>
<td>598</td>
<td>33.3%</td>
<td>27.7-39.4%</td>
</tr>
<tr>
<td>30 to 34</td>
<td>341</td>
<td>23.3%</td>
<td>17.2-30.7%</td>
<td>691</td>
<td>26.0%</td>
<td>21.5-30.9%</td>
</tr>
<tr>
<td>35 to 39</td>
<td>375</td>
<td>23.3%</td>
<td>17.8-29.8%</td>
<td>727</td>
<td>19.3%</td>
<td>14.9-24.6%</td>
</tr>
<tr>
<td>40 to 44</td>
<td>399</td>
<td>17.5%</td>
<td>12.0-24.7%</td>
<td>694</td>
<td>12.4%</td>
<td>9.4-16.2%</td>
</tr>
<tr>
<td>45 to 49</td>
<td>312</td>
<td>10.3%</td>
<td>6.5-16.0%</td>
<td>605</td>
<td>8.7%</td>
<td>6.0-12.6%</td>
</tr>
<tr>
<td>50 to 54</td>
<td>286</td>
<td>14.2%</td>
<td>8.5-22.7%</td>
<td>538</td>
<td>7.5%</td>
<td>4.9-11.2%</td>
</tr>
<tr>
<td>55 to 59</td>
<td>197</td>
<td>6.4%</td>
<td>1.9-19.8%</td>
<td>359</td>
<td>3.0%</td>
<td>1.6-5.6%</td>
</tr>
<tr>
<td>60+</td>
<td>451</td>
<td>4.0%</td>
<td>1.9-8.2%</td>
<td>956</td>
<td>3.7%</td>
<td>2.2-6.3%</td>
</tr>
</tbody>
</table>

Table F.2: Household prevalence statistics and sample sizes by age and sex, 2005
Source: Shisana et al (2005b)