Modelling the impact of HIV in South Africa's provinces: 2022 update

Centre for Infectious Disease Epidemiology and Research working paper

April 2022

Leigh F. Johnson[1]
Rob E. Dorrington[2]

1. Centre for Infectious Disease Epidemiology and Research, University of Cape Town
2. Centre for Actuarial Research, University of Cape Town

Correspondence to:
Dr Leigh Johnson
Centre for Infectious Disease Epidemiology and Research
University of Cape Town
Anzio Road
Observatory
7925
South Africa

Tel: +27 21 406 6981
Fax: +27 21 406 6764
Email: Leigh.Johnson@uct.ac.za

Executive summary

Mathematical models have an important role to play in identifying which interventions are having the greatest impact, and where HIV interventions are most needed. Thembisa is a mathematical model that has been developed to simulate the demographic profile of South Africa and the impact of various HIV prevention and antiretroviral treatment (ART) programmes. The model has previously been applied to each province to identify the factors accounting for HIV prevalence differences across provinces, to compare provinces in terms of how they have progressed in scaling up prevention and treatment programmes, and to assess the differential impact of HIV interventions across provinces.

The previous provincial versions of the Thembisa model were released in March of 2021. This report describes recent updates to the model and presents updates to the results. Since 2021, a number of important changes have been made to the model, which affect the results in all provinces. These include allowing for changes over time in the rates of marriage and divorce, increased ART initiation if people are re-diagnosed after a previous diagnosis, new HIV testing strategies (self-testing and index testing) and linking durations of ART interruption to HIV testing rates.

The updated Thembisa model also includes new provincial data from a variety of sources: monthly numbers of patients receiving ART in 2020-2021, total numbers of HIV tests performed in 2020-2021, numbers of VMMC operations by province, and recent estimates of viral suppression. As before, the provincial models are calibrated to province-specific HIV prevalence data from household surveys and antenatal surveys, antiretroviral metabolite data from the 2012 and 2017 HSRC surveys (survey data on ART coverage), adult vital registration data (all-cause mortality), data on the proportion of adult ART patients who are male, and household survey estimates of HIV prevalence in children. In addition, the new provincial versions of the model are calibrated to data on the age distributions of ART patients, both for adults and children. In most provinces, HIV estimates have not changed substantially relative to the previous version of Thembisa (version 4.4). However, estimates of HIV prevalence have increased slightly in most provinces (relative to Thembisa 4.4 estimates). Estimates of levels of viral suppression have declined slightly in all provinces (relative to version 4.4), which is partly because we have assumed a slower rollout of dolutegravir than assumed previously, and partly due to the assumed higher rates of ART initiation in re-diagnosed individuals (which imply higher rates of ART initiation in later stages of disease, when individuals are less likely to achieve viral suppression).

Results suggests that the epidemiology of HIV in South Africa's is highly heterogeneous. In 2020/21, HIV incidence rates in adults aged 15-49 varied between 0.51% (95% CI: 0.47-0.57%) in Western Cape and 1.09% (95% CI: 1.00-1.22%) in Eastern Cape. Incidence rates have been consistently declining in all provinces, with incidence declines over the 2010-2020 period being greatest in KwaZulu-Natal (66%) and smallest in Western Cape (33%). Despite the steep incidence declines in KwaZulu-Natal, it remained the province with the highest HIV prevalence. HIV prevalence among 15-49 year olds in 2019 varied between 12.4% (95% CI: 11.8-13.2%) in Western Cape and 26.1% (95% CI: 25.4-26.8%) in KwaZulu-Natal.

Progress towards the 95-95-95 targets has been mixed. The model estimates for 2021 suggest that progress towards the first UNAIDS target (95% of HIV-positive individuals diagnosed by

2025) is good, and that progress has been relatively uniform across provinces, with most provinces at around 93%. However, progress towards the second UNAIDS target (95% of HIV-diagnosed individuals on ART) is generally poor, with this fraction treated varying between 59% in Western Cape and 79% in KwaZulu-Natal. Most provinces are close to reaching the third UNAIDS target (95% of ART patients virally suppressed with viral load <1000 RNA copies/ml), with viral suppression being highest in Western Cape (94%), KwaZulu-Natal (93%) and Free State (93%). However, Limpopo and Eastern Cape have viral suppression rates of 87%, suggesting that there is still room for improvement in some provinces. ART coverage in 2021 varied between 54% in Western Cape and 75% in KwaZulu-Natal. Overall, KwaZulu-Natal is the province that has made the most progress towards the 95-95-95 targets, while Western Cape and Limpopo are the provinces that are lagging farthest behind.

Progress in scaling up HIV prevention programmes also appears to have been variable. Levels of condom use appear to be lowest in Northern Cape, Western Cape and Eastern Cape, and highest in Gauteng. The prevalence of male circumcision has increased steeply in KwaZulu-Natal, Free State and Mpumalanga over the last few years, while there has been almost no change in the prevalence of male circumcision in the Western Cape.

One refinement to the model calibration procedure has been to allow for inter-provincial differences in the natural history of paediatric HIV. This has allowed us to get better to fits to the paediatric ART age distributions than we would otherwise. The model estimates suggest somewhat faster progression to late disease and death in children in the Western Cape, Northern Cape and Eastern Cape provinces (relative to the rest of the country), consistent with the estimates for the adult natural history parameters. However, the model does not yet allow for inter-provincial differences in mortality rates after ART initiation. Another limitation is that the model still needs to be calibrated separately to HIV testing data for each province, instead of assuming (as we do currently) that the age and sex patterns of HIV testing are the same across all provinces. This could help to improve the model fit to the age distributions of ART patients in provinces such as Eastern Cape and Limpopo, where the modelled age distribution of ART patients currently does not match the observed age distribution.

Another important refinement has been to the modelling of inter-provincial differences in marriage and divorce rates. Marriage rates were previously assumed to remain constant over time, but in the latest version of Thembisa we estimate declining rates of marriage over time, in all provinces. This may partly explain why estimates of HIV incidence and prevalence in recent years are slightly higher than in the previous version of Thembisa, since a decline in marriage rates over time should imply a slower decline in HIV incidence rates (on the assumption that married individual engage in fewer sexual risk behaviours than unmarried individuals of the same age and sex).

Although it is encouraging to see high levels of HIV diagnosis and viral suppression in ART patients, it remains concerning that South Africa's progress towards the UNAIDS ART coverage target is poor. The declines in HIV incidence over the 2010-2020 period, although consistently falling short of the 75% UNAIDS target for the 2010-2020 period, are nevertheless consistent with the average incidence decline over the 2010-2020 period in the Eastern and Southern African region (43%). Renewed efforts are needed in order to reach the UNAIDS target of 95% ART coverage in HIV-diagnosed individuals by 2025, and continued innovation in the field of HIV prevention will be critical to ensuring that the National Strategic Plan target of a 50% reduction in HIV incidence over the 2017-2022 period is met. More work is required

to identify the success factors that have enabled provinces like KwaZulu-Natal to make good progress towards the 95-95-95 targets, and to follow these examples in other provinces.

Table of contents

1. Introduction

South Africa has the largest number of HIV infections in the world, and in the past has faced major challenges in rolling out HIV prevention and treatment programmes. Although much progress has been made in reducing HIV incidence [1] and HIV-related mortality [2] in South Africa, many challenges remain. Mathematical models have an important role to play in assessing which interventions are likely to have the greatest impact, which interventions are likely to be most cost-effective, and where HIV interventions should be targeted. The latter is particularly important, given the extreme geographical heterogeneity in HIV prevalence in South Africa [3], and given the likely efficiency gains if HIV prevention programmes are prioritized in the sub-populations in which HIV incidence is highest [4].

Mathematical modelling is also important in retrospectively evaluating programme impact, and in comparing prevention and treatment programmes between provinces. UNAIDS has set the goal of 95% of HIV-positive individuals diagnosed, 95% of diagnosed individuals on antiretroviral treatment (ART) and 95% of treated individuals virally suppressed, by 2025 [5]. Tracking progress towards these '95-95-95' targets at a provincial level will be important in identifying where South Africa is falling short.

The Thembisa model was developed to address these questions in the South African context. The model has previously been used to assess which interventions are likely to be most important in reducing future HIV incidence levels [6-8]. The model has also been used to assess progress towards diagnosis and treatment targets, both at national [6, 9, 10] and provincial [11] levels. The model has also provided insights into the factors that account for variation in HIV prevalence between provinces [12], and has been used to assess the impact that antiretroviral treatment (ART) is having on mortality in South Africa [2, 10]. The first set of Thembisa provincial models was released in September of 2016 [13], and updates were released in September of 2017 [14], August 2018 [15], June 2019 [16], June 2020 [17] and March 2021 [18]. Since the March 2021 release, a number of changes have been made to the model. These include allowing for changes over time in the rates of marriage and divorce, increased ART initiation if people are re-diagnosed after a previous diagnosis, new HIV testing strategies (self-testing and index testing), linking durations of ART interruption to HIV testing rates [19], and revising the model of PrEP uptake. In addition, the model has been recalibrated using more recent HIV programme data (for the 2020-21 year), as well as HIV prevalence data from the 2019 antenatal survey (not previously included in the calibration), and new demographic parameters have been incorporated.

The objective of this report is to describe the updated provincial models and to present their results. For a more complete description of the Thembisa model structure and the calibration to national HIV data, the reader is referred to the most recent report on the Thembisa national model [19]. The focus of this report is limited to the assumptions that differ between provinces, and to the presentation of province-specific results. Sections 2 to 7 of this report describe the province-specific assumptions. Section 8 describes the model calibration procedure and results are presented in section 9. The report concludes with a synthesis of the key findings and a discussion of the strengths and limitations (section 10).

2. Modelling sexual behaviour

The sections that follow describe the sexual behaviour parameters that differ between provinces. As there is some uncertainty regarding the sexual behaviour parameters, we have adopted a Bayesian approach, specifying prior distributions to represent the extent of this uncertainty.

2.1 Proportions of individuals in the high-risk group

The high-risk group is defined to consist of individuals with a propensity for concurrent partners and/or commercial sex activity. In the national version of the Thembisa model, the fraction of the adult population that is high-risk is assumed to be 35% for males and 25% for females, based on a review of studies that have estimated the prevalence of concurrency in South Africa. However, the prevalence of concurrency appears to differ substantially between provinces. Table 2.1 shows the proportion of men reporting currently having more than one partner in the 2009 National HIV Communication Survey, stratified by province, as estimated by Morris and Leslie-Cook [20]. (Although results were reported for women as well, there was believed to be substantial under-reporting of concurrency by women, and these results are therefore not shown here.) The ratio of the provincial concurrency prevalence to the national average ranges from 0.383 in the Northern Cape to 1.591 in KwaZulu-Natal.

Table 2.1: Point prevalence of concurrency reported by men, 2009

Province	EC	FS	GT	KZ	LM	MP	NC	NW	WC	SA
Minimum*	0.098	0.052	0.105	0.166	0.103	0.141	0.04	0.071	0.061	0.103
Maximum*	0.108	0.059	0.129	0.179	0.109	0.141	0.043	0.087	0.065	0.114
Relative to SA										
Minimum	0.951	0.505	1.019	1.612	1.000	1.369	0.388	0.689	0.592	
Maximum	0.947	0.518	1.132	1.570	0.956	1.237	0.377	0.763	0.570	
Average	0.949	0.511	1.075	1.591	0.978	1.303	0.383	0.726	0.581	

* Minimum and maximum values reflect different interpretations of missing values.
Source: Morris and Leslie-Cook [20]

An alternative approach to estimating the proportions of the population in the high-risk group is to examine provincial differences in the proportion of men who report having had more than one partner in the last year. Although this does not correspond to the definition of high-risk used in our model, we would expect men in the high-risk group to have the most partners on average, and the relative differences in the proportions of men reporting multiple partners may therefore serve as a crude approximation to the relative differences in the proportion of men who are high-risk. Table 2.2 shows the proportion of men who reported multiple partners in the last year in the 2012 National HIV Communication Survey [21] and 2016 DHS [22] and the ratios of the provincial proportions to the national average. The ratios are generally quite different from those in Table 2.1, especially for Free State and KwaZulu-Natal.

Table 2.2: Proportion of men reporting multiple partners in the last year

Province	EC	FS	GT	KZ	LM	MP	NC	NW	WC	SA
Reported (2012)	14%	33%	25%	19%	13%	12%	12%	19%	12%	19%
Reported (2016)	18.0%	22.6%	17.3%	13.6%	23.7%	18.1%	8.5%	18.4%	11.3%	15.5%
Relative to SA (2012)	0.74	1.74	1.32	1.00	0.68	0.63	0.63	1.00	0.63	
Relative to SA (2016)	1.16	1.46	1.12	0.88	1.53	1.17	0.55	1.19	0.73	

Source: Johnson *et al* [21], Department of Health [22]

A limitation of both analyses is that neither analysis considers the extent to which differences in the reporting of high-risk behaviour might be attributable to inter-provincial differences in the rate of marriage or differences in the age distribution of the population (both factors are assumed to affect the level of extramarital sex in the Thembisa model). Given this limitation, and given the discrepancies between the ratios estimated in Tables 2.1 and 2.2, we represent the uncertainty around the risk group sizes in each province by assigning prior distributions to the factors by which the national high-risk group proportions are multiplied in each province. The means of these prior distributions are the average of the ratios estimated from the three surveys (as shown in the last row of Table 2.1 and last two rows of Table 2.2). The assumed priors are gamma distributions, all with a coefficient of variation of 0.25, to ensure reasonably wide ranges of prior uncertainty around the high-risk proportion in each province. Figure 2.1 shows that these assumptions yield prior distributions with confidence intervals wide enough to include almost all of the ratios estimated from the 2009 and 2012 National HIV Communication Surveys and 2016 DHS.

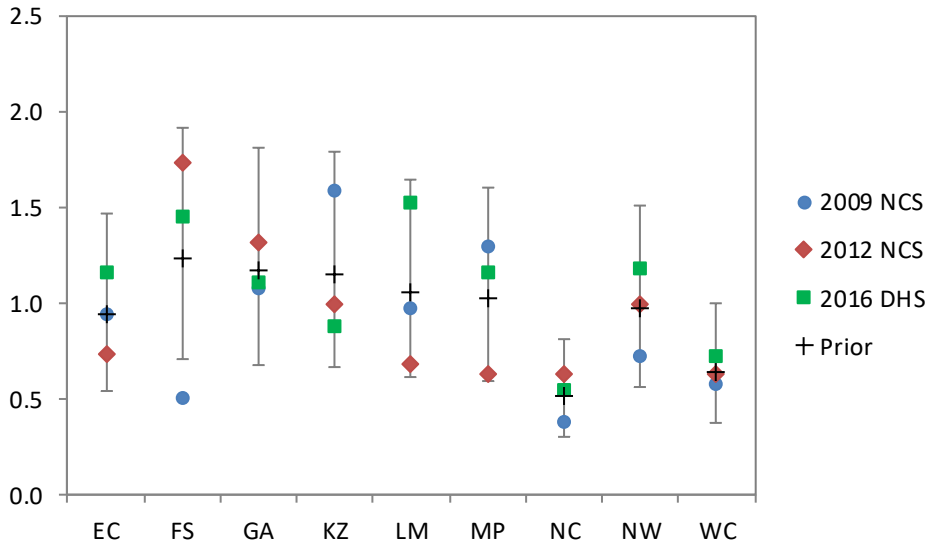

Figure 2.1: Ratio of province-specific high-risk proportion to national average

Error bars represent the range between the 2.5 and 97.5 percentiles of the prior distributions, while crosses represent the means of the prior distributions. DHS = Demographic and Health Survey. NCS = National HIV Communication Survey.

2.2 Marriage and divorce

A full description of the model of marriage and divorce is provided elsewhere [23]. Here we provide a brief description of the model. It is worth noting that the model definition of 'marriage' includes cohabiting relationships, and the model definition of 'divorce' similarly includes separations (even when there is no legal dissolution of the union).

For an individual of age x and sex g, born at time t (measured in years after 1985), the probability that they have never been married is modelled as

$$S_g(x,t) = \frac{1}{1+\left((x-16)exp(-C_g-\beta_g t)\right)^{1/\gamma_g}}$$

for $x > 16$. This assumes that the time to marriage (after age 16) follows a log-logistic distribution, with scale parameter C_g and shape parameter γ_g, with β_g determining the extent of change in marriage rates over time. From this we calculate the probability of entry into marriage over the next year, for an individual currently aged x and born in year t, who has never been married, as

$$p_g(x,t) = 1 - \frac{S_g(x+1,t)}{S_g(x,t)}.$$

The same annual probability is assumed to apply to individuals who have previously been married except in the year in which they become divorced or widowed. The odds or remarriage in the year of divorce or widowhood is assumed to be R_g times the odds of marriage in people of the same age who have never been married.

The probability that an individual of age x and sex g, who is married at the start of year τ, gets divorced in year τ is calculated as

$$D_g(x,t) = 1 - exp\left(-Kd_g(x)G^{\tau-2004}\right)$$

where $d_g(x)$ is an empirically-estimated set of divorce rates in 2004, K is an adjustment factor to correct for possible bias in the empirical estimates, and G is the factor by which divorce rates increase (or reduce) per calendar year.

These parameters were estimated separately for each province by fitting the Thembisa model to census and community survey data in four years (1996, 2001, 2006 and 2016). The census and community survey data were stratified by age and sex, and a likelihood function was specified to represent the model goodness of fit to the census and community survey data. A Bayesian approach was adopted in fitting the model to the data, with prior distributions specified to represent the prior ranges of uncertainty around each of the 10 marriage and divorce parameters. Table 2.3 shows the posterior mean estimates for each of the 10 parameters, by province. A more complete description of the calibration of the model to the province-specific marriage data is provided elsewhere ([24], with manuscript in preparation).

Table 4.3: Marriage and divorce parameters by province

Parameter	Symbol	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Lim-popo	Mpuma-langa	Northern Cape	North West	Western Cape
Log-logistic parameters for 1st marriage										
Constant scale parameter: male	C_1	3.21	2.80	2.89	3.21	3.08	2.96	3.26	3.03	2.82
Constant scale parameter: female	C_2	2.71	2.51	2.49	2.91	2.68	2.57	2.81	2.85	2.67
Birth cohort effect: male	β_1	0.0147	0.0153	0.0164	0.0163	0.0197	0.0165	0.0232	0.0118	0.0134
Birth cohort effect: female	β_2	0.0200	0.0177	0.0193	0.0186	0.0236	0.0179	0.0239	0.0147	0.0212
Shape parameter: male	γ_1	0.53	0.45	0.45	0.48	0.54	0.44	0.61	0.48	0.47
Shape parameter: female	γ_2	0.57	0.68	0.75	0.51	0.68	0.57	0.74	0.74	0.78
Remarriage parameters										
Ratio: male odds of first marriage to odds of remarriage, per year	$1/R_1$	0.020	0.085	0.080	0.028	0.025	0.055	0.042	0.097	0.138
Ratio: female odds of first marriage to odds of remarriage, per year	$1/R_2$	0.709	0.774	0.728	0.650	0.735	0.750	0.787	0.813	0.804
Union dissolution parameters										
Adjustment to empirical estimates	K	1.84	0.73	1.16	1.72	0.89	2.10	0.93	0.78	0.44
Annual change in divorce rates	G	0.9989	0.9751	0.9834	0.9936	0.9783	1.0129	0.9910	0.9865	0.9672

2.3 Sexual mixing

The extent of mixing between high risk and low risk groups is important in determining both the rate at which the epidemic spreads in its early stages and the level at which HIV prevalence ultimately stabilizes. If mixing is highly assortative (i.e. high risk individuals tend to form partnerships mainly with other high risk individuals), HIV prevalence grows more rapidly at first, but ultimately levels off at a lower level than would be expected in the presence of random sexual mixing [25]. The degree of sexual mixing, which defines the extent of the mixing between high and low risk groups, varies between 0 ('completely assortative mixing') and 1 (random mixing, where individuals have no preferences regarding the risk group of their partners). Empirical estimates from high income countries suggest that sexual mixing is generally much closer to random than to completely assortative [26-29]. However, there is a lack of data from African settings, and simulation studies suggest that empirical estimates of the degree of sexual mixing are likely to be biased upward [30]. In previous attempts to fit the Thembisa model to province-specific data we found that the best-fitting degree parameter varied between 0.42 and 0.80 (average 0.57, standard deviation 0.12) [18]. We have therefore assigned a beta prior distribution to represent the uncertainty around the degree of assortative mixing, with a mean of 0.57 and a standard deviation of 0.12.

2.4 Condom usage

Condom usage may be heterogeneous between provinces due to differences in levels of condom distribution, differences in exposure to social marketing campaigns and differences in perceived levels of HIV risk. Multiplicative adjustment factors are applied to the condom usage rates that have been estimated nationally to account for differences between provinces.

In the provincial model, the parameter $\gamma_{2,l}(x, t, p)$ represents the probability that an HIV-negative woman in province p aged x uses a condom in an act of sex with a partner of type l at time t (time is measured in years since 1985). This parameter is calculated as

$$\gamma_{2,l}(x, t, p) = \zeta(t, p)\, \nu^{(x-20)}\, \beta_l\, \theta_r,$$

where $\zeta(t, p)$ represents the time trend in condom use, ν is the factor by which condom use decreases per year of increase in age, β_l is the relative rate of condom use in relationship type l, and θ_r is a scaling parameter, which we include to allow for the possibility of bias in self-reported condom use data (the value depends on the type of reporting, r). The 'base rate' of condom use, $\zeta(t, p)$, relates to women aged 20 who are unmarried ($l = 0$) and reporting on their condom use at last sex ($r = 0$), and the β_0 and θ_0 parameters are therefore both set to 1. The $\zeta(t, p)$ function is a linear combination of a constant term and two cumulative Weibull distribution functions. The constant term ($k_{0,p}$) represents the initial rate of condom usage, prior to the start of the HIV epidemic in South Africa, the first Weibull distribution corresponds to the increase in condom usage following the introduction of behaviour change communication programmes in the mid-1990s, and the second Weibull distribution represents a possible change in condom usage rates in recent years. In mathematical terms,

$$\zeta(t, p) = k_{0,p} + k_{1,p}\left(1 - 0.5^{(t/m_1)^{\phi_1}}\right) + k_{2,p}\left(1 - 0.5^{(t/m_2)^{\phi_2}}\right).$$

The $k_{1,p}$ parameter represents the extent of the increase in condom use following the early phase of the HIV communication programmes, and the m_1 and ϕ_1 parameters represent the median and shape parameters respectively of the first Weibull distribution. The $k_{2,p}$ parameter represents the extent of the change in condom use in recent years (possibly due to changes in funding for behaviour change communication programmes, and possibly due to changes in attitudes towards condom use as ART has become more widely available); the m_2 and ϕ_2 parameters represent the median and shape parameters respectively of the second Weibull distribution.

Most of the parameters in these two equations are not province-specific, and have been fixed at the values estimated when fitting the national model to nationally-representative data on self-reported condom use (as specified elsewhere [23]). However, the $k_{0,p}$, $k_{1,p}$ and $k_{2,p}$ parameters do differ by province. For the sake of simplicity, we specify $k_{0,p} = \alpha_p \times k_0$ and $k_{1,p} = \alpha_p \times k_1$, where k_0 and k_1 represent the national estimates of the corresponding parameters, and α_p represents a constant scaling factor (the ratio of condom use in province p to that nationally, in the early stages of the South African HIV epidemic). For the purpose of setting the α_p parameters, we consider evidence from national surveys conducted up to 2012 (before 2012, the contribution of the second Weibull term is relatively small). Table 2.4 shows the ratio of reported condom usage in each province to the national average. The data from the HSRC household surveys relate to the whole population aged 15 and older, and the ratios presented may therefore differ due to demographic differences between provinces (for example, differences in the age profile or in the fraction of the population that is married). The data from the two DHSs relate only to women aged 15 to 49, and the ratios presented control for differences in type of relationship (but not for differences in age). The only consistent pattern is one of lower rates of condom use in Northern Cape and Western Cape when compared with the rest of the country; for the other provinces the average ratio of provincial condom use to national condom use is close to one.

Table 2.4: Ratio of provincial condom use to national average

	1998 DHS	2002 HSRC	2003 DHS	2005 HSRC	2008 HSRC	2012 HSRC	Average
Eastern Cape	0.75	1.15	0.93	1.01	1.06	1.05	0.99
Free State	1.60	1.29	0.89	0.87	1.05	1.12	1.14
Gauteng	1.20	1.16	1.00	1.06	0.93	0.99	1.06
KwaZulu-Natal	0.79	0.98	1.56	1.03	1.05	1.09	1.08
Limpopo	0.90	1.01	0.82	1.26	1.17	1.09	1.04
Mpumalanga	1.22	0.89	0.75	1.02	1.15	1.09	1.02
Northern Cape	0.64	0.62	0.62	0.54	0.67	0.74	0.64
North West	1.18	0.97	0.90	1.05	1.06	1.13	1.05
Western Cape	0.79	0.78	0.77	0.64	0.77	0.67	0.74

Source: Shisana *et al* [3], Department of Health [31, 32]

Given the limitations of the published statistics, particularly the lack of control for background characteristics, it is difficult to state precisely by what factor condom usage is increased or decreased in each province, relative to the national average. Our approach is therefore to specify prior distributions to represent the ranges of uncertainty around the province-specific

α_p adjustment factors. For each province, the prior is a gamma distribution with a mean equal to the value in the last column of Table 2.4 and a coefficient of variation equal to 0.15. The coefficient of variation was chosen so that the intervals between the 2.5 and 97.5 percentiles of each prior distribution included 96% of the corresponding data points in Table 2.4, thus ensuring reasonable coverage of the 95% confidence intervals.

There is also substantial uncertainty around the $k_{2,p}$ parameters. In fitting the national model we estimated a $k_{2,p}$ value of 0.02, i.e. suggesting no decline in condom use in recent years [23]. Given the lack of recent survey data and the absence of strong evidence of risk compensation, we use the same assumption, $k_{2,p} = 0.02$, in all provinces.

2.5 Rates of short-term partnership formation

In setting the assumed rates of non-marital partnership formation, we follow a five-step process: (1) we specify the rate of non-marital partnership formation in unmarried 'high risk' women who are aged 20; (2) we specify a gamma function that determines the relative rates of non-marital partnership formation that apply at all other ages; (3) we specify rates of non-marital partnership formation in the 'low risk' group as a fraction of those in the high risk group; (4) we specify rates of non-marital partnership formation in married individuals, as a fraction of those in unmarried individuals; and (5) we derive rates of non-marital partnership formation in men from the assumptions made for women. Steps (1)-(5) are essentially the same for all provinces, and have been described previously for the national model [19]. However, the model allows for differences across provinces in the age pattern of non-marital sexual activity, and this section describes how these age differences are accounted for.

We define $c_{g,i,l}(x)$ to be the annual rate of non-marital partnership formation in individuals aged x, of sex g and marital status l, who are in risk group i. The female rates of partnership formation at different ages are modelled using a scaled gamma density of the form

$$c_{2,i,l}(x) = c_{2,i,l}(20)\frac{\lambda^{\alpha}(x-10)^{\alpha-1}\exp(-\lambda(x-10))}{\lambda^{\alpha}10^{\alpha-1}\exp(-10\lambda)},$$

where the λ and α parameters determine the mean and variance of the gamma distribution, and the offset of 10 years is included to prevent sexual activity below age 10. The $c_{2,i,l}(20)$ value is 3.3 for women in the high risk group ($i = 1$) who are unmarried ($l = 0$), based on previous modelling of rates of partnership formation in South Africa [33]. The mean and standard deviation of the gamma density are uncertain; in the previous version of Thembisa (version 4.4), different posterior estimates were obtained for each of the 9 provinces. The average of the gamma means was 37.8 years (standard deviation 3.9 years), and the average of the gamma standard deviations was 20.7 years (standard deviation 2.4 years) [18]. We have therefore represented the uncertainty regarding the gamma mean and standard deviation using gamma prior distributions, with means and standard deviations equal to those estimated from the previous provincial fits. For each sampled value of the gamma mean and standard deviation, λ and α parameters are calculated to be consistent with these values.

2.6 Rates of male contact with sex workers

The Thembisa model assumes that only men in the high-risk group have contact with sex workers, and that their rates of contact depend on their age and marital status. To the extent that the high-risk proportions, age distributions and marriage rates are assumed to differ across provinces, the Thembisa does already make implicit allowance for inter-provincial differences in male rates of sex worker contact. However, there may be other factors that account for inter-provincial differences in rates of sex worker contact. Two important factors that we have not previously considered in parameterizing the provincial models are the effects of urbanization and sex ratios. Previous studies suggest that male contact with sex workers tends to be more frequent in urban settings than in rural settings [34]. Evidence also suggests that male contact with sex workers tends to be more frequent in settings in which there is a high sex ratio (i.e. more men than women) [35]. However, the two factors are somewhat conflated, as sex ratios tend to be higher in urban areas, and relatively few analyses have controlled for both urbanization and local sex ratios when assessing predictors of male contact with sex workers.

To assess local evidence for these two factors, we analyse data from the 2016 Demographic and Health Survey (DHS) [22]. 3.0% of sexually experienced men aged 15-59 reported having paid for sex in the 12 months prior to the survey. This is likely to be an under-estimate of the true fraction of men who have sex worker contact, as self-reported data on sex worker contact is known to be unreliable [36]. Nevertheless, the data may be useful in assessing the *relative* significance of different factors affecting male rates of sex worker contact. Table 2.5 shows the results of a multivariable logistic regression model fitted to the DHS data. In this analysis, the sex ratio was calculated as the number of men per 100 women, in the province in which the male was interviewed. These sex ratios were calculated from the Thembisa version 4.2 outputs for 2016, in the population aged 15-49. In addition to the local sex ratio and urban/rural location, we controlled for age, marital status and multiple partners in the last year (as a proxy for high-risk group activity), to be consistent with the factors that are already controlled for in the Thembisa model.

Table 2.5: Factors associated with male contact with sex workers in the last 12 months

Variable	Level	Adjusted OR (95% CI)
Age	15-24	1
	25-34	4.17 (1.82-9.55)
	35-44	5.76 (2.36-14.06)
	45-59	2.87 (0.94-8.71)
Sex ratio (men per 100 women)	Per unit	1.02 (0.99-1.06)
Location	Rural	1
	Urban	1.36 (0.85-2.18)
Marital status	Unmarried	1
	Married/cohabiting	0.41 (0.21-0.79)
Partners in last 12 months	1 or none	1
	2 or more	3.57 (2.07-6.14)

Source: Author's analysis, based on 2016 DHS data [22].

The results are roughly consistent with previous analyses. The self-reported data suggest that male contact with sex workers is relatively infrequent in the 15-24 age group, and peaks in the 35-44 age group; this is consistent with the assumptions in Thembisa, which yield a peak in male contact with sex workers around age 37 [37]. Men who are married or cohabiting are

estimated to have a lower rate of sex worker contact (aOR 0.41, 95% CI: 0.21-0.79), roughly consistent with the Thembisa assumption that the relative rate of sex worker contact is 0.25 in married men, though not as low as the rate observed in Zambia (OR 0.17) [38]. The finding of a significantly increased rate of sex worker contact in men who report multiple sexual partners is also consistent with the Thembisa assumption that contact with sex workers is limited to the high-risk group. The results also suggest a positive effect of urban location and the provincial sex ratio on the odds of sex worker contact – although neither effect is statistically significant. Although attempts were made to control for province (instead of controlling for urbanization and sex ratio), the resulting provincial odds ratios had extremely wide confidence intervals around them, owing to the small numbers of men reporting contact with sex workers. Controlling only for sex ratio and urbanization is therefore a more parsimonious approach to assessing geographical variation.

To our knowledge, only two previous studies have assessed the effects of both urban location and the sex ratio on men's rates of sex worker contact. The results of these two studies are compared with the results of our South African analysis in Table 2.6; these other two studies also found positive effects of urbanization and sex ratio on the rate of sex worker contact. In all three studies, other important predictors of male contact with sex worker contact were also controlled for (most significantly age and marital status). Pooling the results of the three studies in a meta-analysis, the odds of sex worker contact are on average 1.54 (95% CI: 1.14-2.08) times higher in urban areas than in rural areas, and the odds increase by a factor of 1.009 (95% CI: 1.004-1.015) per unit increase in the number of men per 100 women. The meta-analysis results are heavily weighted to the analysis of South *et al* [39] in India, as this study yielded the most precise estimates.

Table 2.6: Comparison of estimates of effects of urban location and sex ratio (men per 100 women), on male rates of sex worker contact

Study	Location	Effect of urban location (aOR, 95% CI)	Effect of sex ratio (aOR, 95% CI)
Present study	South Africa	1.36 (0.85-2.18)	1.024 (0.988-1.062)
South & Trent [39]	China	3.00 (0.93-9.74)	1.020 (1.000-1.040)*
South *et al* [40]	India	1.56 (1.03-2.36)	1.008 (1.002-1.014)*
Meta-analysis		1.54 (1.14-2.08)	1.009 (1.004-1.015)

* In the original studies, the sex ratio was reported as the number of women per 100 men, and we have therefore inverted the odds ratios estimated from these studies in order to be consistent with the more conventional definition of the sex ratio (number of men per 100 women).

For the purpose of estimating parameters in Thembisa, we use the odds ratios from the meta-analysis to estimate the relative rates of male contact with sex workers in each province (relative to the national average), and multiply these by the assumed rates of sex worker contact in the national model. Results of these calculations are shown in Table 2.7. The odds ratio in the Eastern Cape, for example, is calculated as $1.54^{(0.48-0.645)} \times 1.009^{(90.7-98.9)} = 0.86$, and the annual number of contacts with sex workers, for unmarried sexually experienced high-risk men in the Eastern Cape, aged 21, is $3.5 \times 0.86 = 3.03$ (where 3.5 is the assumed number at a national level). The resulting rates of male contact with sex workers are lowest in the Limpopo and Eastern Cape provinces (which have relatively low levels of urbanization and low sex ratios) and highest in Gauteng (the most urbanized province and also a province with a relatively high sex ratio). Although confidence intervals around these estimates are not estimated, we found when substituting the lower and upper confidence interval limits from the meta-analysis into

the calculations that these resulted in only modest changes in the relative rates of sex worker contact across provinces.

Table 2.7: Relative rates of male contact with sex workers in different provinces

Province	% urban*	Sex ratio†	OR for SW contact (relative to national)	Annual contacts with SWs, for unmarried high-risk men aged 21
Eastern Cape	48.0%	90.7	0.86	3.03
Free State	85.2%	97.1	1.08	3.77
Gauteng	96.5%	107.4	1.24	4.34
KwaZulu-Natal	47.3%	93.6	0.89	3.10
Limpopo	20.0%	90.6	0.77	2.68
Mpumalanga	40.1%	99.7	0.91	3.17
Northern Cape	74.9%	102.7	1.08	3.79
North West	50.2%	110.8	1.05	3.66
Western Cape	94.8%	99.5	1.15	4.01
South Africa	64.5%	98.9	-	3.50

* Estimated from the 2016 South African General Household Survey (author's own calculations). † Number of men per 100 women in 2016, in the 15-49 age group, as estimated by Thembisa version 4.2. SW = sex worker.

3. Modelling HIV survival

3.1 Rates of HIV disease progression pre-ART

The course of untreated HIV disease was previously assumed to be the same across South Africa's provinces. However, there is evidence of differences in CD4 distributions in HIV-negative adults when comparing Gauteng [41] and rural KwaZulu-Natal [42], and there are substantial differences across provinces in the incidence of tuberculosis [43], which is a major cause of death in people living with HIV. There are also major socioeconomic differences between provinces, and to the extent that socioeconomic factors influence HIV mortality [44], differences in HIV survival might be expected. For simplicity, we assign the same prior distribution to represent our uncertainty around the average adult survival time in all provinces; this gamma prior distribution has a mean of 12 years and a standard deviation of 1 year. The model also includes parameters that represent the effect of age and sex on the average adult HIV survival time, but in the interests of simplicity, we do not allow these parameters to vary in the calibration process.

3.2 Rates of ART initiation in adults

Suppose we define $\rho_g(t)$ as the monthly rate of ART initiation in projection year t in HIV-diagnosed individuals of sex g, with CD4 <200 cells/μl (excluding individuals who are newly diagnosed in the current month). We adopt a relatively parsimonious approach to estimating these parameters, which is to specify average rates of ART initiation in women over the mid-2000 to mid-2004 period (before the launch of the national public sector programme, when ART access was limited to the private sector), over mid-2004 to mid-2005 (the first year of the public sector roll-out), over mid-2010 to mid-2011 (the last year before the first major change in ART eligibility criteria), over mid-2011 to mid-2016 (when adult ART eligibility changed to CD4 count < 350 cells/μl and later to < 500 cells/μl), and over the period after mid-2016 (when there was universal ART eligibility). Over the 2005-2010 period, rates of ART initiation are interpolated linearly from those specified in the 2004-05 and 2010-11 years. We also assume that over the period from April 2020 to June 2021 rates of ART initiation were reduced by a proportion C, to reflect the impact of the COVID-19 epidemic on ART services [45]. To present this in more mathematical terms:

$$\begin{aligned}
&\rho_g(t) = 0 && \text{for } t < 2000\\
&\rho_g(t) = \rho_g(2000) && \text{for } 2000 < t \leq 2003\\
&\rho_g(t) = \rho_g(2004) + (\rho_g(2010) - \rho_g(2004))(t - 2004)/6 && \text{for } 2004 < t < 2010\\
&\rho_g(t) = \rho_g(2011) && \text{for } 2011 < t < 2016\\
&\rho_g(t) = \rho_g(2016) && \text{for } t > 2016\\
&\rho_g(t) = \rho_g(2016)\,(1 - C) && \text{for April 2020-June 2021}
\end{aligned}$$

To further simplify the model, we assume that the ratio of ART initiation in men to that in women ($M \equiv \rho_1(t)/\,\rho_2(t)$) is constant over time. There are thus seven parameters that need to be estimated for each province: $\rho_2(2000)$, $\rho_2(2004)$, $\rho_2(2010)$, $\rho_2(2011)$, $\rho_2(2016)$, C and M.

We adopt a Bayesian approach in estimating these parameters, assigning gamma distributions to represent the uncertainty around these parameters that exists *a priori*. For the first five parameters, we assign the prior distributions based on the mean and standard deviation of the rates estimated in the previous version of Thembisa [18], i.e. taking account of plausible levels of variation across provinces. These are gamma distributions with means of 0.0039 for $\rho_2(2000)$, 0.0154 for $\rho_2(2004)$, 0.0576 for $\rho_2(2010)$, 0.0622 for $\rho_2(2011)$, and 0.0591 for $\rho_2(2016)$ (standard deviations are 0.0011, 0.0117, 0.0202, 0.0196 and 0.0240 respectively). The prior distribution assigned to the *C* parameter is a beta prior distribution with a mean of 0.28 and standard deviation of 0.10, based on a recent analysis of South African data on new ART initiations, which found 28% fewer new ART initiations in 2020 than in 2019 [45].

For the purpose of setting the prior distribution on the *M* parameter, we review South African studies on relative rates of linkage to HIV care/ART initiation after HIV diagnosis, when comparing men to women (Table 3.1). Results of these studies are highly variable, with relative rates of linkage/ART initiation varying between 0.34 and 1.11 (median 0.84). These empirical estimates may under-estimate the *M* parameter, as the *M* parameter effectively combines the effect of male sex on rates of linkage to care after diagnosis and the effect of male sex on ART initiation after linkage to care, whereas most of the estimates in Table 3.1 represent only one of these two components. To represent the uncertainty around *M* we assign a gamma prior distribution with a mean of 0.7 (roughly equal to 0.84 × 0.84, i.e. taking into account the compounding effect of male sex on linkage and ART initiation) and a standard deviation of 0.1 (2.5 and 97.5 percentiles of 0.52 and 0.91 respectively).

Table 3.1: Relative rates of linkage to care/ART initiation in HIV-diagnosed men (compared to women)

Study	Outcome	RR for men (95% CI)
Dorward *et al* [46]	Linkage to care within 1 year of diagnosis	0.86 (0.76-0.98)
Boyer *et al* [47]	ART initiation within 1 month of linkage	1.0 (0.7-1.3)
Larson *et al* [48]	Return to clinic within 1 year of pre-ART care	0.83 (0.58-1.20)
Lessells *et al* [49]	Repeat CD4 within 13 months of pre-ART care	0.71 (0.53-0.95)
Kranzer *et al* [50]	CD4 within 6 months of HIV diagnosis:	
	STI patients	0.93 (0.79-1.09)
	Clients seeking HIV testing	1.10 (1.01-1.33)
Lurie *et al* [51]	Linkage to pre-ART care	0.79 (0.69-0.90)
	ART initiation after linkage	0.34 (0.30-0.38)
Maughan-Brown *et al* [52]	Linkage to HIV services 12 weeks after diagnosis	0.99 (0.51-1.93)
Osler *et al* [53]	ART initiation after linkage	0.79 (0.77-0.80)

Rates of ART initiation are also adjusted to allow for differences between CD4 categories in rates of ART initiation. Let $J_s(t)$ be the relative rate of ART initiation in stage *s* relative to that in the CD4 <200/μl category ($s = 5$, i.e. implying $J_5(t) = 1$). In most periods $J_s(t)$ will be zero for $s < 5$, since South African ART guidelines have only recently changed to allow for ART initiation at higher CD4 counts. When all individuals are eligible for ART, we set $J_s(t)$ to 0.40 for CD4 of 500 or higher, 0.50 for CD4 of 350-499, 0.70 for CD4 of 200-349 and 1 for CD4 <200. (These assumptions are based primarily on the observed relative rates of ART initiation

in ART-eligible individuals in different CD4 categories [54, 55], and are consistent with the relative rates at which individuals enrolled in pre-ART care return for regular CD4 testing [48, 49].)

Further details regarding the model calibration to ART programme data is provided in sections 8.1.4-8.1.5.

3.3 Rates of viral suppression on ART

The description that follows is mostly the same as that in Appendix F of the national report [19], but as it contains a number of province-specific assumptions, we repeat it here for convenience.

We adopt a Bayesian approach to estimating the true rate of viral suppression from a number of data sources. For each province, we assume that the odds of viral suppression is proportional to the odds of viral suppression estimated using data from South African cohorts participating in the International epidemiology Databases to Evaluate AIDS (IeDEA) collaboration [56]. This constant of proportionality (or odds ratio) is assumed to differ by province. In addition, we assume for each province that there is an odds ratio relating the odds of viral suppression in patients with unrecorded viral loads to that in patients with recorded viral loads. We follow a two-step Bayesian updating approach in estimating these two parameters. In the first step (described in section 3.3.1), only the TIER data are used in the model fitting. In the second step (described in section 3.3.3), the posterior estimate of the IeDEA bias from the first step becomes the prior distribution for the IeDEA bias in the second step, and the Thembisa model is fitted to other HIV data sources. Note that in the second step it is not necessary to include the uncertainty regarding the bias due to missing data, because this bias affects only the interpretation of the TIER data, and the TIER data on viral suppression are not included in the second step.

3.3.1 Viral suppression in adults

A model of the following form is fitted to IeDEA-SA data on viral suppression in adults:

$$\text{logit}(I_{t,s}) = C + \beta_t + \gamma_s, \qquad (3.1)$$

where $I_{t,s}$ represents the proportion of patients who are virally suppressed in year t, in patients who started ART in baseline CD4 count category s. Full details of the IeDEA-SA dataset and the procedures followed in defining viral suppression are provided elsewhere [56]. The results of the model are summarized in Table 3.2. Consistent with the previous Thembisa estimates [37], the results suggest a substantial decline in rates of viral suppression after 2009, followed by a gradual increase in viral suppression after 2013. The results also suggest substantially higher rates of viral suppression in patients who start ART at higher CD4 counts, consistent with previous studies [57-60].

Table 3.2: Predictors of viral suppression in IeDEA-SA cohorts

Factor	Symbol	Odds ratio (95% CI)
Constant	exp(C)	7.24 (6.22-8.42)
Effect of calendar year (ref. 2005)		
2006	$\exp(\beta_{2006})$	1.05 (0.89-1.24)
2007	$\exp(\beta_{2007})$	0.95 (0.81-1.11)
2008	$\exp(\beta_{2008})$	1.04 (0.88-1.22)
2009	$\exp(\beta_{2009})$	0.91 (0.77-1.07)
2010	$\exp(\beta_{2010})$	0.71 (0.60-0.83)
2011	$\exp(\beta_{2011})$	0.47 (0.40-0.54)
2012	$\exp(\beta_{2012})$	0.53 (0.45-0.62)
2013	$\exp(\beta_{2013})$	0.52 (0.45-0.61)
2014	$\exp(\beta_{2014})$	0.63 (0.54-0.73)
2015	$\exp(\beta_{2015})$	0.82 (0.70-0.96)
2016	$\exp(\beta_{2016})$	0.63 (0.54-0.74)
2017	$\exp(\beta_{2017})$	0.91 (0.78-1.06)
2018	$\exp(\beta_{2018})$	0.93 (0.79-1.09)
Baseline CD4 category (ref. <200 cells/μl)		
200-349 cells/ μl	$\exp(\gamma_{200})$	1.49 (1.44-1.54)
350-499 cells/ μl	$\exp(\gamma_{350})$	1.73 (1.63-1.83)
≥500 cells/ μl	$\exp(\gamma_{500})$	1.92 (1.81-2.05)

We define $V_{t,s}(p)$ as the estimate of the true rate of viral suppression in patients on ART in year t, who started ART in CD4 category s and who currently live in province p. This is calculated as

$$\text{logit}(V_{t,s}(p)) = \text{logit}(I_{t,s}) + \lambda_p, \qquad (3.2)$$

where $\exp(\lambda_p)$ is the odds ratio relating the odds of viral suppression in province p to that in the IeDEA-SA cohorts. The logit transformation is applied to avoid situations in which the regression model predicts a rate of viral suppression <0% or >100%.

The λ_p terms are unknown, and we therefore specify a prior distribution to represent the uncertainty around these parameters. In the analysis of IeDEA-SA adult viral load data, it was noted that the average rate of viral suppression (at a threshold of <400 RNA copies/ml) was 85.7%, and that the rates in each year were consistently between 1% and 5% higher than those reported by the Department of Health at a national level (using the same threshold) [56]. This suggests a prior mean for λ_p of around -0.23 (logit(0.827) – logit(0.857), where 0.827 = 0.857 – 0.03 and 0.03 is the midpoint of the 0.01-0.05 range). We have set the standard deviation of the prior distribution to 0.27, based on fitting the same regression model as shown in equation (3.2) to IeDEA-SA data, but allowing for additional terms to represent differences across ART cohorts. The standard deviation of these cohort-specific terms determines the prior standard deviation of 0.27. The prior distribution is thus a normal distribution with a mean of -0.23 and a standard deviation of 0.27.

For the purpose of calibrating the model estimates ($V_{t,s}(p)$) to routinely reported rates of viral suppression from TIER, it is also necessary to allow for uncertainty regarding the bias due to missing viral load data. We define θ_p as the ratio of the odds of viral suppression in patients with missing viral load measurements to that in patients with recorded viral loads, in province

p. Based on a review of empirical estimates (presented more fully in Appendix F of the national report [61]), we assign a gamma prior to represent the uncertainty around the θ_p parameter, with mean 0.96 and standard deviation of 0.25. This distribution has 2.5 and 97.5 percentiles of 0.53 and 1.51 respectively, roughly consistent with the range of empirical estimates that we have identified.

Suppose that $R_t(p)$ is the reported rate of viral suppression in province *p* at time *t*, and that the corresponding testing coverage on which this is based is $\delta_t(p)$. For the purpose of defining a likelihood function, we define

$$\text{logit}\left(\delta_t(\text{p})R_t(\text{p}) + \frac{1-\delta_t(\text{p})}{1+\dfrac{1\text{-}R_t(\text{p})}{R_t(\text{p})\theta_\text{p}}}\right) = \text{logit}\left(\sum_{s=1}^{4}\pi_{t,s}(p)V_{t,s}(p)\right) + \varepsilon_t(p), \qquad (3.3)$$

where $\pi_{t,s}(p)$ is the model estimates of the fraction of ART patients who started ART in CD4 category *s*, and $\varepsilon_t(p) \sim \text{N}(0, \sigma^2)$. Our assumption is thus that if the model provides a good fit to the data, the difference on the logit scale between the reported rates of viral suppression (after adjustment for missing viral load data) and the modelled rates of viral suppression (after adjustment for differences in viral suppression across baseline CD4 categories) should be close to zero on average, with relatively small variance (i.e. low σ^2). The $\pi_{t,s}(p)$ terms are approximated from version 4.4 of Thembisa.

The variance term is approximated using the maximum likelihood formula:

$$\sigma^2 = \frac{1}{n_p}\sum_{\text{t}\in\text{T}_\text{p}}\left[\text{logit}\left(\delta_t(\text{p})R_t(\text{p}) + \frac{1-\delta_t(\text{p})}{1+\dfrac{1\text{-}R_t(\text{p})}{R_t(\text{p})\theta_\text{p}}}\right) - \text{logit}\left(\sum_{s=1}^{4}\pi_{t,s}(p)V_{t,s}(p)\right)\right]^2,$$

where T_p is the set of time points for which we have viral suppression estimates and n_p is the number of $R_t(p)$ data points for province *p*.

The model is fitted to several datasets for each of the nine provinces (and for the country as a whole). The following sources are combined in creating the dataset for each province:

- Province-specific viral load data from the TIER database for 2013-14, considering patients who had been on ART for 6 months and 48 months, i.e. a total of 18 data points (information on viral suppression was not available for other years or for other ART durations) [62].
- National viral load data from the TIER database for patients who had been on ART for 6 months (for each year from 2005-2014) and for 48 months (for each year from 2009-2014), i.e. a total of 16 data points [62].
- National viral load data from the TIER database for patients on ART in 2013, reported at 12-month intervals from 12 to 108 months after ART initiation, i.e. a total of 9 data points [63].

- National and provincial viral load data from the TIER database, for patients who had been on ART for 6 months, for each of the four quarters in the 2017-18 to 2020-21 fiscal years, i.e. a total of 160 data points (Thapelo Seatlhodi, personal communication).

Other viral load data are available, but this analysis is limited to those data sources for which there was information on both the proportion of patients with viral load measurements and the proportion of those measurements that were suppressed. Suppression was defined in all cases as a viral load of <400 RNA copies/ml.

Having specified the prior distributions and the likelihood function, the final step in the Bayesian analysis is the simulation of the posterior distribution. We follow a Sampling Importance Resampling approach to approximate the posterior [64]. Since there are only two parameters being estimated in the analysis (λ_p and θ_p), it is sufficient to use a small sample size (1000) in both the sampling and resampling steps.

Table 3.3 summarizes the results of the Bayesian analysis for each province (and for the country as a whole). The standard deviation of the model errors is included as a measure of 'goodness of fit'. In most provinces, the standard deviation is around 0.2 or lower. In the Free State, KwaZulu-Natal and Western Cape, levels of viral suppression appear to be close to the rates estimated from the IeDEA data (i.e. estimates of λ_p are close to zero), but in all other provinces viral suppression appears to be lower than that in the IeDEA-SA cohorts. Finally, the posterior estimates of the θ_p parameter are in most cases not very different from the prior mean (0.96).

Table 3.3: Posterior estimates of viral suppression parameters

	Sample Size (n_p)	Standard deviation of model errors (σ)	Difference in viral suppression relative to IeDEA (λ_p)	OR for viral suppression if VL not recorded (θ_p)
EC	18	0.093	-0.47 (-0.62 to -0.34)	1.07 (0.68-1.48)
FS	18	0.230	0.03 (-0.19 to 0.20)	0.97 (0.57-1.37)
GT	18	0.300	-0.40 (-0.59 to -0.22)	1.12 (0.72-1.65)
KZ	18	0.242	-0.01 (-0.24 to 0.16)	0.79 (0.48-1.24)
LM	18	0.177	-0.60 (-0.74 to -0.48)	1.00 (0.64-1.43)
MP	18	0.198	-0.33 (-0.49 to -0.18)	1.10 (0.69-1.66)
NC	18	0.206	-0.31 (-0.57 to -0.05)	1.22 (0.75-1.79)
NW	18	0.231	-0.32 (-0.52 to -0.14)	1.00 (0.60-1.63)
WC	17	0.162	0.13 (-0.10 to 0.39)	1.01 (0.57-2.03)
SA	41	0.232	-0.28 (-0.44 to -0.12)	0.82 (0.57-1.08)

Figure 3.1 shows an example of the model calibration to the data at a national level. The data shown in Figure 3.2 are unadjusted since the posterior mean of θ_p is close to 1 at a national level. The model results are calculated for the posterior mean (λ_p = -0.28) and the 2.5 and 97.5 percentiles (λ_p = -0.44 and -0.12 respectively, as represented by the dashed lines). The model results generally appear consistent with the data, with one notable outlier in 2013 (which related to patients who had been on ART for 9 years, i.e. a relatively small cohort since the public-sector ART programme only started in 2004). The model is also validated by two data points from 2016 (not included in the model calibration because there was no associated information on the testing coverage). Both the model and the data suggest a dip in rates of viral suppression around 2010-2011, followed by a gradual increase thereafter.

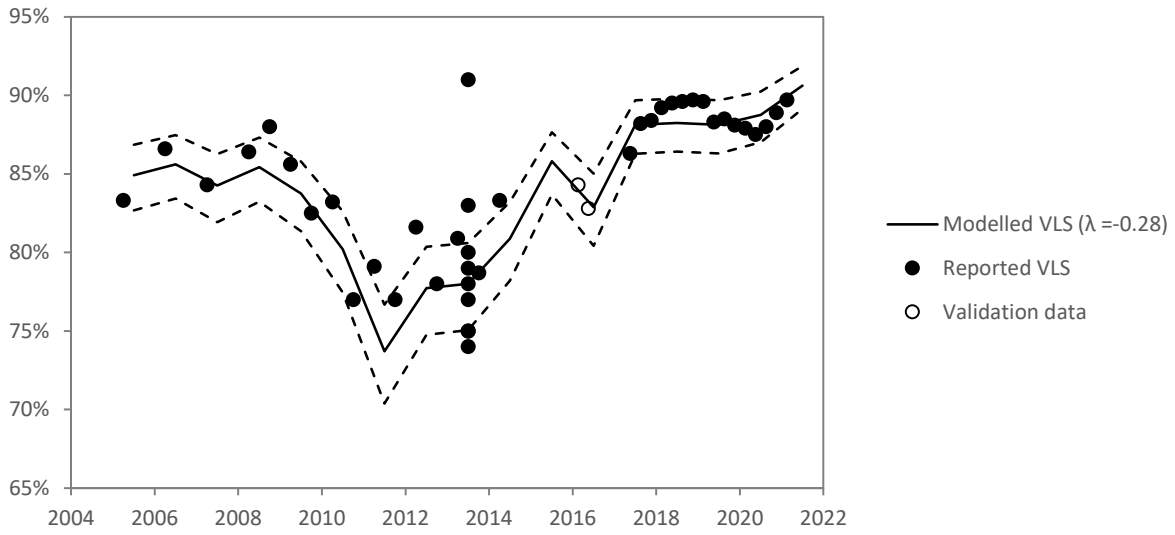

Figure 3.1: Viral suppression trends (% of ART patients with VL <400 RNA copies/ml)
In 2013 there are several data points, as rates of viral suppression were reported at a number of different ART durations; in most other periods the data relate to viral load testing around 6 months after ART initiation. Dashed lines represent 95% confidence intervals calculated using the 2.5 and 97.5 percentiles of λ_p. Model results are calculated using the regression coefficients in Table 3.3, adjusted by λ_p (equation 3.2) and weighting by the Thembisa estimates of the fraction of ART patients starting ART in each CD4 category in the corresponding year.

3.3.2 Viral suppression in children

For children, we lack reliable nationally representative data on rates of viral suppression. The District Health Information System (DHIS), which summarizes the data from TIER, reports overall rates of viral suppression, but not disaggregated by age group. Our approach is therefore to use IeDEA paediatric ART data to estimate time trends in viral suppression [65], and then to adjust these rates using the same adjustment factors that we use in adults (i.e. using the same formula as in equation 3.2). This means that we are not attempting to 'fit' routine viral suppression data, in the way we do for adults. Table 3.4 summarizes the IeDEA-SA estimates of viral suppression in children, and shows the implied rates of viral suppression for different values of λ_p.

Table 3.4: Viral suppression in children (<400 RNA copies/ml)

	IeDEA-SA	Model estimate		
Year	estimate	λ_p = -0.28	λ_p = -0.44	λ_p = -0.12
2005	73.7%	68.0%	64.3%	71.4%
2006	78.2%	73.1%	69.7%	76.1%
2007	78.0%	72.9%	69.5%	75.9%
2008	82.0%	77.5%	74.5%	80.1%
2009	79.5%	74.6%	71.3%	77.5%
2010	73.5%	67.8%	64.1%	71.2%
2011	67.9%	61.6%	57.6%	65.3%
2012	68.0%	61.7%	57.7%	65.4%
2013	64.3%	57.7%	53.6%	61.5%
2014	70.4%	64.3%	60.4%	67.9%
2015	67.7%	61.4%	57.4%	65.1%
2016	66.2%	59.7%	55.6%	63.5%
2017	71.9%	66.0%	62.2%	69.5%
2018	70.6%	64.5%	60.6%	68.1%

3.3.3 Updating the assumptions about viral suppression in Thembisa

In the calibration of the Thembisa model to provincial HIV prevalence data, we allow for uncertainty regarding the λ_p parameter, as the extent of viral suppression influences the trends in HIV incidence. The posterior distributions in Table 3.3 become the prior distributions in this calibration process, i.e. using a Bayesian updating process to estimate λ_p. (In the first step, the estimates of λ_p are determined by province-specific or national viral suppression data, and in the updating step the λ_p parameters are determined by the province-specific HIV prevalence trends.)

We define $V'_{t,s}(p)$ as the Thembisa estimate of the true rate of viral suppression in patients on ART in year t, who started ART in CD4 category s and who currently live in province p. This is calculated as

$$\text{logit}(V'_{t,s}(p)) = \text{logit}(0.5 \times [I_{t,s} + I_{t+1,s}]) + \lambda_p,$$

for $t \geq 2005$ and $t \leq 2018$. This is similar to the equation in (3.2), but we average across calendar years t and $t + 1$ for the purpose of calculating the Thembisa estimates because the projection years in Thembisa run from mid-year to mid-year. In the period from 2020 onward, rates of viral suppression are assumed to increase as a result of the introduction of dolutegravir to replace efavirenz and nevirapine in first-line ART regimens. A recent network meta-analysis estimates that patients receiving dolutegravir are significantly more likely to achieve viral suppression than patients receiving efavirenz (OR 1.87, 95% CI: 1.34-2.64) [66]. We have used this odds ratio to determine the rates of viral suppression in 2021 and subsequent periods, since efavirenz has been the main first-line antiretroviral drug in South Africa up to 2019. In the 2019 projection year (which runs from mid-2019 to mid-2020) we assume that on average 20% of adult ART patients are on dolutegravir; this proportion is then assumed to increase to 40%, 60%, 80% and 100% in the 2020, 2021, 2022 and 2023 projection years respectively (the limited available data suggest a very slow transition toward dolutegravir). No dolutegravir

adjustments are made in the case of children, as this drug is only recommended for individuals over 20 kg.

3.4 Rates of ART interruption

In earlier versions of Thembisa, rates of ART interruption were assumed to be the same across provinces, with this parameter being fixed at a rate of 0.25 per year. However, in some provinces this led to an implausible 'levelling off' in ART coverage in recent years. We have therefore allowed this parameter to vary in the model calibration, at a provincial level. To represent the prior uncertainty around this parameter, we assign a gamma prior distribution, with a mean of 0.25 interruptions per annum, and a standard deviation of 0.10. This distribution has 2.5 and 97.5 percentiles of 0.094 and 0.481 respectively. This is based on an analysis of data from three different South African studies of ART interruption rates [67-69], which estimated ART interruption rates of 0.06-0.19 per annum, though these were noted to be under-estimates [70]. The same prior distribution is used in all provinces.

4. Initializing HIV prevalence and modelling HIV transmission

In this analysis, HIV transmission probabilities per act of sex are assumed to be the same across all provinces (after controlling for differences by age, sex, risk group, HIV disease stage and relationship type). Although there is evidence of differences in the incidence of STIs between South Africa's provinces [71, 72], which might be expected to lead to differences in HIV transmission probabilities, it is not clear whether these STI prevalence differentials would remain after controlling for the behavioural differences that are already allowed for in the model.

It is assumed that initial HIV prevalence levels in 1985 differ between provinces. Although almost all of the earliest AIDS cases in South Africa were from the urban centres in Western Cape, Gauteng and KwaZulu-Natal, most of these early AIDS cases occurred in white men who reported having sexual contact with other men [73], and these statistics are therefore unlikely to be representative of the early HIV epidemic in the heterosexual population. There is thus substantial uncertainty regarding relative levels of HIV prevalence in the early HIV epidemic.

In the national version of the Thembisa model, the epidemic was seeded by specifying the HIV prevalence level among women aged 15-49 in the high-risk group in 1985. A prior distribution, uniform on the range from 0 to 0.2%, was specified to represent the uncertainty regarding this initial HIV prevalence parameter. The upper limit of 0.2% was calculated by estimating the rate of growth in antenatal HIV prevalence in the first three national antenatal surveys (1990-92), and back-projecting the antenatal HIV prevalence to 1985, supposing that all of the initial HIV infections occurred in the high-risk group, and noting that antenatal HIV prevalence in 1985 would probably not have been less than the HIV prevalence among women in the general population. In the provincial versions of the Thembisa model, we attempt to apply the same methodology to setting the priors on the initial HIV prevalence levels, but with a few minor modifications. Firstly, because the antenatal survey sample sizes are smaller at the provincial level, it is more difficult to estimate the epidemic growth rate precisely, and we therefore use the data from the first five antenatal surveys (1990-94) to back-project the antenatal HIV prevalence in 1985. Secondly, because the high-risk proportion is assumed to differ between provinces, we multiply the female high-risk proportion for the country as a whole (25%) by the province-specific adjustment factor (the prior means shown in Figure 2.1). In the case of the North West province, it was not possible to apply this methodology due to the extreme variability in HIV prevalence in this province in the first few antenatal surveys. Figure 4.1 shows the back-projected estimates of the HIV prevalence in pregnant high-risk women in 1985, in the remaining eight provinces. In all provinces except KwaZulu-Natal, the prevalence was well below the previously-assumed upper bound of 0.2%.

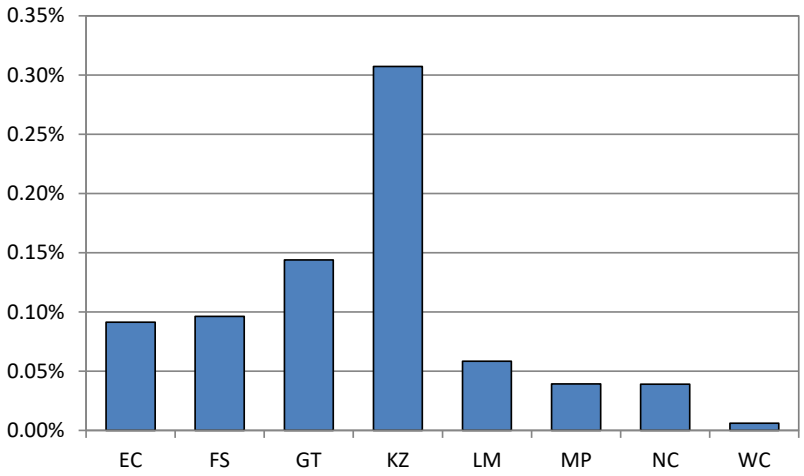

Figure 4.1: Back-projected HIV prevalence in high-risk pregnant women in 1985

Although these back-projected prevalence estimates are generally considered to be upper bounds on the prevalence in non-pregnant high-risk women, the uncertainty regarding the high-risk proportions in each province and the relatively small sample sizes associated with the early antenatal surveys mean that there are substantial margins of uncertainty around these 'bounds'. We have therefore used the same prior distribution as assumed nationally for Eastern Cape, Free State, Limpopo, Mpumalanga, Northern Cape and North West, i.e. uniform on the range from 0 to 0.2%. For KwaZulu-Natal, we have assumed a prior distribution that is uniform on the range 0 to 0.4%, and for Gauteng the prior distribution is uniform on the range from 0 to 0.3%. For Western Cape, initial HIV prevalence appears to have been much lower than in the other provinces, and we therefore assign a prior distribution that is uniform on the range from 0 to 0.05%.

5. Paediatric HIV assumptions

5.1 Breastfeeding by undiagnosed HIV-positive mothers

In the national model, we assume that 86.7% of undiagnosed HIV-positive mothers initiate breastfeeding after birth, and that the duration of breastfeeding by these undiagnosed mothers is Weibull-distributed (with a median of 18 months and a shape parameter of 2). There are likely to differences in breastfeeding durations across provinces, as women tend to breastfeed for longer durations in rural areas than in urban areas [31]. We therefore specify a multiplicative adjustment to the national median of 18 months, for each province, to allow for inter-provincial differences in the breastfeeding duration.

We estimate these multiplicative adjustments from the 1998 and 2003 DHSs, using data on both the proportions of all women who breastfeed and the median duration of breastfeeding. Although we also have data from the 2016 DHS, we have not included these because in 2016 a substantial fraction of mothers had been diagnosed positive, but it is not possible to tell which mothers had been diagnosed positive, and including these mothers would bias the estimate of the median breastfeeding duration because HIV-diagnosed mothers tend to breastfeed for shorter durations. In 1998 and 2003, relatively few mothers would have been diagnosed positive, and we assume that the breastfeeding durations of the HIV-positive mothers were similar to those in HIV-negative mothers. Table 5.1 summarizes the data that are used to estimate the province-specific multiplicative adjustments. The table also shows the ratio of the provincial breastfeeding estimates to the corresponding national average, and the average ratio for each province. In most provinces the average ratio is close to 1, suggesting little difference in breastfeeding durations relative to the national average. However, breastfeeding durations appear to be somewhat longer in Limpopo (a mainly rural province) and shorter in North West province.

Table 5.1: Breastfeeding practices, by province

	1998 DHS		2003 DHS	Ratio or provincial BF to national BF			
	Ever BF	Median BF*	Ever BF	1998 ever	1998 median	2003 ever	Average
EC	90.0%	16.6	77.4%	1.04	1.06	0.95	1.02
FS	89.5%	15.0	78.8%	1.03	0.96	0.97	0.99
GT	83.1%	14.3	82.4%	0.96	0.92	1.01	0.96
KZ	87.5%	14.9	76.7%	1.01	0.96	0.94	0.97
LP	95.5%	19.5	93.0%	1.10	1.25	1.14	1.16
MP	91.8%	16.5	87.8%	1.06	1.06	1.08	1.06
NC	86.3%	15.0	90.1%	1.00	0.96	1.11	1.02
NW	61.5%	14.1	54.4%	0.71	0.90	0.67	0.76
WC	85.4%	10.4	87.1%	0.99	0.67	1.07	0.91
SA	86.7%	15.6	81.5%				

* Median in months. BF = breastfeeding.

To represent the uncertainty around the multiplicative adjustments, we assign gamma prior distributions. The mean of the gamma prior distribution is set to the value estimated in the final

column of Table 5.1, and we assume in each province that the prior standard deviation is 10% of the mean. This 10% coefficient of variation was chosen such that the prior 95% confidence intervals included approximately 95% of the ratios estimated in Table 5.1. Figure 5.1 shows the prior means and 95% confidence intervals, together with the ratios from Table 5.1.

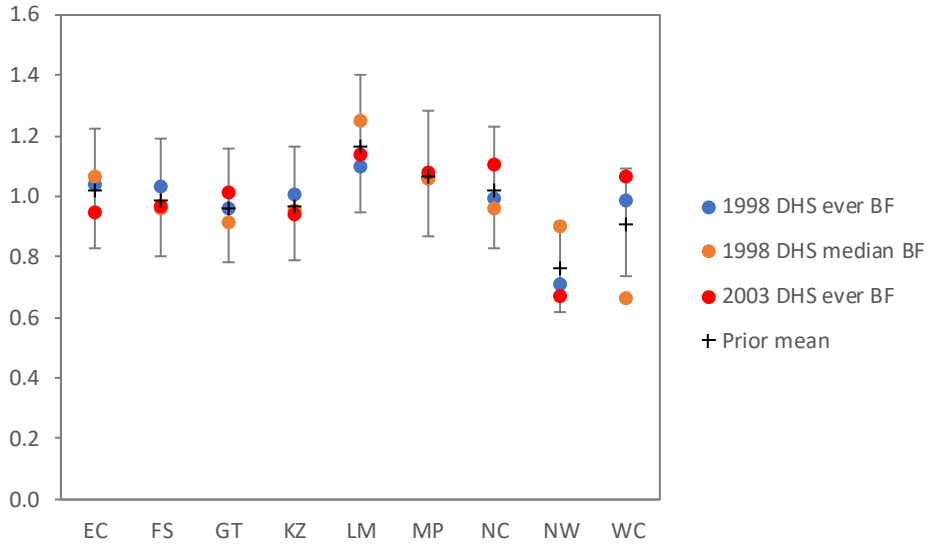

Figure 5.1: Relative rates of breastfeeding, by province

5.2 Prevention of mother-to-child transmission

The assumptions regarding the proportion of HIV-positive mothers who receive HIV testing in each province are the same as those in the ASSA2008 model up to 2007/08 [74]. These assumptions were set based on the data from the District Health Barometer reports from 2004 [75], after correcting errors in some districts; the results of two health facility surveys over the 2002/03 period [76, 77]; and an initial assessment of the PMTCT pilot programme in 2001/02 [78]. The assumed rates for 2008/09 have been adapted from the rates reported in the 2008/09 District Health Barometer [79]. The rates from the subsequent District Health Barometer reports were judged to be implausible, as estimates of HIV testing rates exceeded 100% in many districts. We have therefore set the assumed rates for 2009/10 and 2010/11 by interpolating between the assumed rates for 2008/09 and an ultimate rate of 98% in 2011/12. 98% is the ultimate rate assumed in the national model, and the same rate is assumed to apply in all provinces in all years following 2011/12. Figure 5.2 shows the assumed provincial rates of HIV testing in antenatal clinics over the 1999-2011 period.

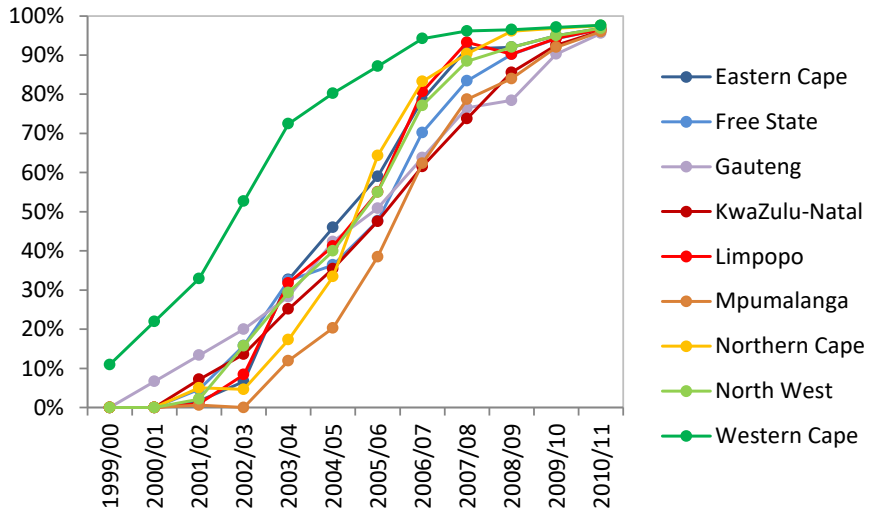

Figure 5.2: Proportion of pregnant women tested for HIV

The proportions of women receiving single-dose nevirapine who are assumed to receive AZT are also the same as those assumed in the ASSA2008 model. These assumptions have been set so that the proportions are zero prior to 2007/08 in all provinces other than the Western Cape. (In the Western Cape a dual regimen of AZT and nevirapine was provided to pregnant mothers from May 2004 [80], but in other provinces this change occurred only in early 2008 [81].) By 2010/11 it is assumed that the proportion of those receiving NVP who also received AZT had risen to 90% in all provinces, consistent with the data from the national PMTCT survey conducted in 2010 (Kate Kerber, personal communication). Following the introduction of Option B in 2013, it is assumed that the provision of short-course AZT and single-dose nevirapine was phased out in all provinces.

The proportion of mothers who chose not to breastfeed varied substantially between provinces in the period between 2002 and 2011, when free formula milk was offered to HIV-positive mothers. Studies estimated high rates of formula feeding in the Western Cape (average 86%, range 66-99%), and intermediate rates in Eastern Cape (average 63%, range 35-95%) and KwaZulu-Natal (average 52%, range 19-79%) [82-89]. However, none of these studies systematically sampled health facilities, and the results therefore cannot be considered representative at a provincial level. The only representative data on infant feeding practices by HIV-positive mothers, prior to the phasing out of free formula milk, are the data from the national PMTCT survey conducted in 2010 [90]. Table 5.2 shows the proportion of mothers of HIV-positive infants who reported not breastfeeding their infants when interviewed at 4-8 weeks after birth. The actual proportion of HIV-positive mothers who never breastfed is likely to be somewhat lower, given that many HIV-positive mothers initiate breastfeeding at birth but discontinue soon afterwards. Given the default Thembisa assumptions about rates at which HIV-positive mothers discontinue breastfeeding, we have calculated the proportions never breastfeeding that would need to be assumed at birth in order to yield estimates consistent with the proportions not breastfeeding at 6 weeks (shown in the last row of Table 5.2). These proportions are roughly consistent with the previously-quoted means estimated from other studies.

Table 5.2: Proportion of diagnosed HIV-positive mothers who were not breastfeeding (pre-2011)

	EC	FS	GT	KZ	LM	MP	NC	NW	WC
At 4-8 weeks*	67.3%	58.6%	67.9%	61.3%	53.3%	59.1%	31.7%	50.8%	85.9%
At birth	62.0%	51.9%	62.7%	55.0%	45.7%	52.4%	20.6%	42.8%	83.6%

* Source: Data from 2010 national PMTCT survey provided by Kate Kerber

For the period prior to 2011, the model assumptions about the fraction of HIV-positive mothers who never breastfed have been set equal to the values in the last row of Table 5.2. It is assumed that following the phasing out of free formula milk from August of 2011, this proportion decreased linearly to 20% by mid-2013, in each province.

5.3 Paediatric HIV disease progression and mortality

In the absence of ART, HIV-positive children are assumed to progress through two HIV stages of disease prior to AIDS mortality: early disease and late disease, the latter defined in terms of the ART eligibility criteria previously used in the 2006 WHO paediatric ART guidelines [91]. Consistent with the approach adopted for adults, we allow for inter-provincial differences in the rates of HIV disease progression and mortality in the absence of ART. We do this by multiplying the rates of disease progression and mortality in the national model (for children with untreated HIV disease) by a province-specific adjustment factor.

Several studies have suggested that there are substantial inter-regional differences in untreated paediatric HIV disease progression and mortality. For example, studies conducted in high-income settings prior to the availability of ART found that vertically-infected infants had a 6-26% probability of death in the first year of life [92-95], compared to a rate of around 33% in African settings [96]. High levels of immune activation among African children [97], high levels of malnutrition [98-100] and high incidence of other infectious diseases in African settings are possible explanations for the relatively high disease progression and mortality rates. However, we lack reliable data on the natural history of HIV in South African children, and we therefore assign the same prior distribution in all provinces when representing the uncertainty around the province-specific adjustment factor. This prior is a gamma distribution with a mean of 1 (implying no difference, on average, between the provincial and the national average disease progression and mortality rates) and a standard deviation of 0.15.

5.4 Paediatric HIV diagnosis

Rates of early infant diagnosis (EID) differ substantially between provinces [101, 102]. We have set the assumed proportion of HIV-exposed infants who are diagnosed at 6 weeks after birth to be the same as the rates of PCR coverage estimated in the 2012 Annual Health Statistics report [101], and using NHLS data from the 2013/14 District Health Barometer [103]. As data are missing for the 2012-13 year, the model assumption for this year is set by interpolating between the 2011/12 and 2013/2014 estimates. Figure 5.3 shows the assumed fractions of HIV-exposed infants who are diagnosed 6 weeks after birth. Rates of diagnosis have generally been highest in Western Cape and Gauteng, the two provinces that are most urbanized. Although the

Department of Health estimates relate only to the period from 2008 to 2011, we have assumed that the proportion was zero in 2003 and increased in proportion to the national growth rates estimated by Sherman *et al* [102] over the 2006-2008 period. Public sector data in 2014/15 are not plausible (the reported rate of PCR testing at birth is 100.6% [104]), and the fraction of HIV-exposed infants who are tested at 6 weeks is therefore assumed to have increased to 92% in 2014/15, in all provinces.

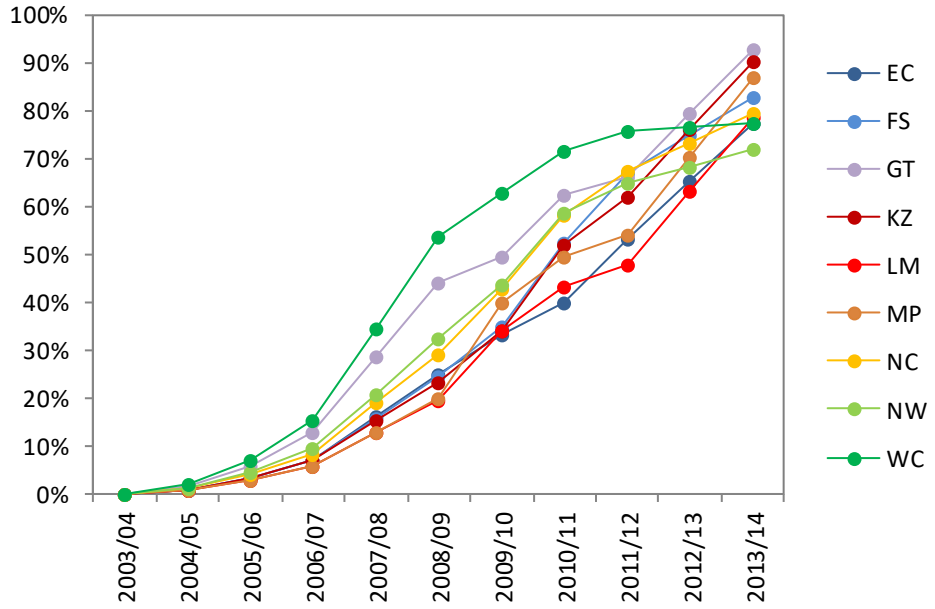

Figure 5.3: Fraction of HIV-exposed infants who are tested and receive their test results soon after birth

In 2015, guidelines on EID changed to recommend that all HIV-exposed infants be tested at birth and again at 10 weeks (if not diagnosed positive at birth), with no testing scheduled at 6 weeks. The model has been extended to allow for testing at both time points, although the model time step is not short enough to distinguish 6 weeks and 10 weeks. The change in guideline has created difficulties in terms of monitoring and evaluation, as the monitoring system is not based on a unique patient identifier, and this makes it impossible to determine how many of the individuals tested at 10 weeks had already been tested at birth and how many were tested at 10 weeks for the first time [105]. However, it is possible to estimate the fraction of HIV-exposed infants who were tested at birth from NHLS data, as shown in Table 5.3.

Table 5.3: Fraction of HIV-exposed infants tested for HIV at birth in 2015-16

	EC	FS	GT	KZ	LM	MP	NC	NW	WC
At birth	48.9%	56.7%	71.3%	82.7%	61.1%	60.6%	61.6%	69.1%	68.7%

Source: NHLS data, as reported in 2015/16 District Health Barometer [105]

More recent data suggest that coverage of birth testing increased to over 90% nationally in 2016-17 [106], and we therefore assume that the same level of coverage (90%) is achieved in all provinces in 2016-17 and subsequent years. Due to the lack of reliable estimates of the fraction of infants who are tested at 10 weeks, the assumed fraction tested at 10 weeks is set to

80% in 2015-16 and subsequent years, the same assumption as has been made in the national model based on Western Cape data [107].

Guidelines also recommend testing at 18 months, but provinces have differed in their approach to 18-month testing: some appear to test only the children born to HIV-diagnosed mothers, while others aim to test all children (including those whose mothers have no HIV diagnosis).

5.5 Paediatric ART initiation

The model allows for both ART initiation within the same month as diagnosis ('immediate' ART initiation) and deferred ART initiation. As in the national version of the model, we assume that the probability of ART initiation in the month of diagnosis is 10 times the monthly probability of linkage thereafter, and we assume that the rate of ART initiation after diagnosis in early HIV disease is 0.5 times that in late disease. The $\rho_1(t)$ parameter represents the monthly rate of ART initiation in children in late HIV disease, who were diagnosed previously and did not start ART in the month of diagnosis, in year t.

We adopt a similar approach to that described for adults, which is to specify $\rho_1(t)$ for different periods, defined by ART availability and ART eligibility criteria. ART initiation rates are assumed to be piecewise-constant over the mid-2000 to mid-2004 period (before the launch of the national public sector programme, when ART access was limited to the private sector), over mid-2004 to mid-2005 (the first year of the public sector roll-out), over mid-2009 to mid-2010 (the last year before the first major change in paediatric ART eligibility criteria), over mid-2010 to mid-2016 (a period of expanding paediatric ART eligibility criteria), and over the period after mid-2016 (when there was universal ART eligibility). Over the 2005-2009 period, rates of ART initiation are interpolated linearly from those specified in the 2004-05 and 2009-10 years. To present this in more mathematical terms:

$$\rho_1(t) = 0 \qquad \text{for } t < 2000$$
$$\rho_1(t) = \rho_1(2000) \qquad \text{for } 2000 < t \leq 2003$$
$$\rho_1(t) = \rho_1(2004) + (\rho_1(2009) - \rho_1(2004))(t - 2004)/5 \qquad \text{for } 2004 < t < 2009$$
$$\rho_1(t) = \rho_1(2010) \qquad \text{for } 2010 < t < 2016$$
$$\rho_1(t) = \rho_1(2016) \qquad \text{for } t > 2016$$

We specify gamma prior distributions for the $\rho_1(t)$ parameters, with means and standard deviations estimated from the provincial estimates obtained previously in Thembisa version 4.4 [70]. Prior means are set at 0.003 for 2000, 0.043 for 2004, 0.177 for 2009, and 0.110 for 2020 and 2016. Corresponding standard deviations are 0.001, 0.075, 0.193 and 0.099 respectively.

For children who are in the early stage of disease and eligible to start ART, the monthly rate of ART initiation (if they did not start ART at the time of diagnosis) is $\theta\ \rho_1(t)$, where θ is the relative rate of ART initiation in early disease (relative to late disease). One might expect a lower rate of ART initiation in children who are not yet symptomatic or severely immunosuppressed, but we lack reliable local data to inform this parameter. We therefore assign a uniform (0, 1) prior distribution to represent the uncertainty in this parameter, and estimate the parameter separately for each province.

6. HIV testing, male circumcision and PrEP

6.1 HIV testing and counselling

The method used to model the adult uptake of HCT in the national version of the Thembisa model has been described elsewhere [9]. Briefly, it is assumed that adults can be tested for HIV in three different ways: HIV testing during pregnancy, HIV testing in patients with HIV-related opportunistic infections (OIs), and HIV testing for other reasons. As described in section 5.1, the model allows for provincial variation in rates of antenatal HIV testing, based on data from the District Health Barometer reports. We currently lack province-specific data on the rate of HIV testing in patients with OIs, and have therefore set the assumed proportions of OI patients tested to be the same as assumed in the national model, for all provinces. However, we do have reasonable estimates of the total numbers of HIV tests performed in each province, and have therefore set the province-specific rates of testing for other reasons in such a way that the model estimate of the total number of HIV tests is consistent with the estimates derived from other data sources.

The set of province-specific estimates of total HIV tests were derived by disaggregating our previously-estimated total numbers of HIV tests for the country as a whole [9]. These totals were derived for the public health sector, medical schemes, the life insurance industry, and other private providers of HIV testing (e.g. workplace HIV testing programmes). Most of the public health sector statistics include provincial disaggregation, and these were used to calculate the numbers of individuals tested for HIV in the public sector in each province [108]. In the case of medical schemes, data on the provincial profile of individuals tested was not directly available. However, information on provincial differences in rates of HIV testing in 2011 is available for the Discovery medical scheme, the largest medical scheme in the country [109]. We assume that the province-specific rates of HIV testing are the same in other medical schemes, and combining these rates with data on the total number of medical scheme beneficiaries in each province at the end of 2011 [110] we calculate the expected fraction of all medical scheme HIV testing in each province. Information is also available on the provincial profile of HIV testing by insurers [111] and other private providers [112]. For all three private sector data sources, the fraction of HIV tests in each province that was estimated was assumed to apply in all years, due to the lack of information on temporal changes in provincial distributions.

Figure 6.1 shows the resulting estimates of the numbers of HIV tests performed in each province in each year. Gauteng and KwaZulu-Natal, the two provinces with the largest populations, account for the greatest numbers of HIV tests performed in most years. Testing numbers reached their highest level in 2019-20, then dropped substantially over the 2020-21 year following the start of the COVID-19 epidemic. The proportionate reductions in total numbers of tests performed, when comparing 2020-21 to the previous year, were greatest in Northern Cape (35%) and Gauteng (34%) and smallest in North West (1%).

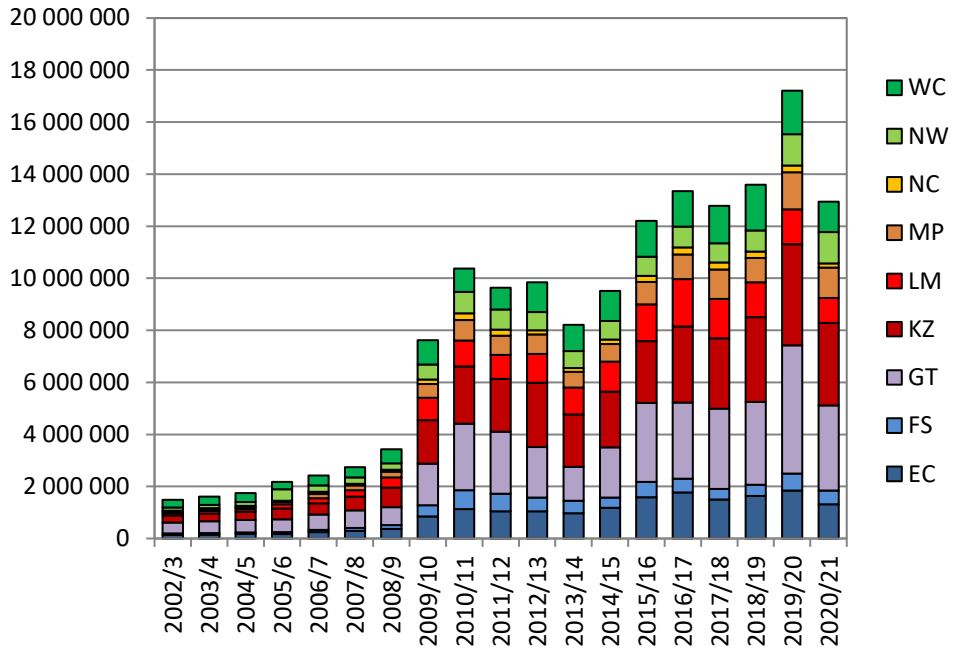

Figure 6.1: Numbers of HIV tests performed in each province, 2002-2019

In the period prior to 2002, we lack data on the number of tests performed. We therefore make the same assumption as that made in the national model, i.e. that the number of HIV tests performed in each province increased linearly over time, from zero in 1990 to the level in 2002, as shown in Figure 6.1.

The data and assumptions presented previously all relate to HIV tests performed by health workers. In Thembisa version 4.5 we have also included self-testing, and the model allows for five different strategies for the distribution of self-testing kits. Programme data from the HIV Self-Testing Africa (STAR) Initiative were used to estimated numbers of people who received self-testing kits in each province (Mohammed Majam, personal communication), and for each distribution strategy. These data are summarized in Table 1. These could be under-estimates, as the data exclude sales of self-testing kits through pharmacies and other private providers. Overall, the numbers of self-testing kits distributed are quite small relative to the numbers of tests performed by health workers (Figure 6.1), although Gauteng is an exception.

Table 6.1: Assumed numbers of self-testing kits distributed

Province	2017-18	2018-19	2019-20
EC	0	2565	4756
FS	2354	31776	41560
GT	249425	393887	167734
KZ	3662	11177	13652
LP	0	1935	1935
MP	26385	60926	45158
NC	0	0	0
NW	8495	10810	13878
WC	0	0	0
Total	290320	513074	288672

In the period after 2019-20, we lack data on the numbers of self-testing kits distributed. In the national version of the model, we have assumed that the annual self-testing rates in the period after 2020 are the same as the average rates calculated over the 2017-2020 period [113]. In the provincial version of the model we use the same rates for the post-2020 period for all provinces, and set these equal to the rates estimated in the national version of the model.

6.2 Male circumcision

6.2.1 Uptake of male circumcision prior to the promotion of male circumcision as an HIV prevention strategy

In the national version of the Thembisa model, the proportion of men aged x who were circumcised prior to the promotion of MMC is modelled using the function

$$p(x) = a + (b-a)\left(1 - 0.5^{(x/m_1)^{\phi}}\right),$$

where a is the proportion of males who are circumcised soon after birth, b is the proportion of males who ever get circumcised, m_1 is the median age at circumcision in men who get circumcised after infancy, and ϕ is the shape parameter that controls the variance of the distribution of ages at circumcision after infancy. At a provincial level, we estimate these parameters by fitting the above functional form to age-specific estimates of the prevalence of male circumcision in 2007, prior to the rollout of VMMC promotion campaigns, as estimated by Thomas *et al* [114]. The estimates of Thomas *et al* are based on statistical models fitted to self-reported data on circumcision status and age at circumcision in nationally-representative household surveys. Because the estimates rely on self-reported data, and self-reporting of male circumcision status is known to be unreliable, we adjust the estimates of Thomas *et al* using an assumed sensitivity and specificity of 96.4% and 88.4% respectively, the same adjustments as used in the national version of the model.

Table 6.2 shows the resulting estimates for the model parameters, and Figure 6.2 shows the model calibration to the adjusted estimates from Thomas *et al* [114].

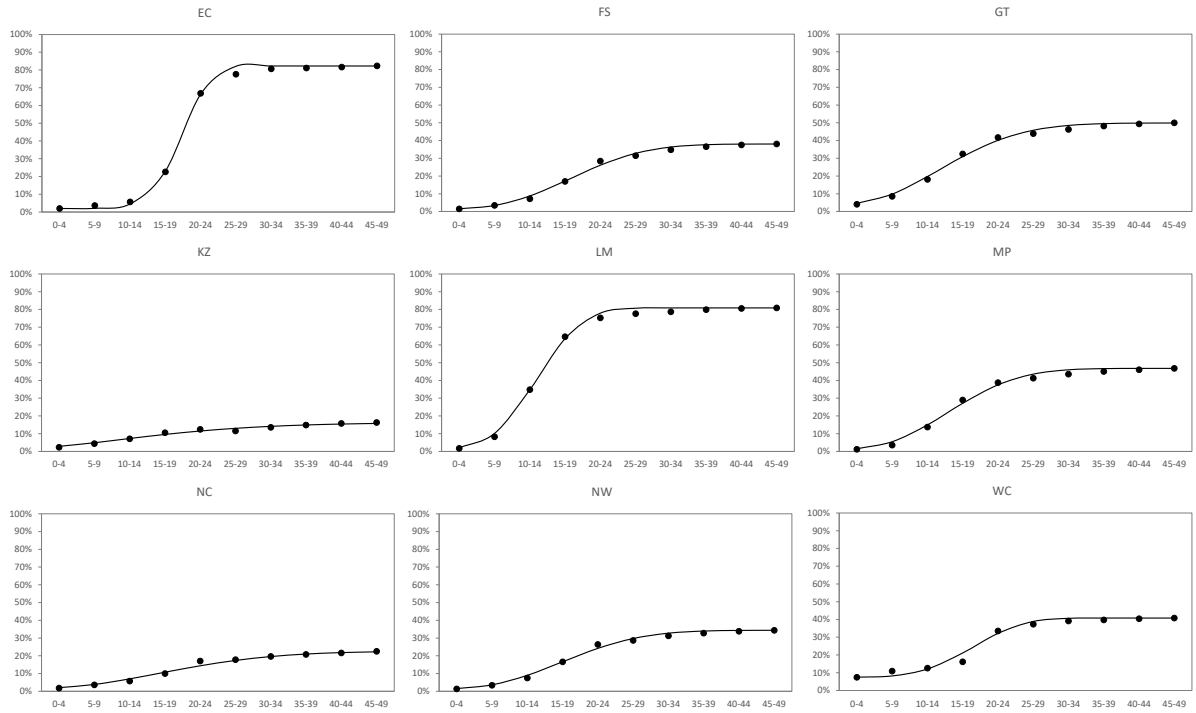

Figure 6.2: Proportions of men circumcised, prior to the promotion of MMC

Dots represent estimates of prevalence of male circumcision in 2007, from Thomas *et al* [114], after adjusted for assumed sensitivity and specificity of self-reporting of 96.4% and 88.4% respectively in males aged 25+. In males aged <15 no adjustment is made, as circumcision in these younger age groups is rare and published estimates of the reliability of self-report, based on adult men, are unlikely to be applicable.

Table 6.2: Male circumcision parameters for different provinces

Province	Proportion circumcised at birth a	Proportion ultimately circumcised b	Median age of circumcision after infancy m_1	Shape parameter ϕ
Eastern Cape	2.0%	82.3%	19.8	6.65
Free State	1.5%	38.0%	18.9	2.74
Gauteng	4.1%	49.9%	15.7	2.22
KwaZulu-Natal	2.3%	16.3%	16.9	1.52
Limpopo	1.7%	80.9%	13.6	3.07
Mpumalanga	1.2%	46.8%	16.2	2.54
Northern Cape	1.8%	22.5%	19.4	2.00
North West	1.2%	34.4%	18.3	2.55
Western Cape	7.4%	40.8%	18.8	3.74

6.2.2 Uptake of medical male circumcision

In the national Thembisa model, annual numbers of medical male circumcisions are specified as inputs, for each year from 2008-09 to 2020-21. For the purpose of parameterizing the provincial models, these numbers are distributed between provinces, making use of disaggregated data as far as possible. In the three years for which the provincial disaggregation is missing (2008, 2009 and 2012), the national MMC total has been allocated between provinces in proportion to the average provincial totals over the 2010, 2011 and 2013-2015

years. Table 6.3 shows the assumed annual numbers of MMC operations in each year, by province. There have been dramatic reductions in MMC numbers in 2020-21 as a result of COVID-19, with numbers dropping by more than two thirds in all provinces except Gauteng (where there appears to have been only a 14% drop). The numbers assumed in Limpopo and Mpumalanga provinces are slightly lower than assumed in previous versions of the Thembisa model, over the 2015-2019 and 2013-2019 periods respectively. This is because some of the MMC operations performed in these provinces were in fact medical circumcisions in traditional settings, and since these circumcisions are already included in the 'background' circumcision rates (as specified in section 6.2.1), we would be double-counting if we included them here. These numbers of medical circumcisions in traditional settings are based on the estimates of Thomas *et al* [114], and are 4-16% of the annual MMC totals in Limpopo and 3-9% of the annual totals in Mpumalanga.

Table 6.3: Assumed numbers of medical male circumcisions occurring in each province

Year	EC	FS	GT	KZ	LM	MP	NC	NW	WC	SA
2008/09	310	317	1381	1595	473	600	62	306	146	5190
2009/10	548	560	2439	2818	835	1060	110	540	258	9168
2010/11	27913	10001	18253	43460	22972	5223	1157	1726	413	131117
2011/12	74079	26541	48441	115338	60966	13862	3070	4579	1097	347973
2012/13	25256	25799	112319	129780	38452	48841	5080	24861	11874	422262
2013/14	4929	13712	115732	107132	10357	44483	4086	19637	9906	331668
2014/15	10533	26260	146275	139046	29481	84064	8043	41658	20042	508404
2015/16	16333	44480	126742	161329	68174	37875	4863	35598	17215	518130
2016/17	13306	33232	121586	131419	54170	37552	2694	30392	12576	446678
2017/18	8789	35303	112699	200463	46398	75477	5252	28039	16557	540327
2018/19	11397	31250	102331	209709	82021	77313	10970	43373	14218	595006
2019/20	10722	20616	75789	139910	35774	49822	8586	48071	18124	417138
2020/21	3527	2474	65038	26985	1006	16169	847	9582	3961	129587

In the national model, it is further assumed that after mid-2021, the rate at which uncircumcised males aged 10-14 get circumcised reverts to the average of the annual probabilities estimated over the period from mid-2015 to mid-2020 (i.e. before the impact of COVID-19). In the provincial versions of the Thembisa model, we adopt a similar approach, calculating the average rate of male uptake of MMC in the 10-14 age group over the period from mid-2015 to mid-2020 and assuming that this is the average rate that applies from mid-2021. These rates that apply after 2021 are shown in Table 6.4. Rates of MMC over the 2015-20 period having been highest in Limpopo and Mpumalanga, and lowest in the Western Cape and Northern Cape.

Table 6.4: Assumed long-term annual MMC probabilities in males aged 10-14

Province	EC	FS	GT	KZ	LM	MP	NC	NW	WC
Probability	0.359	0.387	0.334	0.358	0.898	0.439	0.092	0.202	0.048

6.3 Pre-exposure prophylaxis (PrEP)

Our approach to modelling PrEP uptake is described more fully in the national modelling report [19]. Most parameters relating to PrEP eligibility and relative rates of PrEP uptake in different risk groups are assumed to be the same across provinces, and the values are specified in section 4.8 of the national report. However, there some exceptions are made in the case of MSM

(slightly higher rates of uptake in Gauteng and Western Cape and lower rates in Limpopo) in order to avoid inconsistencies with PEPFAR programme data on the numbers of MSM/males on PrEP. In addition, we assume a slightly later start to PrEP rollout among adolescent girls and young women (AGYW) in Limpopo, Northern Cape and Western Cape, to avoid AGYW on PrEP estimates that are out of line with PEPFAR programme data. The relative rate of PrEP uptake in other risk groups (outside of AGYW, MSM and FSWs) is also assumed to be lower over the 2019-2021 period in three provinces (Eastern Cape, KwaZulu-Natal and Limpopo) in order to avoid estimates of men on PrEP that were very inconsistent with PEPFAR programme data. Because the PEPFAR programme data are not representative of all PrEP recipients, this adjustment process was subjective, and the modelled PrEP estimates by risk group should be treated with caution (especially at a provincial level).

Annual numbers of people receiving PrEP, which are used to determine the rates at which HIV-negative sex workers initiate PrEP, are assumed to differ by province, and are shown in Table 6.5. Most of the early data were not disaggregated by province, and the estimates for the early years should therefore be treated with caution. Since 2020, public sector PrEP numbers have been disaggregated by province, and private sector data for 2020 are also disaggregated by province (Kerensa Govender, personal communication).

Table 6.5: Assumed numbers of people receiving PrEP at the middle of each year

Province	EC	FS	GT	KZ	LM	MP	NC	NW	WC
2016	21	8	142	66	5	21	1	11	30
2017	297	107	2031	940	77	298	21	159	431
2018	701	253	4796	2219	183	703	49	375	1018
2019	2369	856	16204	7496	618	2377	165	1266	3438
2020	6372	2474	29196	22757	1453	9840	274	2182	5615
2021	25176	15217	78914	84411	5360	50513	414	11438	11256

7. Demographic assumptions

A brief description of the derivation of the demographic assumptions included in the model is given below. While these assumptions could be refined further, any adjustments in future are not expected to have significant effect on the estimates of the unfolding of the HIV/AIDS epidemic and the likely success or otherwise of various interventions.

7.1 Population

The initial population numbers in 1985, by sex and individual age from 0 to 89 and open interval 90+, were set the same as those from the ASSA2008 provincial models [74].

7.2 Fertility

Estimates of the number of births by calendar year were estimated from the same sources as used to produce the national estimates, except that births recorded by DHIS were not used prior to 2006, as they appeared to be particularly under-registered prior to that in some provinces. The best estimate was produced using the numbers of births implied by census 2011 counts or school enrolment numbers up to the early to mid-2000s (if these estimates appeared more plausible), and after that DHIS data, and from 2011 projections from the CARe model[1] were included. The average of these estimates was then used to measure the completeness of the vital registration (VR) and the estimates were adjusted to ensure (to the extent that it is possible) that the completeness of the deaths reported up to and including a given year of registration was not implausible.

The estimates of the numbers of births from the different sources were less consistent with one another for the provinces than nationally, and so in some instances judgement was required to give preference to one source over another, or bring the one source in line with the others by assuming change in the level of reporting over time. This process of reconciliation proved particularly problematic for two provinces (Northern Cape and Limpopo), so there remains some uncertainty about the reliability of the estimates of the numbers of births in these provinces.

TFRs for the projection years (from the middle of one year to the middle of the next) were linearly interpolated from the estimates by census year (i.e. from census anniversary in one year to the census anniversary in the next).

TFRs for the period from the middle of 1985 to middle of 1996 were estimated to be the linear trend from the ASSA2008 estimate for 1985 to the estimate for the year starting at the middle of 1996, produced above. Kinks in the estimates in the period 1998 to 2003 were smoothed to produce a smooth trend over time.

[1] The CARe workbook is a simplified model which concentrates on the demographic impact of HIV/AIDS. It uses the same demographic assumptions as the Thembisa model together with output from the Thembisa model that allows for the incorporation of the impact of HIV/AIDS on births and deaths into the simplified model.

The age-specific fertility rates were set by applying proportions of the TFR at each individual age to TFRs for each year from 1985 to 2011.

The proportions of the TFR attributable to single ages for 1996, 2001, 2006, 2011 and 2015 were derived from the proportions attributable to five-year age groups reported by women in each census and the 2007 and 2016 Community Surveys using Beers subdivision. The proportions for individual years between the census/survey years were derived by linear interpolation. For the projection years 1985-1995, the ASSA2008 ASFRs were rescaled to the TFRs estimated above.

For the version 4.5 update, the estimates for five of the provinces (EC, FS, GT, KZ and NW) up to and including projection year 2016 were assumed to be the same as in the previous versions of the model (version 4.1-4.4) and for the years after that it was assumed that fertility remained level for projection years 2017-2020, which brought the numbers of births projected by the model more in line with the estimates of the true number of births over this period. For MP previous assumptions were accepted to 2015 and assumed to remain level from 2016-2020, and for WC the previous assumptions were accepted to 2011 and the fertility was assumed to remain unchanged after that.

The remaining provinces (NC and LM) were more problematic. For both, the numbers of births estimated by the model have been tracking below (approximately 5-15% for LM and 10-15% for NC) the estimate of the true number of births since the 2011 census. This implies that either the number of women aged 15-49 is too low or the fertility rates in the model are too low (which is a possibility in NC), or both. At this stage, it is difficult to make sense of why these problems might exist in these provinces (particularly LM, unless there are migrants not being captured in the census but giving birth to children being captured by the DHIS). It was therefore decided, for both these provinces, to accept previous estimates up to and including 2011 and assume that the fertility did not decline between 2011 and 2020.

Beyond the 2020 projection year, age-specific fertility rates are assumed, for all provinces, to decline to a common set of ultimate fertility rates at annual rates of decline. The assumed ultimate rates and annual rates of decline are both the same as assumed in the ASSA2008 models.

7.3 Non-HIV mortality rates

The age-specific probabilities (q_x) of non-HIV/AIDS mortality for 1997-2010 were derived from the central mortality rates ($_nm_x$) for all-cause and HIV-specific mortality from the 2010 National Burden of Disease (NBD) study [115]. First m_0, $_4m_1$, $_5m_5$, … $_5m_{80}$, and m_{85+} were derived by subtracting the HIV/AIDS-specific rates from the all-cause rates. Next, because of the erratic nature of the rates at the older ages, the rates above age 65 were smoothed to follow the curve of the average rates by age over the period, scaled to the level of the rates in each year. Following this Beers interpolation was applied to the rates from $_4m_1$ to m_{85+} to produce rates at individual ages from 2 to 87. These rates were then converted to probabilities of death for ages 2 to 80. Probabilities above age 80 were derived from extrapolated central mortality rates assuming that rates followed a Gompertz curve, increasing by 9% per year of age, to further reduce fluctuation over time and age. Finally q_0 was set equal to 1-exp(-0.983m_0) and

q_1 was set equal to 1-exp(-3.985$_4m_1$)/[(1-q_2)(1-q_3)(1-q_4)], where m_0 and $_4m_1$ were the rates derived from the NBD estimates.

An initial set of non-HIV/AIDS mortality rates incorporating the impact of the SARS-CoV-2 epidemic on mortality for projection years 2019-2021 was derived using estimates of HIV/AIDS mortality from the Thembisa 4.3 model. The numbers of HIV/AIDS deaths from Thembisa 4.3 were reduced to reflect that some would have died of COVID instead. The rates were then developed by first deriving a set of all-cause mortality, less excess deaths[2] rates for projection years 2018-2021 for age groups 0, 1-4, 5-9, …80-84, 85+. The adjusted HIV/AIDS mortality rates from Thembisa 4.3 were then deducted from these rates, and smoothed single-age mortality rates derived from the results using abridged life tables for these years and a full single-age life table of non-HIV/AIDS for 2015. Finally, single-age specific excess death rates were then added back to the non-HIV/AIDS mortality rates for projection years 2018-2021. Rates for projection years 2016 and 2017 were set by interpolating between rates for 2015 and 2018.

These estimates were then used as input to the Thembisa 4.5 model in order to generate age and sex-specific HIV/AIDS mortality. These were then used together with the estimates of all-cause age and sex-specific mortality rates from the Thembisa 4.5 model to produce revised estimates of non-HIV/AIDS mortality rates in a similar process, except without the need for the adjustment to allow for a reduction in HIV/AIDS due to the increase in deaths from COVID-19 and other causes of excess deaths.

Two problems were apparent from the first iteration, namely, (a) that the estimates of the excess deaths from the weekly RMS appeared to be too high for the Northern Cape and Limpopo and, to a lesser extent, too low for Gauteng, and (b) interpolation between projection years 2015 and 2018 produced too sharp a change in rates over time. Thus, it was decided to reduce the level of the non-HIV/AIDS mortality rates for projection years 2019-2021 for Northern Cape and Limpopo and increase them for Gauteng to bring them more in line with the trend in rates for the provinces in prior years. In addition, it was decided to use 2007 instead of 2015 as the base-year for smoothing and interpolate the rates for projections year 2008 to 2017.

Inspection of the results identified ‘kinks’ in the progression by age for age 1 in the Eastern Cape for males and females in 2020 and Western Cape for males aged 1 in 2020, and in the Northern Cape ages 76-86 for males and 62-78 for females in 2020 and 2021. These were smoothed by interpolation between rates either side of these irregularities.

Although there is great uncertainty as to the impact of the SARS-CoV-2 epidemic on non-HIV mortality in future years, it does not seem appropriate to assume a general long-term decline in mortality from projection year 2023 onwards. Thus, it was decided to use the following estimates of mortality rates (all years referring to projection years):
Rates for 2022 = 0.25 × (3 × 2021 rates + 2018 rates)
Rates for 2023 = 0.50 × (2021 rates + 2018 rates)
Rates for 2024 = 0.25 × (2021 rates + 3 × 2018 rates)
Rates for 2025 = 2018 rates.

[2] From the data underlying “Report on weekly deaths in South Africa” prepared by the SAMRC and the Centre for Actuarial Research (CARe), UCT: https://www.samrc.ac.za/reports/report-weekly-deaths-south-africa?bc=254.

Rates for 2011 to 2015 were set as those projected using the CARe_3.2 provincial models, assuming no HIV/AIDS.

Probabilities of death for 1985 were set to those from the ASSA2008 model and for 1986 to 1996, the probabilities of death were determined by linear interpolation between the estimates for 1985 and 1997.

Beyond projection year 2026, non-HIV/AIDS mortality rates are assumed to decline to a common set of ultimate rates at age-specific rates of decline. The ultimate rates and annual rates of decline are the same as assumed in the ASSA2008 models.

7.4 Migration

An important feature of these projections is that the population is according to the provincial boundaries as at the middle of the year in question and not (in the pre-2011 period) according to the boundaries at the time of the 2011 census. This is the same approach as used for ASSA2008 and the alternative mid-year estimates [116] and thus estimates of migration from those sources incorporate the major boundary changes by treating the change in population as migration in the year in which the change occurred. (This differs from the approach adopted in the official mid-year population estimates produced by Stats SA, which backdates boundary changes to the start of any projection series.)

The numbers of migrants were set in two stages. Initially the numbers of migrants by sex and single age for each year 1985 to 2000 were set equal those from the ASSA2008 models. The numbers for 2001-2010 were set as per those used to produce the alternative mid-year estimates [116]. These numbers were derived from the change in the numbers of people by place of birth (province or outside South Africa) between censuses, less an estimate of the number of South African-born emigrants as captured by censuses in the main countries of destination (UK, Australia, New Zealand, USA and Canada), scaled to match the total numbers recorded in the official mid-year estimates [117].

After this, these numbers were adjusted by an age-specific number (fixed over time) so that the projection of the population to the middle of 1996, 2001 and 2011 matched the census counts (approximately, for example, allowing for differences that might be expected due to errors in the census, such as undercounting of children or age exaggeration at the old ages). Although adjustment of migration of those born in the intercensal period was avoided as far as possible, there are probably some instances where the adjustment of migration compensated for errors in fertility. These adjustments were made at a provincial level, with national net immigration being the sum of the resultant provincial net in-migration.

Generally, these adjustments were determined by subtracting the projected numbers in five-year age groups before adjustment from the census count in these age groups. ${}_5M_{x\text{-}5}$ was set to $({}_5P^c_x - {}_5P_x)/10$, ${}_5M_0$ to $({}_5P^c_5 - {}_5P_5)/5$ and M_{85+} to ${}_5M_{80}$, where ${}_5P_x$ represents the number of people in the population aged between x and x+5, the superscript c represents the census count and ${}_5M_x$ represents the additional number of migrants aged between x and x+5 required for the adjustment. The age range requiring adjustment for each census was limited to that needed to correct for major deviations in one census from what would be expected given the other two,

on the assumption that the estimates of migration reported by census questions are likely to be less accurate than the census counts.

The extent of adjustments varied by province as follows:

- *Northern Cape*: No adjustment was made since the comparison of the projections to the census numbers were inconclusive about whether any of the censuses were more reliable than the projected populations.
- *Western Cape, Mpumalanga and Limpopo*: The 2001-2011 migration numbers (aged 10+, 10+ and 15+ respectively), by five-year age group, were adjusted to reproduce the 2011 census numbers.
- *Free State and KwaZulu-Natal*: No adjustment to the migration for 2001-2011 but, for Free State, extensive adjustment (1985-1996 migration at all ages was adjusted to reproduce the numbers in the 1996 census, and increased in the 1996-2001 at ages 0-14 to reproduce the numbers in the 2001 census) and, for KwaZulu-Natal, less extensive adjustment (1985-1996 migration adjusted to match the numbers 0-9 in the 1996 census, and the 1996-2001 migration to match the numbers 5-14 in the 2001 census).
- *Eastern Cape*: Increased the 1996-2001 migration to match the numbers 5-9 in 2001, and the 2001-2011 migration numbers to match the numbers 10+ in 2011.
- *North West*: Adjusted the 1996-2001 migration to match the numbers 0-24 in 2001 and the 2001-2011 migration to match the numbers at all ages in the 2011 census.
- *Gauteng*: This province required the most extensive adjustments to migration to reproduce the census numbers. The 1991-1996 migration was adjusted to reproduce the numbers in all age groups of a re-estimate of the population in 1996. The re-estimated population numbers were derived as an average of the estimate from ASSA2008 and estimates derived by back projecting the numbers from the 2001 census. In addition the 2001-2011 migration was adjusted to match the numbers at all ages in the 2011 census.

The numbers at each age for 2011 to 2015 were set equal to those for 2010.

In the absence of data and sufficient research on the impact of the COVID-19 epidemic and interventions to manage it on migration, we adopt the crude approach of simply scaling the net numbers of migrants by age nationally and provincially for each sex in projection year 2016 in the same ratios as the net arrivals (i.e., arrivals less departures) recorded at the country's borders for calendar years 2017 to 2021 (as published monthly by Statistics South Africa [118]). In order to estimate the scaling factors the following assumptions were made:

1. The level of national and provincial migration in projection year 2017 was the same as that assumed previously for 2016.
2. The net numbers of arrivals for the first six months of calendar year 2022 was the same as that for the first six months of calendar year 2017.
3. The level of national and provincial migration in projection year 2022 will be the same as that in 2016.

This resulted in the following scaling factors (net arrivals in projection year relative to those in projection year 2017) of 1.0, 1.2, 1.0, 0.1, 0.8 and 1.0 for projection years 2017, 2018, 2019, 2020, 2021 and 2022, respectively.

Beyond projection year 2022, the numbers at each age are assumed to trend asymptotically to zero at a rate of 4.5% per annum.

The model also allows for the possibility that migrants into a province may have a different HIV profile from current residents. The assumptions about relative rates of HIV prevalence in migrants are explained in more detail in Appendix A.

8. Model fitting and uncertainty analysis

The calibration of the Thembisa provincial model follows a two-step process. In the first step, we aim to fit the model to adult HIV data sources, accounting for the uncertainty regarding the key adult HIV parameters. Once the adult parameters have been estimated, we aim to calibrate the model to paediatric HIV data sources, allowing for uncertainty regarding the main paediatric parameters. The calibration process is broadly similar in each step, and we therefore describe the two calibration processes together in the sections that follow.

8.1 Likelihood function

Five data sources are used when fitting the adult model:

1. the HIV prevalence levels in the antenatal clinic surveys (1994-2015 and 2017),
2. the HIV prevalence levels in the HSRC household surveys (conducted in 2005, 2008, 2012 and 2017) and the 2016 Demographic and Health Survey (DHS),
3. the numbers of ART patients (2000-2020),
4. the household survey estimates of the fraction of HIV-positive individuals on ART (2012 and 2017),
5. the proportions of ART patients who are male (2012 and 2017-2019), and
6. the recorded numbers of deaths in adults.

Two data sources are used when fitting the paediatric model:

1. the HIV prevalence levels in the HSRC household surveys (conducted in 2005, 2008, 2012 and 2017), and
2. the numbers of ART patients (2000-2020).

The sections that follow explain how the likelihood function is defined for each of these data sets.

8.1.1 Antenatal clinic HIV prevalence data

The antenatal likelihood is calculated by comparing model estimates of HIV prevalence in pregnant women and corresponding survey estimates. The model is calibrated to the age-specific HIV prevalence data, which are available for the period 1994-2015 and 2017. Suppose that $H_{i,x,t}(\boldsymbol{\varphi})$ is the model estimate of HIV prevalence in pregnant women in province *i*, in age group *x* and year *t*, where the vector $\boldsymbol{\varphi}$ represents the values of the model input parameters. The corresponding prevalence of HIV actually measured in the antenatal survey is represented by $y_{i,x,t}$. It is assumed that if $\boldsymbol{\varphi}$ is the true set of parameter values, then the difference between the logit-transformed model estimate and the logit-transformed observed prevalence is normally distributed with zero mean. The variance of the distribution is assumed to be composed of a 'survey error' term, representing the uncertainty around the survey estimate due to binomial variation and cluster variation in the survey, and a 'model error' term, which represents the possible effect of mis-specified fertility adjustments in HIV-positive women. More formally, it is assumed that

$$\log\left(\frac{y_{i,x,t}}{1-y_{i,x,t}}\right) = \log\left(\frac{H_{i,x,t}(\boldsymbol{\varphi})}{1-H_{i,x,t}(\boldsymbol{\varphi})}\right) + m_{i,x,t} + \varepsilon_{i,x,t},$$

where $m_{i,x,t} \sim N\left(0, \sigma_i^2\right)$ and $\varepsilon_{i,x,t} \sim N\left(0, \sigma_{i,x,t}^2\right)$. The latter two terms represent the model error and the survey error respectively. The logit transformations ensure that the error terms are closer to normality and that the model error terms are roughly independent of the level of HIV prevalence.

In the previous versions of Thembisa, the above equation was modified to include an 'antenatal bias' term, which represented the difference in HIV prevalence between pregnant women and women of the same age in the sexually experienced population, after controlling for HIV effects on fertility. This bias arose in part because of differences in HIV prevalence between pregnant women attending private and public antenatal services (since surveys represent only the public sector) but also because of unmeasured behavioural confounding (independently of the direct effect of HIV on fertility and sexual activity, HIV-positive women might be expected to have different sexual behaviour from HIV-negative women). In the new version of Thembisa, we attempt to account for these sources of bias more explicitly, by specifying relative levels of HIV prevalence in private and public antenatal clinic attenders, effects of HIV diagnosis on fertility, and differences in fertility between undiagnosed HIV-positive women in early HIV infection and HIV-negative women (for a more complete description, see sections 6.2.2 and 7.2.1 of the national report [70]). Whereas we previously specified a prior distribution to represent the uncertainty around the antenatal bias parameter in each province, we now instead specify a prior distribution to represent the uncertainty around the ratio of fertility in undiagnosed HIV-positive women in early-stage HIV infection to fertility in sexually-experienced HIV-negative women of the same age. This prior distribution is a gamma distribution with a mean of 1.30 and a standard deviation of 0.13 (which has 2.5 and 97.5 percentiles of 1.06 and 1.57 respectively). We based this prior on previous arguments that the true value of this ratio was likely to lie between 1.12 and 1.50, based on relative rates of pregnancy in HIV-negative and HIV-positive women in the Western Cape [70, 119].

The $\sigma_{i,x,t}^2$ parameters have been estimated from the published 95% confidence intervals around the antenatal survey estimates, in 1998 and subsequent years. Prior to 1998, the published 95% confidence intervals were calculated on the assumption of simple random sampling (SRS), i.e. not reflecting the clustering associated with the sampling of antenatal clinics. As these confidence intervals would have exaggerated the precision associated with the prevalence estimates, we recalculated the standard errors by inflating the published standard errors up to 1998 by province-specific adjustment factors. The adjustment factor was calculated as the average ratio of the published standard error to the SRS standard error over the 2003-2005 period (the only period for which we had sufficient data to calculate both standard error estimates in all provinces). The $\sigma_{i,x,t}^2$ parameters up to 1998 were then estimated from these inflated standard error estimates. Confidence intervals have not been reported for the 2014 and 2015 survey estimates, and the standard errors in these two years have therefore been assumed to be the same as in 2013.

The variance of the model error (σ_i^2) is assigned a value of 0.3^2, which was chosen in order to achieve an appropriate weighting of the antenatal HIV prevalence data and other data sources (discussed below) while avoiding overly narrow confidence intervals around the model HIV prevalence estimates.

The likelihood in respect of the antenatal data is then calculated based on the assumption that the error terms are normally distributed:

$$L_i(\mathbf{y}_i \mid \boldsymbol{\varphi}) = \prod_x \prod_t \left(2\pi\left(\hat{\sigma}_i^2 + \sigma_{i,x,t}^2\right)\right)^{-0.5} \exp\left[-\frac{\left(\text{logit}\left(y_{i,x,t}\right) - \text{logit}\left(H_{i,x,t}(\boldsymbol{\varphi})\right)\right)^2}{2\left(\hat{\sigma}_i^2 + \sigma_{i,x,t}^2\right)}\right],$$

where $\mathbf{y}_i$ represents the vector of $y_{i,x,t}$ values, across calendar years 1994 to 2015 and 2017.

8.1.2 Household survey HIV prevalence data

The approach followed in defining the likelihood in respect of the HSRC household survey data is similar to that for the antenatal data, with a few key differences. Firstly, we calculate the likelihood separately for 15-24 year olds and adults aged 25 and older, as these are the published age disaggregations (in the paediatric calibration, we consider only the prevalence in 2-14 year olds). Secondly, the model error term ($m_{i,x,t}$) is omitted from the expression for the observed prevalence, because any error introduced by mis-specified assumptions about fertility in HIV-positive women would have minimal effect on estimates of HIV prevalence in the general population.

The approach taken in defining the likelihood function for the 2016 DHS data is the same as that for the HSRC survey data. The 2016 DHS did not include HIV testing in children, and is therefore not used in the paediatric analysis.

For validation purposes, we have also included the results of a 2003 national survey, conducted among youth aged 15-24 [120], although the data from this survey are not included in the definition of the likelihood function.

8.1.3 Numbers of ART patients

The interpretation of public sector ART statistics is challenging because ART reporting systems have changed a number of times since 2009, and the way in which totals are reported is not always consistent across provinces (or even within provinces). The Comprehensive Care, Management and Treatment (CCMT) reporting system initially reported *cumulative* numbers of patients started on ART for all provinces other than the Western Cape (which has always reported numbers of patients *currently* on ART). However, late in 2009 most provinces switched to reporting numbers of patients currently on ART. Since 2011 most government publications have quoted ART statistics from the District Health Information System (DHIS) and the TIER system, which are supposed to report the numbers of patients currently on ART [63]. However, because of delays in switching to the new reporting system, many clinics were

not included in the early reports of DHIS statistics, and the DHIS statistics are therefore probably not representative of the whole public health sector prior to 2012.

To account for the change in reporting in 2009 (from cumulative to current enrolment) and the uncertainty around the speed of the change in reporting, we define τ_p as the time up to which all reported public sector totals in province *p* are known to represent cumulative ART enrolment (in most provinces this will be some time in 2009). After time τ_p, there is uncertainty as to whether the reported public sector totals represent cumulative enrolment, current enrolment, or some combination of the two. We define θ_p to be the annual change in the fraction of public ART services that report current enrolment, after time τ_p. In other words, if $f(t, p)$ is the fraction of public ART services that report cumulative enrolment at time *t*, then

$$f(t,p) = \begin{cases} 1 & \text{if } t < \tau_p \\ \exp\left(-\theta_p\left(t-\tau_p\right)\right) & \text{if } t \geq \tau_p \end{cases}$$

For the purpose of this analysis, we fix the τ_p and θ_p values at the values estimated previously [16]. These are summarized in Table 8.2. Note that results are not shown for Western Cape because this province has always reported current enrolment rather than cumulative enrolment (i.e. $f(t, p) = 0$ for all *t* in the Western Cape).

Table 8.2: Transitions from cumulative enrolment to current enrolment

	EC	FS	GT	KZ	LP	MP	NC	NW
τ_p	August 2009	August 2009	August 2009	August 2009	August 2009	August 2009	August 2009	June 2010
$\exp(-\theta_p)$	0.1401	0.0001*	0.0001*	0.5059	0.4952	0.5449	0.0001*	0.3223

* Values of 0.001 imply immediate transition from reporting cumulative enrolment to current enrolment.

In the current report, we update the calibration using public sector statistics for each province and each month in the April 2018 to March 2019 period (personal communication, Thapelo Seatlhodi), and using 2018-2019 private sector statistics provided by the Council for Medical Schemes (personal communication, Bilia Luwaca, South African National AIDS Council).

The likelihood function represents how well the model fits the reported numbers of patients on ART. For the purpose of calculating the likelihood, we assume that the error terms (the differences between the modelled numbers of patients on ART and the corresponding reported numbers of ART patients, on a log scale) are normally distributed with zero mean and variance σ_m^2.

Firstly, we define the reported ART total at time *t*, in province *p*, to be the sum of the totals reported for the private and public sectors:

$$\Omega(t,p) = R_0^0(t,p) + R_0^1(t,p).$$

where $R_c^s(t,p)$ represents the reported level of ART enrolment at time *t*, in group *c* (0 for adults and children combined, 1 for children only) in sector *s* (0 for public, 1 for private/NGO), and in province *p*. Because the private/NGO numbers are small relative to the public sector numbers, it is the latter that we are most interested in when calibrating the model. The private

sector totals are therefore included in the calibration only at the mid-year time points prior to the start of the public sector ART rollout in 2004, and at the time points for which public sector statistics are reported in subsequent periods (for time points at which a public sector total is reported but there is no corresponding reported private sector total, the private sector total is approximated by linearly interpolating between the nearest reported private sector totals). We use the symbol $n_c(p)$ to represent the number of time points for which we have reported ART enrolment statistics for the public sector (or private sector pre-2004), and we use the symbol $T_c(p)$ to represent the set of these time points.

Secondly, we define $G(\boldsymbol{\Theta}_p, t)$ to be the model estimate of the number of patients we would *expect* to be reported as on ART at time *t*, if parameter combination $\boldsymbol{\Theta}_p$ represents the 'true' set of model parameters. This estimate depends on the assumed fraction of ART services that report cumulative enrolment, which in turn depends on the fraction of patients receiving ART through the public sector (since only the public sector facilities report cumulative enrolment). The Thembisa model does not directly simulate the fraction of patients receiving ART through the public sector, so we approximate this fraction by the quantity

$$\phi(t, p) = 1 - R_0^1(t, p) / N_0(t, p),$$

where $N_0(t, p)$ represents the model estimate of the number of patients currently receiving ART at time *t* in province *p*. If $\phi(t, p) > 0$, then we calculate $G(\boldsymbol{\Theta}_p, t)$ as

$$G(\boldsymbol{\Theta}_p, t) = R_0^1(t, p) + f(t, p) M_0(t, p) \phi(t, p) + (1 - f(t, p)) N_0(t, p) \phi(t, p),$$

where $M_0(t, p)$ represents the model estimate of cumulative ART enrolment in province *p* up to time *t*. This simplifies to $G(\boldsymbol{\Theta}_p, t) = N_0(t, p)$ if $f(t, p) = 0$. If $\phi(t, p) \leq 0$, this implies that all ART is provided through the private sector, which reports current ART enrolment. Thus we use the same simplified formula, $G(\boldsymbol{\Theta}_p, t) = N_0(t, p)$, if $\phi(t, p) \leq 0$.

As noted previously, we calculate the likelihood by assuming that the difference between $G(\boldsymbol{\Theta}_p, t)$ and $\Omega(t, p)$, on a log scale, is normally distributed with zero mean and variance σ_m^2. The variance is set to 0.1^2, which is equivalent to assuming a 10% coefficient of variation to represent the error in the model estimates. The likelihood function is then calculated as

$$L(\boldsymbol{\Omega}_p \mid \boldsymbol{\Theta}_p) = \prod_{t \in T_0(p)} \frac{1}{\sqrt{2\pi}\hat{\sigma}_m} \exp\left(- \frac{\left(\ln\left(G(\boldsymbol{\Theta}_p, t)\right) - \ln\left(\Omega(t, p)\right)\right)^2}{2\hat{\sigma}_m^2} \right),$$

where $\boldsymbol{\Omega}_p$ represents the vector of $\Omega(t, p)$ values, for all $t \in T_0(p)$.

8.1.4 Household survey ART coverage data

For each province, we calculate a likelihood to represent the goodness of model fit to 2012 and 2017 household survey estimates of the proportions of HIV-positive adults who are on ART. In both household surveys, the proportion of HIV-positive adults on ART was estimated based on testing for antiretroviral metabolites (efavirenz, nevirapine, lopinavir and other less commonly used drugs, i.e. accounting for most first- and second-line ART regimens) [3, 121]. Although the survey also collected self-reported data on ART use, we have not used these data in calibration, in part to be consistent with the methods used in the HSRC survey reports, and in part because there were high levels of non-response to questions about ART use (Jeffrey Eaton, personal communication). Estimates of coverage were also obtained separately for men and women in each province, so that the model was calibrated to a total of 4 data points in each province. We calculated the likelihood on the assumption that the difference between the survey estimate of ART coverage and the modelled ART coverage, on a logit scale, was normally distributed with zero mean and variance calculated from the 95% confidence interval around the survey estimate.

A number of other sources provide data on ART coverage based on self-reported receipt of ART; for example, the DHIS provides data on the proportion of HIV-positive pregnant women who report having started ART prior to their current pregnancy. We have not included these data in the model calibration, as we do not consider self-reporting of ART status to be reliable. For example, in the 2017 national antenatal survey, viral load testing was conducted in all HIV-positive pregnant women, and viral loads of less than 1000 RNA copies/ml were detected in 39% of women who reported not being on ART [122]; such a high rate of viral suppression does not seem plausible in HIV-positive women who are truly untreated. Several other South African studies have found a substantial prevalence of detectable antiretroviral metabolites in HIV-positive individuals who report being undiagnosed [9, 123-125], although one South African study found only minimal disagreement between self-reported ART coverage and the ART coverage based on self-report [126].

8.1.5 Proportions of adult ART patients who are male

A significant limitation of the DHIS monitoring system is that it does not provide a male-female disaggregation when reporting numbers of adults on ART. However, it is possible to infer relative rates of ART initiation in men and women from a number of alternative sources. Firstly, the ART coverage estimates from household surveys (described in section 8.1.4) are disaggregated by sex to provide rough estimates of relative rates of ART coverage in men and women. Secondly, a number of alternative data sources provide data on the proportion of ART patients who are male. Data obtained from the CCMT reporting system in 2012, prior to the transition to DHIS reporting, provide a male-female disaggregation for all provinces [101]. We also obtained data from the TB/HIV information system (THIS) on the proportion of adult ART patients who are male in 2018, for each province except the Western Cape. Separately, we obtained data on the proportions of adult ART patients who are male in the Western Cape, for each year from 2007 to 2019 (Themba Mutemaringa and Alexa Heekes, personal communication). We also obtained data from the National Health Laboratory Service in 2017 and 2019, on the proportion of patients receiving viral load tests who are male.

For the purpose of defining the model likelihood, we assume that the difference between the reported proportion of ART patients who are male and the model estimate for the corresponding year is normally distributed with a zero mean and a standard deviation of 0.01, over the 2017-19 period (and for the Western Cape the same standard deviation is used for all years). In the case of the 2012 CCMT data, a greater standard deviation is used (0.02), to reflect the greater uncertainty regarding the quality of the CCMT reporting, and also due to concerns that at this earlier stage in the ART rollout, the private sector accounted for a greater proportion of ART patients [127], which may have biased the model estimates for the population as a whole (relative to the public sector).

8.1.6 Recorded numbers of deaths in adults

Suppose that $N_g(x, t)$ represents the model estimate of the number of deaths (due to all causes) in adults of sex g, in age group x, in year t in the province of interest. Let $R_g(x, t)$ be the corresponding number of recorded deaths. We define $\gamma_g(x, t)$ as the completeness of death recording assumed at a national level (i.e. the assumed proportions of deaths that are recorded through South Africa's vital registration system). These completeness assumptions are described elsewhere [61]. Briefly, completeness is assumed to have increased over the 1997 to 2004 period [128], and then to have remained stable over time [129]. Completeness in adults is also assumed to increase with respect to age, and is assumed to be lower in men than in women, particularly at younger ages [129].

For the purpose of calibrating the provincial models to recorded death data, we further define the province-specific adjustment factor $P_g(t)$, which represents the ratio of the rate of death reporting in the province of interest to the rate of death reporting at a national level. This ratio is estimated by comparing the recorded number of deaths at ages 60 and older to the modelled number of deaths at ages 60 and older, after adjusting for the national level of vital registration completeness. The reason for limiting the analysis to ages 60 and older is that the contribution of HIV mortality to total mortality is likely to be small at these older ages, and thus any bias due to the misspecification of HIV in the initial model is likely to be minimal.

Mathematically, we calculate $P_g(t)$ as

$$P_g(t) = \frac{\sum_{x=60} R_g(x,t)}{H_g(t) \sum_{x=60} N_g(x,t)\gamma_g(x,t)},$$

where $H_g(t)$ is a national-level adjustment that we include to correct for differences between the assumed levels of completeness and the actual levels of completeness in older adults:

$$H_g(t) = \frac{\sum_{x=60} R'_g(x,t)}{\sum_{x=60} N'_g(x,t)\gamma_g(x,t)}.$$

$R'_g(x, t)$ and $N'_g(x, t)$ represent the national numbers of recorded and modelled deaths respectively (as distinct from the province-specific totals defined previously). Although $H_g(t)$

should in theory be close to 1 (if our national completeness assumptions are reasonable), the ratios are in fact consistently greater than 1, suggesting that our model under-estimates the completeness of death reporting at older ages, at a national level (Figure 8.1). We therefore include the $H_g(t)$ adjustment in order to 'correct' this error.

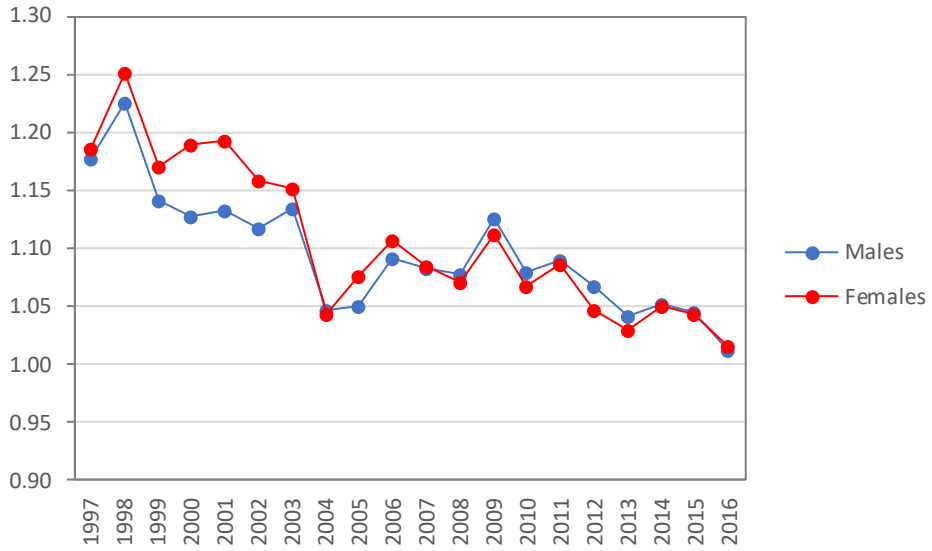

Figure 8.1: Ratio of actual completeness to assumed completeness in adults aged 60 and older, at a national level

Figure 8.2 shows the $P_g(t)$ adjustment factors estimated using the above equation. In the most urbanized provinces (Gauteng and Western Cape) these adjustments are relatively high in 1997, and then drop towards the national average (i.e. a ratio of 1), suggesting a 'catching up' in vital registration in the other provinces. In some provinces, the trend is irregular; for example in the Eastern Cape, vital registration appears to improve relative to the national average over the 1997-2005 period, but deteriorates thereafter. The opposite trend is seen in North West province.

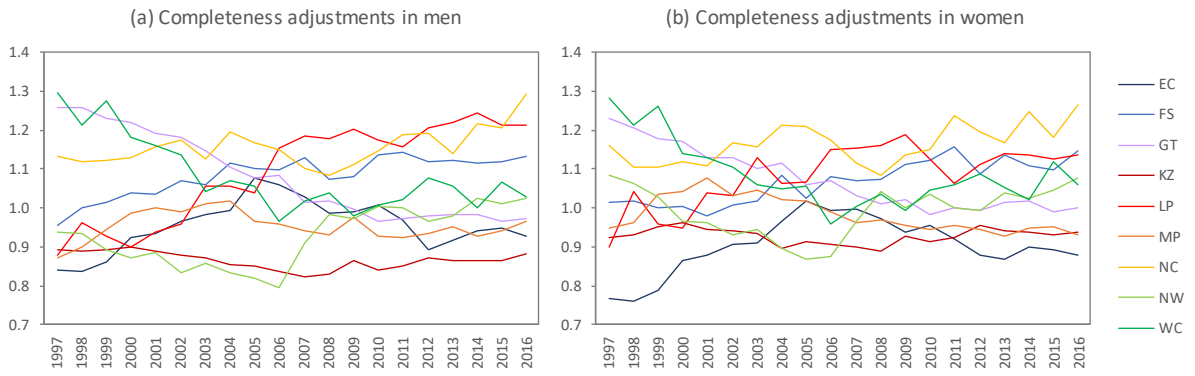

Figure 8.2: Ratio of death reporting in each province to national death reporting, in adults aged 60 and older

As in the national model, we consider deaths in each five-year age group, from 20-24 up to 55-59, over the 1997-2016 period. The recorded deaths are obtained from Statistics South Africa

reports of the numbers of deaths recorded in each province [130]. (We have tabulated the numbers by the province in which the death was recorded rather than the province of residence of the deceased. Although the latter might be considered more appropriate, information on the province of residence is missing for 13% of deaths, and the overall distribution of deaths by province of residence, after removing missing data, does not appear to differ substantially from the distribution of deaths by province of recording.)

Let $c_g(x, t)$ be the province-specific completeness/reporting adjustment. This is calculated as the product of $\gamma_g(x, t)$ and $P_g(t)$.

Suppose $A_g(x, t)$ represents the model estimate of AIDS deaths in individuals of sex g and age group x, in year t, and that $B_g(x, t)$ represents the model estimate of non-AIDS deaths (due to 'background' mortality), i.e. $A_g(x, t) + B_g(x, t) = N_g(x, t)$. In calculating the likelihood for the recorded death data, we assume that the difference between the modelled deaths and the recorded deaths (after completeness adjustment), on a natural log scale, is normally distributed with zero mean. The variance of this normal distribution is assumed to be of the form

$$\begin{aligned}\mathrm{Var}[\ln(N_g(x, t))] &\approx \mathrm{Var}[N_g(x, t)] / N_g(x, t)^2 \\ &= \mathrm{Var}[\ln(A_g(x, t))]\,(A_g(x, t) / N_g(x, t))^2 + \mathrm{Var}[\ln(B_g(x, t))]\,(B_g(x, t) / N_g(x, t))^2\end{aligned}$$

We have set the variance of the AIDS mortality term, $\mathrm{Var}[\ln(A_g(x, t))]$, to 0.3^2, to be consistent with the weight given to the antenatal survey data in the calibration procedure. The variance of the non-AIDS mortality, $\sigma_b^2 \equiv \mathrm{Var}[\ln(B_g(x, t))]$, has been estimated from the difference between the model predictions of mortality and the recorded levels of mortality at ages 60 and older (after adjusting for completeness), since the contribution of AIDS mortality to total mortality is expected to be small at these older ages.

$$\sigma_b^2 = \frac{1}{240}\sum_{x=60}\sum_{g}\sum_{t}\left(R_g(x,t) - N_g(x,t)c_g(x,t)\right)^2 ,$$

where 240 is the number of squared differences across which we are averaging (20 years, 2 sexes and 6 five-year age groups). Table 8.3 shows the estimated values of σ_b^2 for each province.

Table 8.3: Variance of non-HIV mortality (on natural log scale)

	EC	FS	GT	KZ	LP	MP	NC	NW	WC	ZA
Variance	0.034	0.023	0.018	0.025	0.038	0.035	0.035	0.033	0.016	0.020
SD	0.184	0.150	0.133	0.157	0.194	0.186	0.186	0.182	0.125	0.141

SD = standard deviation

As noted previously, the likelihood is calculated on the assumption that the difference between the log-transformed recorded number of deaths (after application of the completeness adjustment) and the log-transformed model estimate of deaths is normally distributed with zero mean and a variance of σ^2. More formally,

$$\ln\left(R_g(x,t)/c_g(x,t)\right) = \ln\left(N_g(x,t)\right) + \varepsilon_g(x,t)$$

where $\varepsilon_g(x,t) \sim$ N(0, Var[ln($N_g(x, t)$)]). The likelihood is calculated for all years from 1997-2016, for both sexes, and for each five-year age group, from 20-24 up to 55-59. We do not calculate likelihood values for ages 60 and older, partly because we expect relatively few AIDS deaths at the older ages (and hence these deaths would have little effect on the model calibration), and partly because we are already using the data at ages 60+ to estimate the completeness adjustments and the variance of the non-AIDS mortality terms.

8.2 Prior distributions

The prior distributions have been described and motivated in previous sections. Tables 8.4 and 8.5 summarizes the prior distributions assumed for all 17 parameters included in the adult uncertainty analysis, for each of the nine provinces. (Table 8.5 does not show parameters separately for each province, as the prior distributions for the ART initiation and ART interruption rates are the same for all provinces.)

Table 8.6 summarizes the prior distributions assumed for the eight parameters included in the paediatric uncertainty analysis, for each of the nine provinces. With the exception of the duration of breastfeeding (relative to the national average), the prior distributions are the same in all provinces.

Table 8.4: Prior distributions in adult calibration (mean and standard deviation in brackets)

Parameter	High risk adjustment factor	Sexual mixing parameter	Condom use adjustment factor	Mean age of female non-marital sex activity	SD age of female non-marital sex activity	Mean untreated HIV survival	OR viral suppression relative to IeDEA-SA	Initial HIV prevalence in high risk women, aged 15-49	Relative fertility in early undiagnosed HIV
Prior type	Gamma	Beta	Gamma	Gamma	Gamma	Gamma	Gamma	Uniform	Gamma
Report section	2.1	2.3	2.4	2.5	2.5	3.1	3.3.1	4	8.1.1
Prior mean (SD)									
EC	0.95 (0.24)	0.53 (0.12)	0.99 (0.149)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.63 (0.051)	0.10% (0.048)	1.30 (0.13)
FS	1.24 (0.31)	0.53 (0.12)	1.14 (0.171)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	1.03 (0.107)	0.10% (0.048)	1.30 (0.13)
GT	1.17 (0.29)	0.53 (0.12)	1.06 (0.159)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.67 (0.060)	0.15% (0.071)	1.30 (0.13)
KZ	1.16 (0.29)	0.53 (0.12)	1.08 (0.163)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.99 (0.097)	0.20% (0.095)	1.30 (0.13)
LM	1.06 (0.27)	0.53 (0.12)	1.04 (0.156)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.55 (0.035)	0.10% (0.048)	1.30 (0.13)
MP	1.03 (0.26)	0.53 (0.12)	1.02 (0.153)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.72 (0.054)	0.10% (0.048)	1.30 (0.13)
NC	0.52 (0.13)	0.53 (0.12)	0.64 (0.096)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.74 (0.093)	0.10% (0.048)	1.30 (0.13)
NW	0.97 (0.24)	0.53 (0.12)	1.05 (0.157)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	0.73 (0.068)	0.10% (0.048)	1.30 (0.13)
WC	0.65 (0.16)	0.53 (0.12)	0.74 (0.110)	37.8 (3.9)	20.7 (2.3)	12.0 (1.00)	1.14 (0.127)	0.03% (0.012)	1.30 (0.13)

OR = odds ratio, SD = standard deviation.

Table 8.5: Prior distributions in adult calibration: ART initiation and interruption rates

Parameter	Monthly rate of ART initiation in HIV-diagnosed women, with CD4 <200					RR male ART initiation	Reduction in ART start due to COVID-19	Rate of ART interruption
	2000	2004	2010	2011	2016			
Prior type	Gamma	Gamma	Gamma	Gamma	Gamma	Gamma	Beta	Gamma
Report section	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.4
Prior mean	0.0039	0.0154	0.0576	0.0622	0.0591	0.70	0.28	0.25
Standard deviation	0.0011	0.0117	0.0202	0.0196	0.0240	0.10	0.10	0.10

RR = relative rate.

Table 8.6: Prior distributions in paediatric calibration (mean and standard deviation in brackets)

Parameter	Relative duration of breastfeeding	Relative disease progression	Monthly rate of ART initiation in previously-diagnosed children in late disease					Relative ART initiation in early disease
			2000	2004	2009	2010	2016	
Prior type	Gamma	Gamma	Gamma	Gamma	Gamma	Gamma	Gamma	Uniform
Report section	5.1	5.3	5.5	5.5	5.5	5.5	5.5	5.5
Prior mean (SD)								
EC	1.02 (0.102)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
FS	0.99 (0.099)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
GT	0.96 (0.096)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
KZ	0.97 (0.097)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
LM	1.16 (0.116)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
MP	1.06 (0.106)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
NC	1.02 (0.102)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
NW	0.76 (0.076)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)
WC	0.91 (0.091)	1.00 (0.15)	0.003 (0.001)	0.043 (0.075)	0.177 (0.193)	0.110 (0.099)	0.110 (0.099)	0.50 (0.29)

8.3 Posterior analysis

The posterior distributions, representing the parameter sets that were most consistent with both the province-specific HIV data and the prior beliefs about the most plausible parameter values (Tables 8.4-8.6) were approximated using the Incremental Mixture Importance Sampling (IMIS) method [131]. In the adult analysis, an initial set of 10 000 parameter combinations were sampled from the prior distributions, for each province, and the likelihood function was calculated for each parameter combination. In subsequent IMIS steps, the regions of the parameter space with the highest posterior density were sampled more heavily, with an additional sample of 1000 parameter combinations being evaluated in each IMIS step. The procedure was repeated until a sufficiently mixed posterior sample was generated, containing 1 000 parameter combinations. The same procedure was followed in the paediatric analysis, except that a smaller initial sample (5 000 parameter combination) and smaller additional sample in each subsequent IMIS step (500) were used, due to the lower number of parameters in the paediatric analysis.

The 1 000 sampled values of the 17 adult parameters of interest were then merged with 1 000 sampled values of the six paediatric parameters, to generate a set of 1 000 combinations of the 19 parameters, for each province. All posterior means and 95% confidence intervals are calculated from this sample of 1 000 parameter combinations. The 95% confidence intervals around the model outputs thus reflect both the uncertainty in the adult parameters and the uncertainty in the paediatric parameters.

9. Results

9.1 Comparison of prior and posterior distributions

Figure 9.1 compares the prior and posterior distributions for the adult parameters included in the uncertainty analysis (excluding the ART initiation rates). In the case of the adjustment to the high-risk proportion (panel a), the posterior means are generally similar to the prior means, although the posterior is substantially higher than the prior mean in Mpumalanga, and substantially lower than the prior mean in Western Cape. The high-risk proportion is estimated to be at high levels in Mpumalanga and KwaZulu-Natal (1.3 times the national average) and at relatively low levels in Northern Cape and Western Cape (0.3-0.4 times the national average).

In the case of the sexual mixing parameter (panel b), posterior means are generally between the 2.5 and 97.5 percentiles of the prior distribution, although in Limpopo sexual mixing appears to be somewhat more random than at a national level.

In previous versions of the Thembisa model, estimates of the relative rate of condom use in different provinces were mostly close to 1 (implying levels of condom use close to the national average), except in Western Cape and Northern Cape, where the ratio was usually below 1, and Gauteng, where the ratio was usually above 1 [12, 18]. In Thembisa version 4.5, however, there appear to be a number of other provinces in which the ratio is well below 1: Eastern Cape, Limpopo and Mpumalanga. The explanation for this is unclear, although these are relatively poor, rural provinces.

In all provinces, the posterior estimate of the mean age at female non-marital sexual activity is within the 2.5-97.5 percentile range of the prior distributions (panel d). The posterior mean is highest in Limpopo (44 years) and lowest in Western Cape (33 years). The standard deviation of the age distribution of non-marital sexual activity is generally close to the prior mean (panel e).

Panel f compares the posterior estimates of the mean adult HIV survival time (in the absence of ART) across provinces. The mean duration of untreated survival varies between 9.0 years in Northern Cape and 13.3 years in Gauteng. Some of this variation may be attributable to differences in TB incidence: Northern Cape and Western Cape, which have historically had the highest TB incidence rates, are among the provinces with the shortest HIV survival time [43]. The relatively high mean survival time in Gauteng might be a reflection of the high average socio-economic status in this province, and better HIV survival in individuals of higher socio-economic status [44].

The posterior estimates of the relative rate of ART initiation in men are consistently below the prior mean, except in Eastern Cape and Western Cape (panel g). It is noteworthy that the three provinces with the highest posterior estimates of the relative rate of ART initiation in men (viz. Western Cape, Eastern Cape and Northern Cape) are also the provinces with the shortest posterior estimates of untreated HIV survival (panel f).

The odds of viral suppression, relative to IeDEA-SA cohorts, are highly variable across provinces, varying between 0.51 in Limpopo and 1.11 in Western Cape (panel h). In general,

the posterior means are close to the prior means, suggesting that calibration to province-specific HIV prevalence data does not substantially change estimates of viral suppression.

Posterior estimates of annual rates of ART interruption, tend to be below the prior mean (0.25), except in Western Cape and Limpopo (panel i). These results need to be interpreted with a degree of caution, as the ART interruption rate parameter is the only parameter varied in the calibration process that influences the extent of mortality on ART. It might be preferable to allow for uncertainty in the ART mortality rates more directly in the calibration process, which would probably lead to different posterior estimates of the ART interruption rates.

Initial HIV prevalence levels (in 1985) are estimated to have been highest in Mpumalanga and Gauteng provinces (panel j). In contrast, early HIV prevalence levels were relatively low in the Northern Cape and Western Cape.

Posterior estimates of the relative rate of fertility in HIV-positive women in the early stages of HIV infection (before immune decline and before HIV diagnosis), when compared to HIV-negative women, are in most cases close to the prior mean of 1.30 (panel k). The relative rates vary between 1.15 in Northern Cape and 1.43 in Gauteng.

Finally, posterior estimates of the reduction in the rate of ART initiation after diagnosis, following the start of the COVID epidemic, are generally close to the posterior mean (a 28% reduction) (panel l). However in Northern Cape province there appears to have been a particularly remarkable 49% drop in the ART initiation rate.

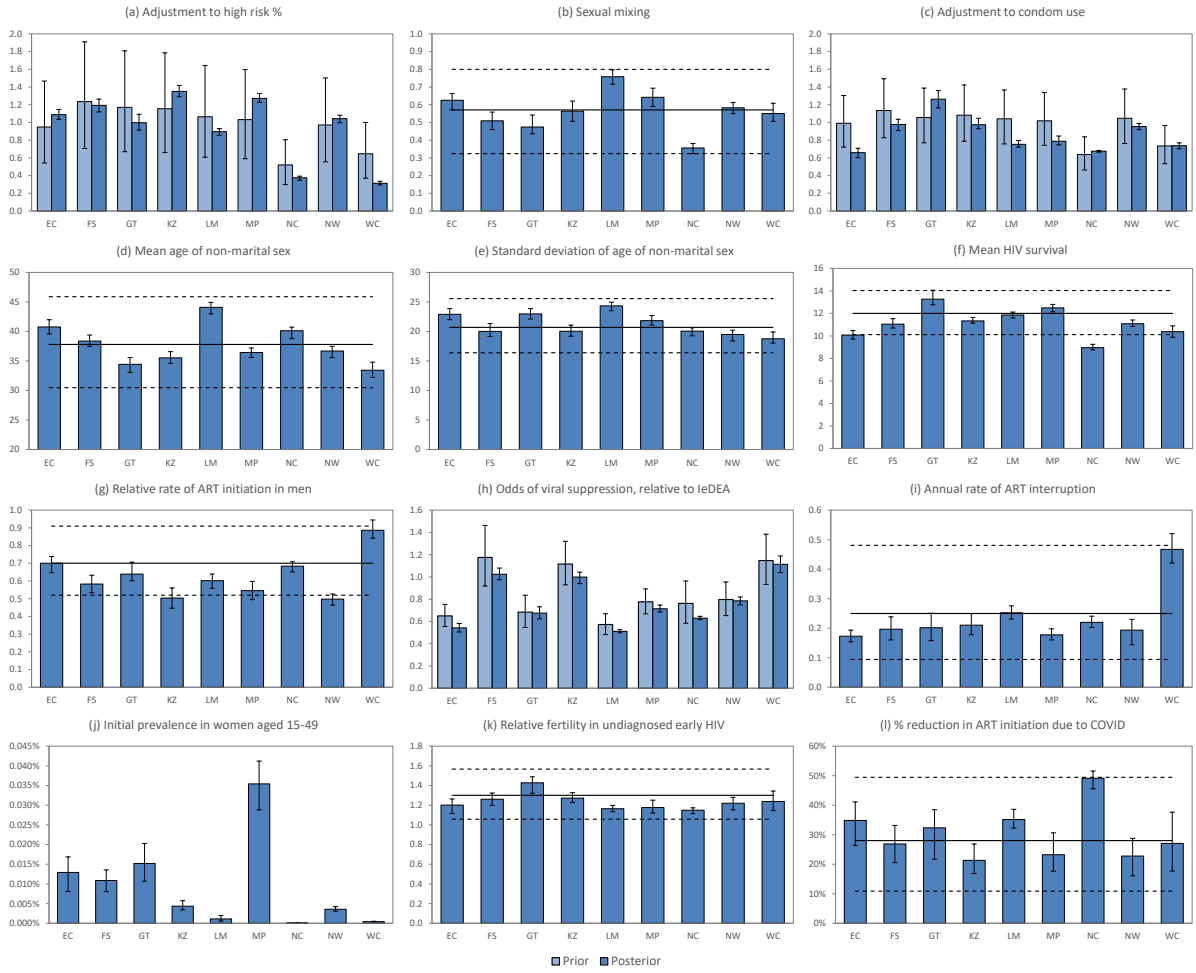

Figure 9.1: Comparison of prior and posterior distributions for adult parameters (excluding ART initiation parameters)

Bar heights represent means and vertical error bars represent 2.5 and 97.5 percentiles (95% confidence intervals) of the relevant distributions. In panels (b), (d)-(g), (i), (k) and (l) the prior distributions are the same for all provinces, and the horizontal lines therefore represent the prior mean and 95% confidence interval.

Table 9.1 shows the posterior estimates for the ART initiation rates in previously diagnosed women with CD4 counts of <200 cells/µl. Posterior estimates of the ART initiation trends are quite different across provinces. For example, in Western Cape ART initiation rates started at relatively high levels (compared to other provinces) over the 2004-2010 period, the early phase of the public sector ART rollout, but then gradually declined. In contrast, ART initiation rates have steadily increased over time in provinces such as KwaZulu-Natal and Mpumalanga.

Table 9.1: Posterior estimates of ART initiation rates in previously-diagnosed females with CD4 <200 (means and 95% confidence intervals in brackets)

	2000	2004	2010	2011	2016
Prior	0.004	0.015	0.058	0.062	0.059
	(0.002-0.006)	(0.001-0.045)	(0.025-0.103)	(0.030-0.106)	(0.022-0.115)
EC	0.005	0.022	0.039	0.031	0.027
	(0.005-0.006)	(0.019-0.028)	(0.036-0.043)	(0.028-0.034)	(0.023-0.03)
FS	0.004	0.007	0.040	0.037	0.072
	(0.004-0.005)	(0.006-0.008)	(0.037-0.043)	(0.034-0.041)	(0.063-0.082)
GT	0.005	0.018	0.044	0.044	0.038
	(0.004-0.005)	(0.014-0.021)	(0.038-0.049)	(0.04-0.051)	(0.033-0.044)
KZ	0.005	0.014	0.049	0.063	0.077
	(0.005-0.006)	(0.012-0.017)	(0.044-0.055)	(0.055-0.07)	(0.066-0.085)
LM	0.006	0.019	0.064	0.041	0.035
	(0.006-0.007)	(0.018-0.022)	(0.06-0.068)	(0.037-0.044)	(0.031-0.039)
MP	0.006	0.011	0.040	0.047	0.057
	(0.005-0.006)	(0.009-0.013)	(0.037-0.043)	(0.043-0.05)	(0.052-0.065)
NC	0.003	0.020	0.018	0.030	0.031
	(0.003-0.003)	(0.017-0.022)	(0.015-0.02)	(0.029-0.032)	(0.028-0.036)
NW	0.004	0.021	0.040	0.025	0.036
	(0.003-0.004)	(0.020-0.024)	(0.035-0.044)	(0.021-0.028)	(0.033-0.039)
WC	0.003	0.054	0.066	0.048	0.044
	(0.002-0.003)	(0.047-0.063)	(0.056-0.078)	(0.044-0.054)	(0.039-0.05)

Figure 9.2 compares the prior and posterior estimates of the paediatric parameters, by province. Posterior estimates of the relative rate of ART initiation in early paediatric disease (relative to advanced disease) appear highly heterogeneous across provinces (panel a), which may be a reflection of differences in HIV testing and linkage practices in HIV-positive children who do not have advanced disease. Posterior estimates of the paediatric HIV disease progression and mortality adjustment are generally close to 1 (panel b), although the adjustments are notably greater than 1 in Western Cape, Northern Cape and Eastern Cape. These are provinces that have historically had very high TB incidence rates [43], and it is possible that the high TB incidence rates in these provinces are responsible for the relatively high rates of disease progression. The results are also quite consistent with those obtained for adults (Figure 9.1f), which suggest relatively poor HIV survival rates in the Northern Cape, Eastern Cape and Western Cape provinces. Posterior estimates of breastfeeding duration adjustments (panel c) are generally close to 1, implying breastfeeding durations close to the national average.

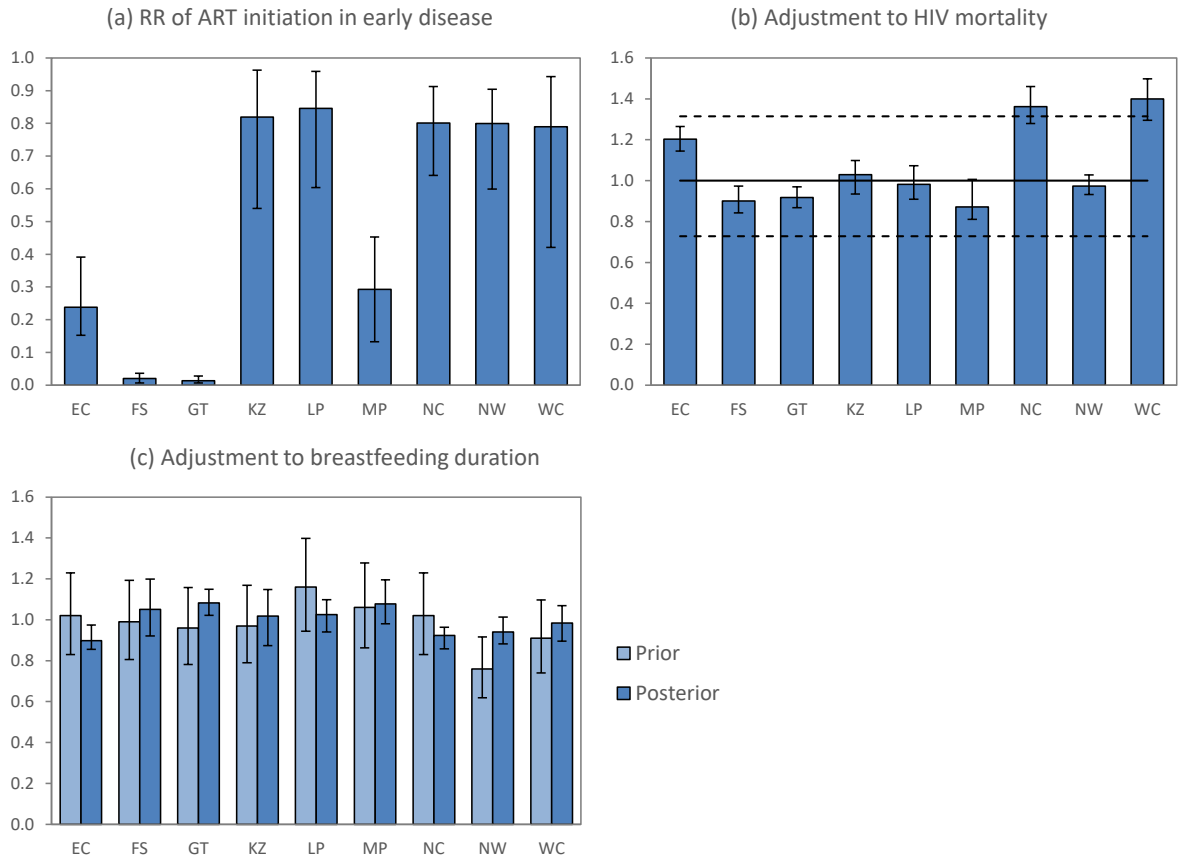

Figure 9.2: Prior and posterior estimates of the paediatric HIV parameters (excluding ART initiation parameters)

Bar heights represent means and vertical error bars represent 2.5 and 97.5 percentiles (95% confidence intervals) of the relevant distributions. In panels (a) and (b) the prior distributions are the same for all provinces, and the horizontal lines in panel (b) represent the prior mean and 95% confidence interval.

Table 9.2 summarizes the posterior estimates of the paediatric ART initiation rates. ART initiation rates over the 2000-2003 period were consistently between 0.001 and 0.002 per month, then increased to around 0.01 per month in 2004, following the start of the public sector ART programme (though Western Cape was a notable exception, with extremely rapid early rollout of paediatric ART). By 2009-2010, ART initiation rates increased to 0.02-0.10 per month in most provinces, but were at substantially higher rates in the highly urbanized provinces (Gauteng and Western Cape). In most provinces (with the exception of Free State) there were reductions in ART initiation rates in later years, as ART eligibility criteria were expanded.

Table 9.2: Posterior estimates of ART initiation rates in previously-diagnosed children in late disease (means and 95% confidence intervals in brackets)

	2000	2004	2009	2010	2016
Prior	0.003 (0.001-0.006)	0.043 (0.000-0.260)	0.177 (0.002-0.703)	0.110 (0.005-0.372)	0.110 (0.005-0.372)
EC	0.002 (0.002-0.002)	0.011 (0.009-0.013)	0.044 (0.039-0.052)	0.207 (0.140-0.286)	0.077 (0.056-0.100)
FS	0.001 (0.001-0.001)	0.004 (0.003-0.005)	0.023 (0.020-0.026)	0.062 (0.052-0.075)	0.148 (0.087-0.273)
GT	0.002 (0.002-0.002)	0.017 (0.015-0.019)	0.043 (0.035-0.050)	0.392 (0.189-0.723)	0.017 (0.013-0.022)
KZ	0.002 (0.001-0.002)	0.006 (0.005-0.007)	0.043 (0.037-0.049)	0.081 (0.065-0.122)	0.015 (0.012-0.022)
LM	0.001 (0.001-0.001)	0.007 (0.006-0.008)	0.033 (0.028-0.038)	0.040 (0.034-0.047)	0.027 (0.023-0.035)
MP	0.001 (0.001-0.001)	0.003 (0.003-0.004)	0.017 (0.015-0.019)	0.035 (0.030-0.048)	0.053 (0.036-0.090)
NC	0.001 (0.001-0.002)	0.015 (0.013-0.017)	0.097 (0.080-0.119)	0.040 (0.031-0.048)	0.060 (0.047-0.072)
NW	0.001 (0.001-0.001)	0.005 (0.004-0.006)	0.040 (0.037-0.044)	0.026 (0.021-0.032)	0.014 (0.011-0.017)
WC	0.002 (0.002-0.002)	0.828 (0.591-1.099)	0.407 (0.132-0.859)	0.280 (0.140-0.477)	0.050 (0.038-0.070)

9.2 Calibration to HIV prevalence and ART data

9.2.1 Eastern Cape calibration

Figure 9.3 shows the comparison of the Eastern Cape provincial model to the HIV prevalence data from the antenatal surveys. The model does not fit the age-specific HIV prevalence data well, especially in the 1990s, when the model tends to estimate a higher HIV prevalence in pregnant women than the survey data suggest. (This is probably because the model is 'forced' to estimate a higher HIV prevalence in the 1990s in order to match the rising levels of mortality observed in the late 1990s.) In addition the model under-estimates HIV prevalence in the two most recent surveys (2017 and 2019).

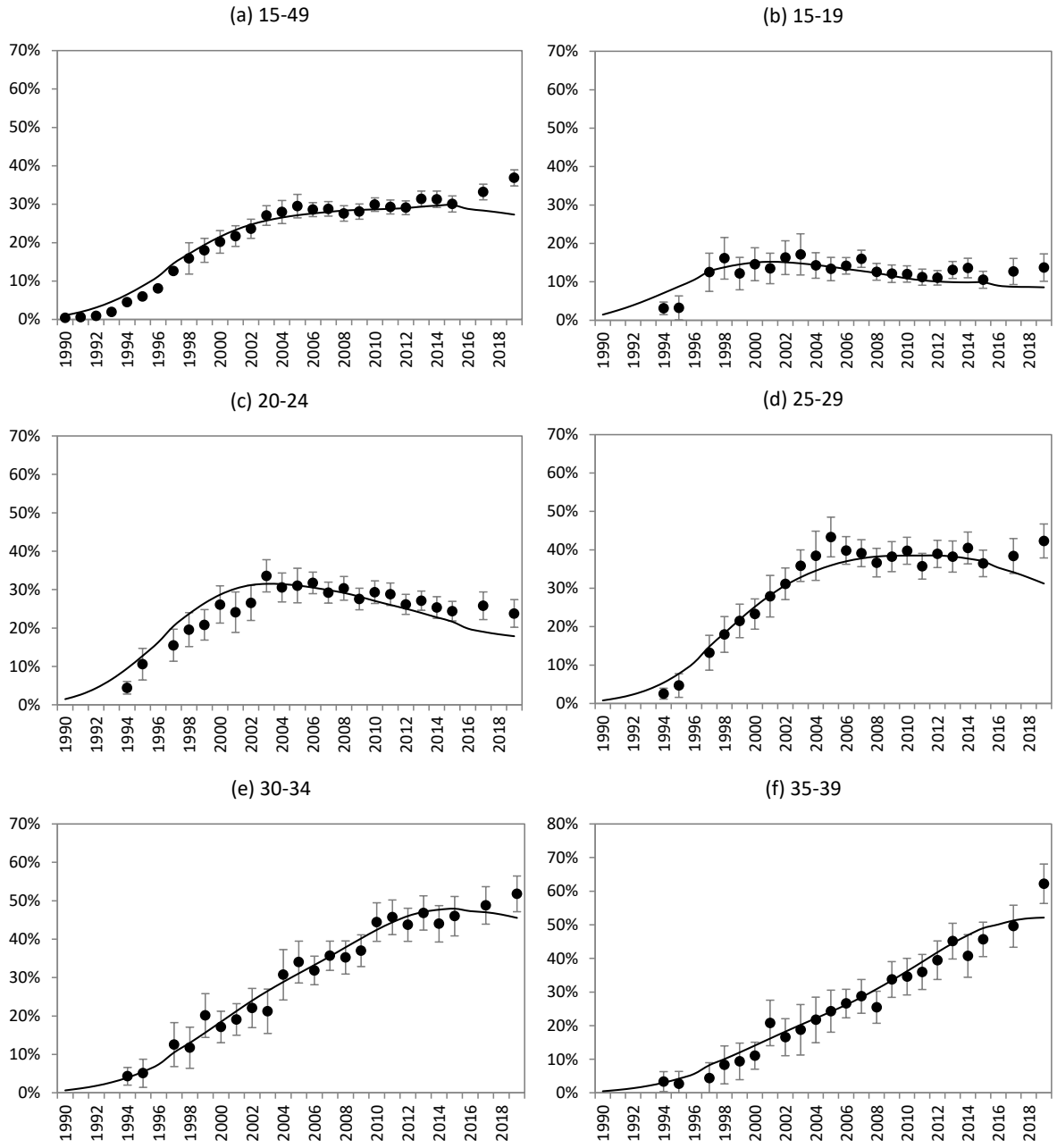

Figure 9.3: HIV prevalence levels in pregnant women attending public antenatal clinics in Eastern Cape

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.4 shows the comparison of the HIV prevalence data from the household surveys to the Eastern Cape model. The model is generally consistent with the surveys, except in the case of the 2017 survey (and more specifically youth in the 2017 survey), where the model under-estimates HIV prevalence.

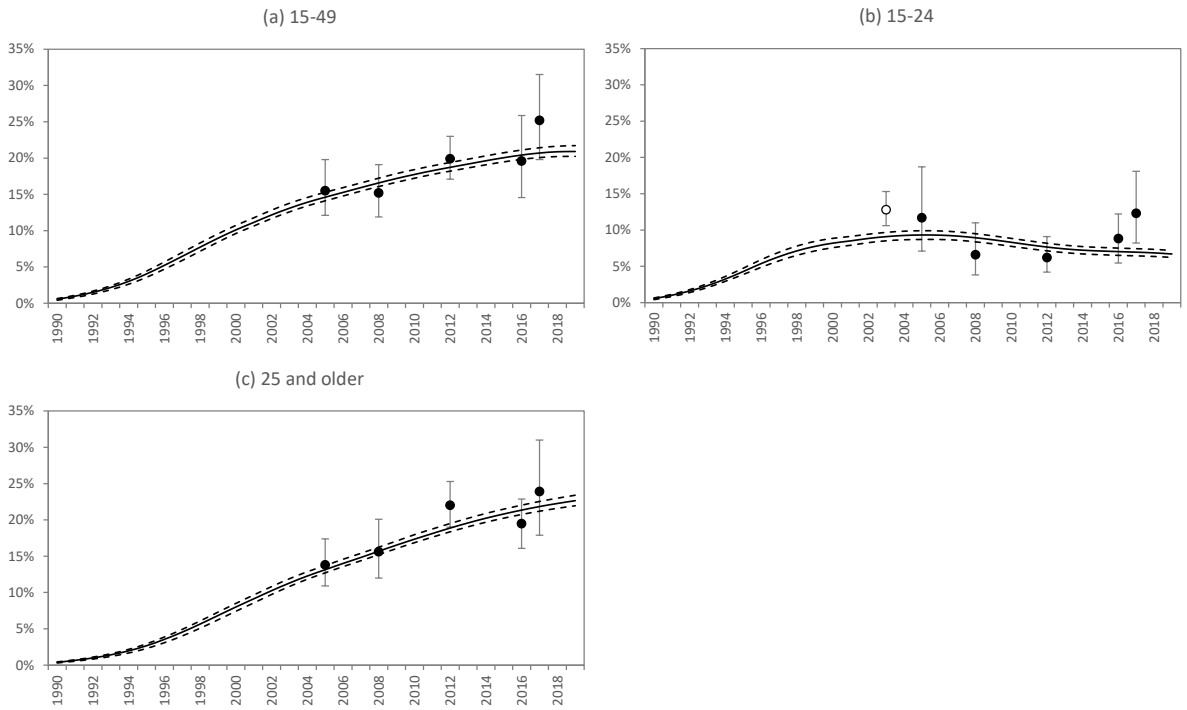

Figure 9.4: HIV prevalence in the general adult population of the Eastern Cape
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.5 shows the Eastern Cape model fit to ART coverage data. The model matches the total number of ART patients quite well in the period after 2010, but is slightly lower in the period before 2010 (panel d), because the reported numbers in the period before 2010 reflect cumulative enrolment rather than current enrolment. The model matches the 2012 HSRC survey estimates of ART coverage, but under-estimates the ART coverage in the 2017 HSRC survey (panels a and b). It is worth noting that it would not be possible to improve the fit to the 2017 data points without either (a) reducing HIV prevalence (which would give a worse fit to the data in Figure 9.4) or (b) increasing numbers of adults on ART (which would give a worse fit to the data in panel d). Available data on the proportion of adult ART patients who are male are quite inconsistent in 2012 and over 2015-2021, but the model fits the more recent data well (panel c). Finally the model does not fit the age distributions of ART patients well (panels e and f). In particular the modelled age distribution in women is too old, while the model under-estimates the numbers of older men on ART.

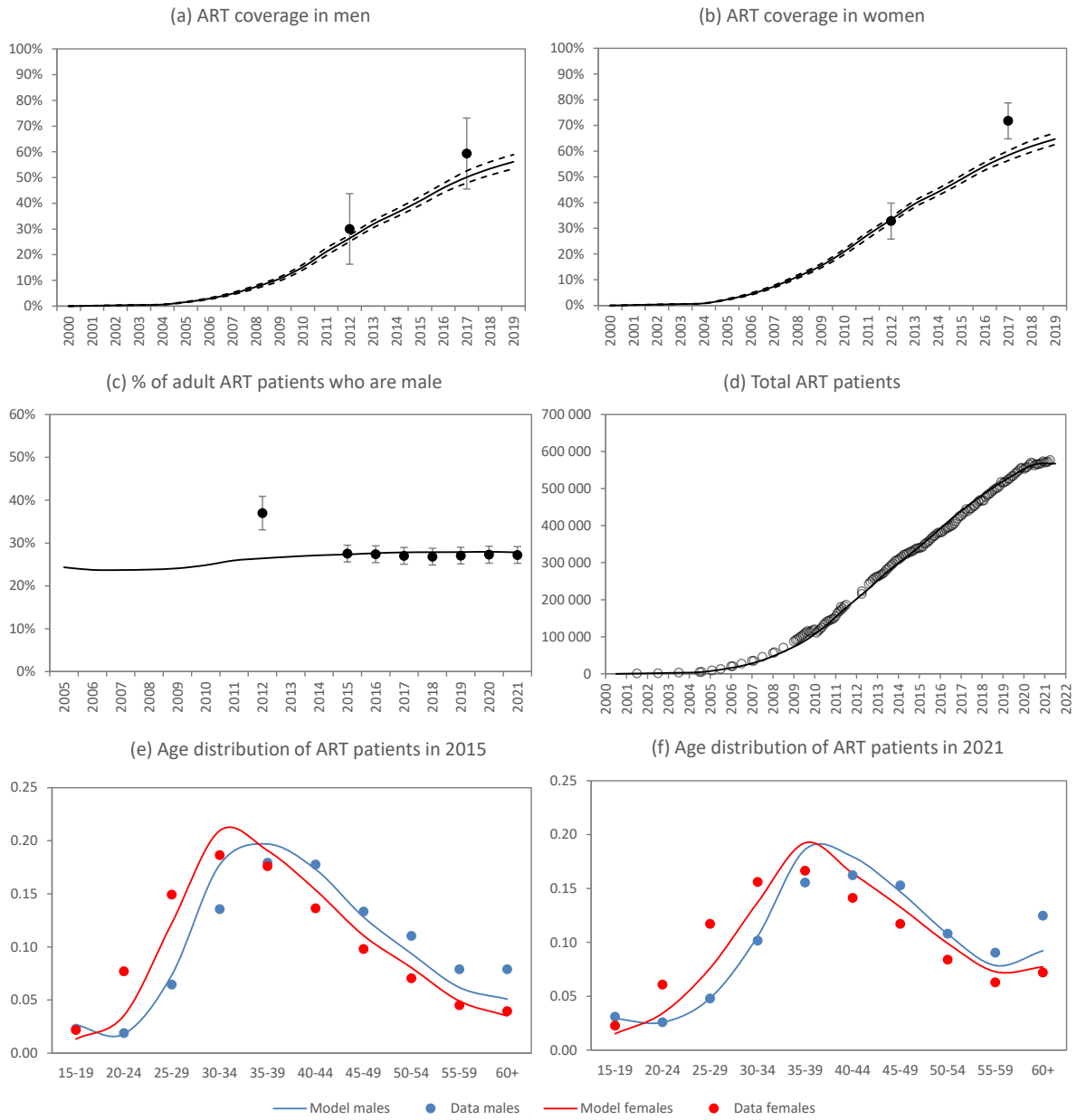

Figure 9.5: Adult ART calibration in the Eastern Cape
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.6 shows the model calibration to the recorded death data. Model estimates of mortality are roughly consistent with recorded levels, although the modelled male mortality estimates appear slightly too high in the late 1990s.

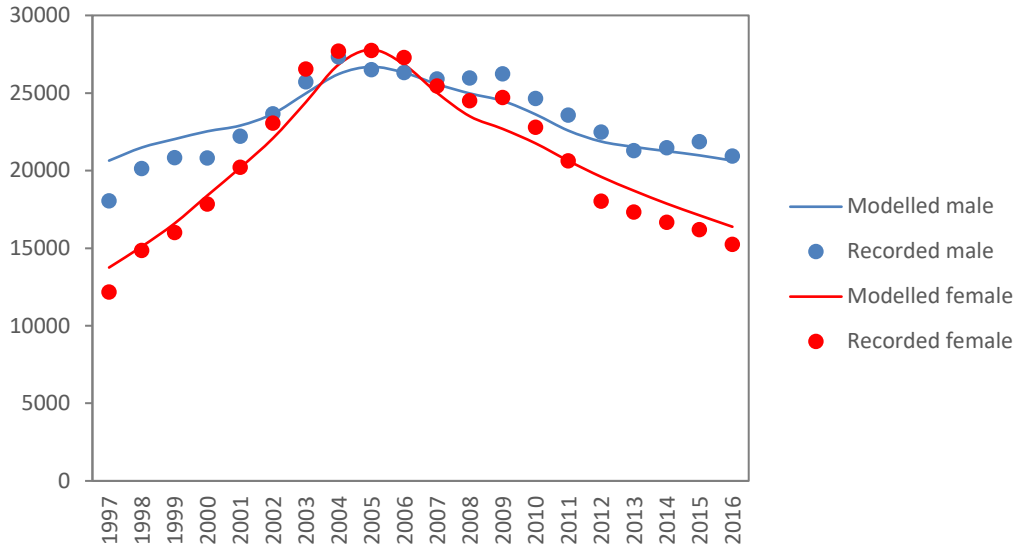

Figure 9.6: Deaths due to all causes in adults aged 20-59, Eastern Cape
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.7 shows the model fit to the paediatric data sources. The model is consistent with recent paediatric ART data (panel a), but the model tends to under-estimate the reported totals in the period before 2010 (as in adults, this is largely because the reported totals before 2010 reflected cumulative enrolment). The model estimates of HIV prevalence in children are also consistent with household survey estimates of HIV prevalence (panel b) and the modelled age distribution of children on ART is consistent with programme data (panel c).

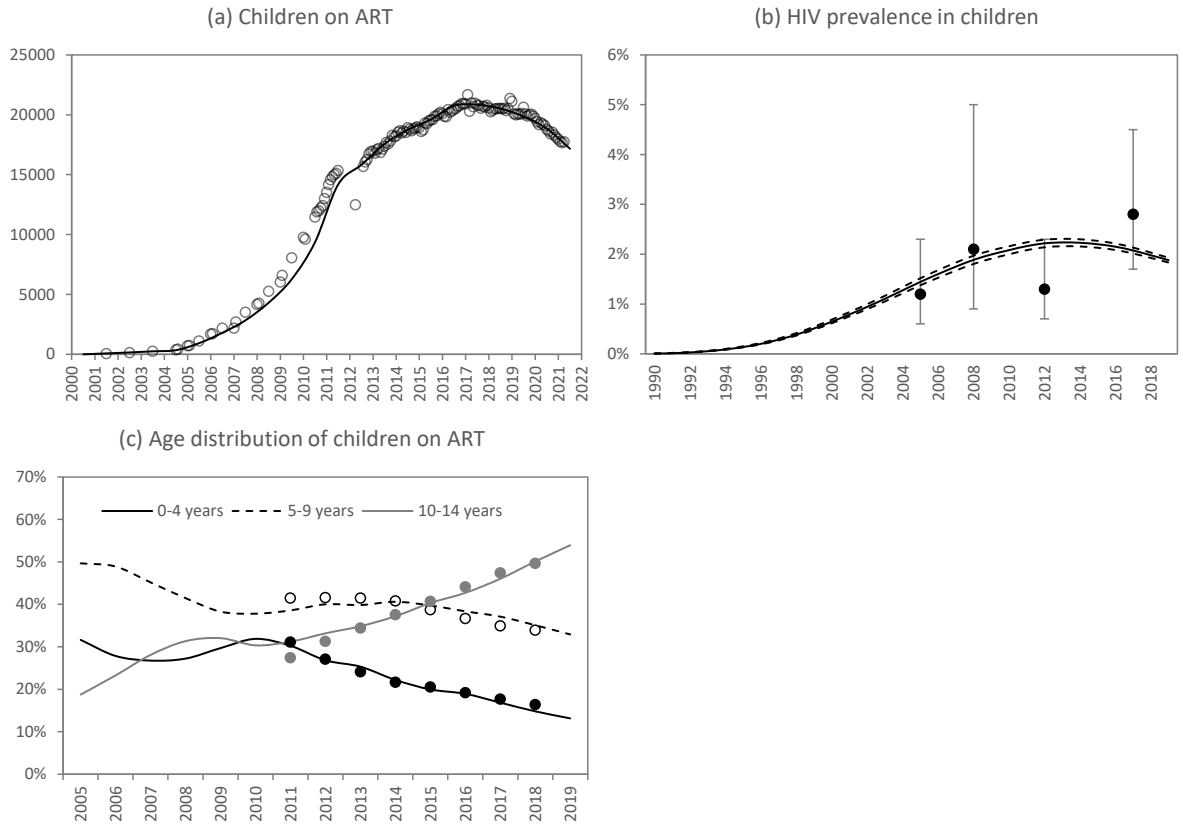

Figure 9.7: Paediatric calibration in the Eastern Cape
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

9.2.2 Free State calibration

Figure 9.8 shows the model fit to the antenatal prevalence data from Free State. The model fits the survey data well, although the model tends to over-estimate prevalence in the early 1990s.

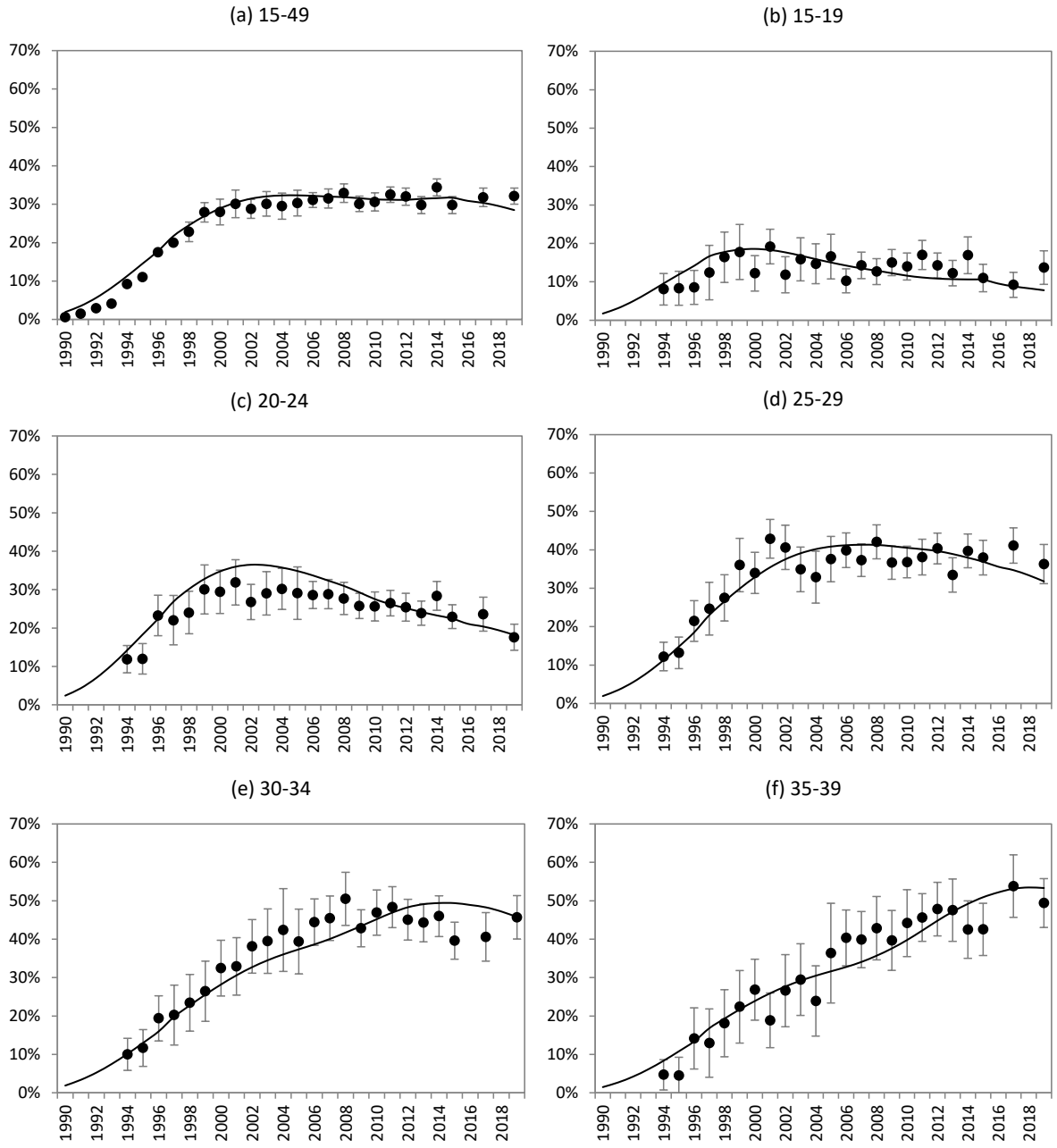

Figure 9.8: HIV prevalence levels in pregnant women attending public antenatal clinics in Free State

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.9 shows the model fit to the household survey prevalence data. On the whole the model fits the data well, although in 2017 the model under-estimates the survey prevalence at ages 25 and older (panel c), and in 2008 and 2012 the model over-estimates the survey prevalence in the 15-24 age group (panel b).

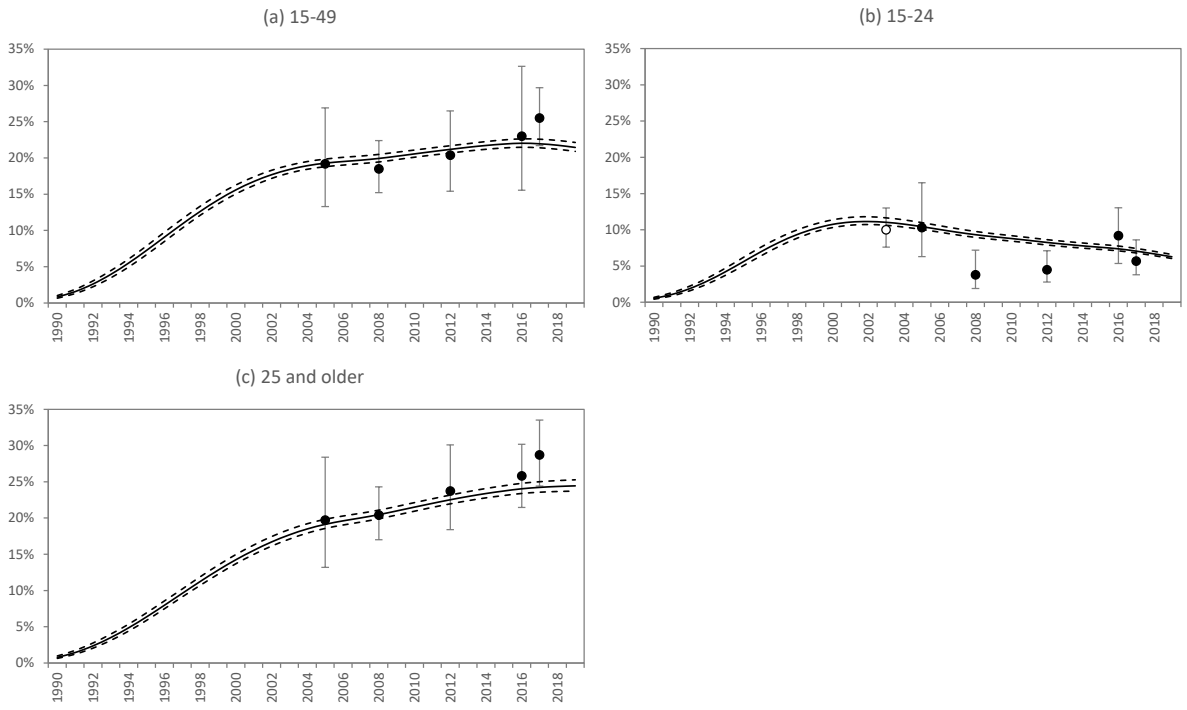

Figure 9.9: HIV prevalence in the general adult population of Free State
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.10 shows the model fit to the adult ART data. The model appears to be fairly consistent with both the programme data and the HSRC household survey data. However the modelled age distribution of patients on ART is not quite consistent with the programme data: the model over-estimates the proportion on male ART patients who are aged 35-39, while under-estimating the proportion of female ART patients who are aged 20-24.

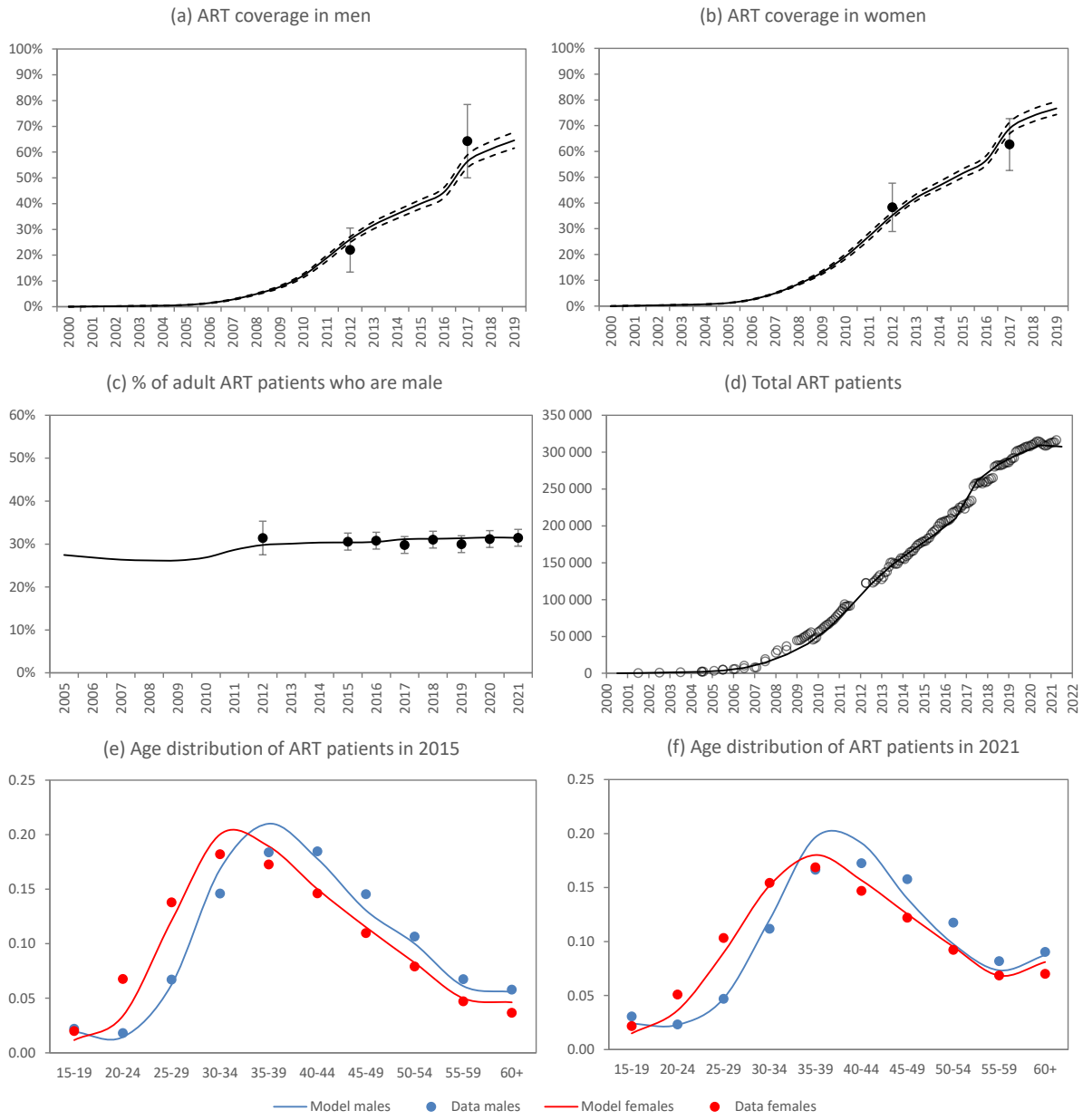

Figure 9.10: Adult ART calibration in Free State
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.11 shows that the model fits the recorded death data reasonably well, although the model slightly over-estimates the number of male deaths in the period 1997-2007.

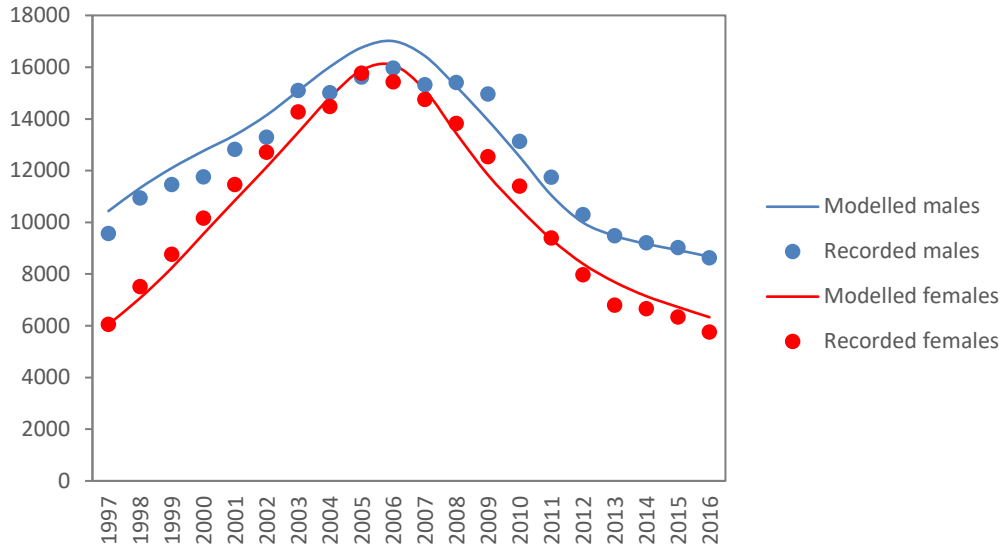

Figure 9.11: Deaths due to all causes in adults aged 20-59, Free State
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.12 shows the model calibration to Free State paediatric data. The model matches most of the paediatric ART data, although there appear to be problems with the quality of the data over the 2010-2013 period (panel a). Survey estimates of HIV prevalence in children have extremely wide confidence intervals around them; nevertheless, the model appears to over-estimate HIV prevalence in 2012 (panel b). The modelled age distribution of children on ART is reasonably consistent with programme data (panel c).

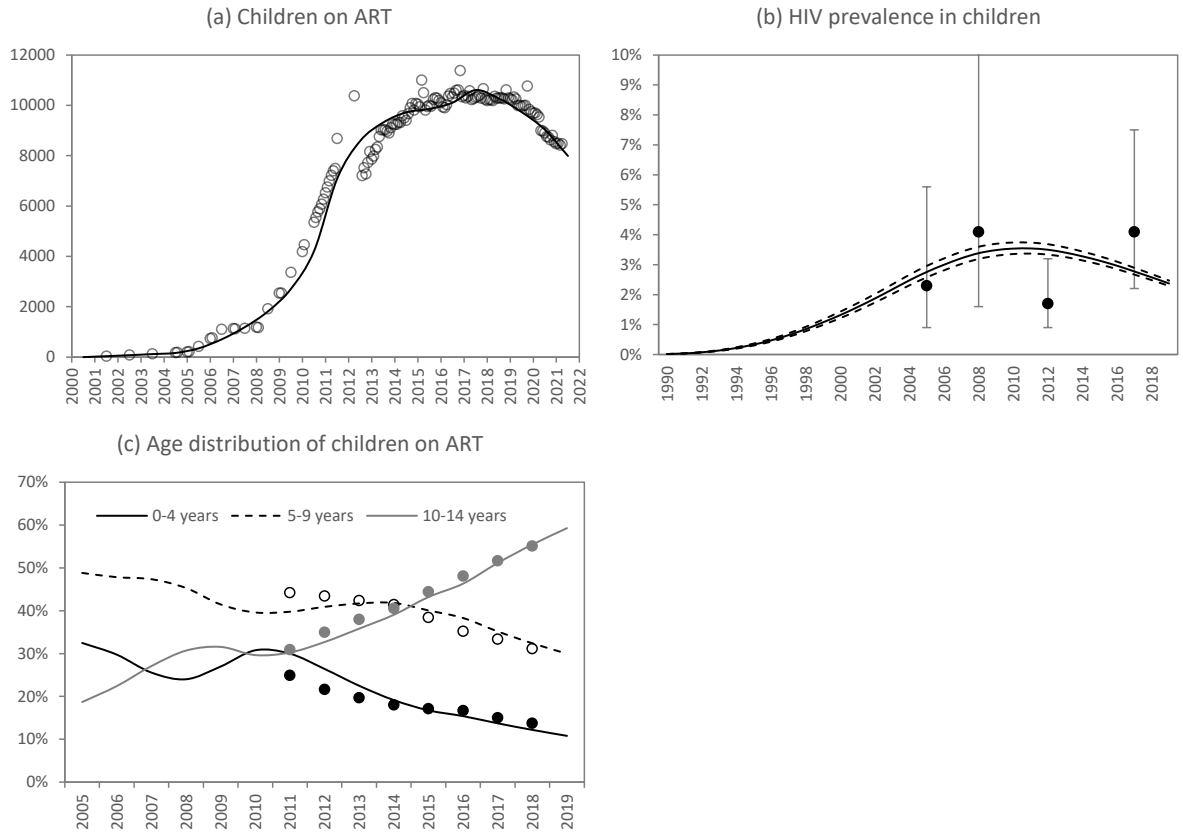

Figure 9.12: Paediatric calibration in Free State
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.3 Gauteng calibration

Figure 9.13 shows the model fit to the antenatal prevalence data from Gauteng. Although the model fits the overall survey data reasonably well, the model tends to under-estimate HIV prevalence in the 25-34 age group over the 2000-2006 period. The model also tends to over-estimate HIV prevalence in the early 1990s, although these data are not used in calibration.

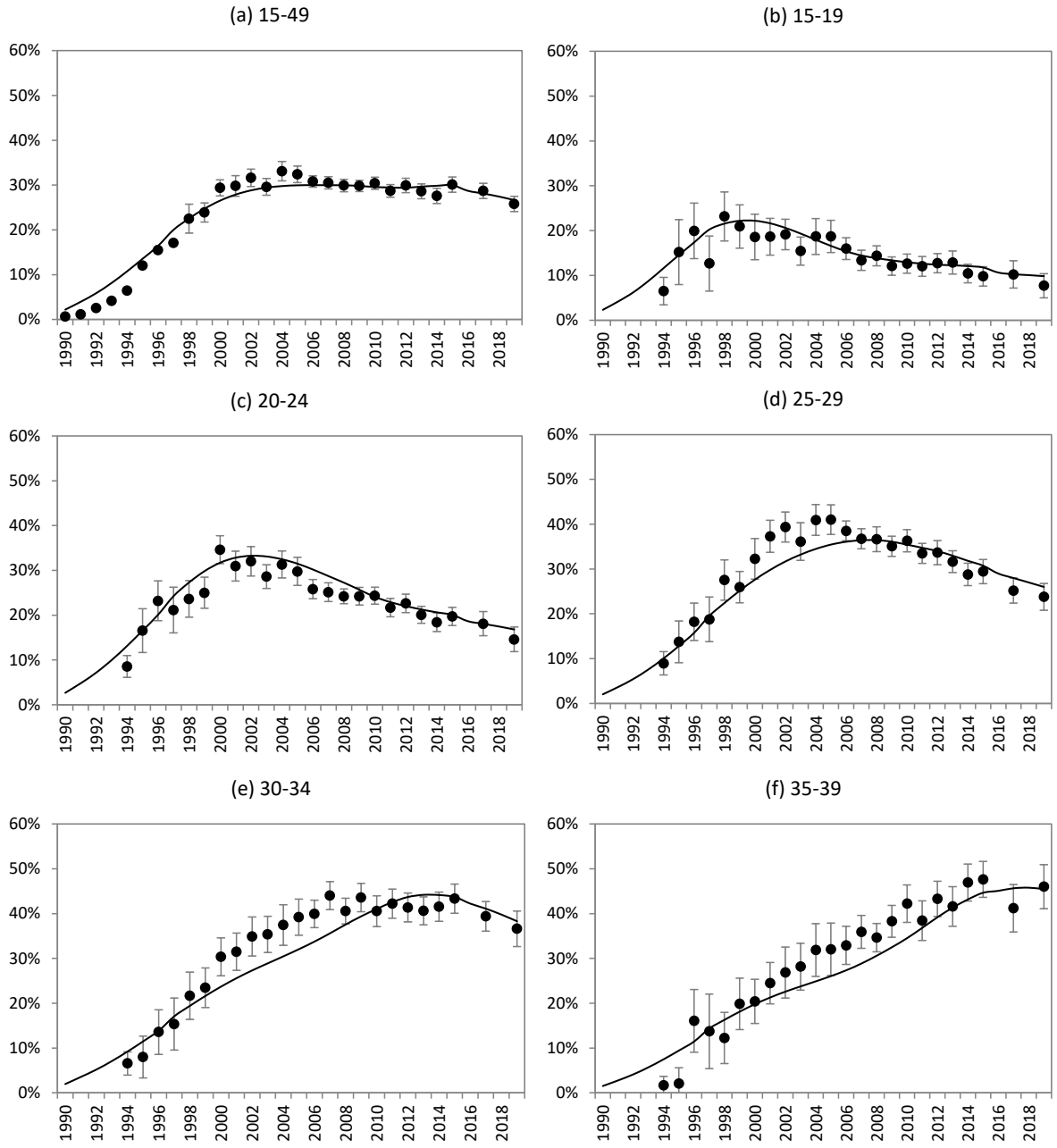

Figure 9.13: HIV prevalence levels in pregnant women attending public antenatal clinics in Gauteng

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.14 shows the model fit to the household survey HIV prevalence data. In general, the model fits the data reasonably well, although the DHS prevalence estimate in 2016 is significantly higher than the model estimate (particularly in the 25+ age group).

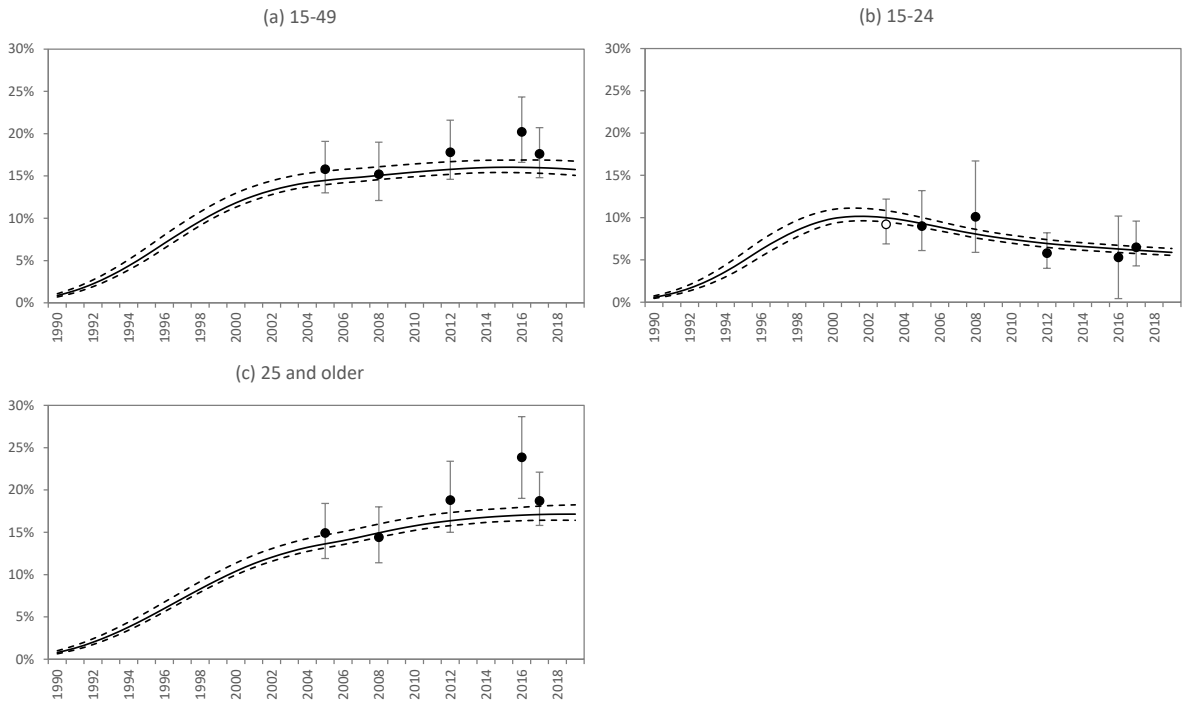

Figure 9.14: HIV prevalence in the general adult population of Gauteng
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.15 shows the model fit to Gauteng ART data. Overall there is good consistency between the model estimates and the survey data and the ART programme data.

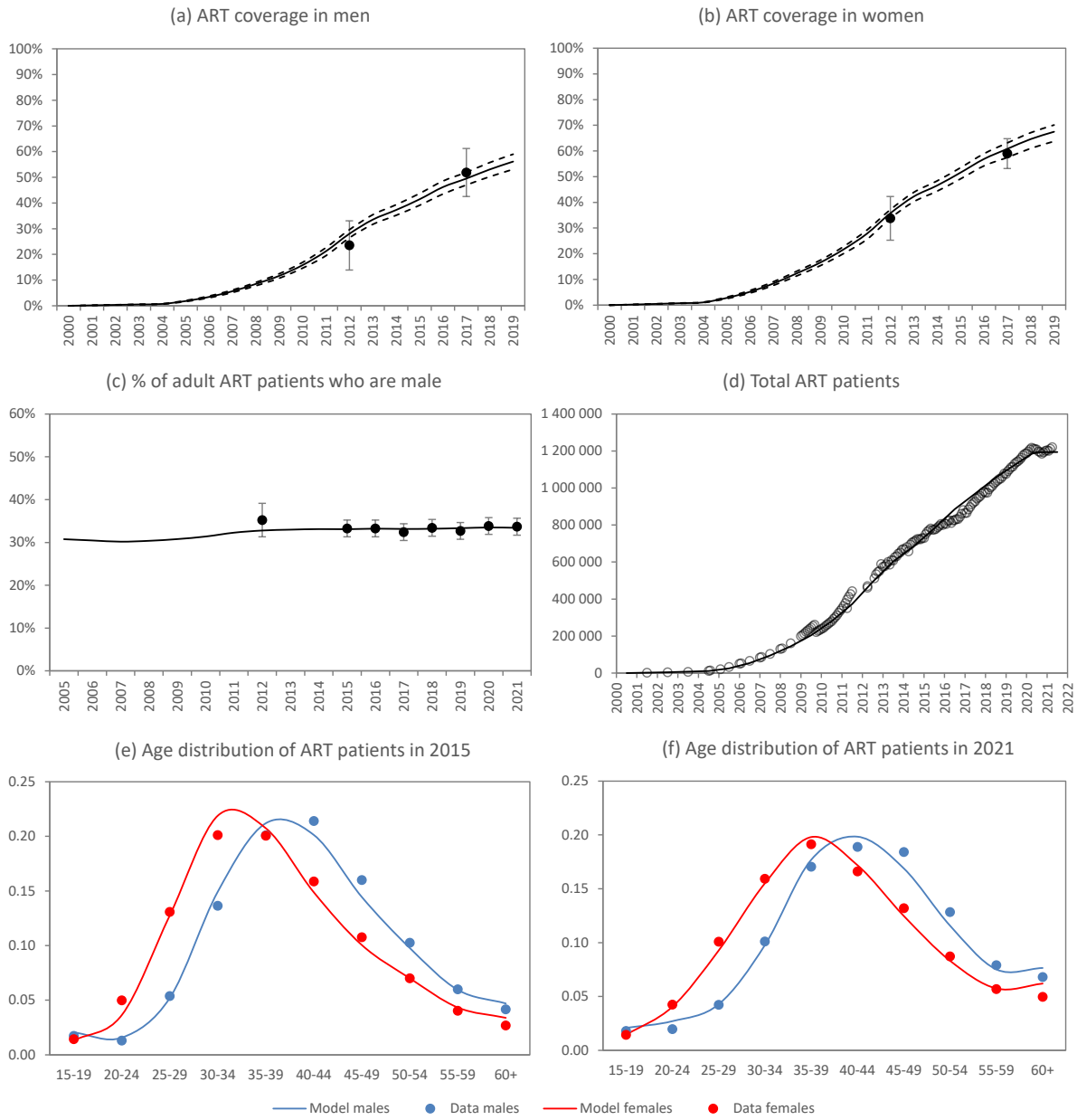

Figure 9.15: Adult ART calibration in Gauteng
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.16 shows the model fit to the recorded death data. Although the model is roughly consistent with the data, the data suggest a slightly lighter mortality peak (around 2007-8) than the model does (around 2005-6).

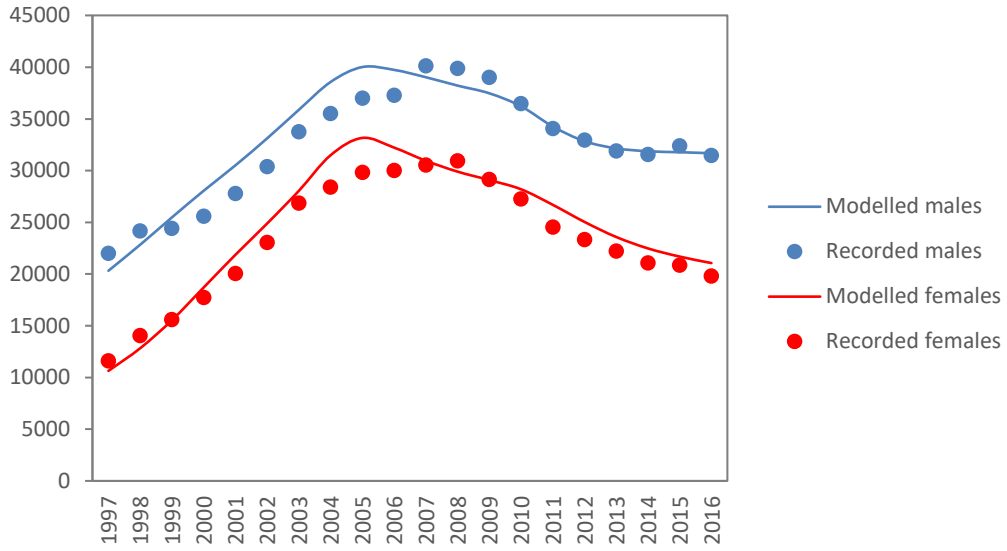

Figure 9.16: Deaths due to all causes in adults aged 20-59, Gauteng
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.17 shows the model fit to the paediatric data sources. The fit to the paediatric ART data in this province is generally poor, in part because of the noise in the data; the model falls below the reported numbers in 2010-2011, but above the reported number in 2016-2017 (panel a). Although the model estimates of HIV prevalence in children are within the 95% confidence intervals around the survey estimates, the confidence intervals are generally very wide. The modelled age distribution of children on ART is not very consistent with the data: the model tends to over-estimate the fraction aged 5-9 while under-estimating the fraction aged 10-14.

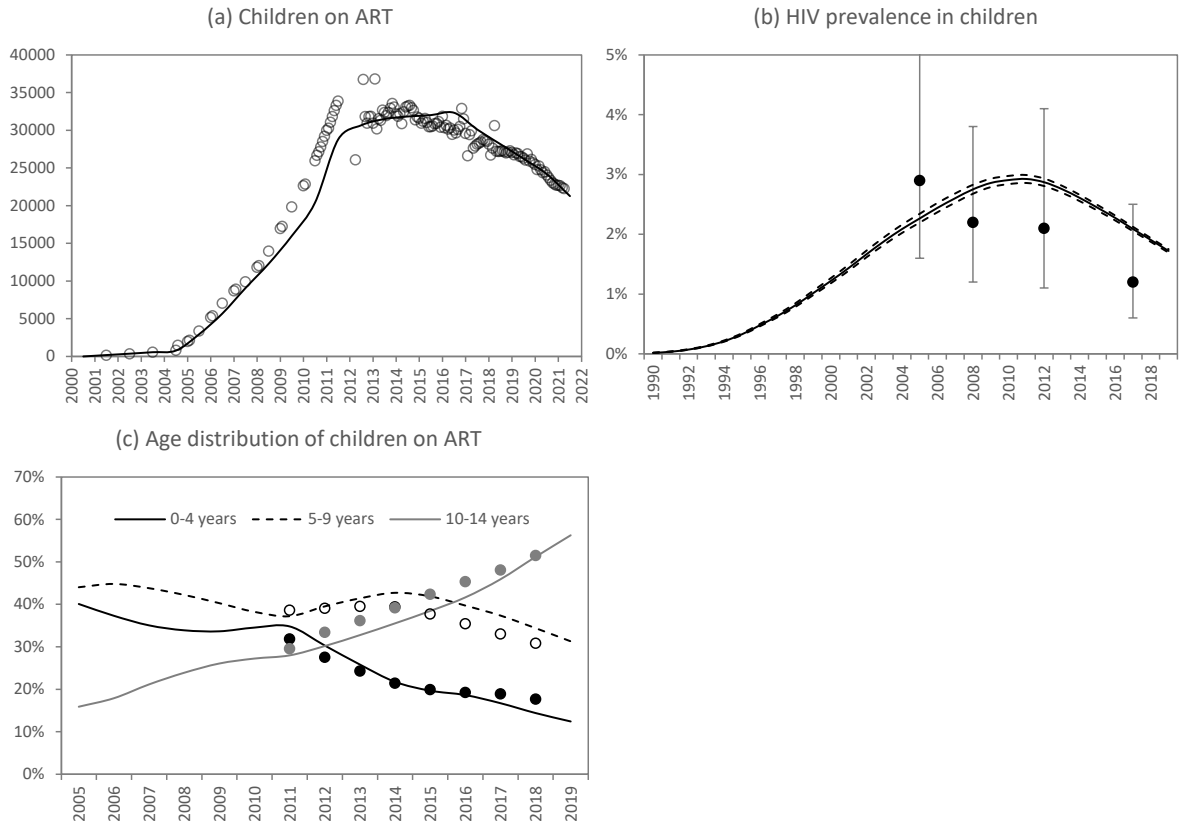

Figure 9.17: Paediatric calibration in Gauteng
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.4 KwaZulu-Natal calibration

Figure 9.18 shows the model fit to the antenatal prevalence data from KwaZulu-Natal. Although the model fits the data acceptably in most years, the model does not match the high HIV prevalence levels observed in 2014 and 2015, and also under-estimates the HIV prevalence in the most recent survey (2019).

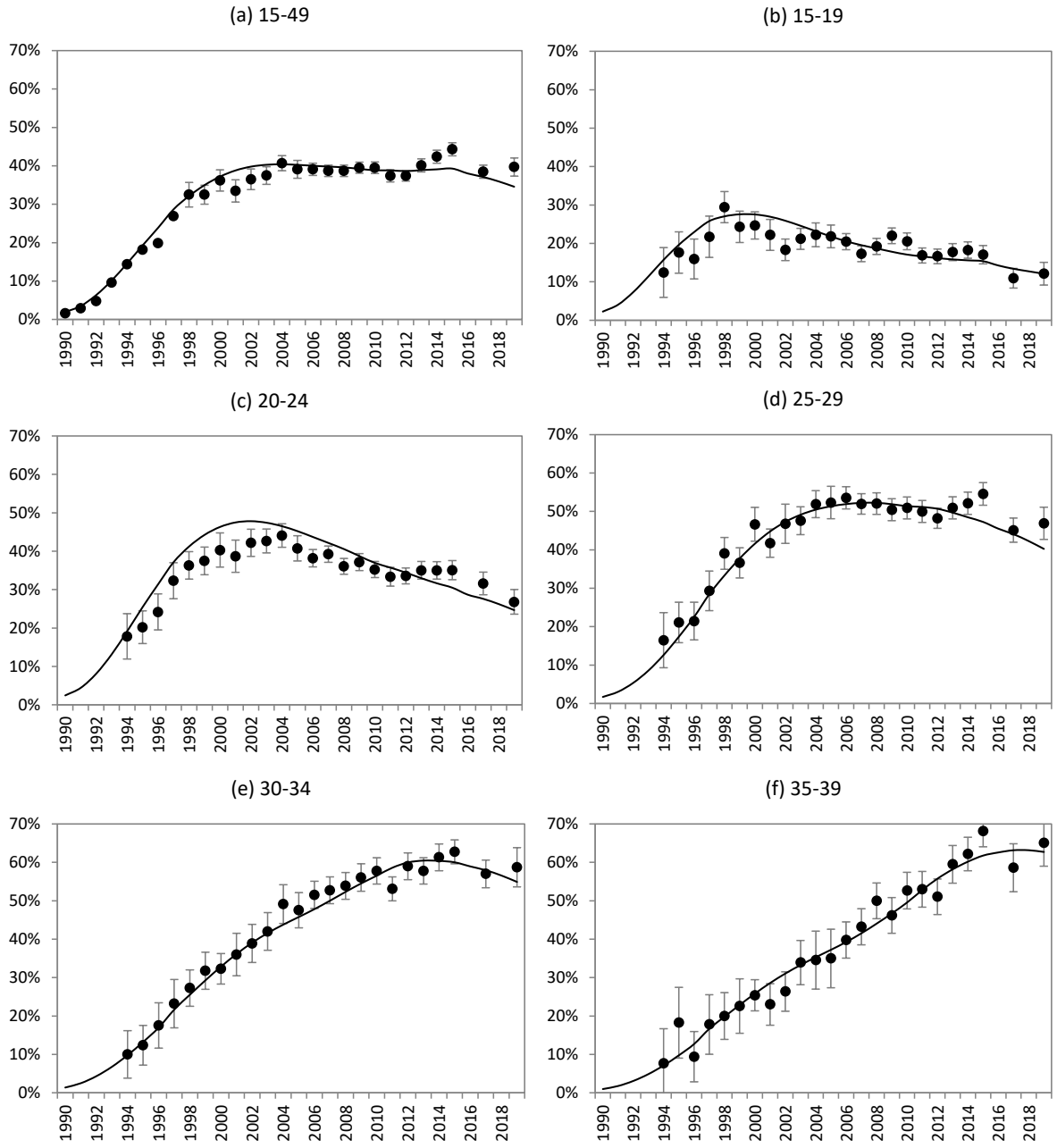

Figure 9.18: HIV prevalence levels in pregnant women attending public antenatal clinics in KwaZulu-Natal

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.19 shows the model fit to the HIV prevalence data from household surveys. Overall the model is reasonably consistent with the household survey data.

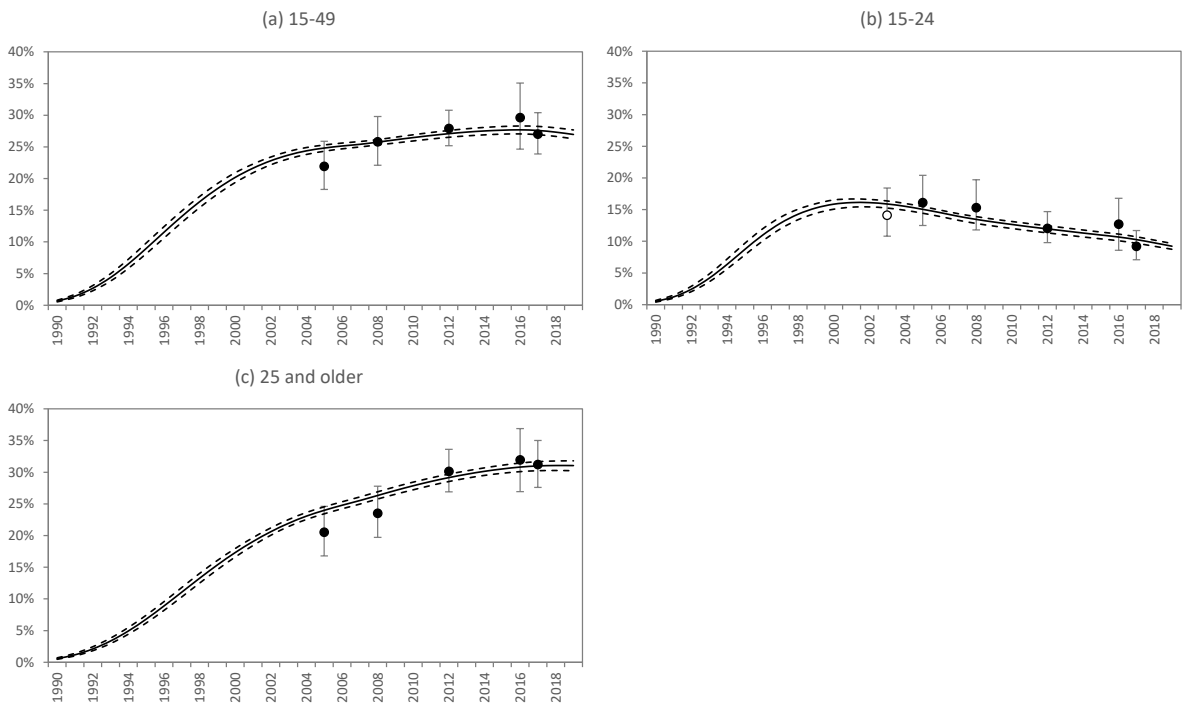

Figure 9.19: HIV prevalence in the general adult population of KwaZulu-Natal
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.20 shows the model fit to the adult ART data. The model fits the reported ART totals reasonably well in the period from 2012 onward (panel d), but the model estimates of ART numbers in the period before 2012 are mostly below the reported numbers, reflecting the previous reporting of cumulative ART enrolment and a slow transition to reporting current ART enrolment over the 2009-2011 period. Estimates of ART coverage in 2012 and 2017 are roughly consistent with the results from the household surveys. The modelled age distributions of ART patients are also reasonably consistent with the routine data (panels e and f)

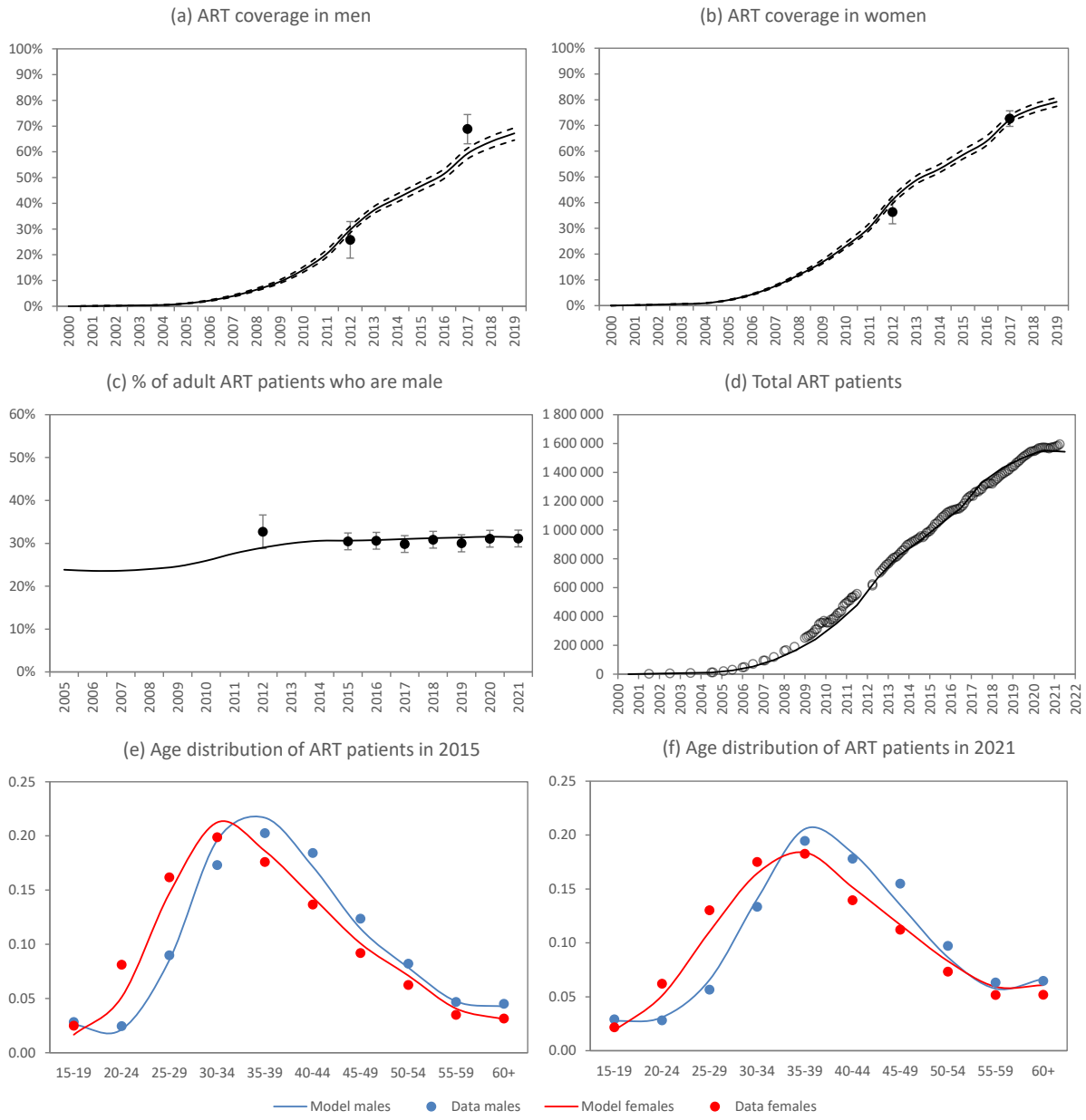

Figure 9.20: Adult ART calibration in KwaZulu-Natal
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.21 shows the model provides a good fit to the recorded death data, although the model slightly over-estimates mortality in women in 2005 and 2006.

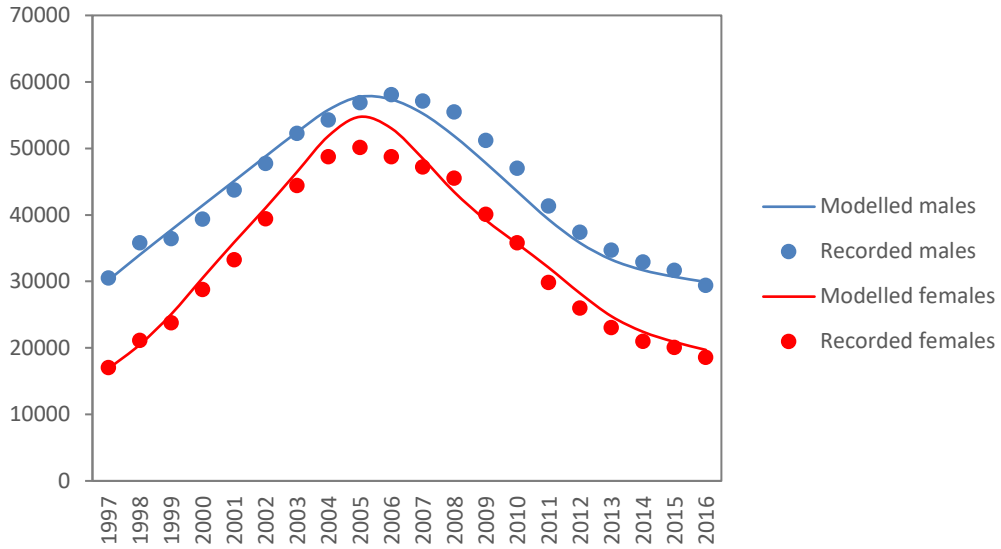

Figure 9.21: Deaths due to all causes in adults aged 20-59, KwaZulu-Natal
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.22 shows the model fit to the paediatric data. The model fits the recorded ART numbers reasonably well in the period from 2014 onward, but the model consistently under-estimates the reported numbers in the period before 2013 (panel a), for the same reasons as in adults. Although the model is not consistent with the exceptionally high HIV prevalence measured in the 2005 survey, it is roughly consistent with the three most recent surveys (panel b). The modelled age distribution of children on ART is also roughly consistent with the data (panel c).

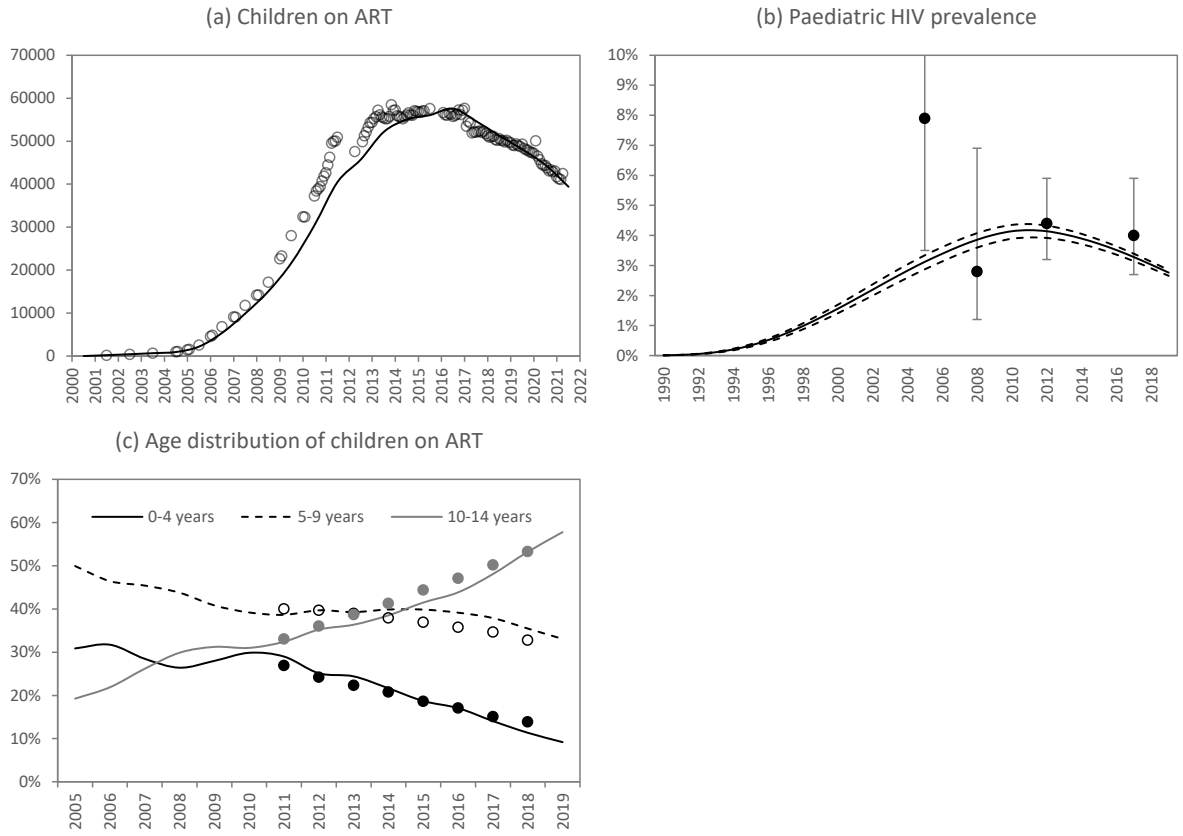

Figure 9.22: Paediatric calibration in KwaZulu-Natal
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.5 Limpopo calibration

Figure 9.23 shows the model fit to the antenatal prevalence data from Limpopo. The model fits the data reasonably well, although the model over-estimates HIV prevalence in the 35-39 age group.

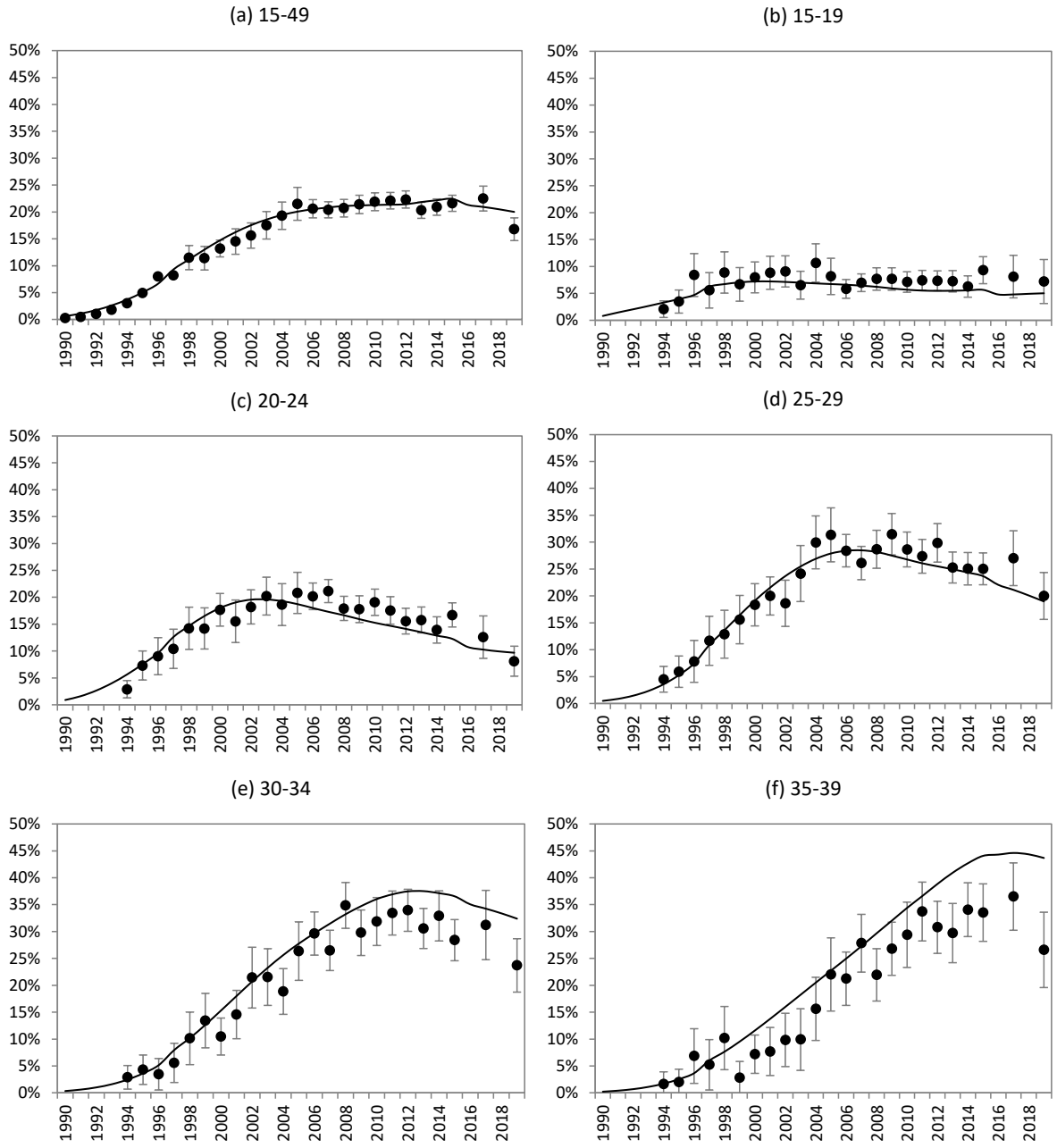

Figure 9.23: HIV prevalence levels in pregnant women attending public antenatal clinics in Limpopo

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.24 shows the model fit to the household survey HIV prevalence data. The model produces a substantially higher estimate of HIV prevalence in 2016 than was measured in the 2016 DHS, especially among youth, but the model estimates are consistent with the data from the other household surveys.

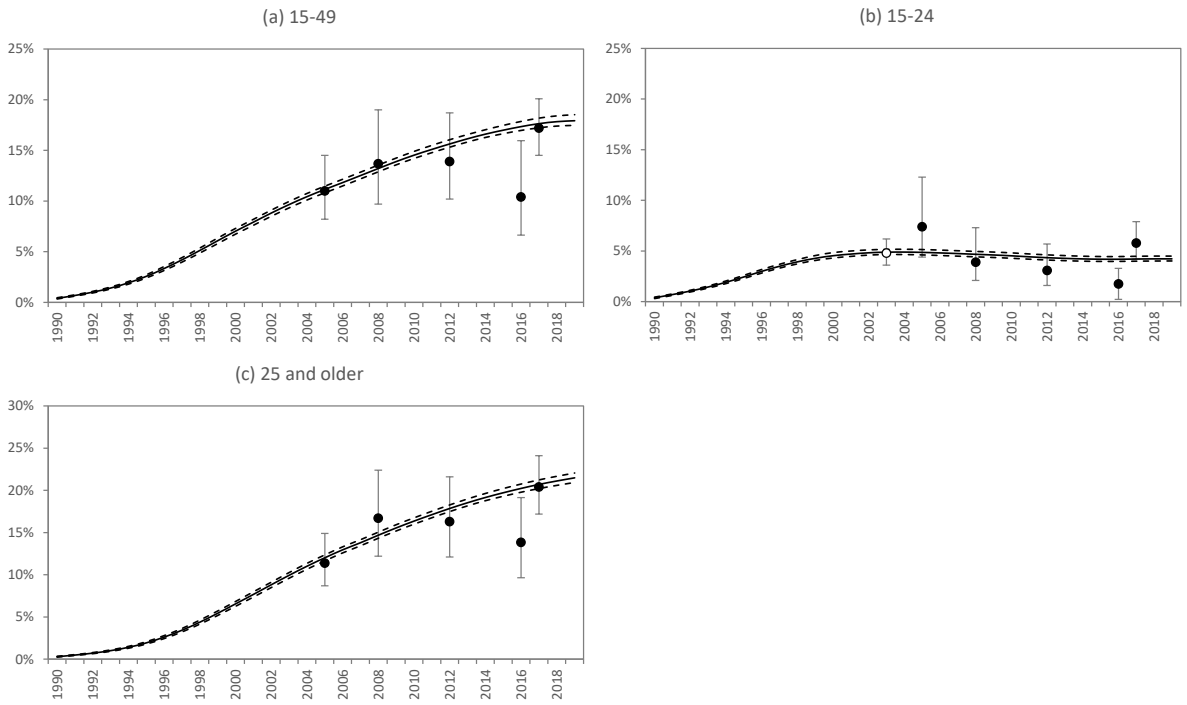

Figure 9.24: HIV prevalence in the general adult population of Limpopo
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.25 shows the model fit to the adult ART data. The model is reasonably consistent with the reported ART totals (panel d). However, the model estimate of ART coverage in women is somewhat higher than that measured in the 2012 HSRC survey but lower than that measured in the 2017 survey (panel b). The model does not match the age distribution of adult ART patients (panels e and f): as in the Eastern Cape, the model under-estimates the proportion of female ART patients in the 20-29 age group, and over-estimates the proportion of male ART patients in the older age groups.

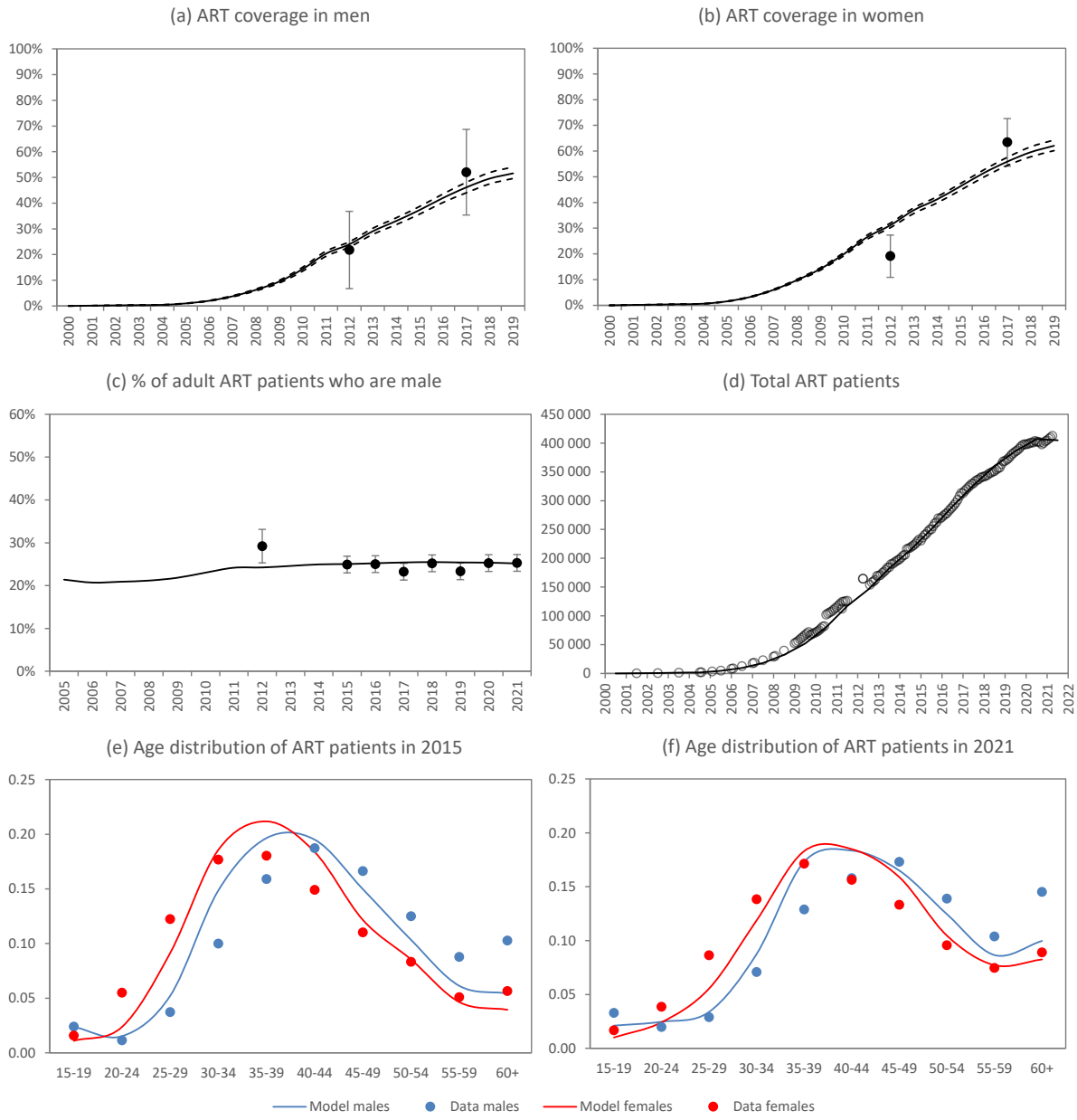

Figure 9.25: Adult ART calibration in Limpopo
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.26 shows that the model does not match the adult mortality data well, especially in men. The data suggest a steeper decline in adult male mortality than is estimated by the model.

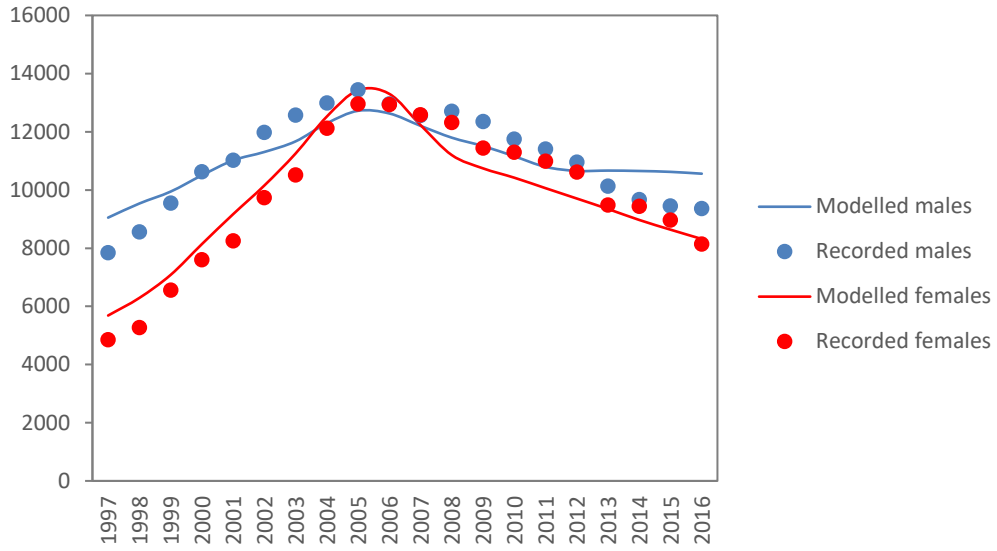

Figure 9.26: Deaths due to all causes in adults aged 20-59, Limpopo
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.27 shows the model fit to the Limpopo paediatric data. Although the model is roughly consistent with the paediatric ART data from 2014 onward, model estimates of numbers of children on ART before 2010 are substantially lower than the programme data before 2010 (panel a). Some of this can be explained by programme data reflecting cumulative rather than current enrolment, but it may also reflect unrealistic model assumptions, and further work is required to improve the calibration procedure. The model estimates of HIV prevalence in children are consistent with the three most recent surveys, but are significantly lower than the prevalence measured in the 2005 HSRC survey (panel b). The modelled age distribution of children on ART is reasonably consistent with the programme data (panel c).

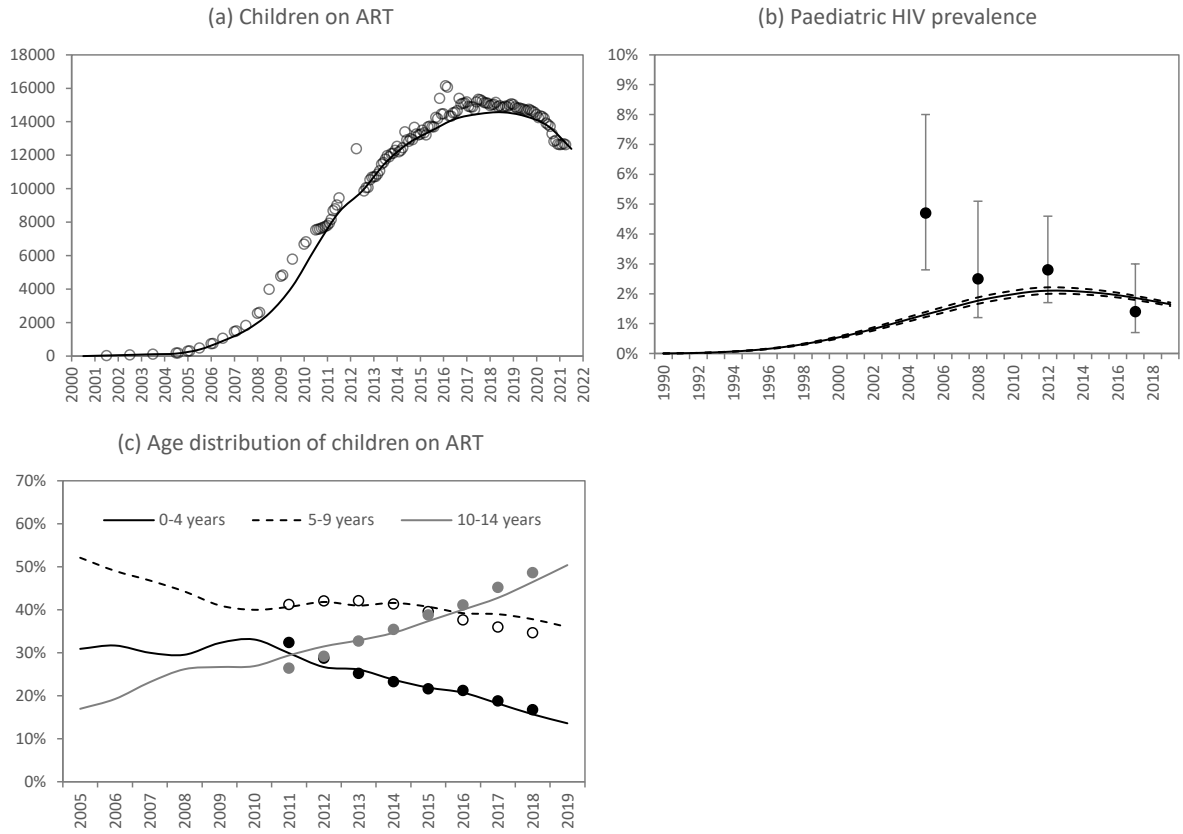

Figure 9.27: Paediatric calibration in Limpopo
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.6 Mpumalanga calibration

Figure 9.28 shows the model fit to the antenatal prevalence data from Mpumalanga. The model appears to fit the data reasonably well.

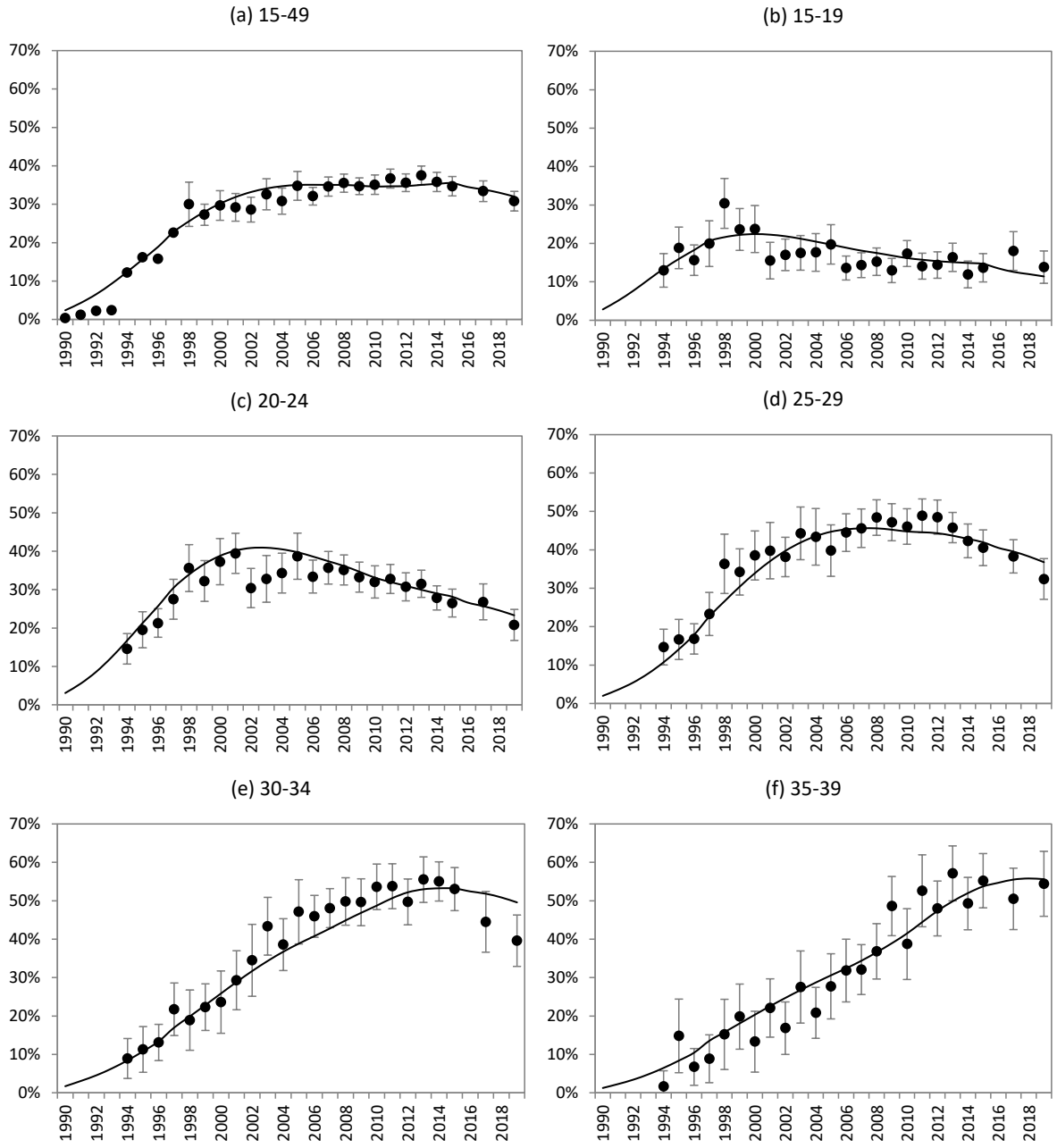

Figure 9.28: HIV prevalence levels in pregnant women attending public antenatal clinics in Mpumalanga

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.29 shows the model fit to the Mpumalanga adult HIV prevalence data from household surveys. The model is reasonably consistent with the surveys.

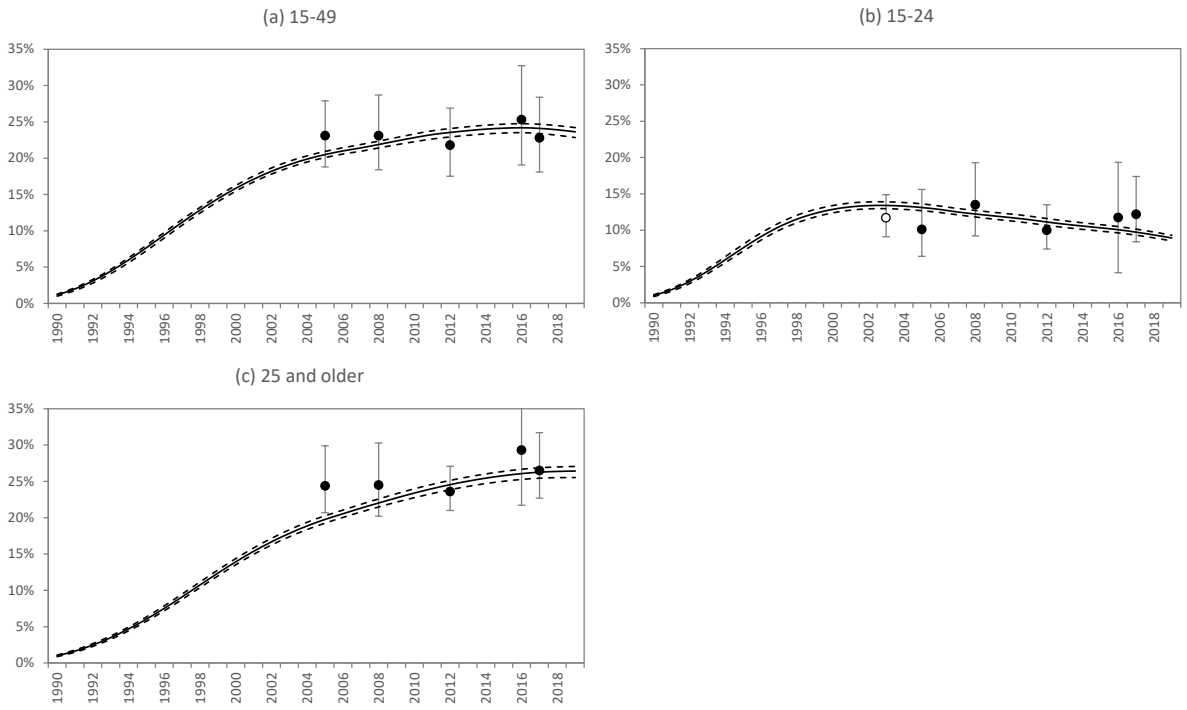

Figure 9.29: HIV prevalence in the general adult population of Mpumalanga
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.30 shows the model fit to the adult ART data. The model is roughly consistent with the programme data, although the model slightly under-estimates the ART numbers in the most recent year (panel d). As in the Eastern Cape, the programme data suggest a substantial reduction in the proportion of adult ART patients who are male between 2012 and 2015-2021, and the model fits the latter set of data points reasonably well (panel c). Survey estimates of ART coverage in men are consistent with the model (panel a), and the modelled age distribution of ART patients is reasonably consistent with programme data (panels e and f).

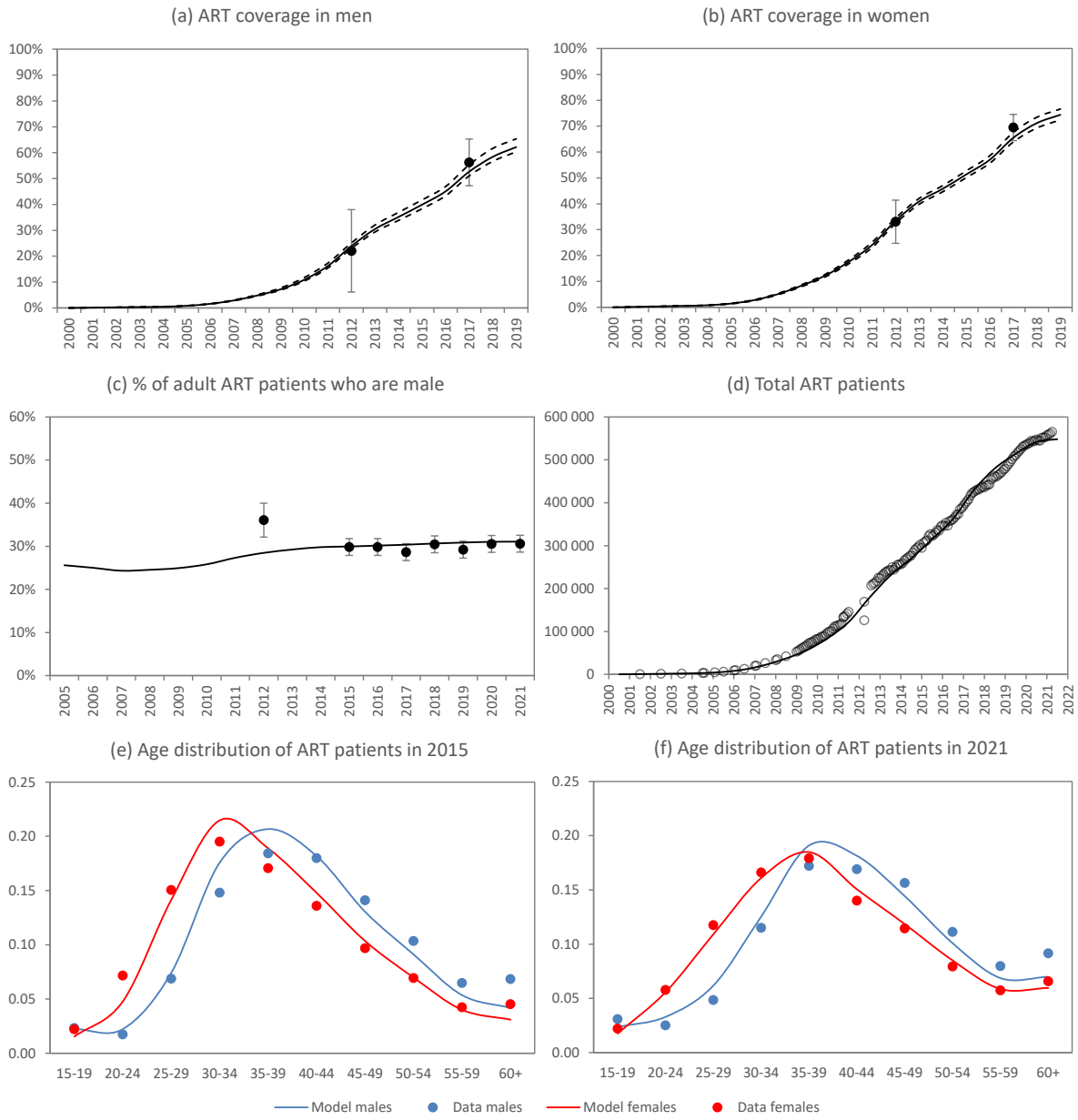

Figure 9.30: Adult ART calibration in Mpumalanga
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.31 shows that the model matches the vital registration data well.

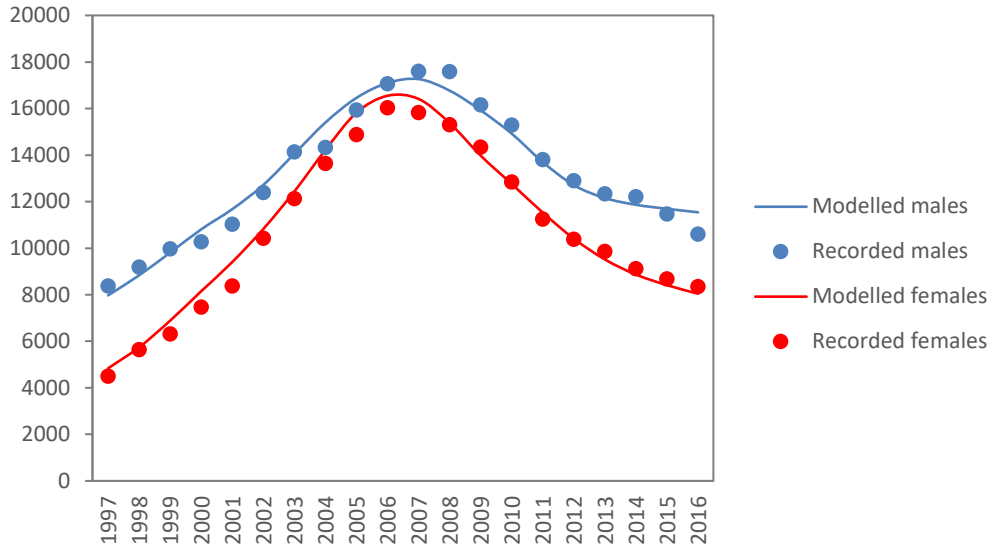

Figure 9.31: Deaths due to all causes in adults aged 20-59, Mpumalanga
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.32 shows the model calibration to the paediatric data. The model estimates of numbers of children on ART in recent years are roughly consistent with the programme data, although there appears to be some noise in the data (panel a). In the period before 2010, the model appears to under-estimate the reported ART totals, again because the latter reflect cumulative ART enrolment. Model estimates of HIV prevalence in children differ substantially from the household survey data (panel b). However, it is difficult to discern a consistent pattern of discrepancy; while the model estimates of HIV prevalence in 2005 and 2017 are substantially lower than the corresponding survey estimates, the model estimate of HIV prevalence in 2012 is substantially higher than the survey estimate. The modelled age distribution of children on ART is consistent with programme data (panel c).

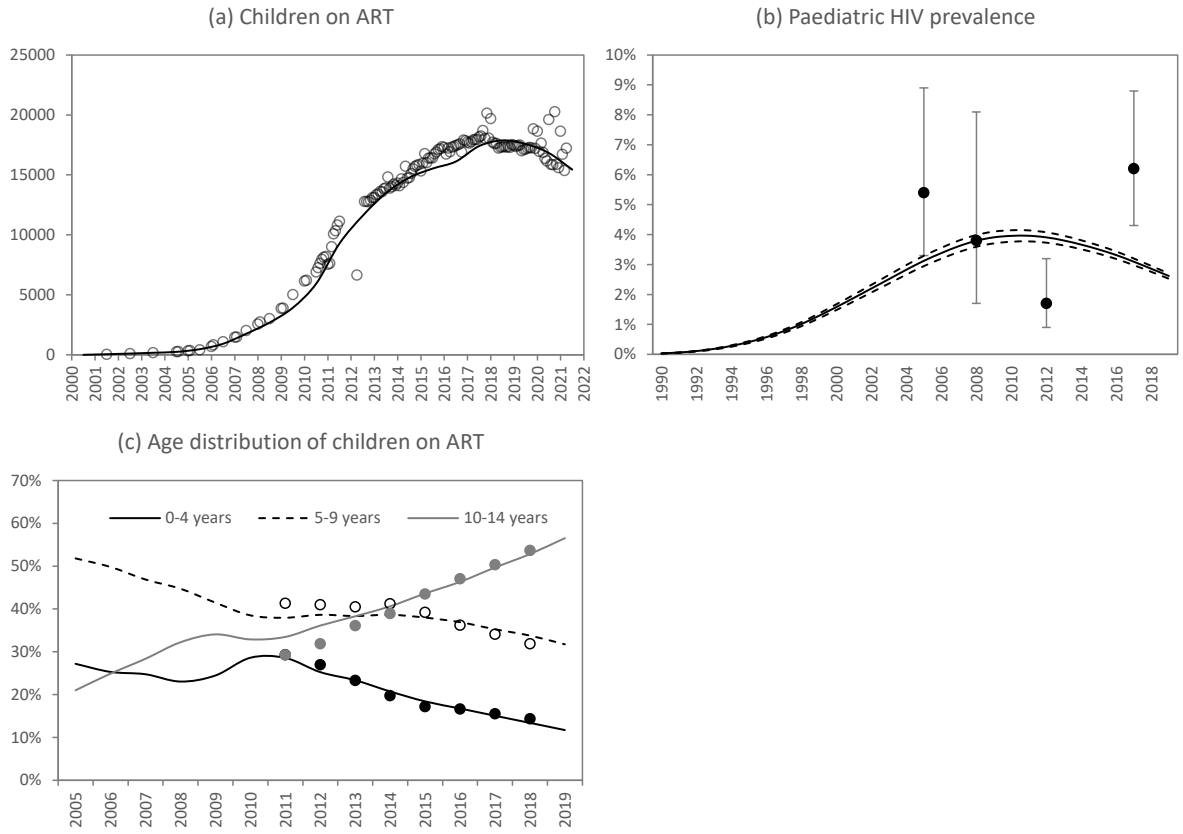

Figure 9.32: Paediatric calibration in Mpumalanga
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.7 Northern Cape calibration

Figure 9.33 shows the model fit to the antenatal prevalence data from Northern Cape. The model fits the data reasonably closely in the younger age groups, but in the period after 2010 the model over-estimates HIV prevalence in the 30-39 age group. The small size of the Northern Cape population means that confidence intervals around the survey estimates are generally wide.

Figure 9.34 shows the model fit to the HIV prevalence data from household surveys. Although the model is consistent with the most recent survey in 2017, the model tends to over-estimate HIV prevalence when compared against earlier surveys, which suggests that estimates of HIV prevalence may be too high (especially in the age group 25 and older). However, reducing the model estimate of HIV prevalence would have the effect of increasing the estimated ART coverage, which would lead to worse fits to the data in Figure 9.29.

Figure 9.35 shows the model fit to the adult ART data. The model appears roughly consistent with the recorded treatment totals, although the model appears to under-estimate the numbers of ART patients over the 2014-2018 period. The model estimates of ART coverage in males are higher than those measured in the surveys, especially in 2017 (panel a).

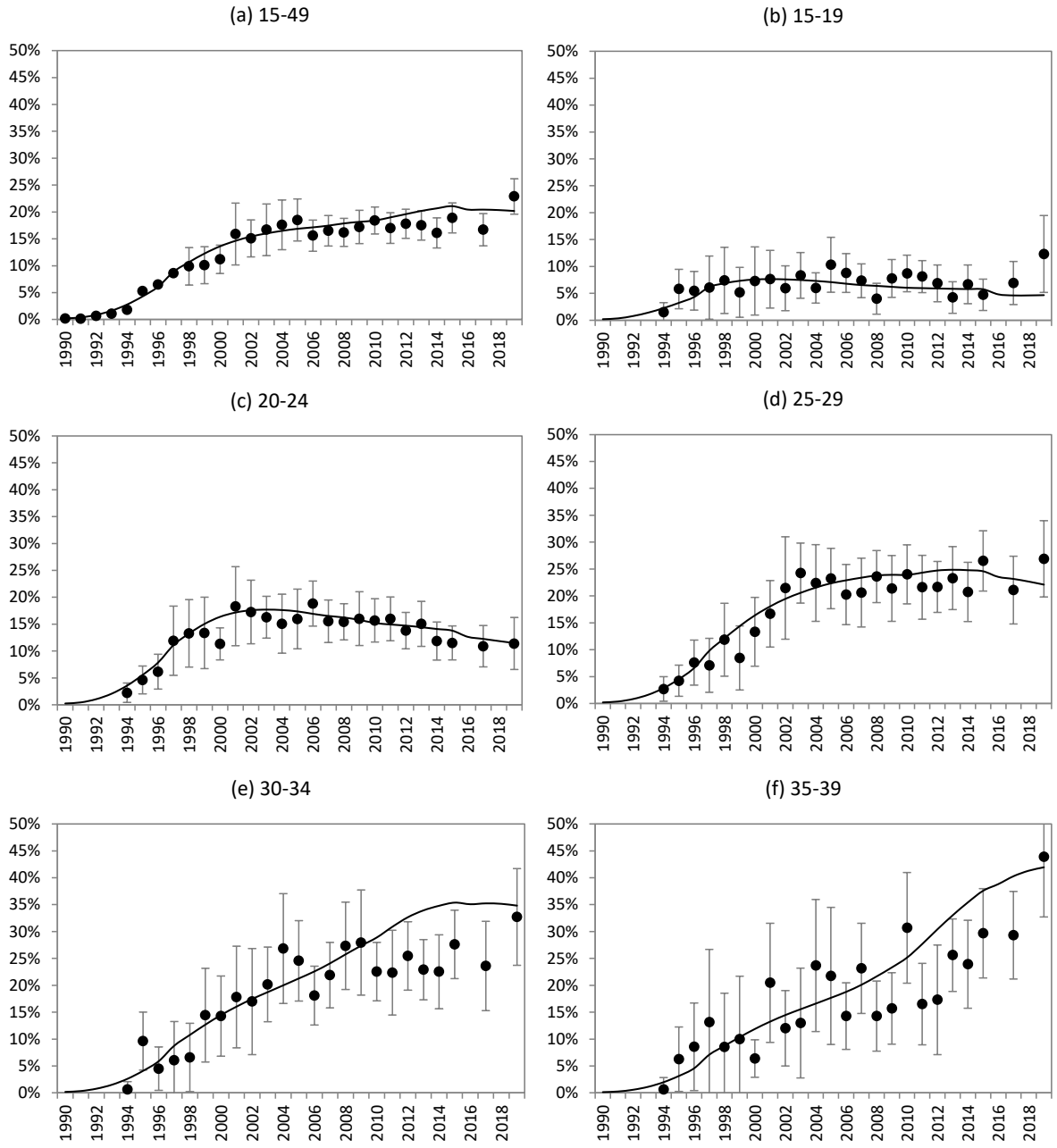

Figure 9.33: HIV prevalence levels in pregnant women attending public antenatal clinics in Northern Cape

Solid lines represent posterior means. Dots represent antenatal survey estimates.

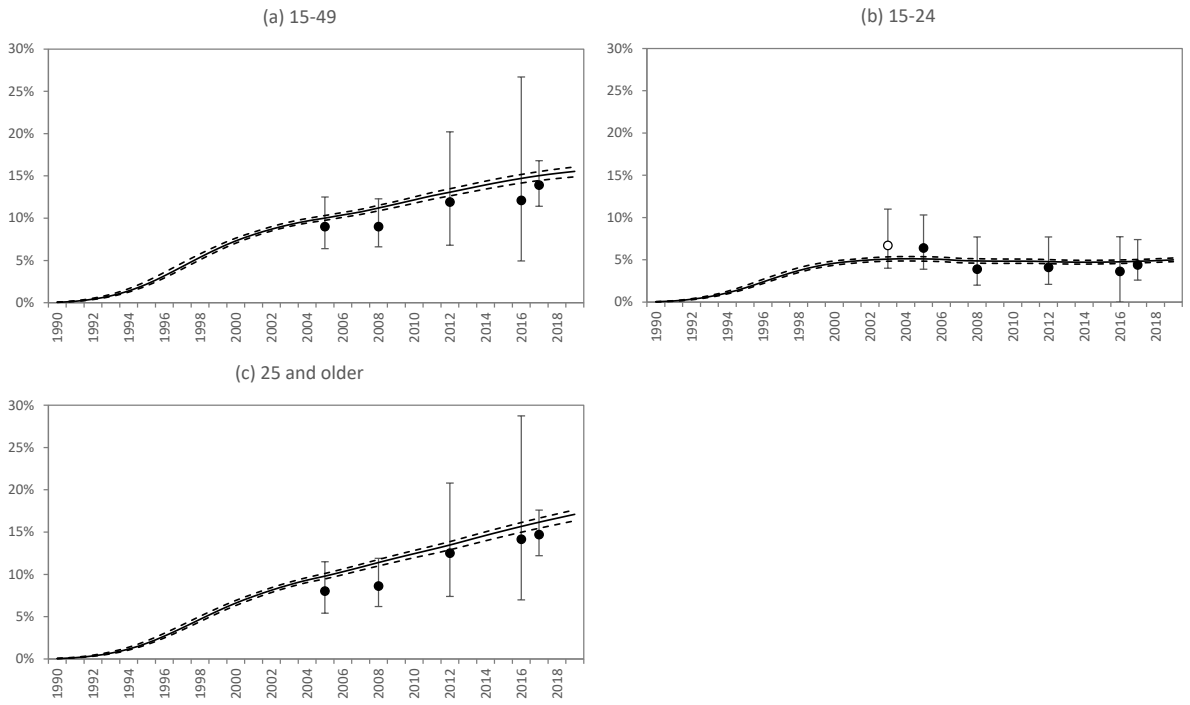

Figure 9.34: HIV prevalence in the general adult population of Northern Cape
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

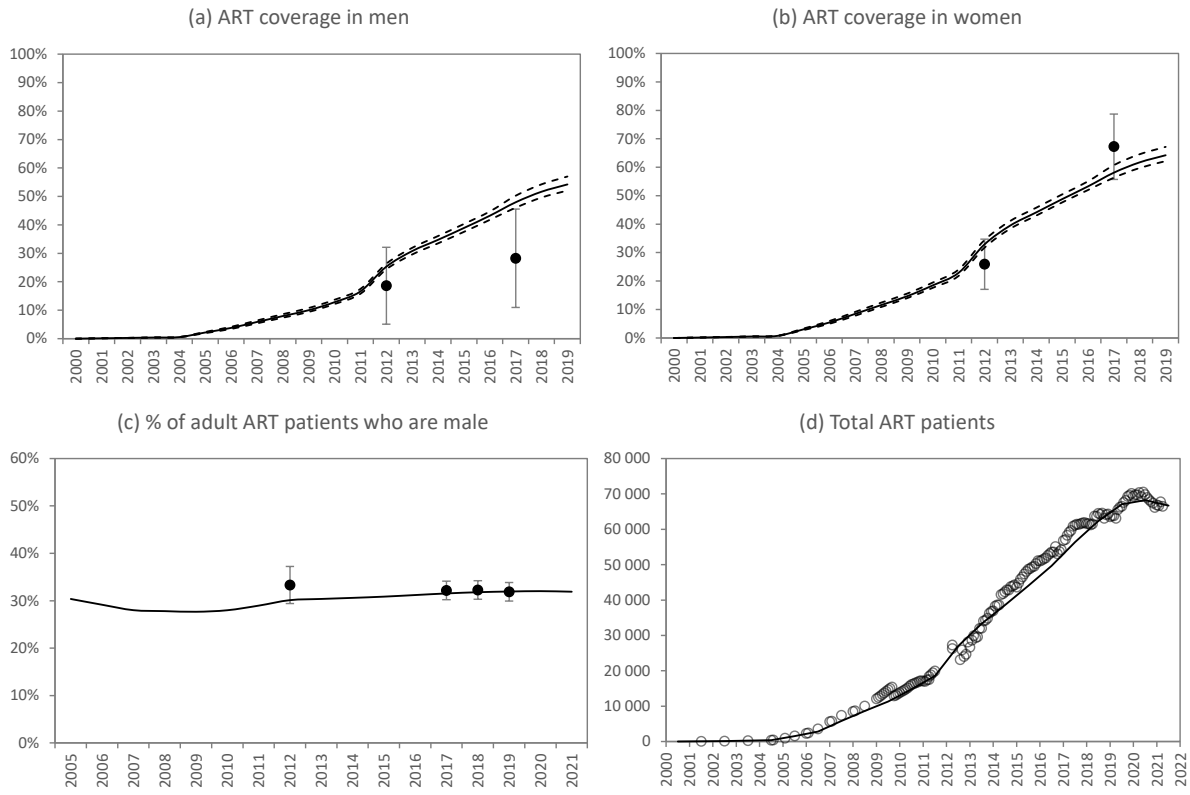

Figure 9.35: Adult ART calibration in Northern Cape

Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.36 shows the model fit to the recorded death data. Although the model is generally consistent with the female data, model estimates for men appear too low in the 2006-2016 period.

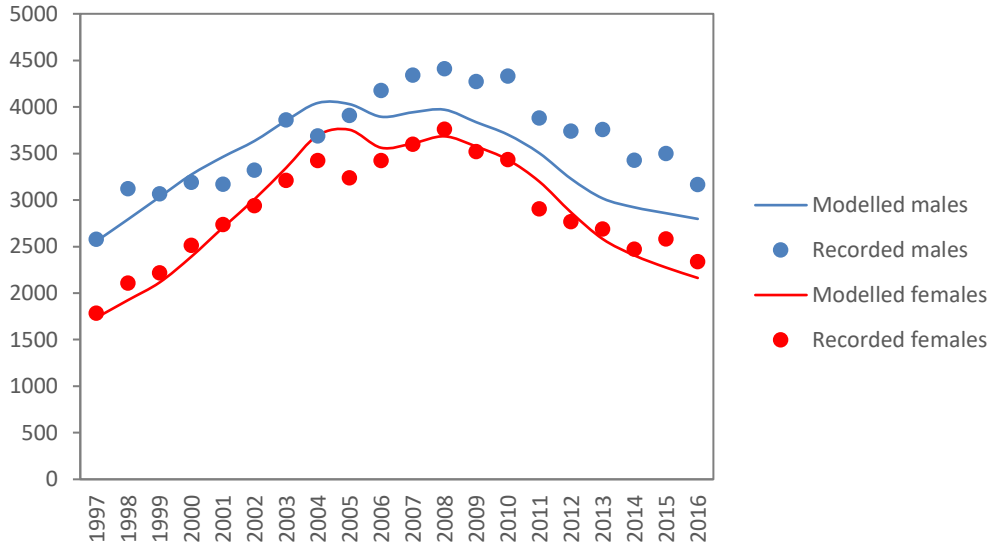

Figure 9.36: Deaths due to all causes in adults aged 20-59, Northern Cape
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.37 shows the model fit to the paediatric data sources. The model estimates of numbers of children on ART are roughly consistent with the programme data in recent years (panel a), although it should be noted that the programme data have been adjusted to correct for the mis-reporting of adults on ART in Kimberley Hospital as the numbers of children on ART, and vice versa (Jeffrey Eaton, personal communication). The quality of reporting in earlier years is clearly also a concern. The model estimates of HIV prevalence in children appear reasonably consistent with the survey data (panel b). The modelled age distribution of children on ART is slightly inconsistent with the data: the model over-estimates the proportion in the 5-9 age group but under-estimates the proportion in the 10-14 age group.

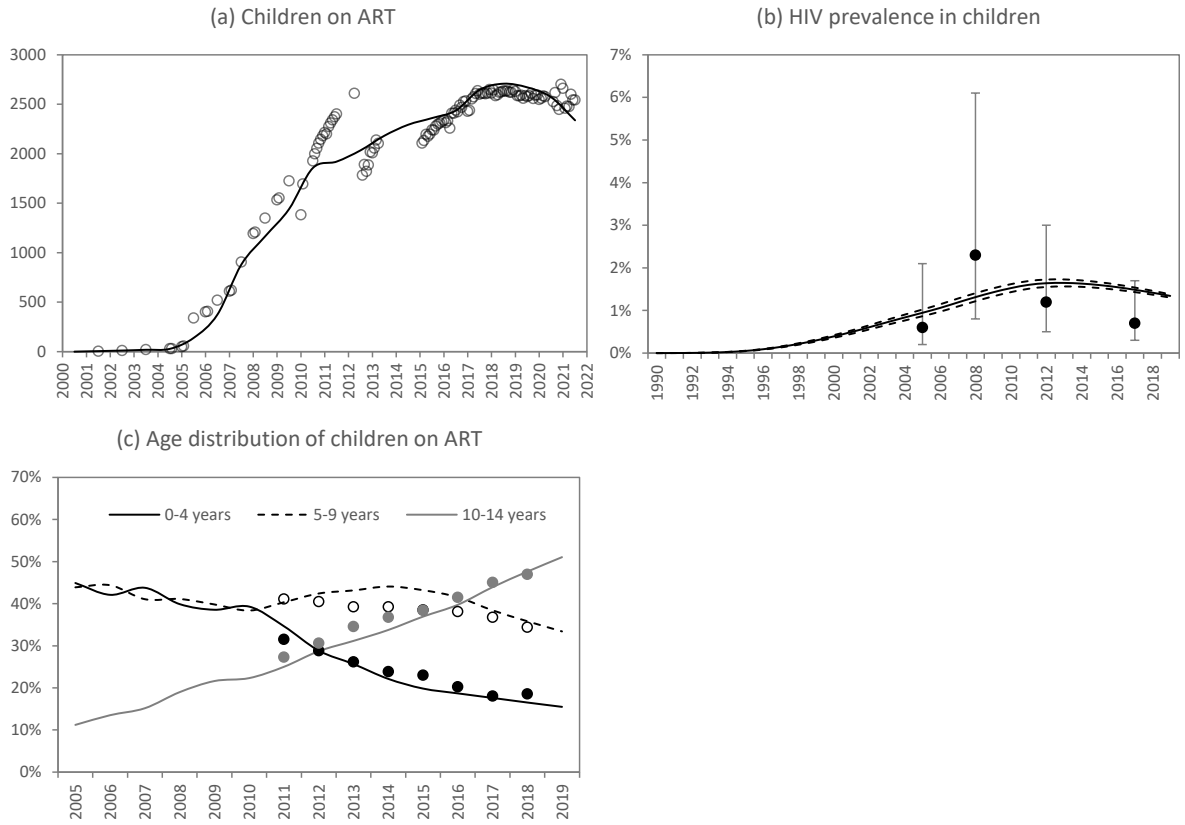

Figure 9.37: Paediatric calibration in Northern Cape
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.8 North West calibration

Figure 9.38 shows the model fit to the antenatal prevalence data from North West. The overall model fit to the data is reasonable, although the model slightly over-estimates HIV prevalence in recent years among women aged 30-39. The early antenatal surveys contain a number of outlier prevalence measurements, although it is only the outlier in 1996 that affects the model calibration, since the pre-1994 data are not age-disaggregated and have therefore not been used in the model calibration.

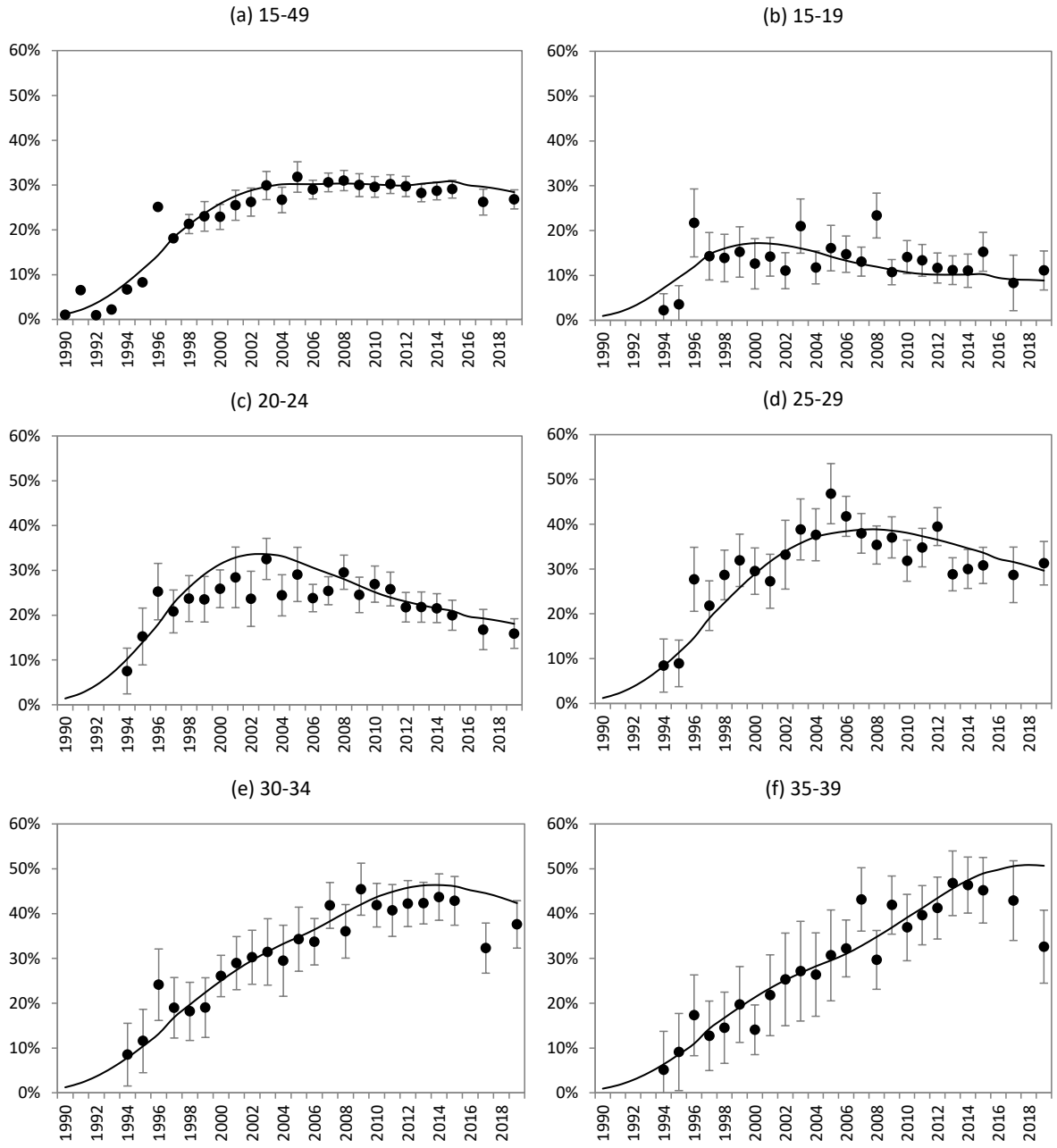

Figure 9.38: HIV prevalence levels in pregnant women attending public antenatal clinics in North West

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.39 shows the model fit to the HIV prevalence data from household surveys. The model appears consistent with the data, although the model estimates of HIV prevalence appear a bit low when compared against the two most recent surveys, especially in the 25+ age group.

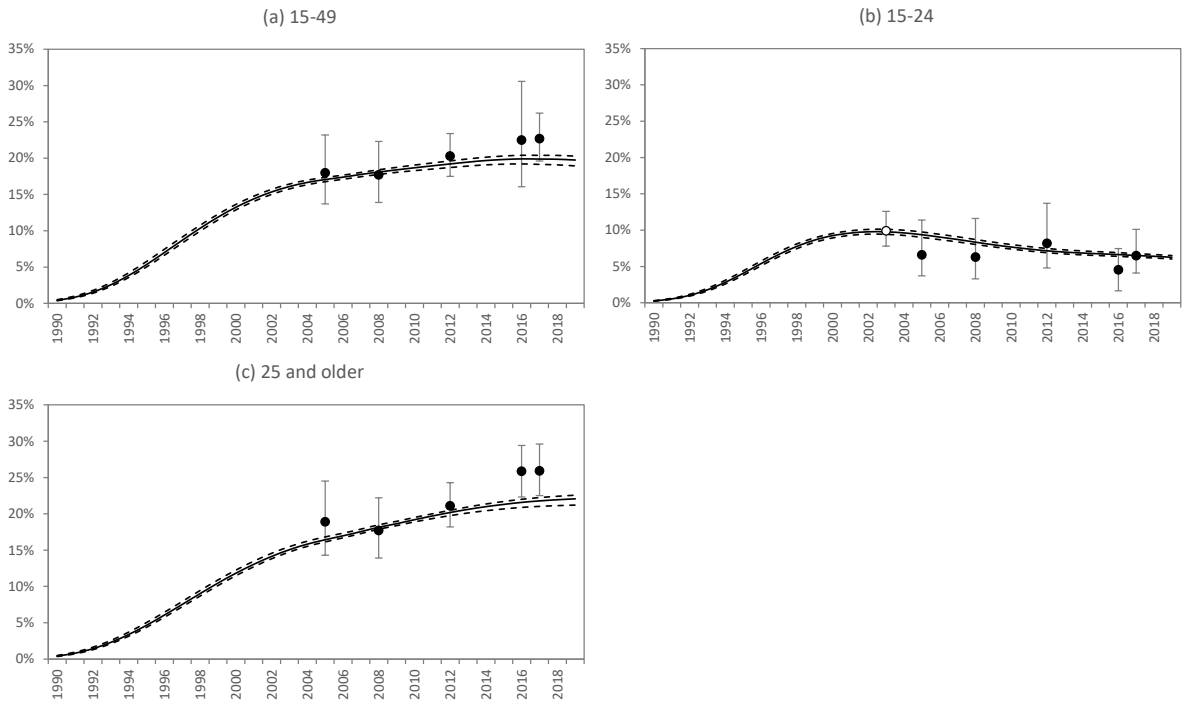

Figure 9.39: HIV prevalence in the general adult population of North West
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.40 shows the model fit to the adult ART data. The model appears roughly consistent with the recorded ART numbers in the period after 2015 (panel d). However, there is poor consistency with the data in the earlier periods, as North West transitioned to reporting current enrolment later than the other provinces, and the reported numbers before 2015 mostly reflect cumulative enrolment (i.e. we would expect these to exceed the model numbers of current enrolment). Model estimates of ART coverage are also consistently lower than survey estimates (panels a and b). However, improving the model fits to these coverage data would only be possible if we either (a) increased the numbers of adults on ART (which would give a worse fit to the data in panel d) or (b) reduced the model estimates of adult HIV prevalence (which would give worse fit to the HIV prevalence data in recent years in Figure 9.39). The modelled age distribution of female ART patients is roughly consistent with the programme data in 2021 but not in 2015. The modelled age distribution in men does not match the data well: the model estimates too many men on ART in the 35-39 age range and not enough men on ART at ages 50 and older.

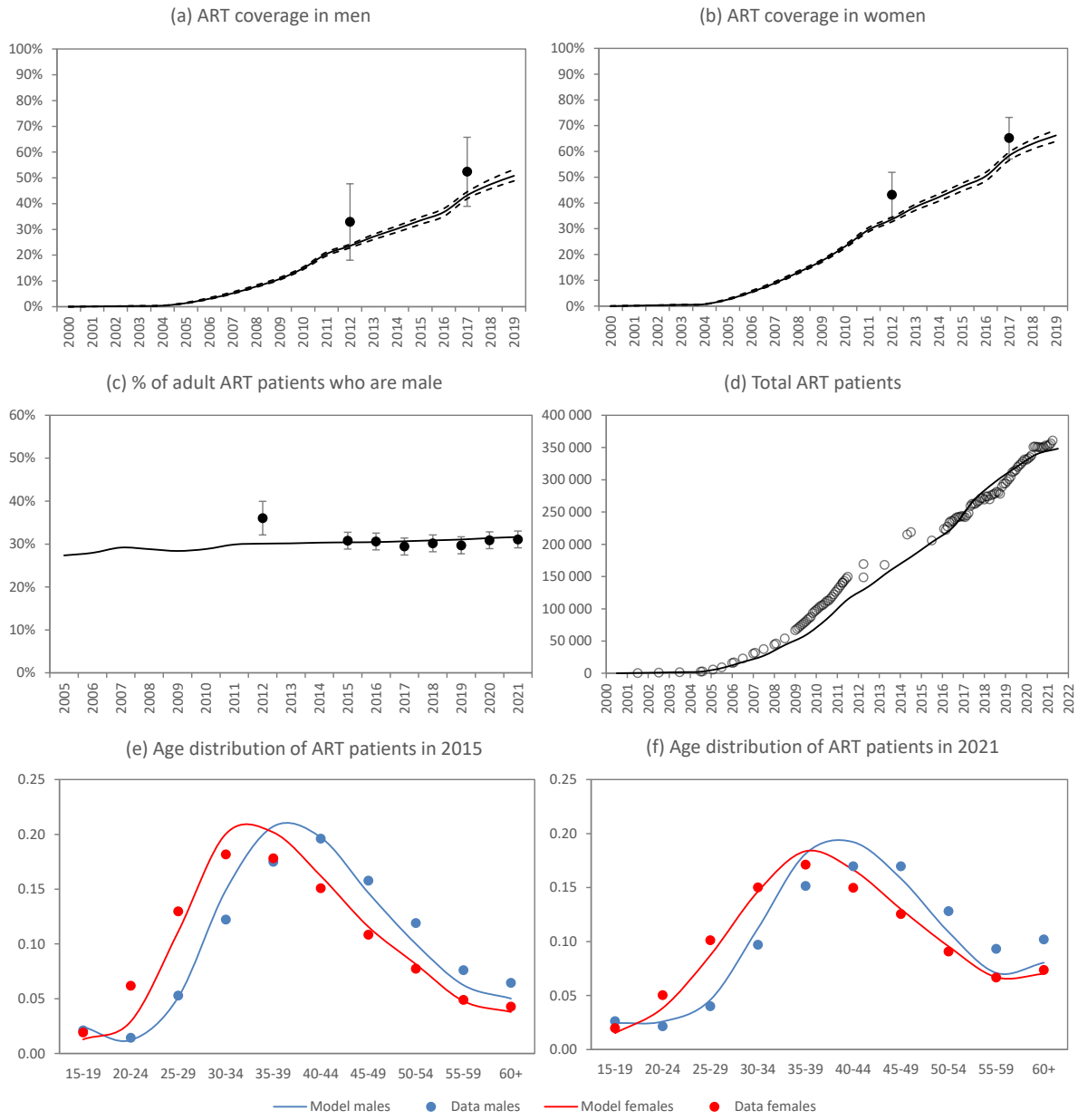

Figure 9.40: Adult ART calibration in North West

Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.41 compares the model estimates of adult mortality with the vital registration data. The model matches the vital registration data reasonably well.

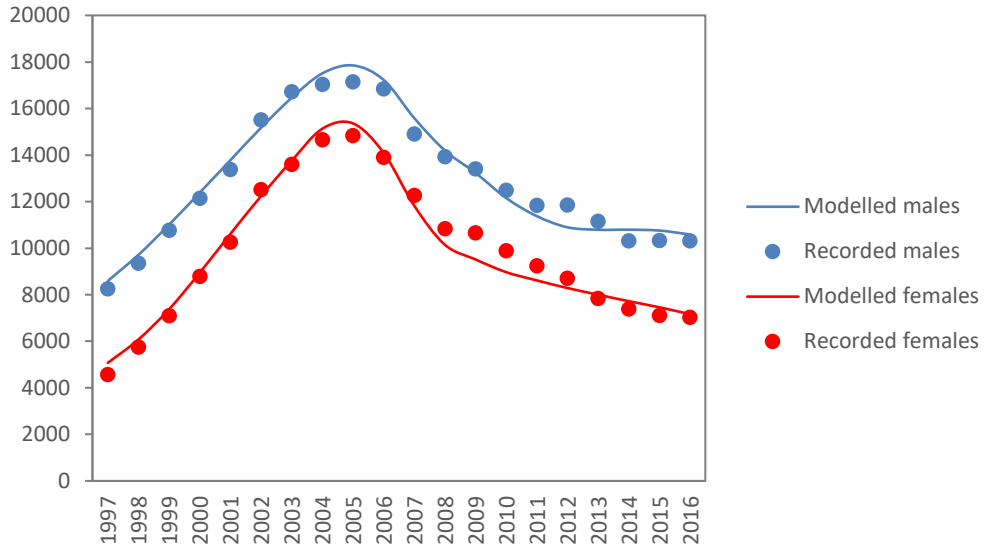

Figure 9.41: Deaths due to all causes in adults aged 20-59, North West
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.42 shows the model fit to the paediatric data. The fit to the programme data is generally poor (panel a). In the period before 2012, the model estimates are mostly below the reported numbers, for the same reason as in adults. However, the model estimates are roughly consistent with the data in the most recent years (note that the data over the 2015-18 period have been corrected, to adjust for an error in one district, which included vertically infected youth over age 14 in the paediatric totals). Model estimates of HIV prevalence in children are consistent with the survey data (panel b). The model estimates of the age distribution of children on ART are also consistent with programme data, although the model slightly over-estimates the proportion of children aged 5-9 in recent years (panel c).

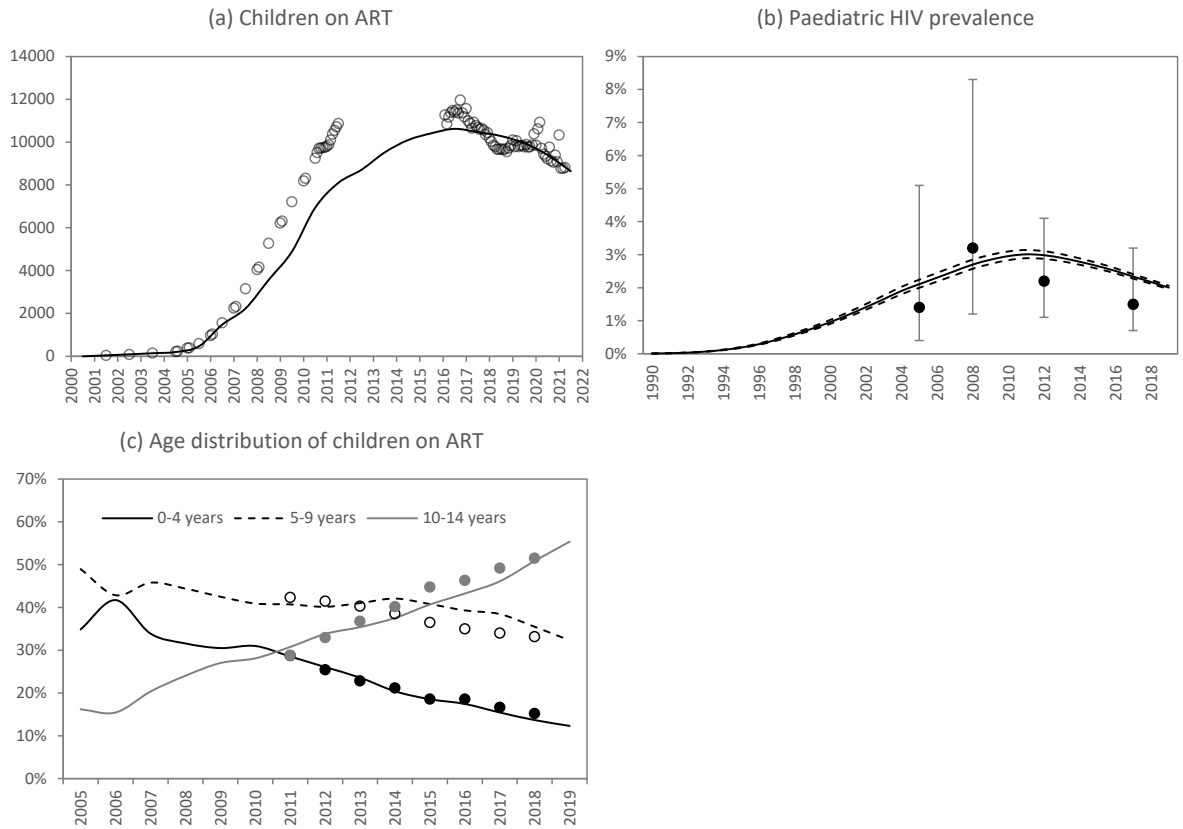

Figure 9.42: Paediatric calibration in North West
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.2.9 Western Cape calibration

Figure 9.43 shows the model fit to the antenatal prevalence data from Western Cape. The model fits the data reasonably closely.

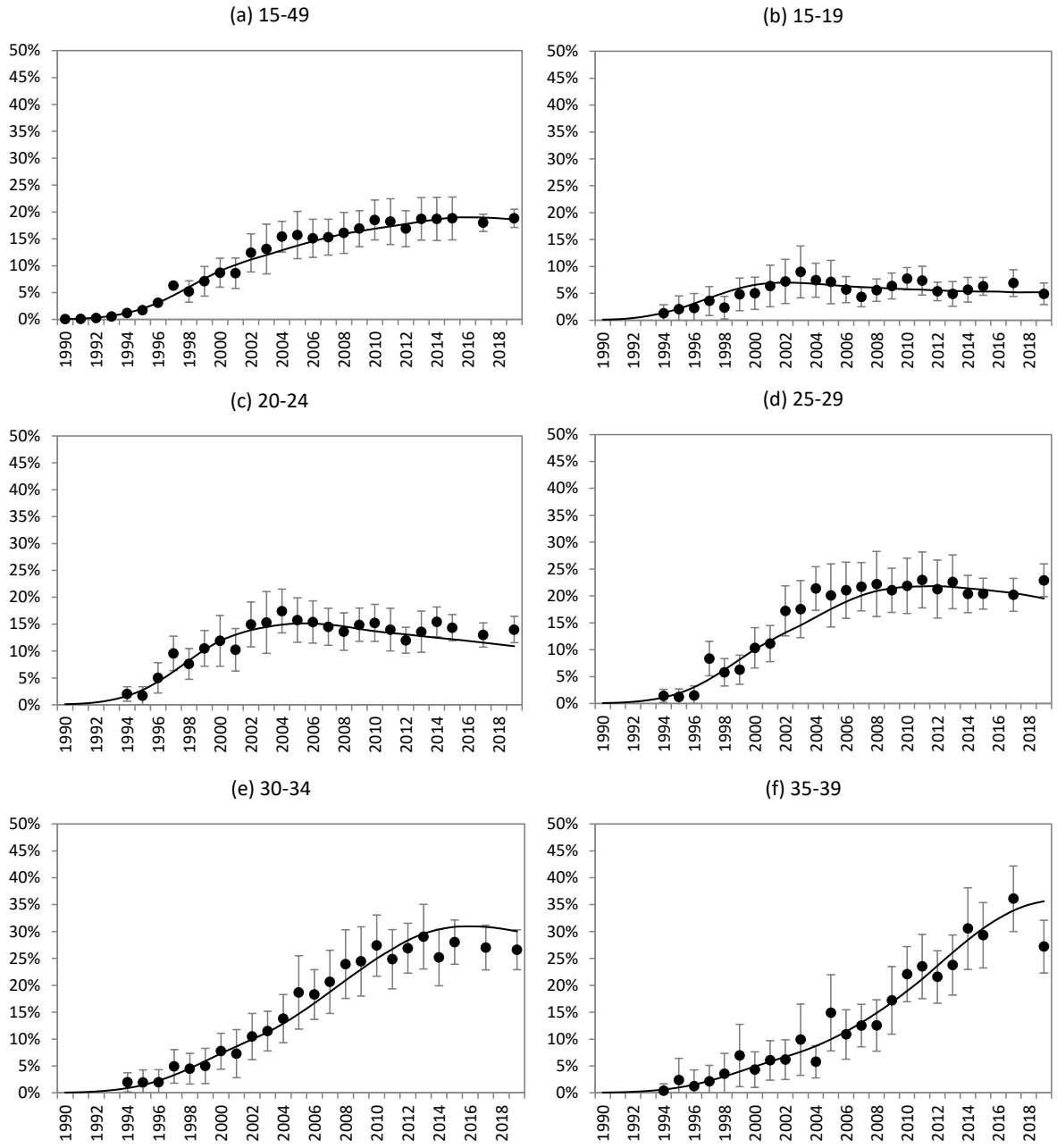

Figure 9.43: HIV prevalence levels in pregnant women attending public antenatal clinics in Western Cape

Solid lines represent posterior means. Dots represent antenatal survey estimates.

Figure 9.44 compares the model estimates of HIV prevalence in the general population with the data from the household surveys. On the whole, the model fit to the data is poor; the data suggest a much steeper increase in HIV prevalence over the 2005-2017 period than the model estimates. However, changing the model to give a steeper increase in prevalence would lead to model estimates of HIV prevalence in pregnant women much less consistent with the data from the antenatal surveys (Figure 9.35).

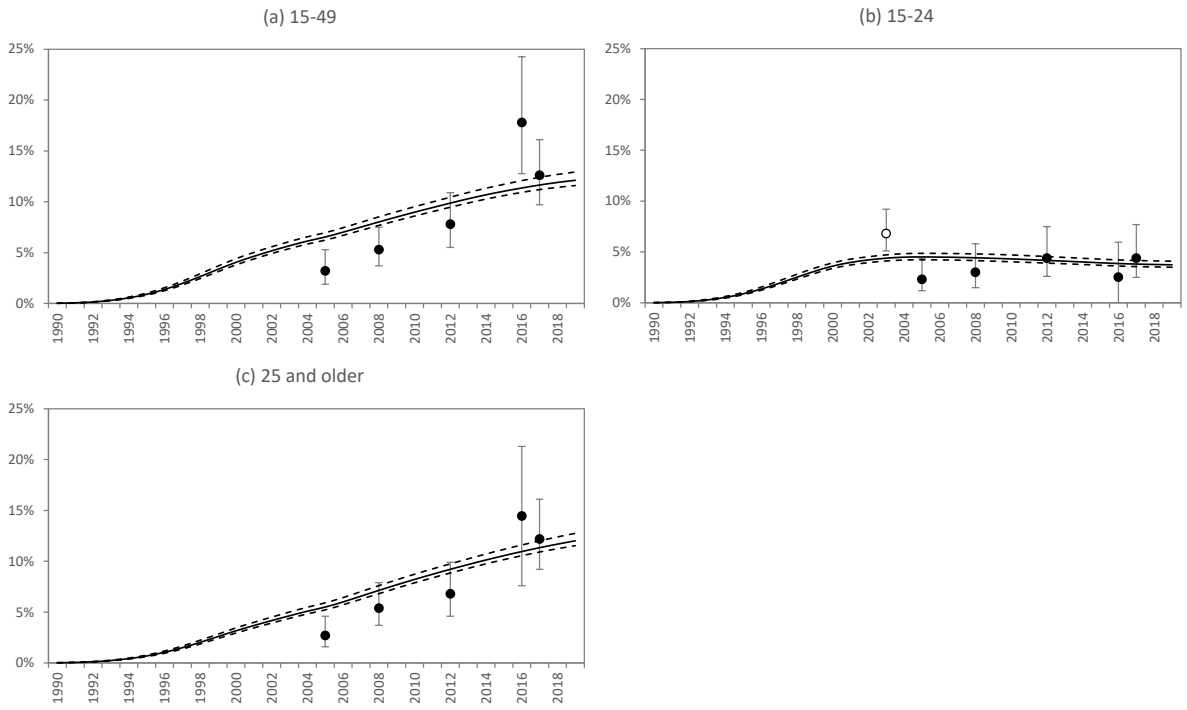

Figure 9.44: HIV prevalence in the general adult population of the Western Cape
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data (the 2003 youth survey data point was not used in calibration).

Figure 9.45 shows the model fit to the adult ART data. The model fit to the ART programme data is good (panel d). Unlike the other provinces, Western Cape has always reported current enrolment, not cumulative enrolment, so we would expect to see consistency between the reported totals and the modelled totals in the period before 2010. The Western Cape also differs from the other provinces in having more data on the proportion of adult ART patients who are male, and the model appears quite consistent with these data (panel c). When compared against the survey estimates, model estimates of ART coverage appear slightly too low in 2017. The modelled age distribution of adults on ART is reasonably consistent with programme data (panels e and f).

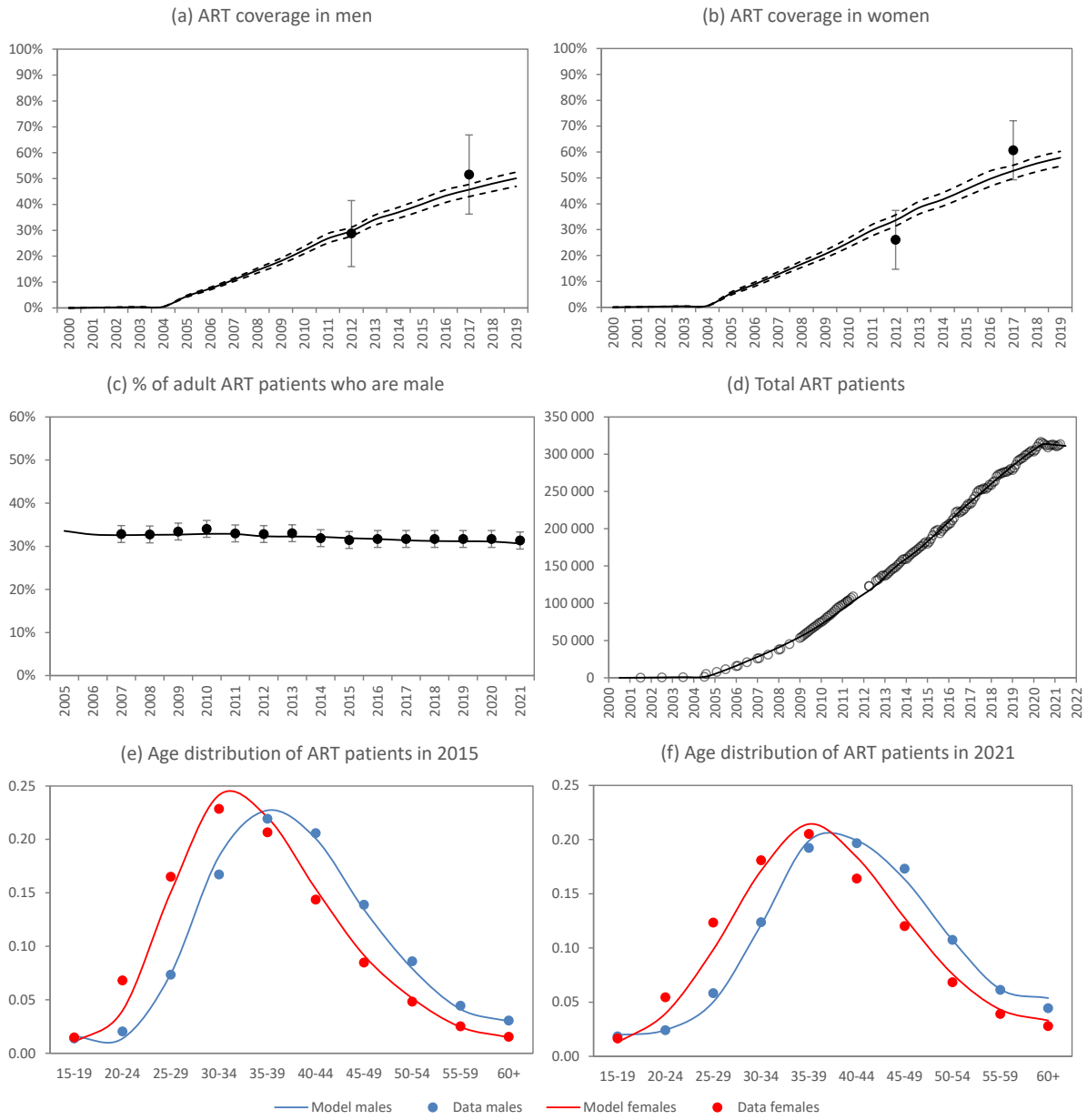

Figure 9.45: Adult ART calibration in Western Cape
Solid lines represent posterior means and dashed lines represent 95% confidence intervals. Dots represent calibration data.

Figure 9.46 shows the model fit to the vital registration data. Although the model matches the female trends reasonably well, estimates of mortality in men tend to be lower than reported in the period after 2005.

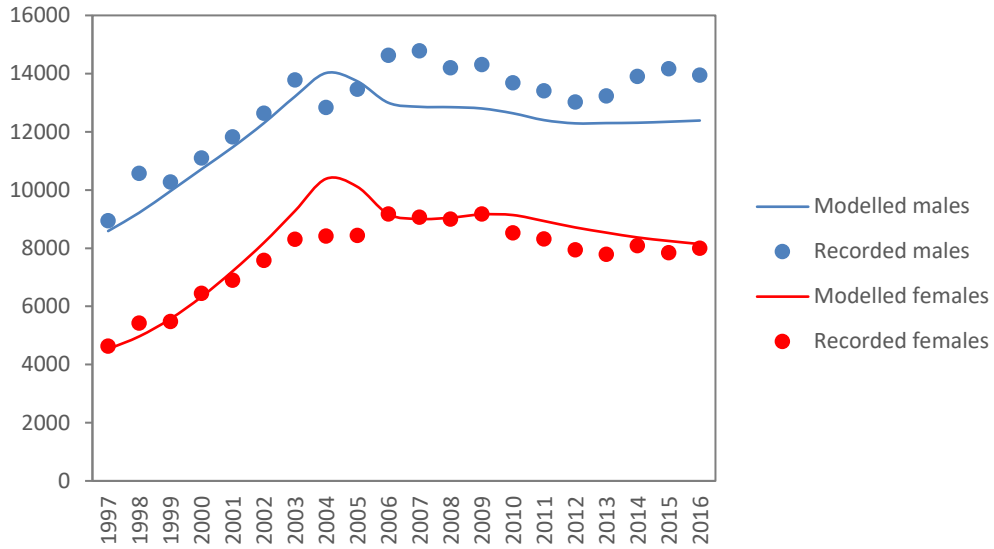

Figure 9.46: Deaths due to all causes in adults aged 20-59, Western Cape
Recorded deaths have been adjusted to reflect incomplete reporting, as described in section 8.1.6. Solid lines represent posterior means.

Figure 9.47 shows the model calibration to the paediatric HIV data sources. The model estimates of paediatric ART enrolment are slightly lower than recorded numbers over the 2010-2013 period, but are otherwise roughly consistent with the reported data (panel a). Model estimates of HIV prevalence in children are consistent with most of the household survey data, but appear too low when compared against the 2017 survey result, which is a clear outlier (panel b). The modelled age distribution of children on ART is very inconsistent with the programme data, with the model substantially over-estimating the proportions of children aged 5-9 and under-estimating the proportions aged 10-14 (panel c)

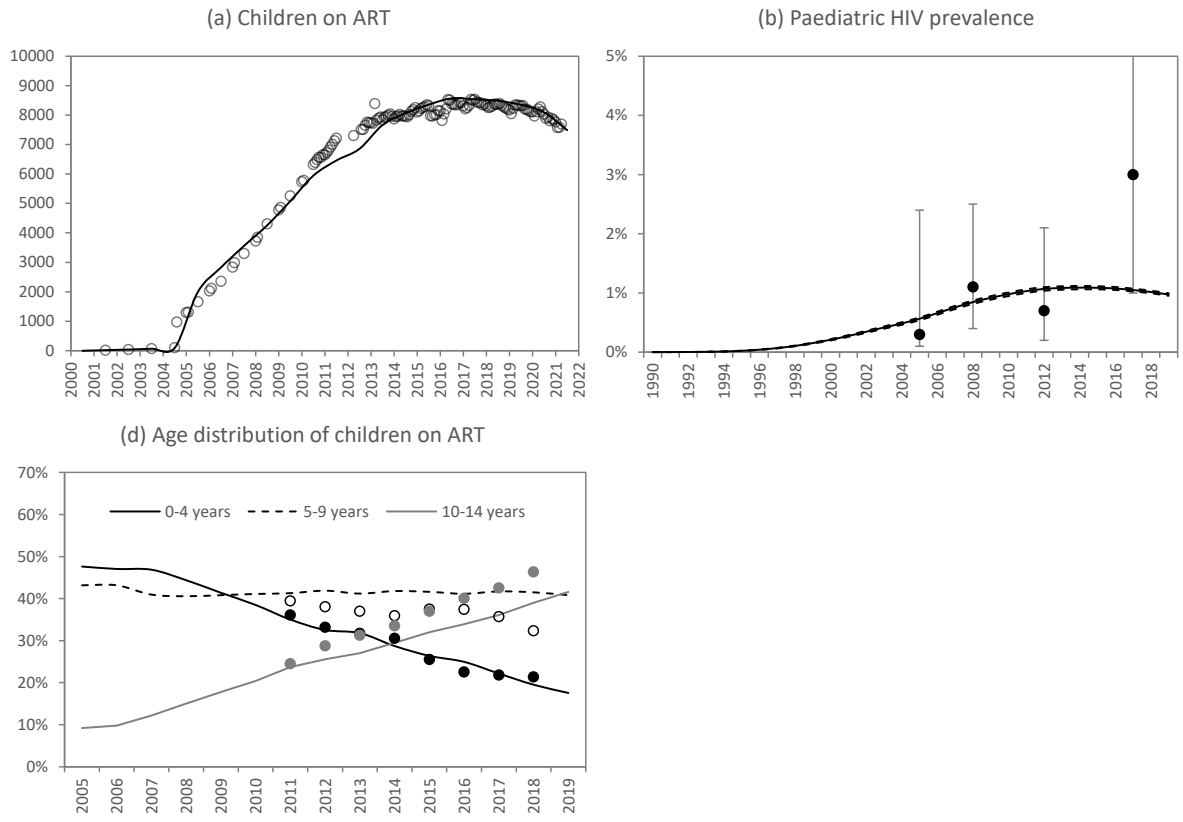

Figure 9.47: Paediatric calibration in Western Cape
Solid lines represent posterior means and dashed lines in panel b represent 95% confidence intervals. Dots represent calibration data.

9.3 HIV incidence outputs

Figure 9.48(a) compares HIV incidence trends by province, in the 15-49 year age group. In all provinces, HIV incidence peaked between 1998 and 2004 (with later peaks in Limpopo, Western Cape and Eastern Cape) and has been steadily declining since then. However, provinces differ in the pace of the HIV incidence decline. The percentage reduction in HIV incidence between the start of 2000 and the start of 2020 is greatest in KwaZulu-Natal (77%) and Free State (74%), but is relatively modest in the Western Cape (33%), Limpopo (44%) and Northern Cape (47%). Similar rankings are observed when considering the HIV incidence decline over the period from the start of 2010 to the start of 2020, with incidence declines being lowest in Western Cape (33%) Northern Cape (39%) and Limpopo (43%) and greatest in KwaZulu-Natal (66%), Free State (62%) and Mpumalanga (61%). Figure 9.48(b) compares the HIV incidence rates in 2020-21 across provinces: incidence rates are highest in Eastern Cape (1.09%, 95% CI: 1.00-1.22%), and lowest in Western Cape (0.51%, 95% CI: 0.47-0.57%) and Gauteng (0.53%, 95% CI: 0.47-0.64%). These incidence estimates should be interpreted with caution, as we do not have any HIV prevalence data after 2019, and estimates of HIV incidence over the 2020-21 period are therefore projections of past trends.

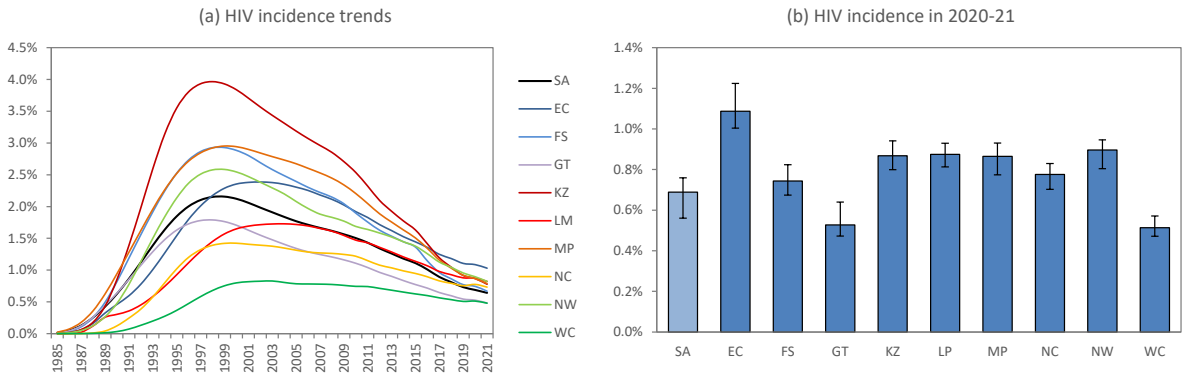

Figure 9.48: HIV incidence among adults aged 15-49
Solid lines in panel (a) and bars in panel (b) represent posterior means. Error bars in panel (b) represent 95% confidence intervals (2.5 and 97.5 percentiles of the posterior distributions).

9.4 HIV prevalence outputs

Figure 9.49 compares HIV prevalence trends by province. In all provinces, the prevalence has been steadily increasing over time, although in some provinces prevalence appears to have peaked (KwaZulu-Natal, Mpumalanga, Free State, North West and Gauteng). Prevalence in the 15-49 age group appears to be increasing more rapidly in the Western Cape, Eastern Cape, Northern Cape and Limpopo than the other provinces. Prevalence has consistently been highest in KwaZulu-Natal and Mpumalanga, followed by Free State and Eastern Cape in recent years. In 2021, HIV prevalence among 15-49 year olds varied between 12.4% (95% CI: 11.8-13.2%) in Western Cape and 26.1% (95% CI: 25.4-26.8%) in KwaZulu-Natal.

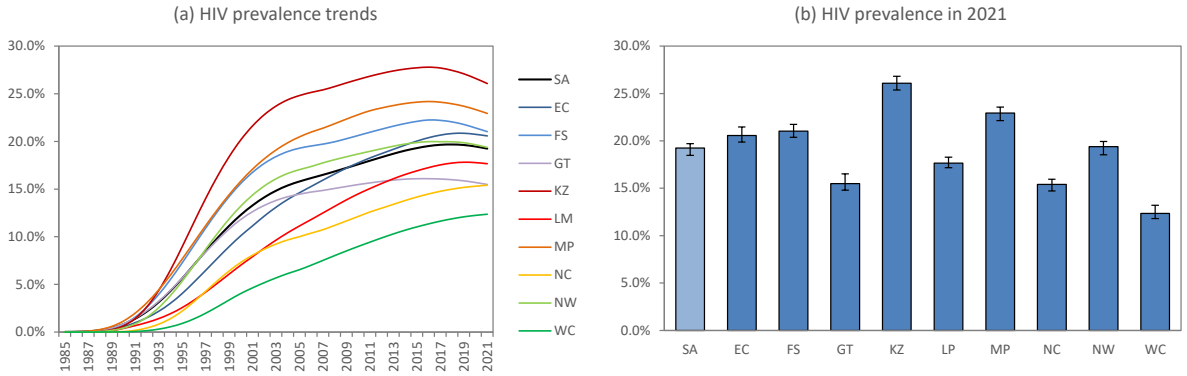

Figure 9.49: HIV prevalence in 15-49 age group, by province
Solid lines represent posterior means.

9.5 AIDS mortality outputs

Figure 9.50 shows crude AIDS mortality rates by province (total AIDS deaths in each year per 100 000 population each year). Although AIDS mortality rates have historically been highest in KwaZulu-Natal and Free State, these provinces have also seen the steepest decline in AIDS

mortality since 2005 (panel a), due to the particularly rapid rollout of ART in these provinces. In 2020-21, AIDS mortality rates were highest in North West and Eastern Cape, and AIDS mortality rates were lowest in Western Cape (panel b). However, as the model has not been calibrated to mortality data in the post-2016 period, these estimates should be treated with caution.

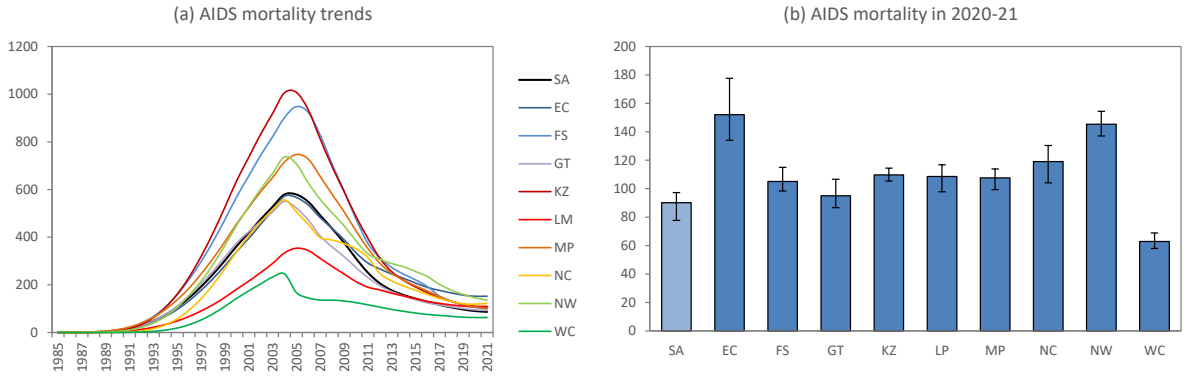

Figure 9.50: Crude AIDS mortality rates (AIDS deaths per 100 000 population)
Solid lines in panel (a) and bars in panel (b) represent posterior means. Error bars in panel (b) represent 95% confidence intervals (2.5 and 97.5 percentiles of the posterior distributions).

9.6 Intervention coverage

Figure 9.51 shows the trends in levels of ART coverage by province. Levels of ART coverage have been heterogeneous between provinces over time. The public sector ART programme in the Western Cape started slightly earlier than that in other provinces, and as a result, ART coverage was generally higher than that in other provinces prior to 2012. However, in the last decade, rates of ART initiation have picked up substantially in KwaZulu-Natal, with the result that ART coverage in 2021 is highest in this province. ART coverage in 2021 varied between 54% in Western Cape and 75% in KwaZulu-Natal. In several provinces ART coverage dropped in 2021 due to the impact of COVID-19.

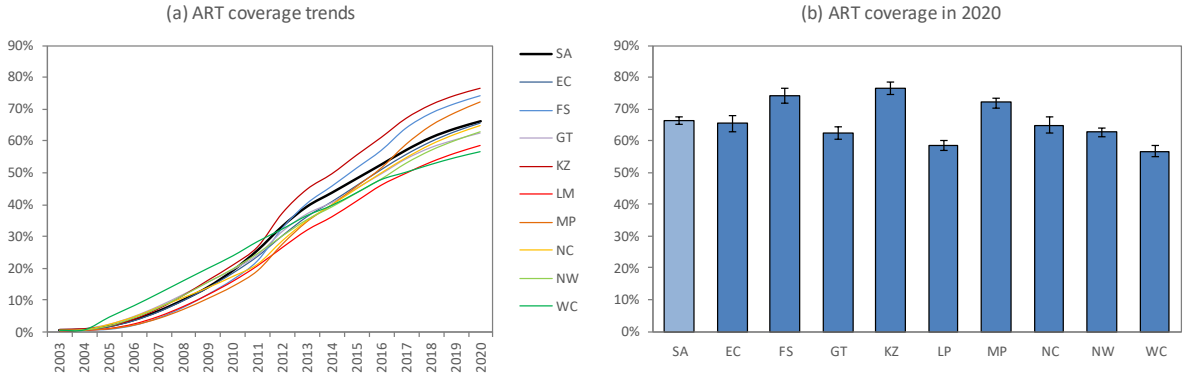

Figure 9.51: Fraction of HIV-positive individuals on ART
Lines represent posterior averages. ART coverage is defined on the assumption that all HIV-positive individuals are eligible for ART.

Figure 9.52 shows progress towards the 95-95-95 targets in 2021. Levels of HIV diagnosis are similarly high across provinces, at around 93%, with only KwaZulu-Natal having reached the 95% target. There is greater variation in the fraction of diagnosed adults who are on ART, ranging from 59% in Western Cape to 79% in KwaZulu-Natal; none of the provinces is close to the 95% target. Rates of viral suppression (<1000 RNA copies/ml) also vary between provinces, from levels of 87% in Limpopo and Eastern Cape to 94% in Western Cape. The target for all three indicators (for 2025) is 95% [5], and there is thus substantial scope for improvement, particularly on the second indicator. Combining all three targets implies a target of 86% of all HIV-positive individuals on ART and virally suppressed by 2020. Figure 9.52(d) shows that all provinces fall short of this combined target, with the net viral suppression rate varying from 50% in Western Cape to 70% in KwaZulu-Natal.

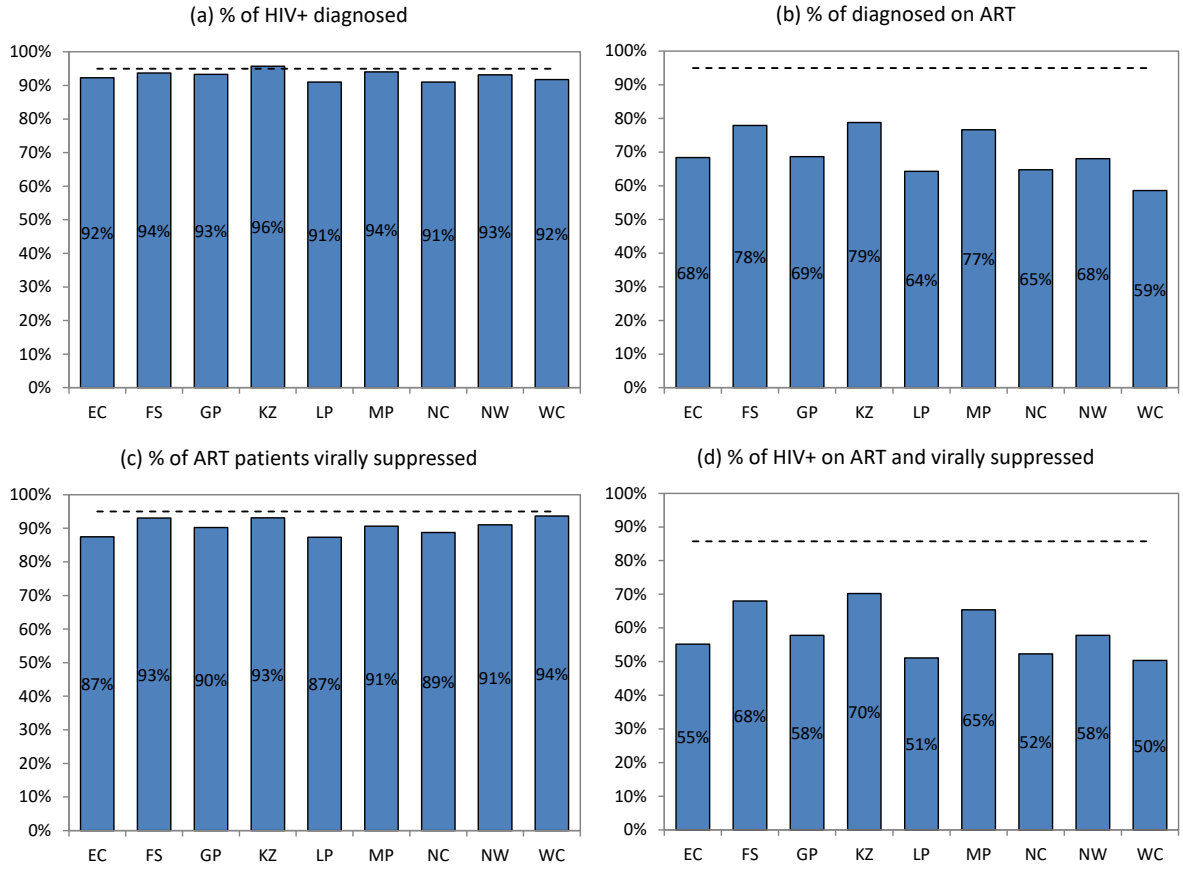

Figure 9.52: Progress towards the UNAIDS 95-95-95 targets in 2021
Bars represent posterior averages. Dashed lines represent UNAIDS targets for 2025.

Figure 9.53 shows the change over time in the proportion of men who are circumcised. In general there has been little change in the fraction circumcised over the 2003-2011 period, with increases only becoming noticeable in 2012-2021. These increases have been most substantial in KwaZulu-Natal, Free State and Mpumalanga. Increases have been small in Western Cape, and modest in the Eastern Cape and Northern Cape. Overall levels of circumcision remain highest in Limpopo.

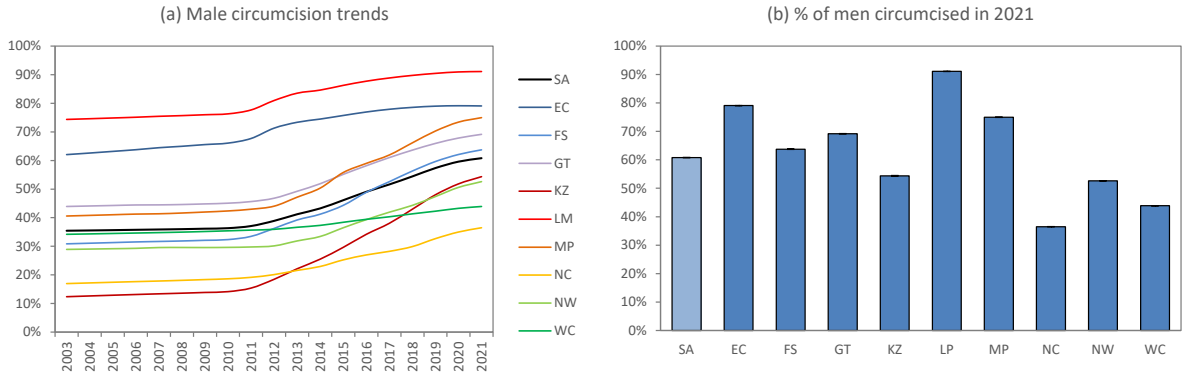

Figure 9.53: Proportion of men aged 15-49 who are circumcised
Lines represent posterior averages.

9.7 Comparison with previous Thembisa estimates

Figure 9.54 compares key indicators from the previous Thembisa provincial modelling report (version 4.4) to the results from the updated analysis (version 4.5), as presented in this report. For the sake of comparison, we have focused on estimates in 2020 (or the 2019-20 projection year), to limit the potential effect of new programme data that were not available at the time of the previous round of estimates. Estimates of adult HIV incidence in adults (15-49) are mostly not significantly different from those estimated previously, although in most provinces (with the exception of Gauteng and Western Cape), version 4.5 gives slightly higher estimates of HIV incidence than version 4.4. Consistent with this, the new estimates of HIV prevalence in the 15-49 age group are mostly higher than in version 4.4 (panel b). It is worth noting that version 4.4 prevalence estimates were mostly lower than in version 4.3, and thus the new model estimates (version 4.5) are more in line with version 4.3. Thembisa 4.5 estimates of HIV prevalence in children are mostly similar to those in Thembisa 4.4 except in Free State (where it is again worth noting that the new estimates are more in line with version 4.3) and Limpopo (where the new model estimates are significantly lower than those in version 4.4).

The new model estimates levels of ART coverage in 2020 tend to be close to those estimated by the previous model in most provinces (panel d). There is also similarity between versions 4.4 and 4.5 when ART coverage estimates are disaggregated by sex (panels e and f). In most provinces, estimates of the ART coverage in children are slightly higher than those estimated previously (panel g), except in the case of Free State (due to the increase in paediatric prevalence shown in panel c). Finally, estimates of viral suppression are consistently lower than those estimated previously (panel h), which is probably because we have assumed a slower rollout of dolutegravir, and also because the changes to the modelling of ART initiation after re-diagnosis mean that relatively more ART patients start ART at low CD4 counts (when there is a lower chance of virological suppression being achieved).

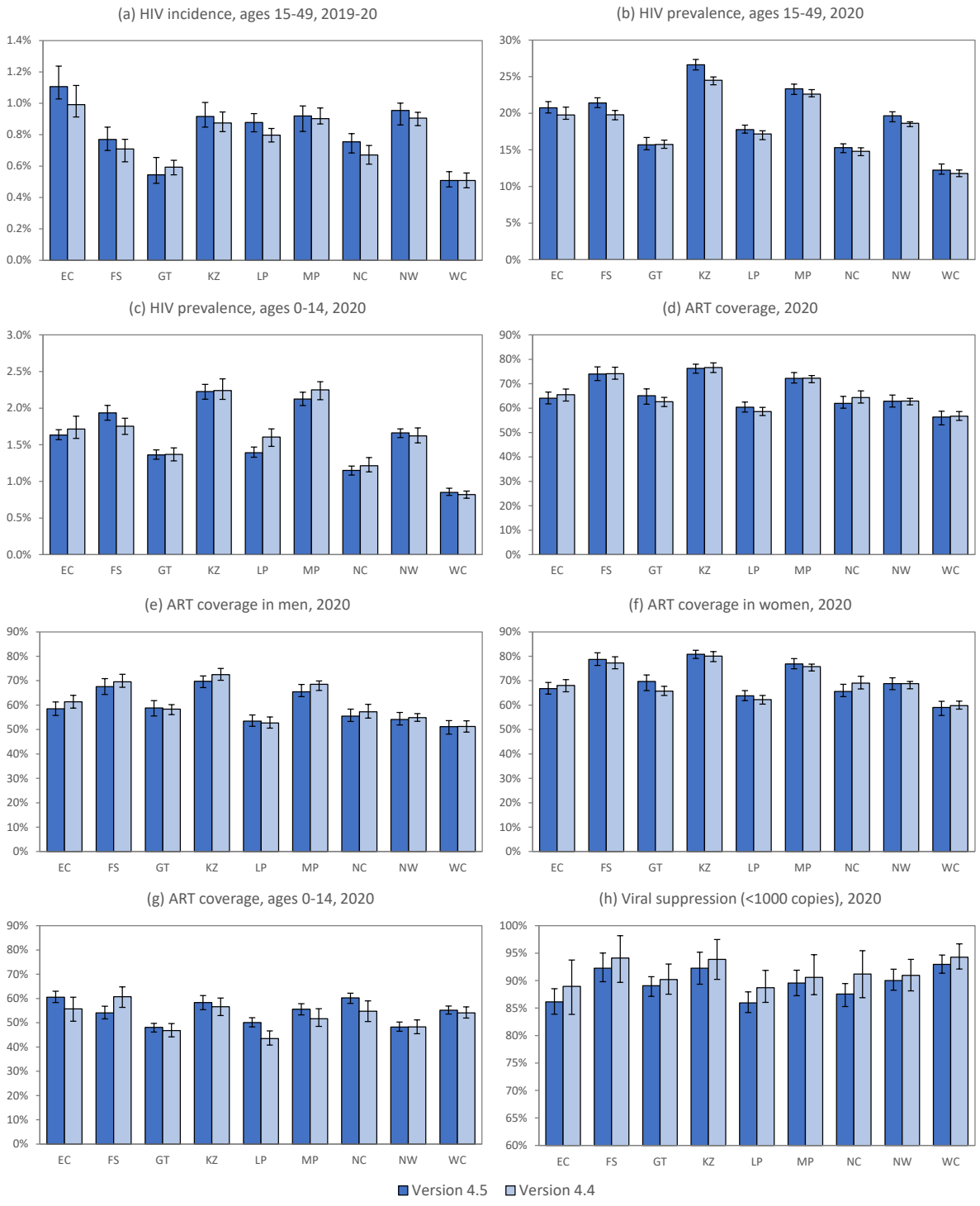

Figure 9.54: Comparison of key epidemic indicators in most recent (version 4.5) and previous (version 4.4) Thembisa models

10. Discussion and conclusions

10.1 Key findings

As in previous versions of Thembisa, we find that most provinces have made good progress towards the first 95% target and the third 95% target, but there is a persistent problem with reaching the second 95% target. No province has come close to the target of 95% of all HIV-diagnosed individuals on ART by 2025, although KwaZulu-Natal and Free State have made reasonable progress, with 79% and 78% of HIV-diagnosed patients on ART, respectively. Western Cape and Limpopo appear to be the poorest performing provinces, in terms of the 95-95-95 targets. It should be noted that since the HIV epidemic in Western Cape and Limpopo emerged later than in the other provinces, a lower ART coverage is to be expected (because there are relatively more HIV-infected individuals in the early stages of HIV disease). In many provinces ART coverage is estimated to have declined over the 2020-2021 period, as a result of COVID-19.

There have been substantial increases in the prevalence of male circumcision in recent years, especially in KwaZulu-Natal, Mpumalanga and Free State. However, there has been virtually no change in male circumcision prevalence in the Western Cape, and growth in male circumcision coverage in Northern Cape and Eastern Cape has been modest. The Eastern Cape, Western Cape and Northern Cape appear to have the lowest levels of condom use in the country, while rates of condom use have generally been highest in Gauteng. The model estimates that over the 2000-2020 period, HIV incidence declines have been smallest in the Western Cape, Limpopo and Northern Cape (between 33% and 47%), which may be partly explained by the relatively low rates of change in condom use and/or MMC uptake in these provinces. In contrast, the HIV incidence decline over the 2000-2019 period has been most substantial in KwaZulu-Natal and Free State (74-77%), which is probably a reflection of the high ART coverage in these province and high uptake of MMC. Importantly, KwaZulu-Natal has transitioned from being the province with the highest HIV incidence rate in the period before 2017, to having an incidence rate close to the national average in 2020-21. Eastern Cape, in contrast, has transitioned from having one of the lowest HIV incidence rates in the country in the 1990s, to have the highest HIV incidence rate in 2020-21. Estimates of HIV incidence need to be treated with caution, however, as we lack recent HIV prevalence data to validate the model.

10.2 Strengths and limitations

The Thembisa version 4.5 provincial model improves on the previous (version 4.4) model in several respects. Firstly, the new model incorporates data on the age distributions of ART patients. In the case of children, these data on age distributions were previously incorporated in the national model calibration [10] but not in the provincial model calibration. The age distributions are useful in informing the likely extent of postnatal transmission (because postnatally infected children are more likely to survive to older ages) and thus the likely durations of breastfeeding. Previously, in version 4.4, posterior estimates of breastfeeding durations varied between 0.80 and 1.32 times the national average, whereas the new posterior

estimates are between 0.90 and 1.08 times the national average. The inclusion of the age distribution data thus suggests greater similarity in breastfeeding durations across provinces than was previously estimated.

In the case of adults, the inclusion of the age distribution data has had mixed results. In some provinces (e.g. Western Cape, Gauteng, KwaZulu-Natal), the modelled age distribution is reasonably consistent with the ART programme data, but in other provinces (most notably Limpopo and Eastern Cape) there are marked discrepancies between the modelled and observed age distributions. We have given relatively little weight to the age distribution data in the calibration (relative to other data sources) and it is therefore not surprising that the model fit to the data is poor in some provinces. Nevertheless, it is concerning that in both Eastern Cape and Limpopo there is a substantially greater difference between males and females in the observed age distributions than in the modelled age distributions. This suggests that the model may be failing to capture important gender dynamics in these predominantly rural provinces. Our model currently assumes that male-female differences in HIV testing rates are the same across provinces, but if in fact there are more extreme differences in rural provinces than in urban provinces [132], this may go some way to explaining the poor model fits in the rural provinces. These and other possible explanations will be explored in future updates to the Thembisa model.

A major structural change in Thembisa 4.5 is to the modelling of marriage/cohabitation and dissolution of marital/cohabiting relationships. Previously, limited effort was made to account for inter-provincial differences in marriage and union dissolution. In addition, Thembisa 4.4 did not allow for changes over time in rates of marriage and rates of union dissolution. The new model allows for changes in these rates over time, and ten different parameters (controlling incidence of marriage and union dissolution, age and sex differences in these rates and changes over time) are estimated for each province, based on calibration to age- and sex-specific marriage prevalence data from four national censuses/community surveys. In all provinces the incidence of marriage is estimated to have declined steeply over time, especially at younger ages. This might partly explain why HIV incidence estimates in recent years are mostly higher than in Thembisa 4.4 (since unmarried individuals generally have higher HIV incidence than married individuals [133-136]).

Another change to the provincial model has been to allow for uncertainty in the HIV natural history parameters in children. Previously we allowed for inter-provincial differences in the adult natural history parameters but not the paediatric natural history parameters. By and large, the inter-provincial differences are consistent for adults and children: the provinces in the southern and western parts of the country appear to have more rapid progression to AIDS and death (in the absence of ART) than the other provinces, both for adults and children. Although we have noted the historically high TB incidence rates in these provinces [43] as a possible explanation, other environmental factors could be significant. For example, levels of vitamin D deficiency tend to be greater in the most southern provinces, due to lower levels of exposure to UV radiation at lower altitudes and higher latitudes [137], and vitamin D deficiency is strongly associated with HIV disease progression [138, 139].

We have also revised our assumptions about male circumcision. In some provinces (Mpumalanga and Limpopo), the previously reported VMMC numbers from the DHIS included significant numbers of medical circumcisions in traditional settings [114], which are accounted for in our model in the ‘background’ male circumcision assumptions. This means

that previous versions of Thembisa were 'double counting' these medical circumcisions in traditional settings. We have corrected the VMMC numbers in these provinces based on recent estimates of the numbers of medical circumcisions in traditional settings [114]. We have also revised our assumptions about the 'background' levels of male circumcision prevalence, at a provincial level, based on this recent analysis of South African data. This has led to notably higher levels of male circumcision prevalence in Eastern Cape (the prevalence in men aged 15-49 in 2005 has increased from 48% in version 4.4 to 63% in version 4.5) and Limpopo (prevalence in 2005 increasing from 64% to 75%). In other provinces the effects of the changes in male circumcision assumptions have been relatively small.

Thembisa version 4.5 includes unpublished data from the 2019 antenatal prevalence survey. In most provinces the model matches the most recent prevalence survey data closely. However, in Eastern Cape and KwaZulu-Natal the model estimates of HIV prevalence in pregnant women in 2019 are substantially lower than those measured in the survey. In Limpopo, the model estimates are higher than measured in the 2019 survey. It is difficult to establish whether these differences are due to the model failing to match recent HIV incidence trends or if they are due to random survey measurement outliers. The Eastern Cape is most concerning, given that the model estimates are also inconsistent with the earlier 2017 survey. The province has experienced severe drought since 2015 [140], and it is possible that this may have led to increases in sexual risk behaviour, as observed in neighbouring Lesotho [141]. KwaZulu-Natal was also severely affected by drought in 2015 and (to a lesser extent) 2017 [142].

The new model also includes more recent HIV programme data. At the time of the Thembisa 4.4 release there was much uncertainty regarding the impact of COVID-19 on HIV services. We can now estimate these impacts with slightly more confidence, and it is clear that although there has been an impact in all provinces, this impact has been heterogeneous. Limpopo and Northern Cape had particularly substantial drops in HIV testing volumes when comparing 2020-21 to 2019-20 (51% and 44% reductions respectively). Northern Cape has also experienced particularly significant reductions in rates of linkage to ART after HIV diagnosis. Numbers of VMMC operations in 2020-21 have reduced by over two thirds relative to 2019-20, in all provinces except Gauteng (in which the reduction was only 14%). A limitation of the model is that it does not consider potential impacts of COVID-19 on HIV testing during pregnancy and rates of linkage to ART after paediatric HIV diagnosis.

Our new model includes updated demographic parameters. The previous version of Thembisa (version 4.4) relied on demographic parameters produced in 2018, and it therefore did not reflect the impact of COVID-19 on non-HIV mortality. The revisions to the demographic estimates have led to higher estimates of non-HIV mortality in the period from 2020 and slightly reduced population growth forecasts. However, the HIV estimates are not generally much affected by the revisions to the demographic assumptions.

A major limitation is that our estimates of ART numbers are particularly uncertain. In this and other recent versions of Thembisa we have relied mainly on data from the DHIS for estimates of ART uptake in the public sector, and on data from the Council for Medical Schemes for estimates of ART uptake in the private sector. However, recent comparisons of DHIS and Tier data suggest that the DHIS may have slightly under-estimated the true numbers of ART patients in the public sector, particularly over 2015-2019 (unpublished data). This could be due to clinics not reporting numbers to DHIS and being accidentally excluded when reporting provincial totals (in contrast to Tier, which is an individual-based record system). In addition,

it has become clear that the medical scheme data could substantially understate the private sector total because there are many 'cash-paying' ART patients in the private sector who are not medical scheme members. Unpublished data suggest that these cash-paying patients are particularly substantial in the Gauteng and Western Cape provinces. Further work will be required to update and improve our provincial ART estimates in future.

Another limitation is that we have not to date allowed for differences across provinces in mortality rates in ART patients (controlling for age, sex, baseline CD4 count and ART duration). Our previous analyses of IeDEA data suggest that even after controlling for these predictors of mortality there can be significant heterogeneity in mortality across cohorts [143, 144]. It is therefore likely that such inter-provincial differences exist, and we hope to explore possible inter-provincial differences in ART mortality rates in future Thembisa updates. It will be particularly important to consider the role of inter-provincial differences in rates of viral suppression, as viral suppression is a major predictor of mortality after the first 6 months on ART [145]. In the interim, model estimates of the extent of the inter-provincial differences in the mean duration of untreated HIV and the ART interruption rate need to be treated with some caution, as it is possible that inter-provincial differences in ART patient mortality are being incorrectly attributed to these other parameters when we calibrate to vital registration data.

There remains significant uncertainty regarding the extent of inter-provincial differences in transmission between sex workers, clients, and men who have sex with men. Previous modelling studies suggest that epidemics that are driven largely by key populations might be easier to bring under control than epidemics that are more generalized [146], and consideration of the relative differences in key population sizes across provinces are therefore important to consider. Unfortunately there are few reliable estimates of key population size at a provincial level, with most previous studies only producing estimates for major metropolitan areas [147]. However, recent data from a national survey provide estimates of HIV prevalence in sex workers in all nine provinces (Jenny Coetzee, personal communication). We aim to calibrate future provincial versions of Thembisa to province-specific data on HIV prevalence in key populations, in order to improve confidence in key population estimates at a provincial level. Our model also assumes that the fraction of men who are MSM is the same across provinces, although it might be expected that the MSM proportion is higher in the more urbanized provinces [148, 149].

We also plan to include province-specific data on the proportion of HIV testers who receive positive results, in future updates to Thembisa. This will allow us to more reliably assess whether there are differences across provinces in progress towards the 95% diagnosis target. The additional HIV prevalence data may also provide indirect evidence on HIV incidence trends, which is important, given that other HIV prevalence data sources only provide prevalence estimates up to 2019. HIV antibody testing data in children might similarly help to improve confidence in the estimates of HIV prevalence in children, at a provincial level.

A final limitation to note is that we have calibrated the model using programme data up to March 2021. The impact of COVID-19 in the period from April 2020 to March 2021 is particularly noticeable, and we have made a number of assumptions about the extent to which HIV services recover after March 2021, which might not be realistic. Early indications are that the recovery in ART enrolment after March 2021 might have been more significant than we have assumed here (Jeffrey Eaton, personal communication), and ART forecasts beyond 2021 should therefore be treated with some caution.

10.3 Conclusion

This update to the previous Thembisa provincial models makes use of more recent data and improved statistical methods to derive updated estimates of HIV incidence and progress in scaling up access to HIV prevention and treatment services in South Africa. Although it is encouraging to see high levels of HIV diagnosis and viral suppression in ART patients, it remains concerning that South Africa's progress towards the UNAIDS ART coverage target is poor. The decline in HIV incidence over the 2010-2020 period varies between 33% in Western Cape and 66% in KwaZulu-Natal, which falls short of the 75% UNAIDS target for the 2010-2020 period. This is nevertheless in line with the average incidence decline over the 2010-2020 period in the eastern and southern African region (43%) [150]. Renewed efforts are needed in order to reach the UNAIDS target of 95% ART coverage in HIV-diagnosed individuals by 2025, and continued innovation in the field of HIV prevention will be critical to ensuring that the National Strategic Plan target of a 50% reduction in HIV incidence over the 2017-2022 period is met [151]. More work is required to identify the success factors that have enabled provinces like KwaZulu-Natal to make good progress towards the 95-95-95 targets, and to follow these examples in other provinces.

Acknowledgements

This research was funded by the Bill and Melinda Gates Foundation, and earlier funding for the work on the provincial Thembisa model was provided by UNAIDS and the South African National AIDS Council.

We are grateful to the following individuals who have provided input and assisted us in obtaining data for the current version of Thembisa: Thapelo Seatlhodi, Riona Govender, Sifiso Phakathi, Mohammed Majam, Jenny Coetzee, Natsai Shoko, Billia Luwaca, Jeffrey Eaton, Gesine Meyer-Rath, Ali Feizzadeh and Sarah Porter.

We are also grateful to the South African Department of Health for sharing data from the District Health Information System and the TB/HIV information system, ICF for sharing the data from the 2016 Demographic and Health Survey, PEPFAR and the Clinton Health Access Initiative for sharing data on PrEP numbers, and the Perinatal Health Research Unit/South African Medical Research Council for sharing the unpublished data from the 2019 national sex worker survey. All other data sources used in calibrating the Thembisa model are from published reports or papers, which have been cited where relevant.

References

1. Human Sciences Research Council. HIV Impact Assessment Summary. 2018. Available: http://serve.mg.co.za/content/documents/2018/07/17/7M1RBtUShKFJbN3NL1Wr_HSRC_HIV_Survey_Summary_2018.pdf. Accessed 18 July 2018
2. Johnson LF, May MT, Dorrington RE, Cornell M, Boulle A, Egger M, *et al.* Estimating the impact of antiretroviral treatment on adult mortality trends in South Africa: a mathematical modelling study. *PLoS Med* 2017; **14**:e1002468.
3. Shisana O, Rehle T, Simbayi LC, Zuma K, Jooste S, Zungu N, *et al.* South African National HIV Prevalence, Incidence, and Behaviour Survey, 2012. Cape Town: Human Sciences Research Council; 2014. Available: http://www.hsrc.ac.za/en/research-outputs/view/6871. Accessed 16 April 2014
4. Anderson SJ, Cherutich P, Kilonzo N, Cremin I, Fecht D, Kimanga D, *et al.* Maximising the effect of combination HIV prevention through prioritisation of the people and places in greatest need: a modelling study. *Lancet* 2014; **384**:249-256.
5. UNAIDS. World AIDS Day Report 2020: Prevailing against pandemics by putting people at the centre. 2020. Available: https://aidstargets2025.unaids.org/. Accessed 2 Dec 2020
6. Johnson LF, Chiu C, Myer L, Davies MA, Dorrington RE, Bekker LG, *et al.* Prospects for HIV control in South Africa: a model-based analysis. *Glob Health Action* 2016; **9**:30314.
7. Jamieson L, Gomez GB, Rebe K, Brown B, Subedar H, Jenkins S, *et al.* The impact of self-selection based on HIV risk on the cost-effectiveness of pre-exposure prophylaxis in South Africa. *AIDS* 2020; **34**:883-891.
8. Joseph Davey DL, Bekker LG, Gomba Y, Coates T, Myer L, Johnson LF. Modelling the potential impact of providing preexposure prophylaxis in pregnant and breastfeeding women in South Africa. *AIDS* 2019; **33**:1391-1395.
9. Johnson LF, Rehle TM, Jooste S, Bekker LG. Rates of HIV testing and diagnosis in South Africa, 2002-2012: successes and challenges. *AIDS* 2015; **29**:1401-1409.
10. Johnson LF, Patrick M, Stephen C, Patten G, Dorrington RE, Maskew M, *et al.* Steep declines in pediatric AIDS mortality in South Africa, despite poor progress towards pediatric diagnosis and treatment targets. *Pediatr Infect Dis J* 2020; **39**:843-848.
11. Johnson LF, Dorrington RE, Moolla H. Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa. *South Afr J HIV Med* 2017; **18**:a694.
12. Johnson LF, Dorrington RE, Moolla H. HIV epidemic drivers in South Africa: a model-based evaluation of factors accounting for inter-provincial differences in HIV prevalence and incidence trends. *South Afr J HIV Med* 2017; **18**:a695.
13. Johnson LF, Dorrington RE, Moolla H. Modelling the impact of HIV in South Africa's provinces. University of Cape Town; 2016. Available: http://www.thembisa.org/content/downloadPage/Provinces2016. Accessed 24 Aug 2016
14. Johnson LF, Dorrington RE. Thembisa version 3.2: A model for evaluating the impact of HIV/AIDS in South Africa. 2017. Available: https://www.thembisa.org/content/downloadPage/Thembisa3_2report. Accessed 29 Sept 2017

15. Johnson LF, Dorrington RE. Modelling the impact of HIV in South Africa's provinces: 2018 update. Centre for Infectious Disease Epidemiology and Research working paper; 2018. Available: https://www.thembisa.org/downloads
16. Johnson LF, Dorrington RE. Modelling the impact of HIV in South Africa's provinces: 2019 update. University of Cape Town; 2019. Available: https://www.thembisa.org/
17. Johnson LF, Dorrington RE. Modelling the impact of HIV in South Africa's provinces: 2020 update. Centre for Infectious Disease Epidemiology and Research, University of Cape Town; 2020. Available: https://www.thembisa.org/
18. Johnson LF, Dorrington RE. Modelling the impact of HIV in South Africa's provinces: 2021 update. Centre for Infectious Disease Epidemiology and Research, University of Cape Town; 2021. Available: https://www.thembisa.org/
19. Johnson LF, Dorrington RE. Thembisa version 4.5: A model for evaluating the impact of HIV/AIDS in South Africa. 2022. Available: https://www.thembisa.org/
20. Morris M, Leslie-Cook A. Basic sexual concurrency analyses. In: *The HIV epidemic in South Africa: what do we know and how has it changed?* Edited by Fraser-Hurt N, Zuma K, Njuho P, *et al.*: World Bank; 2011.
21. Johnson S, Kincaid DL, Figueroa ME, Delate R, Mahlasela L, Magni S. The Third National HIV Communication Survey, 2012. Pretoria: Johns Hopkins Health and Education in South Africa; 2013. Available: http://jhhesa.org/sites/default/files/hiv_survey.pdf. Accessed 19 April 2014
22. Department of Health, Statistics South Africa, South African Medical Research Council, ICF. South Africa Demographic and Health Survey 2016. Pretoria; 2019. Available: https://www.dhsprogram.com/pubs/pdf/FR337/FR337.pdf. Accessed 19 March 2019
23. Johnson LF, Meyer-Rath G, Dorrington RE, Puren A, Seatlhodi T, Zuma K, *et al.* The effect of HIV programmes in South Africa on national HIV incidence trends, 2000-2019. *J Acquir Immun Defic Syndr* 2022; [In press].
24. Mugashu R, Mhlungu W. Inter-provincial differences in marriage and divorce trends in South Africa: Insights from mathematical modelling: University of Cape Town; 2021.
25. Garnett G. An introduction to mathematical models in sexually transmitted disease epidemiology. *Sex Transm Infect* 2002; **78**:7-12.
26. Garnett GP, Hughes JP, Anderson RM, Stoner BP, Aral SO, Whittington WL, *et al.* Sexual mixing patterns of patients attending sexually transmitted diseases clinics. *Sex Transm Dis* 1996; **23**:248-257.
27. Manhart LE, Aral SO, Holmes KK, Foxman B. Sex partner concurrency: measurement, prevalence, and correlates among urban 18-39-year-olds. *Sex Transm Dis* 2002; **29**:133-143.
28. Granath F, Giesecke J, Scalia-Tomba G, Ramstedt K, Forssman L. Estimation of a preference matrix for women's choice of male sexual partner according to rate of partner change, using partner notification data. *Math Biosci* 1991; **107**:341-348.
29. Laumann EO, Gagnon JH, Michael RT, Michaels S. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press; 1994.
30. Ghani AC, Donnelly CA, Garnett GP. Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases. *Stat Med* 1998; **17**:2079-2097.

31. Department of Health. South Africa Demographic and Health Survey 1998: Full Report. 1999.
32. Department of Health. South Africa Demographic and Health Survey 2003: Preliminary Report. Pretoria; 2004. Available: http://www.doh.gov.za/docs/reports/2003/sadhs2003/part2.pdf. Accessed 6 Jan 2012
33. Johnson LF, Dorrington RE, Bradshaw D, Pillay-Van Wyk V, Rehle TM. Sexual behaviour patterns in South Africa and their association with the spread of HIV: insights from a mathematical model. *Demographic Res* 2009; **21**:289-340.
34. Vandepitte J, Lyerla R, Dallabetta G, Crabbé F, Alary M, Buvé A. Estimates of the number of female sex workers in different regions of the world. *Sex Transm Infect* 2006; **82** (Suppl 3):S18-25.
35. Bien CH, Cai Y, Emch ME, Parish W, Tucker JD. High adult sex ratios and risky sexual behaviors: a systematic review. *PLoS One* 2013; **8**:e71580.
36. Morison L, Weiss HA, Buvé A, Caraël M, Abega S-C, Kaona F, *et al.* Commercial sex and the spread of HIV in four cities in sub-Saharan Africa. *AIDS* 2001; **15**:S61-S69.
37. Johnson LF, Dorrington RE. Thembisa version 4.2: A model for evaluating the impact of HIV/AIDS in South Africa. University of Cape Town; 2019. Available: https://www.thembisa.org/
38. Leclerc PM, Garenne M. Clients of commercial sex workers in Zambia: prevalence, frequency and risk factors. *Open Demography J* 2008; **1**:1-10.
39. South SJ, Trent K. Imbalanced sex ratios, men's sexual behavior, and risk of sexually transmitted infection in China. *Journal of Health and Social Behavior* 2010; **51**:376-390.
40. South SJ, Trent K, Bose S. India's 'missing women' and men's sexual risk behavior. *Population Research and Policy Review* 2012; **31**:777-795.
41. Williams BG, Korenromp EL, Gouws E, Schmid GP, Auvert B, Dye C. HIV infection, antiretroviral therapy, and CD4+ cell count distributions in African populations. *J Infect Dis* 2006; **194**:1450-1458.
42. Malaza A, Mossong J, Bärnighausen T, Viljoen J, Newell ML. Population-based CD4 counts in a rural Area in South Africa with high HIV prevalence and high antiretroviral treatment coverage. *PLoS One* 2013; **8**:e70126.
43. Nanoo A, Izu A, Ismail NA, Ihekweazu C, Abubakar I, Mametja D, *et al.* Nationwide and regional incidence of microbiologically confirmed pulmonary tuberculosis in South Africa, 2004-12: a time series analysis. *Lancet Infect Dis* 2015; **15**:1066-1076.
44. Probst C, Parry CD, Rehm J. Socio-economic differences in HIV/AIDS mortality in South Africa. *Trop Med Int Health* 2016; **21**:846-855.
45. Benade M, Long L, Rosen S, Meyer-Rath G, Tucker JM, Miot J. Reduction in initiations of HIV treatment in South Africa during the COVID pandemic. *MedRxiv* 2021.
46. Dorward J, Mabuto T, Charalambous S, Fielding KL, Hoffmann CJ. Factors associated with poor linkage to HIV care in South Africa: secondary analysis of data from the Thol'impilo trial. *Journal of Acquired Immune Deficiency Syndrome* 2017; **76**:453-460.
47. Boyer S, Iwuji C, Gosset A, Protopopescu C, Okesola N, Plazy M, *et al.* Factors associated with antiretroviral treatment initiation amongst HIV-positive individuals linked to care within a universal test and treat programme: early findings of the ANRS 12249 TasP trial in rural South Africa. *AIDS Care* 2016; **28** (Suppl 3):39-51.

48. Larson BA, Brennan A, McNamara L, Long L, Rosen S, Sanne I, *et al.* Early loss to follow up after enrolment in pre-ART care at a large public clinic in Johannesburg, South Africa. *Trop Med Int Health* 2010; **15** (Suppl 1):43-47.
49. Lessells RJ, Mutevedzi PC, Cooke GS, Newell ML. Retention in HIV care for individuals not yet eligible for antiretroviral therapy: rural KwaZulu-Natal, South Africa. *J Acquir Immun Defic Syndr* 2011; **56**:e79-86.
50. Kranzer K, Zeinecker J, Ginsberg P, Orrell C, Kalawe NN, Lawn SD, *et al.* Linkage to HIV care and antiretroviral therapy in Cape Town, South Africa. *PLoS One* 2010; **5**:e13801.
51. Lurie MN, Kirwa K, Callaway J, Cornell M, Boulle A, Bengtson AM, *et al.* Quantifying the HIV treatment cascade in a South African health sub-district by gender: retrospective cohort study. *Trop Med Int Health* 2020; **25**:186-192.
52. Maughan-Brown B, Beckett S, Kharsany ABM, Cawood C, Khanyile D, Lewis L, *et al.* Poor rates of linkage to HIV care and uptake of treatment after home-based HIV testing among newly diagnosed 15-to-49 year-old men and women in a high HIV prevalence setting in South Africa. *AIDS Care* 2020; [In press].
53. Osler M, Cornell M, Ford N, Hilderbrand K, Goemaere E, Boulle A. Population-wide differentials in HIV service access and outcomes in the Western Cape for men as compared to women, South Africa: 2008 to 2018: a cohort analysis. *J Int AIDS Soc* 2020; **23** (Suppl 2):e25530.
54. Mujugira A, Celum C, Thomas KK, Farquhar C, Mugo N, Katabira E, *et al.* Delay of antiretroviral therapy initiation is common in East African HIV-infected individuals in serodiscordant partnerships. *J Acquir Immun Defic Syndr* 2014; **66**:436-442.
55. Bor J, Chiu C, Ahmed S, Katz I, Fox MP, Rosen S, *et al.* Failure to initiate HIV treatment in patients with high CD4 counts: evidence from demographic surveillance in rural South Africa. *Trop Med Int Health* 2018; **23**:206-220.
56. Pillay T, Cornell M, Fox MP, Euvrard J, Fatti G, Technau K, *et al.* Recording of HIV viral loads and viral suppression in South African patients receiving antiretroviral treatment: a multicentre cohort study. *Antivir Ther* 2020; **25**:257-266.
57. Jiamsakul A, Kariminia A, Althoff KN, Cesar C, Cortes CP, Davies MA, *et al.* HIV viral load suppression in adults and children receiving antiretroviral therapy - results from the IeDEA collaboration. *Journal of Acquired Immune Deficiency Syndrome* 2017; **76**:319-329.
58. George S, McGrath N, Oni T. The association between a detectable HIV viral load and non-communicable diseases comorbidity in HIV positive adults on antiretroviral therapy in Western Cape, South Africa. *BMC Infect Dis* 2019; **19**:348.
59. Fatti G, Grimwood A, Nachega JB, Nelson JA, LaSorda K, Zyl GV, *et al.* Better virological outcomes amongst people living with HIV initiating early antiretroviral treatment (CD4 counts >/= 500 cells/microL) in the HPTN 071 (PopART) trial in South Africa. *Clin Infect Dis* 2020; **70**:395-403.
60. Edet A, Akinsola H, Bessong PO. Virologic and immunologic responses of patients on highly active antiretroviral therapy in a rural community health centre in Limpopo, South Africa: A retrospective study. *South Afr J HIV Med* 2019; **20**:818.
61. Johnson LF, Dorrington RE. Thembisa version 4.3: A model for evaluating the impact of HIV/AIDS in South Africa. 2020. Available: https://www.thembisa.org/.
62. Department of Health. ART Programme Analysis: Reviewing the ART programme from April 2004 to March 2014. Pretoria; 2015.
63. Department of Health. Health Indicators Update: Antiretroviral Indicators. 2013. Available: http://www.health.gov.za/reports.php. Accessed 14 May 2014

64. Smith AFM, Gelfand AE. Bayesian statistics without tears - a sampling resampling perspective. *Am Stat* 1992; **46**:84-88.
65. Pillay T. Determinants of HIV viral suppression in South Africans receiving antiretroviral treatment. Cape Town: University of Cape Town; 2019.
66. Kanters S, Vitoria M, Doherty M, Socias ME, Ford N, Forrest JI, *et al.* Comparative efficacy and safety of first-line antiretroviral therapy for the treatment of HIV infection: a systematic review and network meta-analysis. *Lancet HIV* 2016; **3**:e510-e520.
67. Kaplan SR, Oosthuizen C, Stinson K, Little F, Euvrard J, Schomaker M, *et al.* Contemporary disengagement from antiretroviral therapy in Khayelitsha, South Africa: A cohort study. *PLoS Med* 2017; **14**:e1002407.
68. Kranzer K, Lewis JJ, Ford N, Zeinecker J, Orrell C, Lawn SD, *et al.* Treatment interruption in a primary care antiretroviral therapy program in South Africa: cohort analysis of trends and risk factors. *J Acquir Immun Defic Syndr* 2010; **55**:e17-23.
69. Bassett IV, Govindasamy D, Erlwanger AS, Hyle EP, Kranzer K, van Schaik N, *et al.* Mobile HIV screening in Cape Town, South Africa: clinical impact, cost and cost-effectiveness. *PLoS One* 2014; **9**:e85197.
70. Johnson LF, Dorrington RE. Thembisa version 4.4: A model for evaluating the impact of HIV/AIDS in South Africa. Centre for Infectious Disease Epidemiology and Research, University of Cape Town; 2021. Available: https://www.thembisa.org/
71. Funani I, Sonko R, Marumo E, Odugwu S, Hammelmann C. STIs: routine monitoring and clinical sentinel surveillance of sexually transmitted infections. In: *South African Health Review 2003*. Edited by Ijumba P, Day C, Ntuli A: Health Systems Trust; 2004.
72. Department of Health. The 2012 National Antenatal Sentinel HIV and Herpes Simplex Type-2 Prevalence Survey in South Africa. 2014. Available: http://www.health.gov.za/reports.php. Accessed 14 May 2014
73. Department of Health. AIDS in South Africa: Reported AIDS cases as on 30 November 1995. *Epidemiol comments* 1995; **22**:233-234.
74. Actuarial Society of South Africa. ASSA2008 AIDS and Demographic Model. 2011. Available: http://aids.actuarialsociety.org.za. Accessed 5 April 2011
75. Day C, Barron P, Monticelli F, Sello E. District Health Barometer 2007/08. Health Systems Trust; 2009. Available: http://www.hst.org.za/publications/850. Accessed 10 July 2009
76. Ramkissoon A, Kleinschmidt I, Beksinska M, Smit J, Hlazo J, Mabude Z. National Baseline Assessment of Sexually Transmitted Infection and HIV Services in South African Public Sector Health Facilities. Durban: Reproductive Health Research Unit; 2004. Available: http://www.rhru.co.za. Accessed 13 February 2004
77. Reagon G, Irlam J, Levin J. The National Primary Health Care Facilities Survey 2003. Durban: Health Systems Trust; 2004. Available: http://www.hst.org.za/publications/617. Accessed 6 Aug 2010
78. McCoy D, Besser M, Visser R, Doherty T. Interim findings on the national PMTCT pilot sites: lessons and recommendations. Durban: Health Systems Trust; 2002. Available: http://www.hst.org.za/publications/478. Accessed 9 April 2006
79. Day C, Monticelli F, Barron P, Haynes R, Smith J, Sello E. District Health Barometer: Year 2008/09. Durban: Health Systems Trust; 2010. Available: http://www.hst.org.za/publications/864. Accessed 25 June 2010

80. Draper B, Abdullah F. A review of the prevention of mother-to-child transmission programme of the Western Cape provincial government, 2003 - 2004. *S Afr Med J* 2008; **98**:431-434.
81. Department of Health. Policy and Guidelines for the Implementation of the PMTCT programme. 2008. Available: http://www.doh.gov.za/docs/policy/2008/pmtct.pdf. Accessed 6 Jan 2012
82. Jackson DJ, Chopra M, Doherty TM, Colvin MS, Levin JB, Willumsen JF, *et al.* Operational effectiveness and 36 week HIV-free survival in the South African programme to prevent mother-to-child transmission of HIV-1. *AIDS* 2007; **21**:509-516.
83. Coetzee D, Hilderbrand K, Boulle A, Draper B, Abdullah F, Goemaere E. Effectiveness of the first district-wide programme for the prevention of mother-to-child transmission of HIV in South Africa. *Bull WHO* 2005; **83**:489-494.
84. Hilderbrand K, Goemaere E, Coetzee D. The prevention of mother-to-child HIV transmission programme and infant feeding practices. *S Afr Med J* 2003; **93**:779-781.
85. Peltzer K, Mosala T, Dana P, Fomundam H. Follow-up survey of women who have undergone a prevention of mother-to-child transmission program in a resource-poor setting in South Africa. *J Assoc Nurses AIDS Care* 2008; **19**:450-460.
86. Bera E, Jwacu N, Pauls F, Mancotywa T, Ngcelwane N, Hlati Y. Risk factors for perinatal HIV-1 transmission in pregnant women requiring lifelong antiretroviral therapy: A longitudinal study at a tertiary hospital in South Africa. *S Afr J Obstet Gynaecol* 2010; **16**:6-13.
87. Rispel LC, Peltzer K, Phaswana-Mafuya N, Metcalf CA, Treger L. Assessing missed opportunities for the prevention of mother-to-child HIV transmission in an Eastern Cape local service area. *S Afr Med J* 2009; **99**:174-179.
88. Horwood C, Vermaak K, Butler L, Haskins L, Phakathi S, Rollins N. Elimination of paediatric HIV in KwaZulu-Natal, South Africa: large-scale assessment of interventions for the prevention of mother-to-child transmission. *Bull WHO* 2012; **90**:168-175.
89. Orie EF, Songca PP, Moodley J. An audit of PMTCT services at a regional hospital in South Africa. *S Afr Fam Pract* 2009; **51**:492-495.
90. Goga AE, Dinh TH, Jackson DJ. Evaluation of the Effectiveness of the National Prevention of Mother-to-Child Transmission (PMTCT) Programme Measured at Six Weeks Postpartum in South Africa, 2010. South African Medical Research Council, National Department of Health of South Africa and PEPFAR/US Centers for Disease Control and Prevention; 2012. Available: http://www.doh.gov.za/docs/reports/2012/pmtcteffectiveness.pdf. Accessed 12 June 2012
91. World Health Organization. Antiretroviral therapy for HIV infection in infants and children: towards universal access. Recommendations for a public health approach, 2006. Geneva; 2007. Available: http://www.who.int/hiv/pub/guidelines/art/en/ Accessed 26 Jan 2009
92. Diaz C, Hanson C, Cooper ER, Read JS, Watson J, Mendez HA, *et al.* Disease progression in a cohort of infants with vertically acquired HIV infection observed from birth: the Women and Infants Transmission Study (WITS). *J Acquir Immun Defic Syndr* 1998; **18**:221-228.
93. Blanche S, Newell ML, Mayaux MJ, Dunn DT, Teglas JP, Rouzioux C, *et al.* Morbidity and mortality in European children vertically infected by HIV-1. *J Acquir Immun Defic Syndr* 1997; **14**:442-450.

94. HIV Paediatric Prognostic Markers Collaborative Study Group. Short-term risk of disease progression in HIV-1-infected children receiving no antiretroviral therapy or zidovudine monotherapy: a meta-analysis. *Lancet* 2003; **362**:1605-1611.
95. European Collaborative Study. Fluctuations in symptoms in human immunodeficiency virus-infected children: the first 10 years of life. *Pediatrics* 2001; **108**:116-122.
96. Marston M, Zaba B, Salomon JA, Brahmbhatt H, Bagenda D. Estimating the net effect of HIV on child mortality in African populations affected by generalized HIV epidemics. *J Acquir Immun Defic Syndr* 2005; **38**:219-227.
97. Charlebois ED, Ruel TD, Gasasira AF, Achan J, Kateera F, Akello C, *et al.* Short-term risk of HIV disease progression and death in Ugandan children not eligible for antiretroviral therapy. *J Acquir Immun Defic Syndr* 2010; **55**:330-335.
98. Cross Continents Collaboration for Kids. Markers for predicting mortality in untreated HIV-infected children in resource-limited settings: a meta-analysis. *AIDS* 2008; **22**:97-105.
99. van Kooten Niekerk NK, Knies MM, Howard J, Rabie H, Zeier M, van Rensburg A, *et al.* The first 5 years of the family clinic for HIV at Tygerberg Hospital: family demographics, survival of children and early impact of antiretroviral therapy. *J Trop Pediatr* 2006; **52**:3-11.
100. Bobat R, Coovadia H, Moodley D, Coutsoudis A, Gouws E. Growth in early childhood in a cohort of children born to HIV-1-infected women from Durban, South Africa. *Ann Trop Paediatr* 2001; **21**:203-210.
101. Department of Health. Annual Health Statistics 2012. 2013. Available: http://www.hst.org.za/sites/default/files/AnnualHealthStatistics2012_Aug2013.pdf. Accessed 20 Sept 2016
102. Sherman GG, Lilian RR, Bhardwaj S, Candy S, Barron P. Laboratory information system data demonstrate successful implementation of the prevention of mother-to-child transmission programme in South Africa. *S Afr Med J* 2014; **104**:235-238.
103. Massyn N, Day C, Peer N, Padarath A, Barron P, English R. District Health Barometer 2013/14. Durban; 2014.
104. Massyn N, Peer N, Padarath A, Barron P, Day C. District Health Barometer 2014/15. Durban; 2015. Available: http://www.hst.org.za/sites/default/files/Complete_DHB_2014_15_linked.pdf. Accessed 5 March 2017
105. Massyn N, Peer N, English R, Padarath A, Barron P, Day C. District Health Barometer 2015/16. Durban; 2016. Available: http://www.hst.org.za/publications/district-health-barometer-201516-0. Accessed 5 March 2017
106. Moyo F, Mazanderani A, Barron P, Bhardwaj S, Goga AE, Pillay Y, *et al.* Introduction of routine HIV birth testing in the South African national consolidated guidelines. *Pediatr Infect Dis J* 2018; **37**:559-563.
107. Kalk E, Kroon M, Boulle A, Osler M, Euvrard J, Stinson K, *et al.* Neonatal and infant diagnostic HIV-PCR uptake and associations during three sequential policy periods in Cape Town, South Africa: a longitudinal analysis. *J Int AIDS Soc* 2018; **21**:e25212.
108. Republic of South Africa. Global AIDS Response Progress Report 2012. 2012. Available: http://www.unaids.org/en/dataanalysis/knowyourresponse/countryprogressreports/2012countries/ce_ZA_Narrative_Report.pdf. Accessed 3 May 2014

109. Adonis L, An R, Luiz J, Mehrotra A, Patel D, Basu D, *et al.* Provincial screening rates for chronic diseases of lifestyle, cancers and HIV in a health-insured population. *S Afr Med J* 2013; **103**:309-312.
110. Council for Medical Schemes. Annual Report: 2012-2013. 2013. Available: http://www.medicalschemes.com/Publications.aspx. Accessed 6 May 2014
111. Swiss Re. Swiss Re HIV Testing Survey 2011: Report for the ASSA AIDS Committee. 2012.
112. Bizwell. Bizwell Stakeholder Report. 2012. Available: http://www.bizwell.co.za/documents/newsLinks/Bizwell%20Report%201%20July%202012.pdf. Accessed 5 May 2014
113. Jamieson L, Johnson LF, Matsimela K, Sande LA, d'Elbee M, Majam M, *et al.* The cost effectiveness and optimal configuration of HIV self-test distribution in South Africa: a model analysis. *BMJ Global Health* 2021; **6** (Suppl 4):e005598.
114. Thomas ML, Zuma K, Loykissoonlal D, Dube B, Vranken P, Porter SE, *et al.* A multi-level model for estimating region-age-time-type specific male circumcision coverage from household survey and health system data in South Africa. *arXiv* 2021:2108.09142.
115. Msemburi W, Pillay-van Wyk V, Dorrington RE, Neethling I, Nannan N, Groenewald P, *et al.* Second national burden of disease study for South Africa: Cause-of-death profile for South Africa, 1997-2010. Cape Town: South African Medical Research Council; 2014.
116. Dorrington RE. Alternative South African mid-year estimates, 2013. Centre for Actuarial Research; 2013. Available: http://www.commerce.uct.ac.za/Research_Units/CARE/Monographs/Monographs/Mono13.pdf. Accessed 19 Nov 2013
117. Statistics South Africa. Mid-year population estimates: 2013. Pretoria; 2013. Available: http://www.statssa.gov.za/publications/P0302/P03022013.pdf. Accessed 9 July 2015
118. Statistics South Africa. Tourism and Migration: July 2021. Pretoria; 2021. Available: https://www.statssa.gov.za/publications/P0351/P0351July2021.pdf. Accessed 11 April 2022
119. Johnson LF, Mutemaringa T, Heekes A, Boulle A. The effect of HIV and antiretroviral treatment on pregnancy rates in the Western Cape province of South Africa. *J Infect Dis* 2020; **221**:1953-1962.
120. Reproductive Health Research Unit. HIV and sexual behaviour among young South Africans: a national survey of 15-24 year olds. Joint publication of Reproductive Health Research Unit and loveLife; 2004. Available: www.rhru.co.za/images/Docs/national%20survey%20RHRU.pdf. Accessed 8 May 2004
121. Simbayi LC, Zuma K, Zungu N, Moyo S, Marinda E, Jooste S, *et al.* South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017. Cape Town: Human Sciences Research Council; 2019. Available: https://www.hsrcpress.ac.za/books/south-african-national-hiv-prevalence-incidence-behaviour-and-communication-survey-2017. Accessed 6 Nov 2019
122. Woldesenbet SA, Kufa T, Barron P, Chirombo BC, Cheyip M, Ayalew K, *et al.* Viral suppression and factors associated with failure to achieve viral suppression among pregnant women in South Africa: a national cross-sectional survey. *AIDS* 2020; **34**:589-597.

123. Grobler A, Cawood C, Khanyile D, Puren A, Kharsany ABM. Progress of UNAIDS 90-90-90 targets in a district in KwaZulu-Natal, South Africa, with high HIV burden, in the HIPSS study: a household-based complex multilevel community survey. *Lancet HIV* 2017; **4**:e505-e513.
124. Rohr JK, Xavier Gomez-Olive F, Rosenberg M, Manne-Goehler J, Geldsetzer P, Wagner RG, *et al.* Performance of self-reported HIV status in determining true HIV status among older adults in rural South Africa: a validation study. *J Int AIDS Soc* 2017; **20**:21691.
125. Sandfort TGM, Dominguez K, Kayange N, Ogendo A, Panchia R, Chen YQ, *et al.* HIV testing and the HIV care continuum among sub-Saharan African men who have sex with men and transgender women screened for participation in HPTN 075. *PLoS One* 2019; **14**:e0217501.
126. Huerga H, Shiferie F, Grebe E, Giuliani R, Farhat JB, Van Cutsem G, *et al.* A comparison of self-report and antiretroviral detection to inform estimates of antiretroviral therapy coverage, viral load suppression and HIV incidence in Kwazulu-Natal, South Africa. *BMC Infect Dis* 2017; **17**:653.
127. Johnson LF. Access to antiretroviral treatment in South Africa, 2004-2011. *South Afr J HIV Med* 2012; **13**:22-27.
128. Dorrington RE, Moultrie TA, Timæus IM. Estimation of mortality using the South African Census 2001 data. Centre for Actuarial Research; 2004. Available: http://www.commerce.uct.ac.za/Research_Units/CARE/Monographs/Monographs/Mono11.pdf. Accessed 3 Dec 2015
129. Johnson LF, Dorrington RE, Laubscher R, Hoffmann CJ, Wood R, Fox MP, *et al.* A comparison of death recording by health centres and civil registration in South Africans receiving antiretroviral treatment. *J Int AIDS Soc* 2015; **18**:20628.
130. Statistics South Africa. Mortality and causes of death in South Africa, 2016: Findings from death notification. 2018. Available: http://www.statssa.gov.za/publications/P03093/P030932016.pdf. Accessed 18 Dec 2018
131. Raftery AE, Bao L. Estimating and projecting trends in HIV/AIDS generalized epidemics using Incremental Mixture Importance Sampling. *Biometrics* 2010; **66**:1162-1173.
132. Jooste S, Mabaso M, Taylor M, North A, Shean Y, Simbayi LC, *et al.* Geographical variation in HIV testing in South Africa: Evidence from the 2017 national household HIV survey. *South Afr J HIV Med* 2021; **22**:a1273.
133. Ramjee G, Wand H, Whitaker C, McCormack S, Padian N, Kelly C, *et al.* HIV incidence among non-pregnant women living in selected rural, semi-rural and urban areas in KwaZulu-Natal, South Africa. *AIDS Behav* 2012; **16**:2062-2071.
134. Delany-Moretlwe S, Lombard C, Baron D, Bekker LG, Nkala B, Ahmed K, *et al.* Tenofovir 1% vaginal gel for prevention of HIV-1 infection in women in South Africa (FACTS-001): a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet Infect Dis* 2018; **18**:1241-1250.
135. Tanser F, Bärnighausen T, Grapsa E, Zaidi J, Newell ML. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science* 2013; **339**:966-971.
136. Woldesenbet S, Kufa-Chakezha T, Lombard C, Manda S, Cheyip M, Ayalew K, *et al.* Recent HIV infection among pregnant women in the 2017 antenatal sentinel cross-sectional survey, South Africa: Assay-based incidence measurement. *PLoS One* 2021; **16**:e0249953.

137. Norval M, Coussens AK, Wilkinson RJ, Bornman L, Lucas RM, Wright CY. Vitamin D status and its consequences for health in South Africa. *International Journal of Environmental Research and Public Health* 2016; **13**.
138. Viard JP, Souberbielle JC, Kirk O, Reekie J, Knysz B, Losso M, *et al.* Vitamin D and clinical disease progression in HIV infection: results from the EuroSIDA study. *AIDS* 2011; **25**:1305-1315.
139. Alvarez N, Aguilar-Jimenez W, Rugeles MT. The potential protective role of vitamin D supplementation on HIV-1 infection. *Front Immunol* 2019; **10**:2291.
140. Mahlalela PT, Blamey RC, Hart NCG, Reason CJC. Drought in the Eastern Cape region of South Africa and trends in rainfall characteristics. *Climate Dynamics* 2020; **55**:2743-2759.
141. Low AJ, Frederix K, McCracken S, Manyau S, Gummerson E, Radin E, *et al.* Association between severe drought and HIV prevention and care behaviors in Lesotho: A population-based survey 2016-2017. *PLoS Med* 2019; **16**:e1002727.
142. Ndlovu MS, Demlie M. Assessment of meteorological drought and wet conditions using two drought indices across KwaZulu-Natal province, South Africa. *Atmosphere* 2020; **11**:623.
143. Johnson LF, Mossong J, Dorrington RE, Schomaker M, Hoffmann CJ, Keiser O, *et al.* Life expectancies of South African adults starting antiretroviral treatment: Collaborative analysis of cohort studies. *PLoS Med* 2013; **10**:e1001418.
144. Johnson LF, Anderegg N, Zaniewski E, Eaton JW, Rebeiro PF, Carriquiry G, *et al.* Global variations in mortality in adults after initiating antiretroviral treatment: an updated analysis of the IeDEA cohort collaboration. *AIDS* 2019; **33** (Suppl 3):S283-S294.
145. May M, Sterne JA, Sabin C, Costagliola D, Justice AC, Thiebaut R, *et al.* Prognosis of HIV-1-infected patients up to 5 years after initiation of HAART: collaborative analysis of prospective studies. *AIDS* 2007; **21**:1185-1197.
146. Anderson SJ, Garnett GP, Enstone J, Hallett TB. The importance of local epidemic conditions in monitoring progress towards HIV epidemic control in Kenya: a modelling study. *J Int AIDS Soc* 2018; **21**:e25203.
147. Grasso MA, Manyuchi AE, Sibanyoni M, Marr A, Osmand T, Isdahl Z, *et al.* Estimating the population size of female sex workers in three South African cities: Results and recommendations from the 2013-2014 South African Health Monitoring Survey and stakeholder consensus. *JMIR Public Health and Surveillance* 2018; **4**:e10188.
148. Maleke K, Makhakhe N, Peters RPH, Jobson G, De Swardt G, Daniels J, *et al.* HIV risk and prevention among men who have sex with men in rural South Africa. *Afr J AIDS Res* 2017; **16**:31-38.
149. Oster AM, Sternberg M, Lansky A, Broz D, Wejnert C, Paz-Bailey G. Population size estimates for men who have sex with men and persons who inject drugs. *J Urban Health* 2015; **92**:733-743.
150. Mahy MI, Sabin KM, Feizzadeh A, Wanyeki I. Progress towards 2020 global HIV impact and treatment targets. *J Int AIDS Soc* 2021; **24** (Suppl 5):e25779.
151. South African National AIDS Council. National Strategic Plan for HIV, TB and STIs, 2017-2022. 2017. Available: http://sanac.org.za/wp-content/uploads/2017/05/NSP_FullDocument_FINAL.pdf. Accessed 27 Feb 2018

Appendix A: Differences in HIV profiles between immigrants and residents

Prior to version 4.3, the Thembisa model assumed that the HIV profile of individuals migrating into and out of a province in a given year was that same as that of individuals in the province who did not migrate. This is probably an unrealistic assumption, especially in the provinces with relatively low HIV prevalence, for which most immigrants are likely to be from provinces with higher prevalence levels.

Migration is modelled as occurring at the end of each projection year. Suppose that N_x is the population at age x at the end of the year before adjusting for migration, that I_x is the number of in-migrants at age x at the end of the year, and that E_x is the number of out-migrants at age x at the end of the year. Further suppose that π_x is the HIV prevalence in the population aged x before the migration adjustment is made and θ is the ratio of HIV prevalence in immigrants to that in the receiving population (after controlling for age). It is assumed that the individuals leaving the population have the same age-specific HIV prevalence levels as those who remain in the population. Then the ratio of HIV prevalence in the population aged x after migration to that before migration is

$$R_x = \frac{\pi_x\left(N_x + I_x\theta - E_x\right)}{N_x + I_x - E_x} \Big/ \pi_x = \frac{N_x + I_x\theta - E_x}{N_x + I_x - E_x},$$

which is independent of π_x. This can be calculated as $R_x = 1 + (\theta - 1)J_x$, where $J_x = I_x / \left(N_x + I_x - E_x\right)$ is the rate of in-migration (the fraction of the population that moved into the province over the last year). The J_x values are estimated from the censuses conducted in 1996, 2001 and 2011. The 2001 and 2011 censuses also asked questions about the province (or country) in which the recent migrants lived prior to moving to their current location. Suppose that α_i is the proportion of recent immigrants who are from province/country i, and that θ_i is the ratio of HIV prevalence in province/country i to that in the receiving province. Then we can calculate

$$\theta = \sum_i \alpha_i \theta_i .$$

The α_i values are calculated from the census data. In the case of the 1996 census, there is the complication that previous locations were not recorded by province or country. We have therefore assumed that individuals who reported having migrated to their current location in the last year but who reported a country of birth outside of South Africa were international in-migrants, and all in-migrants who were not classified as international were allocated to different provinces of origin in the same proportions as reported in the 2001 census.

For the purpose of calculating the θ_i ratios, we have taken estimates of HIV prevalence in 15-49 year olds in each year from two sources. Estimates of HIV prevalence in each of South Africa's provinces are taken from the previous version of the Thembisa provincial models [13]. Estimates of HIV prevalence in countries outside of South Africa are taken from UNAIDS (accessed from http://aidsinfo.unaids.org/ on 10 Feb 2017). For the purpose of our analysis, it

is sufficient to take the prevalence estimates from five other African countries that make up the bulk of in-migration into South Africa (Zimbabwe, Mozamique, Lesotho, Malawi and Swaziland). Other international in-migrants are almost all from low-HIV prevalence settings, and the θ_i ratio is therefore set to zero for international in-migrants from other countries.

The resulting R_x values are shown in Figure A1 for each province. In most provinces and in most years, the ratio is very close to 1, indicating that migration has little effect on HIV prevalence. However, in the three provinces with the lowest levels of HIV prevalence (Western Cape, Northern Cape and Limpopo), the ratio is notably above 1, as might be expected when migrants arrive mostly from provinces (or countries) in which there is a relatively high HIV prevalence. In contrast, the ratios tend to be below 1 in the provinces that have the highest HIV prevalence levels (KwaZulu-Natal and Mpumalanga), since most of the migrants into these provinces will be from areas with relatively low prevalence. The ratios are generally furthest from 1 in 1996, when the South African HIV epidemic was in its early stages and there were large differences in prevalence between provinces; over time, these inter-provincial differences in prevalence have reduced. It is also apparent that the ratios are furthest from 1 in the young adult age range (20-35) and in children under the age of 5; these are the age groups in which the most migration occurs, and one would therefore expect the impact of migrant on HIV prevalence to be most substantial in these age groups.

The R_x values calculated from the 1996, 2001 and 2011 census data are assumed to apply in the 1995, 2000 and 2010 projection years. The reason for this is that projection years run from mid-year to mid-year, and the census is conducted in October of each year. Thus the 1995 projection year runs from mid-1995 to mid-1996, and the migration adjustments are made at the end of the projection year (mid-1996). This corresponds most closely to the census data collected in October 1996 regarding migration over the last year. We lack reliable census/survey data on inter-provincial migration before 1996 and after 2011, and have therefore assumed that the R_x values for projection years before 1995 are the same as in 1995, and R_x values for projection years after 2010 are the same as those in 2010. In the projection years 1996-1999, the R_x values are linearly interpolated between the values estimated in the 1995 and 2000 years. Similarly, in the projection years 2001-2009, the R_x values are linearly interpolated between the values estimated in the 2000 and 2010 years.

Suppose that at age x, HIV prevalence is $\pi_{g,x}$ in individuals of sex g. Similarly, $M_{g,x}$ is the migration adjustment factor (the ratio of the population size after immigrants and emigrants have been subtracted/added, to the population size before migration). Both variables are indexed by age at the end of the year. The number of HIV-positive individuals in each HIV disease state is increased by a factor of $M_{g,x}\ R_x$, and the number of HIV-negative individuals is increased by a factor of $M_{g,x}\ (1 - \pi_{g,x}\ R_x)/(1 - \pi_{g,x})$. This adjustment ensures that the population size (for HIV-positive and HIV-negative combined) increases by a factor of $M_{g,x}$.

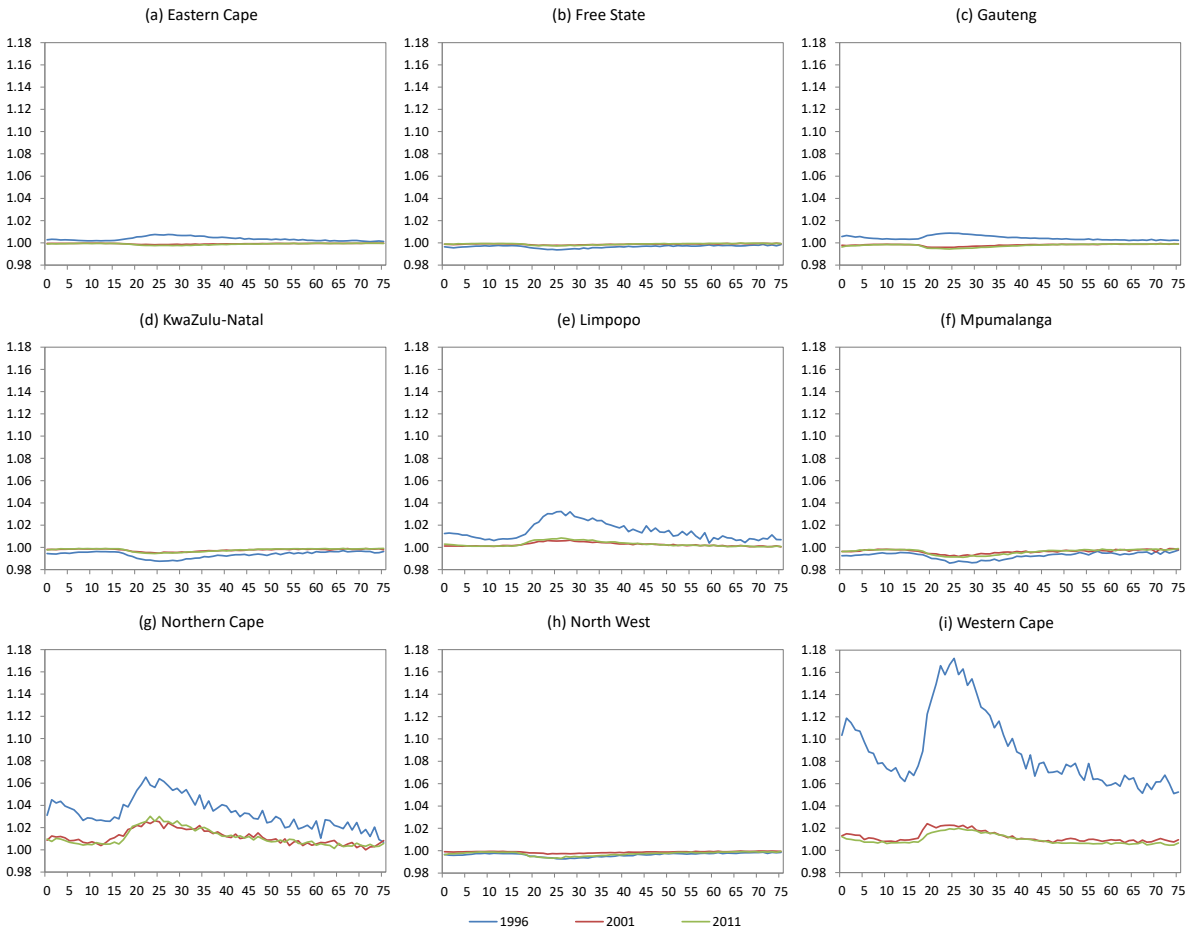

Figure A1: Ratio of HIV prevalence in the population after migration to HIV prevalence before migration