MicroCOSM-HPV: Technical Appendix

This technical appendix provides more details about the MicroCOSM model (version 1) and modifications to this framework to include the transmission of 13 oncogenic human papillomavirus (HPV) types and their progression to cervical cancer.

1. Natural history of HPV

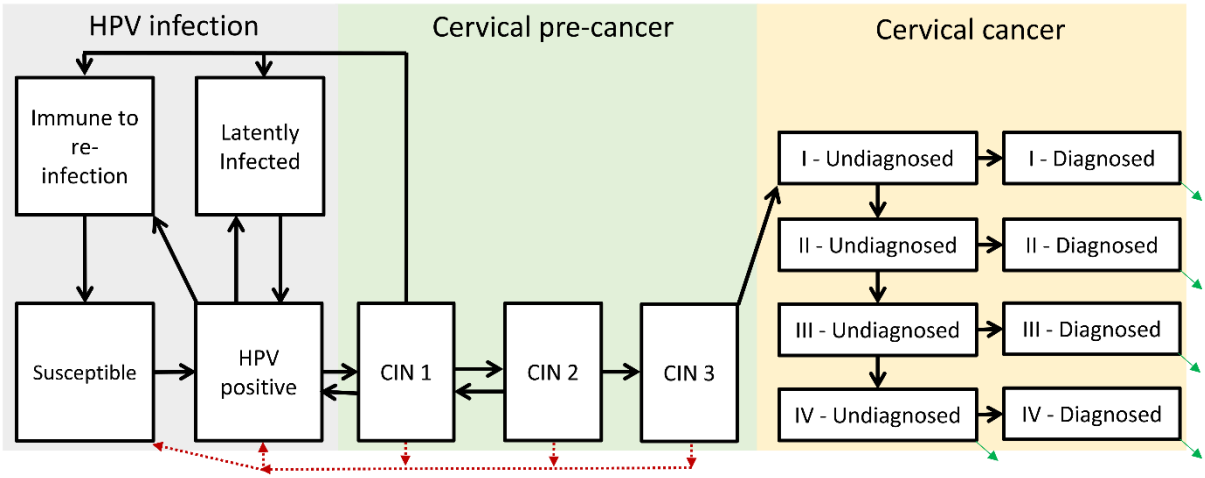


Figure 1 - Model of the natural history of HPV and cervical cancer. Black arrows represent natural movement between states, red arrows represent clearance of disease due to treatment, and green arrows represent excess mortality due to cervical cancer.

The structure of stages representing HPV infection in males and females (shaded in grey in Figure 1) was established in a previous study that showed this model structure (which includes latent infection and reactivation) fits better to HPV type-specific data during calibration, and to vaccine trial results (1). HPV duration and rates of reactivation of latent infection are dependent on HIV stage, and HPV types are simulated independently of each other.

Cervical disease stages shaded in green and yellow apply only to women, who can progress from HPV infection through 3 pre-cancer stages to cervical cancer. Women can naturally regress from the first two pre-cancer stages (CIN1/2), but not from the third. A fraction of women who regress from lower grade lesions will remain HPV infected, while the rest will either become naturally immune or latently infected. In the model, women move directly back to either susceptible or infected after treatment of abnormalities following screening (red dotted lines in Figure 1). Rates of progression and regression are dependent on age, HIV and ART status, but HIV status does not influence transmission probabilities for HPV infection (2). Details about these parameters are shown in Sections 6 and 7 of this document.

2. HIV and other sexually transmitted infections

MicroCOSM simulates the natural history of human immunodeficiency virus (HIV), genital herpes, syphilis, gonorrhoea, chlamydia and trichomoniasis. Assumptions about transitions between disease states are described in detail in the online appendix of Johnson & Geffen (3). Of these STIs, only HIV is simulated in this study. HIV is introduced to the population in 1990 by randomly choosing a fraction of high-risk individuals to be HIV-positive. The parameters with greatest uncertainty driving the spread of HIV in South Africa were estimated in Johnson & Geffen and in this study, we will use the medians of the best fitting parameter combinations (Table A 1).

Table A 1 - Best-fitting HIV model parameter values

Parameter	**Median (IQR)**
Transmission probability per sex act	
M-to-F, non-spousal	0.81% (0.69-0.93%)
F-to-M, non-spousal	0.36% (0.31-0.4%)
M-to-F, spousal	0.19% (0.14-0.24%)
F-to-M, spousal	0.17% (0.13-0.21%)
Relative infectiousness, acute HIV	19.3 (14.9-23.1)
Relative infectiousness, AIDS	6.9 (5.71-8.15)
Initial prevalence in high-risk women	2.31% (1.76-2.6%)
Bias in self-reported condom use	0.638 (0.537-0.773)

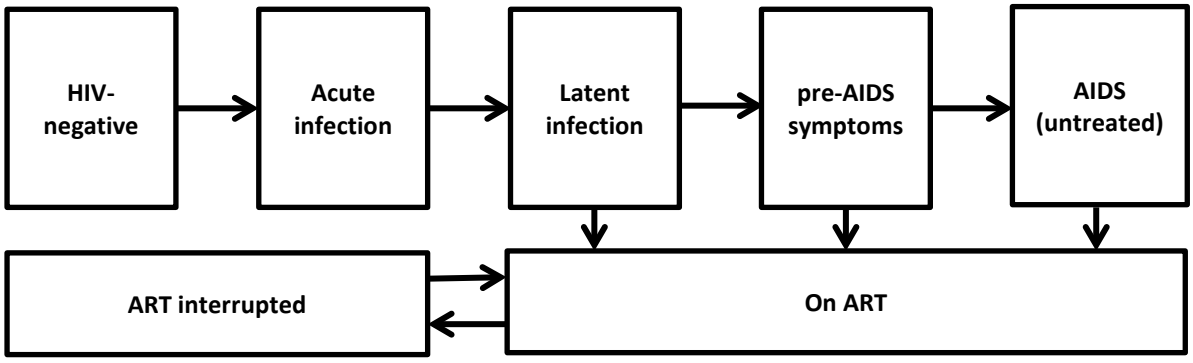


Figure A 1- Model of the natural history of HIV

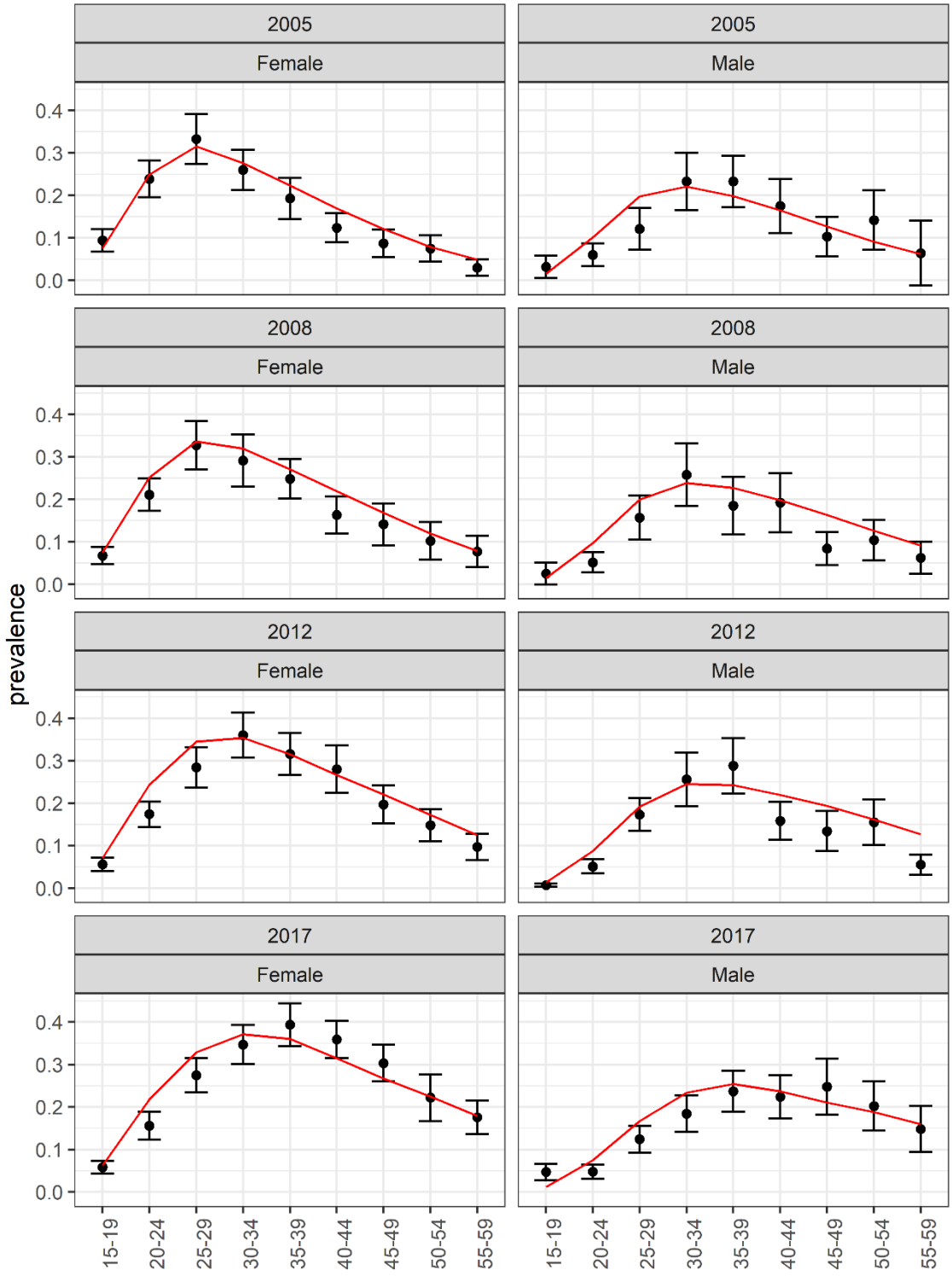


Figure A 2 – HIV prevalence by age, sex and time. Black dots and error bars are data from the HIV prevalence surveys performed by the Human Sciences Research Council (HSRC), and the red lines are estimates from the model.

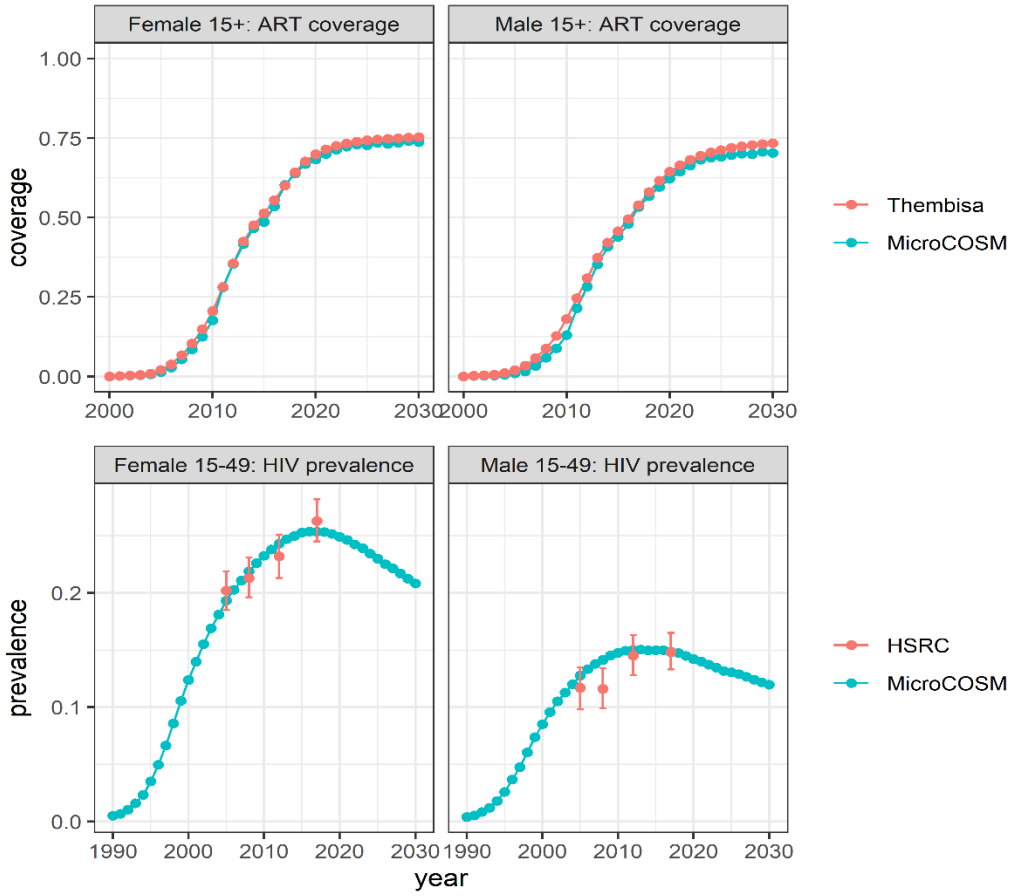


Figure A 3 – ART coverage and HIV prevalence over time.

3. Sexual behaviour

At birth, or in 1985, an individual is randomly assigned the static status of high-risk or low-risk based on the propensity for concurrent partnerships or commercial sex. All individuals become sexually active between ages 10 and 30. At each time step, individuals looking for sexual partners are matched to other individuals looking for partners. Low-risk individuals will be single, looking for one partner. High-risk individuals may be looking for a primary partner, secondary partner or, if male, a contact with a female sex worker. All single individuals will look for a short-term relationship (at rates determined by age, sex and risk group), which may eventually (with an average duration of six months) dissolve or become a long-term relationship (marital or cohabiting). When sexual relationships are formed, the ID of the individual is linked to the ID of the partner which allows us to simulate the sexual network and keep track of transmission of STIs in the population. Relationship dissolution or marriage is based on rates determined by age and sex. It is possible to form relationships with individuals in different age groups or risk groups. Condom usage and frequency of sex acts depend on age, sex, and relationship type. Condom usage is modelled in a similar way as in the Thembisa model (4), with the proportion of sex acts protected by condoms increasing as HIV prevalence increases, and declining again after the scale-up of ART (Figure A 4). Male circumcision is not simulated in this version of MicroCOSM. Rates at which sexual behaviour related events occur have been estimated in previous publications (3,5) and details regarding the wide variety of data sources and the calibration methods to estimate the rates are documented in these publications.

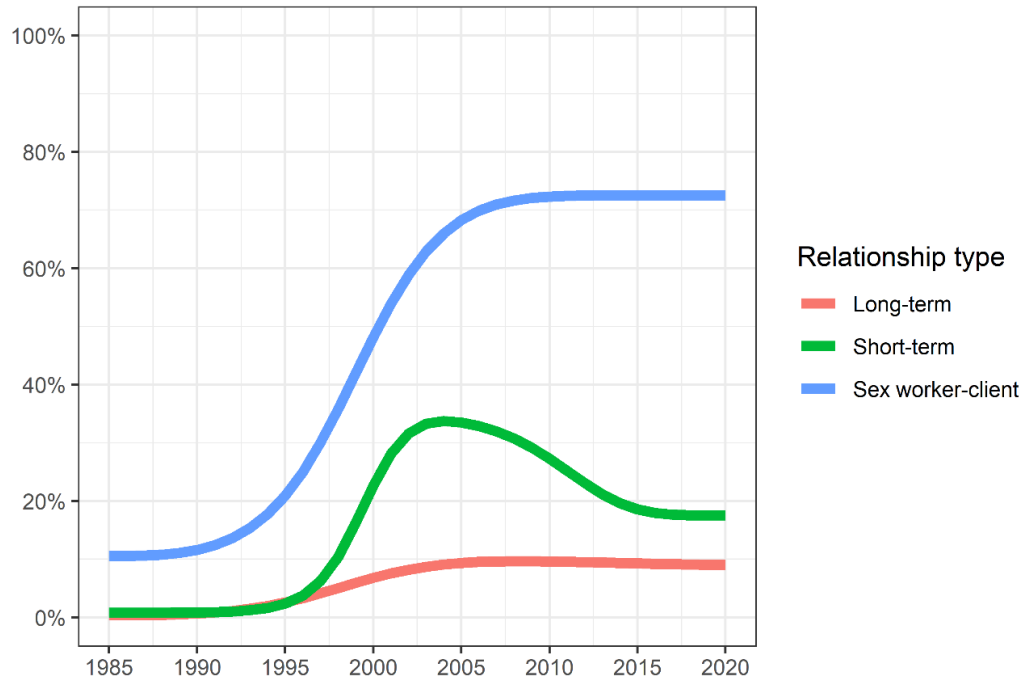


Figure A 4 – The proportion of sex acts protected by condoms among females aged 20-24 over time.

4. Starting conditions

As mentioned in the previous section, the simulation of individuals in MicroCOSM starts in 1985, and HIV is introduced in 1990. Except for condom usage, sexual behaviour does not change over time. Individuals are randomly assigned to HPV stages (for each of 13 types) in 1985 according to their sex, age and sexual behaviour risk category. For men, these stages are HPV naïve (susceptible), HPV infected, latently infected, or naturally immune. For women, the stages are the same as for men, but with three stages of pre-cancer (CIN1-3) and one stage for cancer (Stage I, undiagnosed cancer). The fraction that is assigned to each stage is determined by an iterative process:

1) Calculate crude estimates of the fraction in each stage in 1985 by age, sex, and risk category by using prevalence data.
2) Using the medians of the prior distributions of the parameters and these crude starting conditions, simulate HPV in the population for 50 years, assuming no HIV, no change in condom use and no CC screening.
3) Calculate the fraction of individuals in each HPV stage by age, sex, and risk category at the end of the simulation and update the starting conditions with these values. Repeat steps 2 and 3 until fractions remain stable over time.
4) Calibrate the model to data and after each calibration step, repeat steps 2 and 3 by replacing the medians of the prior distributions with the medians of the best fitting parameter values.

The natural history of HPV – from infection to cancer – is a process that happens over a ~30-year period in immune competent women. By following this process to determine starting conditions, we are therefore not only assuming that sexual behaviour parameters remain constant after 1985, but we are also implicitly assuming that they were the same in the decades preceding 1985. This is a limitation of our study.

5. Screening algorithm

A.5.1 Background on screening in South Africa

The South African National Department of Health published a cervical cancer screening policy in 2000. Before, women in the public sector typically received pap smears only for diagnostic purposes, with some regional exceptions. The policy allows for a pap smear every 10 years, with the first smear after age 30. Screening coverage estimates are published in the annual District Health Barometer (DHB) (6). The numerator is a count of all smears performed, excluding those collected for diagnostic purposes or repeat smears (e.g., following an inadequate smear). Since each woman should be screened once in 10 years, the target population (denominator) every year is a 10^{th} of the population of women aged at least 30.

1) In 2010, NDoH published HIV management guidelines that included a section on cervical cancer screening (7). In this document it was stipulated that all women should receive a Pap smear immediately after testing HIV-positive and every three years thereafter. However, HIV status is not captured well on cytology forms sent to NHLS, and to date the DHB estimate of cervical cancer screening coverage is not separated by HIV status. Therefore, this estimate is biased in at least three ways:
    - The same HIV-positive women can be counted 3 times in a ten-year interval, and they may be younger than 30.
    - Due to the lack of a unique identifier, it is impossible to determine whether women and clinicians follow the ten-yearly interval policy, and some women may be screened more often.
    - Only public sector Pap smears are counted, but the entire population older than 30 (public and private healthcare users) is used as the denominator.

Screening coverage in the WC, as estimated in the DHB, has been close to the national average (Figure A 10) between 2006 and 2017. For this reason, we use data from the Western Cape – available between 2007 and 2018 – as a proxy for national level data in the derivation of the national population level screening algorithm in our model.

The WC is the only province in South Africa that utilises a unique patient identifier at all levels in the public health care sector (8) and therefore this dataset has individual level records of Pap smears performed. The cytology dataset has also been linked to HIV data sources such as laboratory tests and ART programme data and we can estimate HIV and ART experience at the time of the smear.

Figure A5 shows the process of screening that is in general followed in the public health sector of South Africa and that we will simulate in our model. More details regarding every step are explained below.

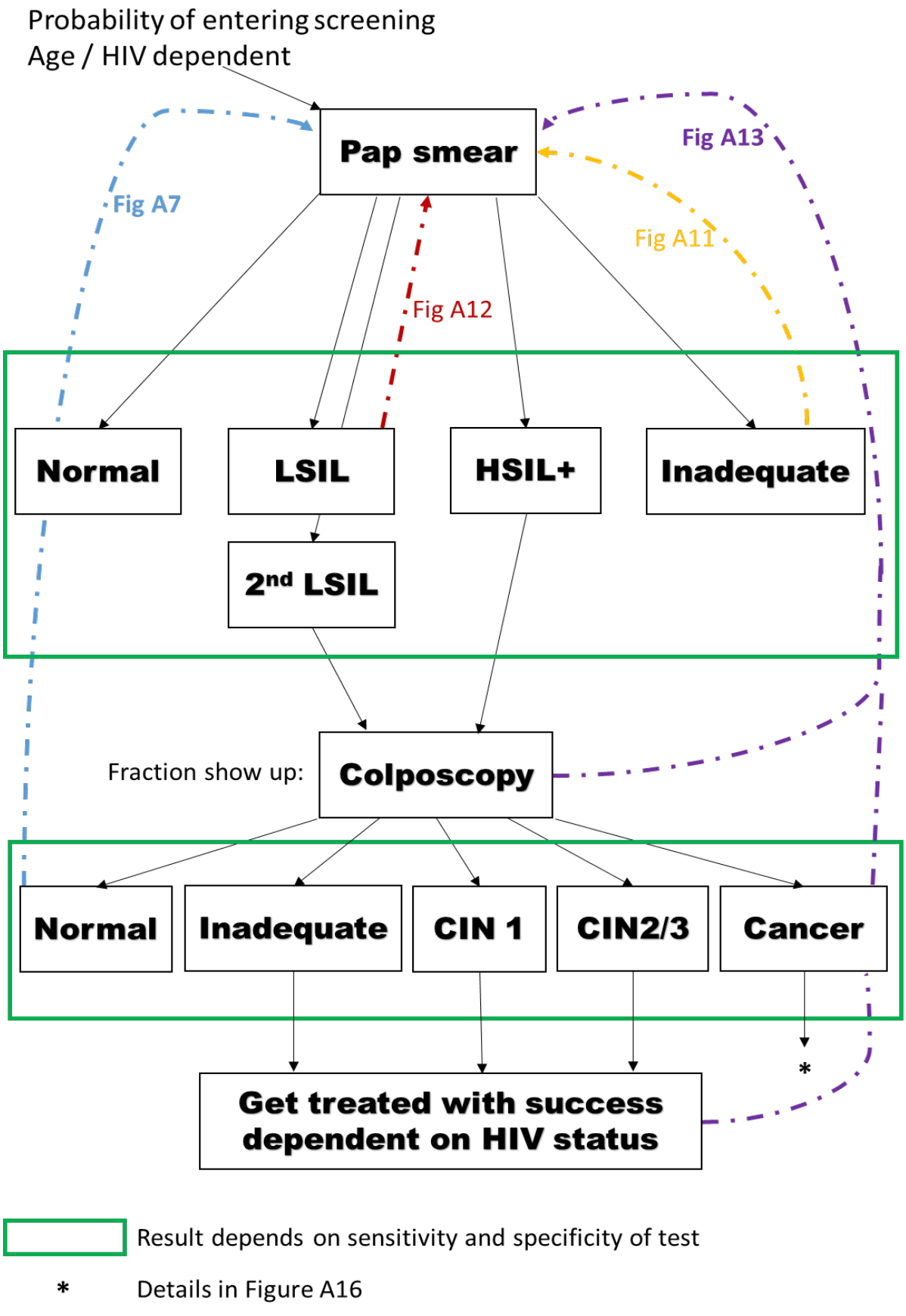


Figure A 5 - Screening algorithm in MicroCOSM

A.5.2 Probabilities of entering screening

To inform the screening algorithm in our model, we built a *separate simulation model* to obtain probabilities of entering cervical cancer screening. This simulation model is described below.

A.5.2.1 Simulating the population of the Western Cape

The first step in this individual-based model is to simulate a model population that represents the female population aged 15 and older in the Western Cape between 2001 and 2018. We use population size, all-cause mortality, migration and HIV estimates from Thembisa 4.2 (9). In the WC screening data, we have found that the time between routine screens is on average similar for women with negative or unknown HIV status and women who are HIV-positive, but ART-naïve. Therefore, we have split the population of our screening simulation model into two parts – women who are ART-naïve (HIV-negative or -positive) and women who are ART-experienced. To match population estimates from Thembisa, we assume that women can initiate ART starting from 2005 and that the yearly rate of initiating ART increases linearly. This does not realistically correspond to the way ART has been rolled out in the last 15 years, but produces credible numbers of women on ART in our simple model world. These rates, by age group, are estimated as the rates that minimise the squared difference between the simulated population size and the Thembisa population size. Women on ART have different mortality rates than other women, who experience all-cause mortality rates (including HIV). Figure A 6 illustrates population sizes in the Thembisa data and simulation model for the best fitting parameters, as shown in Table A 2 - Least square estimate of the slopes of the linear increase in rate of ART initiation. These numbers can be interpreted as follows: In 2005, each woman aged between 30 and 40 initiated ART at a rate of 0.0029, while in 2015, this rate increased to 10*0.0029=0.029.Table A 2.

*Table A 2 - Least square estimate of the slopes of the linear increase in rate of ART initiation. These numbers can be interpreted as follows: In 2005, each woman aged between 30 and 40 initiated ART at a rate of 0.0029, while in 2015, this rate increased to 10*0.0029=0.029.*

Age	**15-30**	**30-40**	**40-50**	**50-90**
Initiate ART	0.07%	0.29%	0.12%	0.02%

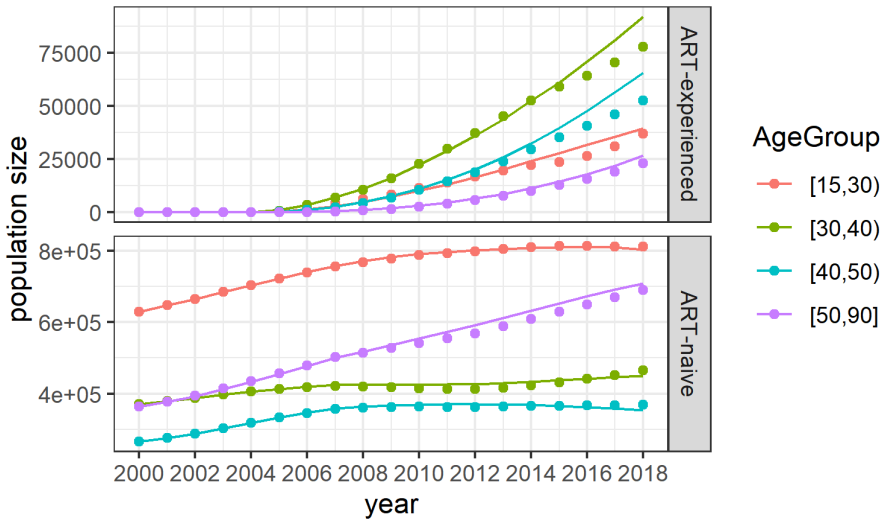


Figure A 6 - The number of ART-naïve women (HIV-positive or -negative) and ART-experienced in each age group in the Western Cape, as estimated by Thembisa (the dots) and the screening simulation model (the lines).

A.5.2.2 Simulating cervical cancer screening in the Western Cape

In the screening simulation model, we make some key assumptions. We assume that screening only happens up to the age of 60, and that HIV-negative and ART-naïve women can start screening at the age of 20 and women on ART at the age of 15. We assume that no one was screened before 2000 and that the yearly rate of entering the screening programme or receiving a first-time screen increases linearly from zero in 1999 and (for HIV-negative and ART-naïve women) plateaus in 2010, the year that the number of screens performed among these women seems to be stabilising. The exception is that rates of entering screening are constant for HIV-negative and ART-naïve women aged 20 to 30 – these women are excluded from the national recommended screening policy, but data show that they are screened at low frequencies. Hence, there are 4 unknown parameters: the constant rate for HIV-negative and ART-naïve women aged 20 to 30 and a slope of the linear increase for the age groups 30-40, 40-50 and 50-60. The yearly rate of entering screening for ART-experienced women of four age groups (including women aged 15 to 30) increases linearly from zero in 2004 and does not plateau, i.e., there are also 4 unknown parameters. The simulation model does not distinguish between routine screens and screens for diagnostic purposes. The rates of entering screening represent screening for any reason.

Intervals between screens are randomly drawn from two Weibull distributions – one for HIV-negative and ART-naïve women and one for ART-experienced women (Figure A 7). These distributions were derived from the WC data and represent the time between a Normal Pap smear result and the following smear. If an HIV-negative or ART-naïve woman receives a Normal Pap smear result after the age of 50, she is not screened again.

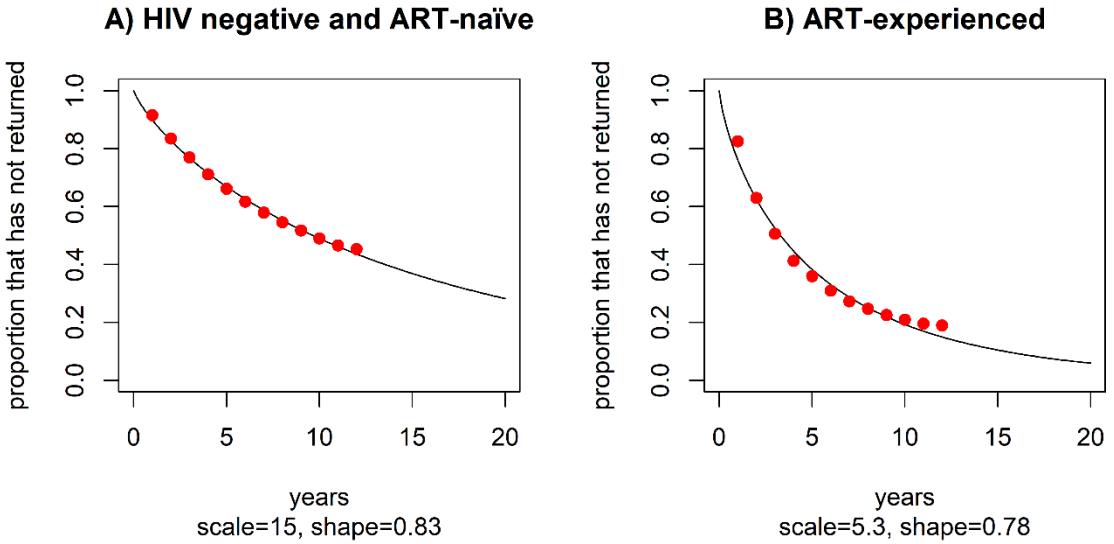


Figure A 7 - Distributions of time between a Pap smear with Normal cytology and next routine Pap smear. Red dots are data from the WC NHLS cytology database and black lines are best fitting Weibull distributions. These figures can be interpreted as follows: In A), around 23% of HIV-negative and ART-naïve women have returned for a next routine screen within 3 years and in B), around 50% of ART-experienced women have returned for a next routine screen within 3 years.

We exclude repeat smears from the WC data. These are smears performed following an inadequate smear (follow-up should be within 3 months) or a follow-up smear following a Pap smear result of lower grade lesions (follow-up should be within 1 year). We do not simulate adequacy or cytology

result and therefore repeat events in the WC data are represented by one event in the simulation model, which has yearly time steps.

The simulation model is stochastic, but seeds are fixed in the calibration phase. We use the Nelder-Mead algorithm to estimate the rate (for women aged 20-30 HIV-negative/ART-naïve) and slopes (for everyone else) that results in the smallest squared difference between the number of screens performed every year between 2007 and 2018 in the simulation model and the WC data. Then, varying the seed for the random number generator, we run the simulation a 100 times with the least-squares estimates of the parameters (Table A 3 and Figure A 8) and show the means in Figure A 9. In Figure A 8 we convert rates to probabilities for ease of interpretation.

*Table A 3 - Least square estimate of the slopes of the linear increase in rate of entering the screening programme. These numbers can be interpreted as follows: ART-naïve women aged 20-30 have a constant rate of entering the screening programme of 0.011 per year. In 2000, each woman aged between 30 and 40 ART-naïve entered the screening programme a rate of 0.008, while in 2010 and every year thereafter, this rate is 11*0.008=0.088.*

Age	**15-30**	**30-40**	**40-50**	**50-60**
HIV-negative and ART-naïve	0.011*	0.008	0.005	0.008
ART-experienced	0.046	0.048	0.043	0.053

*Constant rate for women aged 20-30

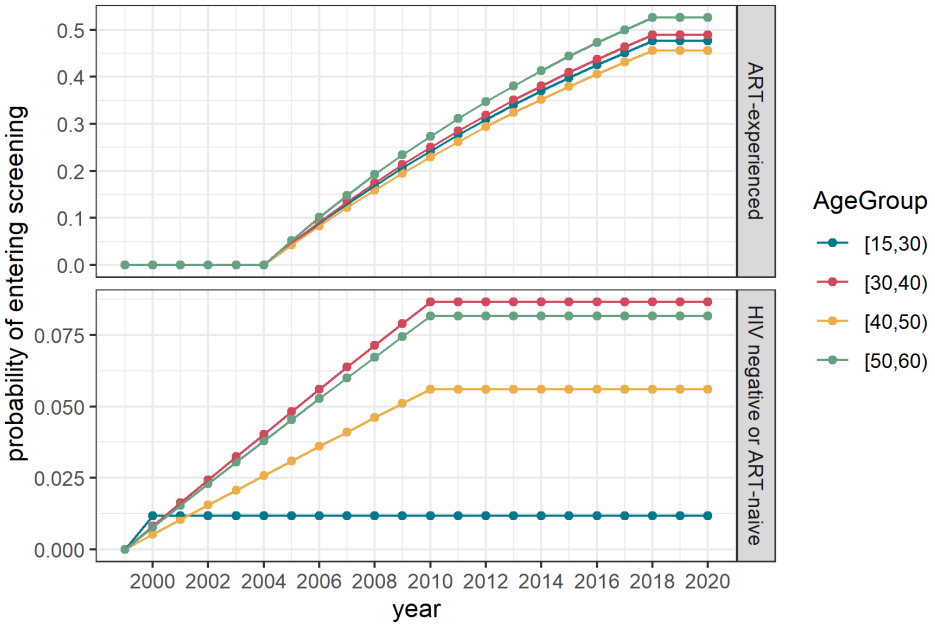


Figure A 8 – Yearly probabilities of entering the screening programme, by age and ART status.

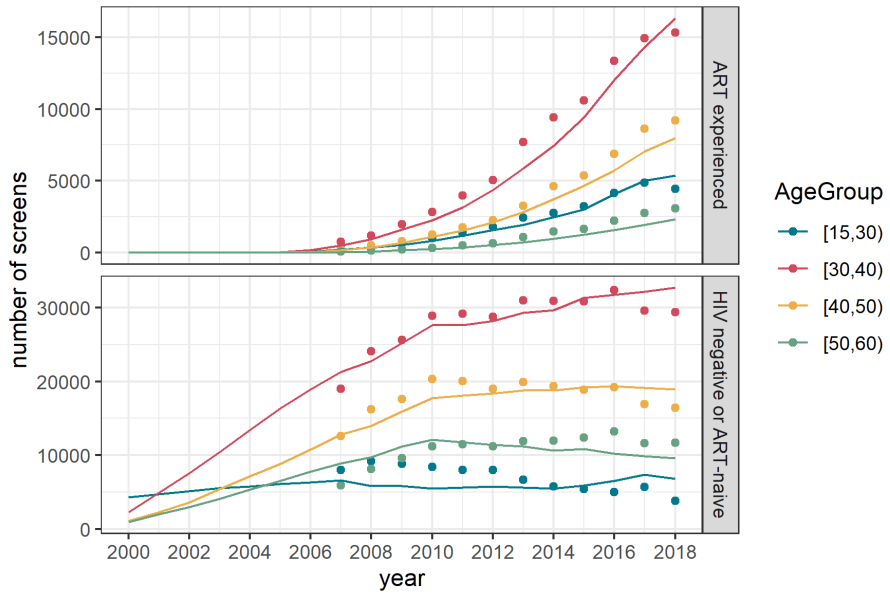


Figure A 9 - The number of ART-naïve women (HIV-positive or -negative) and on ART in each age group receiving routine Pap smears in the public sector of the Western Cape, as counted from the WC NHLS data (the dots) and the screening simulation model (the lines).

We calculate screening coverage from our simulated data using the same method as the estimate published in the District Health barometer (6) and show this estimate over time along with estimates from the WC and South Africa as published in the DHB (Figure A 10). Coverage estimates from the simulation model are slightly lower than the numbers published in the DHB, since we removed some additional repeat smears that were not indicated as such in the data (the authors of the DHB do not have access to the individual-level data and are not able to do this). We implemented this screening algorithm in the individual-based model that simulates the natural history of HPV (MicroCOSM). Coverage from this simulation is also shown in Figure A 10.

Note that although we used the 2007 to 2018 Western Cape data to fit the model, we use this data as a proxy for national level data and start at a coverage of 0% in 1999. For this reason, we intended that the simulation data in Figure A 10 match the South African coverage from 2000 to 2006.

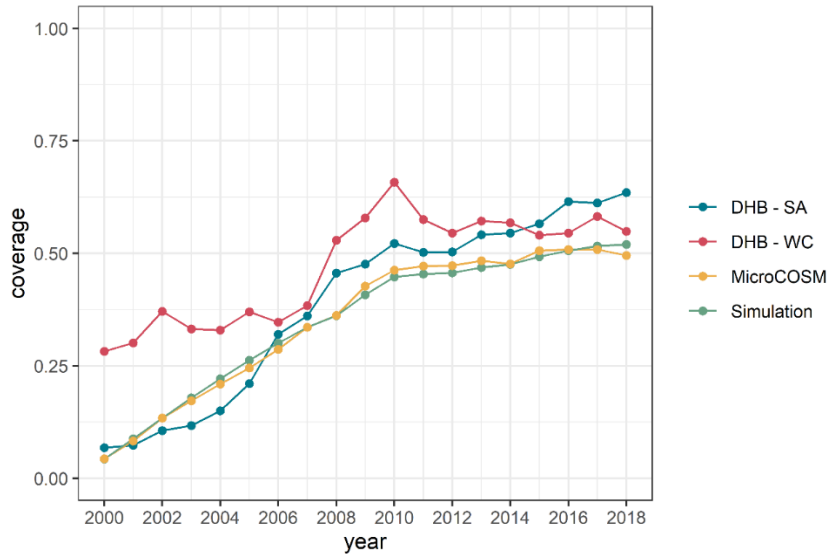


Figure A 10 - Screening coverage defined as the number of routine screens performed for women aged 30 and older, divided by a 10th of the population. In this definition, all screens are counted, except for repeat smears and diagnostic smears.

A.5.3 Adequacy of smears

According to the Western Cape NHLS cytology dataset, around 90% of all screens were adequate for evaluation (contained endocervical cells) between 2007 and 2014. In 2015, after publication of the updated Bethesda system (10), smears without endocervical cells could also be classified as adequate, and the adequacy increased to 97.5% in 2018. In Limpopo, adequacy remained around 98% between 2007 and 2010 (11). However, analyses of adequacy in South Africa as a whole, also using NHLS data, found very different results. Makura *et al.* (12) estimated the median adequacy rate over all the districts in the country in 2013/14 was 47% with an interquartile range of 44 to 56%. Schnippel *et al.* (13) showed that overall national adequacy declined from 80.5% in 2010 to 54.4% in 2014 and they argue that this decline may be attributable to the increase of Pap smear coverage. In a study performed in 2001/2, large discrepancies in adequacy rates were shown among three labs at different time points, varying from 45% to 100% (14). A national level study performed in the early 2000s showed very high overall adequacy in all sites of 95% and above (15). In our model, we assume that adequacy was 90% until 2007, linearly declined to 54.4% in 2014 and stayed constant thereafter.

A woman with an inadequate screen should return for screening within three months. However, between 2007 and 2015, only 9% of women returned within 3 months in the WC. To simulate rescreening for women with inadequate results, we fitted a Weibull distribution through the proportions of women who have not returned (Figure A 111). We use the subset of women with inadequate smears between 2007 and 2015 to maximise follow-up time and reduce right censoring. In the model, a time of rescreening will be randomly drawn from this Weibull distribution.

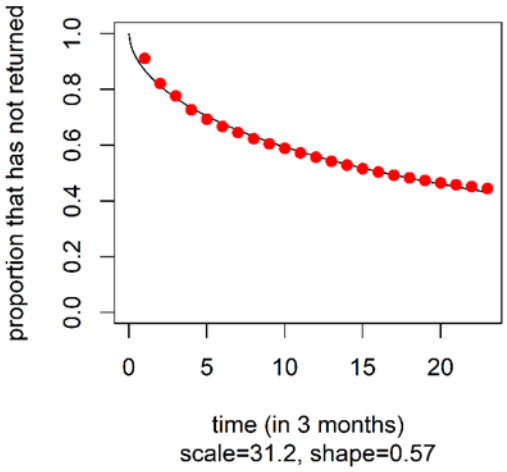


Figure A 11 - Distribution of time between an inadequate screen and the next screen, in three-monthly time steps. Red dots are data from the WC NHLS cytology database and the black line is the best fitting Weibull distribution. This figure can be interpreted as follows: 8.8% of women have returned for rescreening within 3 months and 17.8% within 6 months.

A.5.4 Screening interval following Normal smear

According to South African cervical screening policy, women should be rescreened after 10 years following a Normal pap smear result (16) and an HIV-positive woman should be rescreened after 3 years (7). To investigate the implementation of these guidelines, we use the individual level data from the WC to estimate time to follow-up screen following a Normal result, by HIV and ART status. We use the subset of women screened in 2007/2008 to maximise follow-up time and minimise right censoring and fit Weibull distributions to the proportions of women who have not returned over time. We fit separate distributions according to HIV and ART status and found very similar distributions for HIV-negative/unknown status women and for HIV-positive, but ART-naïve women. For this reason, time to next screen is randomly drawn from the appropriate Weibull according to HIV/ART status as shown in Figure A 7.

A.5.5 Screening interval following LSIL smear

According to the policy, a woman who screens positive for a lower grade lesion should be rescreened after 1 year (16). If the follow-up screen is LSIL or worse, the woman will be referred to colposcopy. This guideline is the same for HIV-negative and positive women. To investigate the implementation of the guideline, we use the individual level data from the WC to estimate time to follow-up screen following a LSIL result.

We use the subset of women screened in 2007/2008 to maximise follow-up time and minimise right censoring and fit Weibull distributions to the proportions of women who have not returned over time. We fit separate distributions according to ART status (Figure A 12). The distributions were similar for HIV-negative/unknown and HIV-positive but ART-naïve, and therefore these data were grouped.

The average time between a LSIL and follow-up screen is 14.7 years for women who are ART-naïve and 5.7 years for women who are ART-experienced.

In the model, time to next screen is randomly drawn from the appropriate Weibull according to ART status.

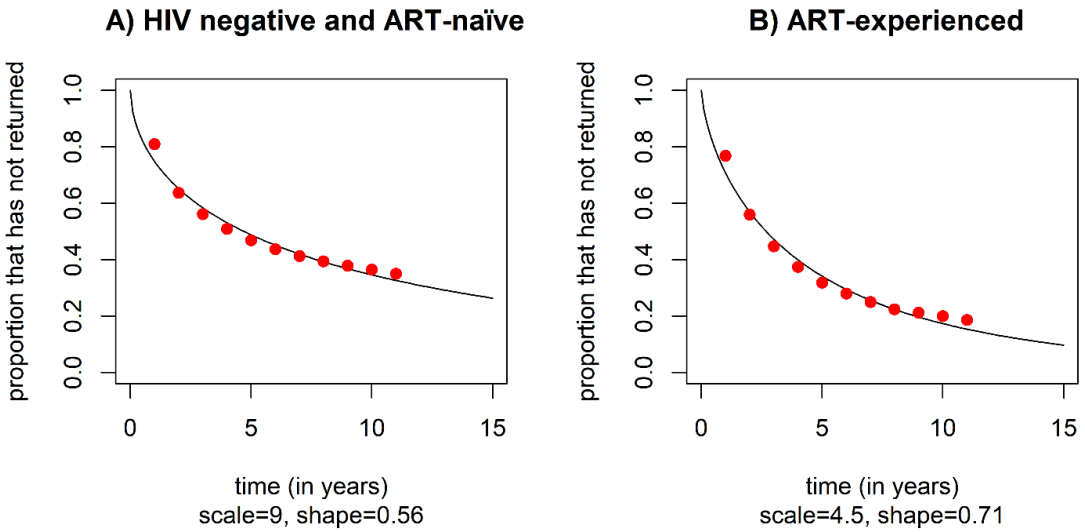


Figure A 12 - Distributions of time between a Pap smear with ASCUS/LSIL cytology and follow-up Pap smear. Red dots are data from the WC NHLS cytology database and black lines are best fitting Weibull distributions. These figures can be interpreted as follows: In A), around 19% of women unknown/negative HIV status or HIV-positive but ART-naïve have returned for the follow-up screen within a year and in B), around 23% of women who are ART-experienced have returned for the follow-up screen within a year.

A.5.6 Sensitivity and specificity of screening tests

A.5.6.1 Conventional cytology

Sensitivity and specificity of Pap smear depends on many factors. In one meta-analysis, the pooled sensitivity/specificity of conventional cytology to detect LSIL+/CIN1+ was 75.6%/81.2% (17). The nine studies included in this meta-analysis included women with recent history of abnormal smear which may result in higher sensitivity and false-positive rate. In another meta-analysis of 71 studies median sensitivity/specificity of LSIL+ cytology to detect CIN1+ was 69%/81% (18). When only considering studies in Nanda *et al.* (18) without verification bias that included only routinely screened non-pregnant women, sensitivity was much lower (weighted average 35%) and specificity much higher (weighted average 96%).

Three South African studies were identified that had sufficient data available to calculate sensitivity of Pap smear to detect each histological state (19–21). The overall sensitivity of a LSIL+ Pap smear to detect CIN1+ in these studies is higher than among the studies without verification bias in (18). These studies were performed among women in Cape Town, in communities with HIV prevalence of around 12% at the time (19–21). For the general screening population, regardless of HIV status we will use the probabilities in Table A 5 in our simulation.

Table A 4 - Sensitivity and Specificity of Pap smear in three South African studies (19–21).

Study	Sample size	Sensitivity	Specificity	Gold Standard
Denny	2922	66.9%	94.4%	Colposcopy for all women positive on either cytology, HPV, cervicogram or VIA (26%) and then colposcopy and biopsy/endocervical curettage
Wright	1352	49%	96.8%	Colposcopy for all women positive on either cytology, HPV, cervicogram or VIA (38.9%) and then colposcopy and biopsy/endocervical curettage
Taylor	2444	64.1%	95.8%	All participants received colposcopy and biopsy/endocervical curettage

Combining the numbers for all 3 studies produced the results in *Table A 5*.

Table A 5 - Pap smear accuracy derived from three South African studies (19–21).

	Histology result		
	Normal	**CIN1**	**CIN2+**
Cytology result			
Normal	6036 (95.4%)	96 (48.7%)	54 (27.8%)
LSIL	213 (3.4%)	69 (35%)	35 (18%)
HSIL+	78 (1.2%)	32 (16.2%)	105 (54.1%)
Total	6327	197	194

In the model we assume that when a woman who has no cervical disease receives a Pap smear, the result will have a 95.4% probability of being Normal, 3.4% probability of LSIL and 1.2% probability of HSIL. If a woman with CIN2+ is screened, she has 54.1% probability of being correctly diagnosed and referred to colposcopy in the simulation and 27.8% of being diagnosed with a normal cervix, and therefore not referred. Eighteen percent of women with CIN2+ will be diagnosed with LSIL and will only be referred if the screening event is a repeat screen. Only two of the studies showed numbers for cervical cancer diagnoses (19) (21). Combining the small numbers from these two studies, we estimate that 6/17 (35%) of HSIL+ Pap results will correctly diagnose asymptomatic cervical cancer.

A.5.6.2 HPV-DNA test

A recent meta-analysis estimated that the sensitivity/specificity of a HPV-DNA test to detect CIN2+ is 95%/92% (22). The majority of the studies included in the meta-analysis were performed in low HIV prevalence settings. Table A 6 shows values from South African studies that estimated diagnostic accuracy of HPV-DNA tests to detect CIN2+.

Table A 6 – Diagnostic accuracy of HPV-DNA testing as screening method.

	HIV-negative		**HIV-positive**			
Study	Sensitivity	Specificity	Sensitivity	Specificity	**Age**	**test**
Kitchener 2007 (23)			100%	33%	median 29	HC2
Firnhaber 2013 (24)			92%	51%	18-65	HC2
McDonald 2014 (25)	85%	81%	99%	52%	17-65	HC2

Segondy 2016 (26)			92%	61%	25-50	careHPV
Segondy 2016 (26)			98%	31%	25-50	INNO-LiPA
Kuhn 2020 (27)	89%	87%	94%	60%	30-65	GeneXpert

It is clear that HPV-DNA testing has low specificity for CIN2+ in HIV-positive women, since they have high levels of infection. We can see that specificity is lowest in the study where the median age of participants was lower than 30 and higher in the study where all the participants were older than 30.

We assume that an HPV-DNA screening test is 95% sensitive, but 100% specific to detect infection with any of the 13 HPV types (regardless of HPV disease status). Using these assumptions, the clinical sensitivity generated by the model for detecting CIN 2+ (among women aged 30-65) is 95%, and specificity is 81% for HIV-negative women and 56% for HIV-positive women.

A.5.6.3 HPV-DNA test with cytology triage

Due to the low specificity of HPV-DNA testing to detect CIN2+, especially in groups with high HIV prevalence, the use of HPV-DNA testing as a referral test will lead to large numbers of women unnecessarily treated, which has psychological and fertility implications. We will investigate the impact of HPV-DNA testing followed by a Pap smear for women testing positive, on both CC incidence and relative referral rates. For South Africa, we could not find data on the diagnostic accuracy of Pap smear when preceded by a positive HPV-DNA test. However, the WHO Guidelines Development Group (2020) performed meta-analyses of studies that estimated accuracy of this screening method. They found that sensitivity/specificity of the finding of any atypical cells (ASCUS+) on the triage Pap smear to predict CIN2+ was 72%/75% in populations with low HIV prevalence (28) and 92%/44% among HIV-positive women [personal communication: Helen Kelly]. We will use these values for HIV-negative and -positive women in our simulations.

A.5.7 Fraction who visits colposcopy clinics

In the Western Cape, large fractions of women who should receive colposcopy services have no record of doing so. The fractions of women who access colposcopy services within two years, by year and HIV/ART status is shown in Table A 7. In the model, a woman who needs colposcopy will have a probability of accessing this service based on the values in Table A 7 and if she does access the service, it will occur at six months post Pap smear. If she does not attend the colposcopy visit, a time to next routine screen will be drawn from the appropriate Weibull distribution in Figure A 7 and Section A.4.4. After 2017, percentages will stay constant at the 2017 values.

Table A 7 - Percentages of women who access colposcopy services within 2 years after indicative Pap smear.

	2007	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
HIV-negative or unknown	33.3	28.3	28.2	29	30.1	34.8	40.9	45.1	49.6	53.7	54.4
HIV-positive – ART-naïve	24	18.9	20.8	21.7	20.4	25.1	28.6	30.6	34.2	40	38.4
HIV-positive – ART-experienced	41.6	24.8	27.7	25.9	24.1	28.2	31	31.4	34.9	42	44.2

It is possible that these low but increasing levels of linkage could be partly explained by poor but improving adherence to electronic data capture requirements. However, health systems in the WC are generally more efficient than in the rest of the country and although these numbers may be an underestimate of linkage in the WC, it may be an overestimate for the rest of the country. The NCR provided us with a list of facilities in the country where cervical cancer was pathology diagnosed – i.e., an approximation of the number of facilities that could perform a colposcopy. Although the number of such facilities per adult woman varies widely across provinces, the number for the WC was close to the average.

A.5.8 Colposcopy visit

Although the first cervical cancer screening policy provided guidelines for referral to colposcopy clinics based on Pap results, the policy provided only two sentences on treatment (16): "If negative on colposcopy and cytology, the patient can be discharged. If positive, treat." The most common treatment method for cervical pre-cancer in South Africa is the LLETZ (large loop excision of the transformation zone). Based on conversations with gynaecologists and medical officers at different colposcopy clinics in South Africa, we conclude that the "look and LLETZ" (29) approach is followed in most facilities. This approach implies that a LLETZ will be performed if the colposcopist can see a CIN2 lesion or higher, *or* if the indicative Pap smear was HSIL. This means that even if the transformation zone cannot be visualised through colposcopy, but Pap smear was HSIL, the LLETZ will be performed. In some clinics, HIV-positive women are treated with LLETZ even if results were LSIL/CIN1 (29). In other clinics, all visualised CIN1 lesions will be treated with cryotherapy or electrode cauterisation. To simplify our model's algorithm, all women with CIN1+ on colposcopy will be treated.

Mitchel *et al.* (30) performed a meta-analysis of studies comparing detection of any abnormalities by colposcopy to the gold standard of colposcopy directed biopsy. Based on this gold standard, colposcopy has high sensitivity (96%), but only 48% specificity. Cantor *et al.* (31) shows that the accuracy of colposcopy directed biopsy versus the true gold standard of biopsy for all participants has high sensitivity of 88%, but that the specificity is again low at 57%. Using these point estimates, we derive sensitivity of colposcopy to detect CIN1+ compared to the true gold standard of biopsy as 91% and specificity of 29%.

In the model, 91% of women with CIN1+ will be diagnosed by colposcopy and 71% of women with healthy cervixes will be incorrectly diagnosed and treated. In addition, women who had a HSIL smear will also be treated. We assume that colposcopy has 100% sensitivity to detect cancer. Time to follow-up visit is drawn from Weibull distributions based on WC data in Figure A 13.

A.5.9 Success of treatment

In the model, women with high grade abnormalities will be treated with LLETZ. According to a Cochrane review published in 2010, this method has a 95% success rate, defined as having no persistent disease 6 months after receiving the treatment (32). A review of South African studies shows quite different treatment success rates, as shown in Table A 8.

Table A 8 - Success of LLETZ in South Africa

	HIV-negative		HIV-positive		
Study	**N**	**% success**	**N**	**% success**	**Definition of success**
Adam 2008 (29)	149	75.2%	266	30.8%	Normal at follow-up Pap smear, median 4 months later
Zeier 2012 (33)	335	73.1%	778	46.2%	Normal at follow-up Pap smear
Batra 2010 (34)	275	77.8%	219	47.9%	Normal Pap at 4 months post treatment
Noël 2015 (35)			259	41.3%	Normal at follow-up Pap smear
Smith 2017 (36)			83	19.7%	Normal/ASCUS Pap 6 months after treatment
Kabir 2012 (37)			571	35.0%	Normal at follow-up Pap smear
Weighted average		75.2%		40.0%	

In the model, a woman will receive treatment if diagnosed with high grade lesions on colposcopy and if HIV-negative has a 75.2% probability of moving to the Normal state. Based on data in Adam *et al.* (29), of the 24.8% whose treatment is not successful, half will move to CIN1 and the other half will remain in CIN2/3. If a woman who has CIN1 gets treated, half will move to Normal and the other half will remain in CIN1. We will assume that treatment with cryotherapy or electrode cauterisation has the same success rates. Around 15-18% of women do not clear the HPV infection after successful treatment (38–40). In the model, 15% of women with successful treatment will remain HPV infected, but without cervical disease and the other 85% will become susceptible to reinfection. After a colposcopy visit, it is recommended that women get a Pap smear again within 6 months. Using data from the WC, we draw the time to the next Pap from Weibull distributions shown in Figure A 13.

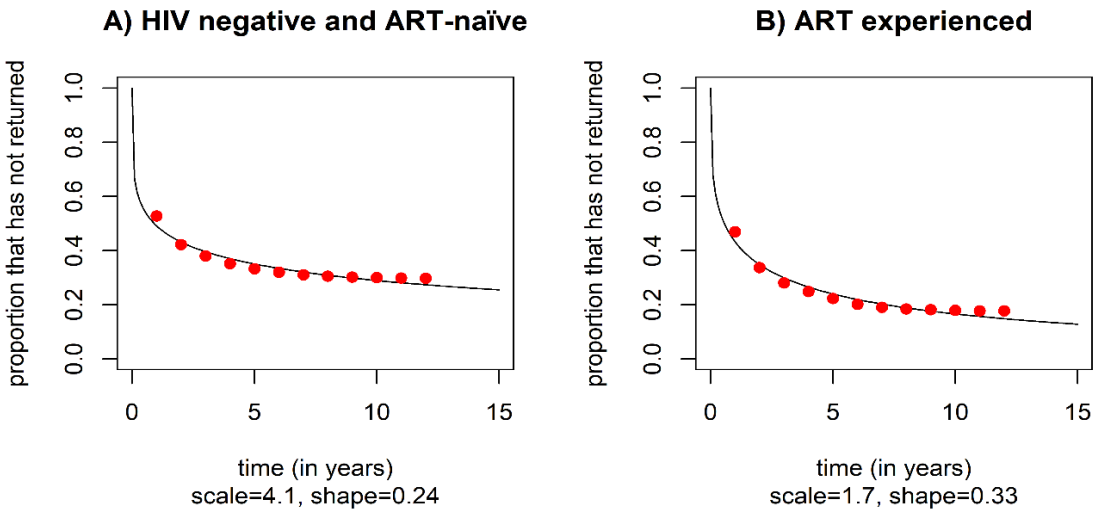


Figure A 13 - Distributions of time between colposcopy visit and follow-up Pap smear. Red dots are data from the WC NHLS cytology database and black lines are best fitting Weibull distributions.

6. Calibration method and data

The HPV infection, cervical pre-cancer and cancer components of the model were calibrated in three steps.

A.6.1 Calibration to HPV prevalence data

First, all stages of cervical disease were collapsed into an "HPV infected" stage and the main parameters that determine type-specific HPV prevalence were varied (more details in Section A.6).

The data in Table A 10 were compiled by performing a review of South African studies. All these studies differ in design, size, location and target population and no single study is nationally representative. Only studies that recruited participants from the general population, at ART clinics or at family planning clinics were included in the calibration process. Studies whose target populations were women who attended clinics (for reasons other than family planning or ART) and studies that only measured HPV seroprevalence were not included in the calibration (we do not simulate antibody response).

Since none of the studies included in the calibration were performed at a nationally representative level, results may be biased if the model is fit to the point estimates of prevalence from the studies, without taking heterogeneity between sites into account. Johnson *et al.* (41) have shown that an approach that considers inter-study variability in the definition of the likelihood function can produce reasonable nationally representative model estimates of STI prevalence. In this approach, the model estimate of prevalence in study i, given parameter combination φ expressed as $M_i(\varphi)$ relates to the true prevalence θ_i in the ith study population by

$$ln\left(\frac{\theta_i}{1-\theta_i}\right) = ln\left(\frac{M_i(\varphi)}{1-M_i(\varphi)}\right) + b_i$$

Where b_i is assumed to be a normally distributed random effect with zero mean and a variance of σ_b^2. After some rearranging, θ_i can be expressed as:

$$\theta_i = \left(1 + \left(\frac{1-M_i(\varphi)}{M_i(\varphi)}\right)e^{-b_i}\right)^{-1}$$

Using a third-order Taylor approximation around $b_i = 0$, the mean and variance of θ_i can be expressed as:

$$\mathrm{E}(\theta_i) = M_i(\varphi) + \sigma_b^2 M_i(\varphi)(1-M_i(\varphi))(0.5-M_i(\varphi))$$

$$Var(\theta_i) = M_i(\varphi)^2\left(1-M_i(\varphi)\right)^2[\sigma_b^2 + (1.5 - 8M_i(\varphi) + 8M_i(\varphi)^2)\sigma_b^4 + 15(1/6 - M_i(\varphi) + M_i(\varphi)^2)^2\sigma_b^6] \quad (1)$$

We assume that the number of positive cases (y_i) in study i of size n_i are binomially distributed and if θ_i was known, the likelihood function could be expressed as:

$$p(y_i|\,\theta_i, x_i) = \binom{n_i}{y_i}\theta_i^{y_i}(1-\theta_i)^{n_i-y_i},$$

where x_i are covariates such as time, study type and age range. The unknown true prevalence θ_i is assumed to be beta distributed:

$$p(\theta_i|\varphi^*, x_i) = \frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)}\theta_i^{\alpha_i-1}(1-\theta_i)^{\beta_i-1}$$

where φ^* is the combination of the model parameters and the variance of the random effect σ_b^2. The likelihood of observing the data is then:

$$
\begin{aligned}
p(y_i|\,\varphi^*, x_i) &= \int_0^1 p(y_i|\,\theta_i, x_i)\, p(\theta_i|\varphi^*, x_i) d\theta_i \\
&= \binom{n_i}{y_i} \int_0^1 \frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \theta_i^{\alpha_i+y_i-1}(1-\theta_i)^{\beta_i+n_i-y_i-1} d\theta_i \\
&\propto \frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \frac{\Gamma(\alpha_i+y_i)\Gamma(\beta_i+n_i-y_i)}{\Gamma(\alpha_i+\beta_i+n_i)} \qquad (2)
\end{aligned}
$$

Using the properties of the beta distribution, the mean and variance of θ_i are $\mathrm{E}(\theta_i) = \alpha_i/(\alpha_i+\beta_i)$ and $Var(\theta_i) = \alpha_i\beta_i/\big((\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)\big)$. By setting these equations equal to the equations in (1), the values for α_i and β_i in (2) can be derived and the likelihood calculated.

Parameter combinations from the prior distributions in Table A 12 are randomly sampled 500,000 times and model estimates of prevalence for each study $M_i(\varphi^*)$ are simulated for each parameter combination. The total likelihood for each of the 500,000 parameter combinations over all 20 studies in Table A 10 is then calculated by multiplying all 20 values of Equation (2). Each parameter combination is assigned a weight by taking the exponent of the difference between the log-likelihood value and the maximum of all the log-likelihood values, and normalising these values to sum to one. The parameter combinations are then resampled 500 times, with replacement, using these values as sampling weights. These parameter combinations are samples from the posterior distributions of each parameter (42).

This approach is only appropriate when using a test with perfect diagnostic accuracy. For HPV infection, we assumed that the DNA tests used in the studies have perfect sensitivity and specificity to detect infected cells from the sample tested when the individual is actively infected (but not when the individual is latently infected, since latent infection is by definition undetectable by current assays). The assumption of perfect sensitivity and specificity has been made in the absence of studies that have evaluated DNA tests relative to a superior gold standard.

A.6.2 Calibration to cervical pre-cancer data

For the second calibration step, we fix the HPV infection parameters for all 13 high-risk HPV types at the medians of the posterior distributions and vary the main parameters that determine cervical pre-cancer (described in Section A.6.2).

We use the prevalence data shown in Table A 9 and Table A 10 for this calibration step. We include type-specific HPV prevalence data in this step, since the progression parameters should result in type-specific prevalence in the model that is consistent with data. We define prevalence of cervical pre-cancer in three ways: 1) prevalence of abnormalities of any grade, 2) prevalence of high-grade lesions, given any abnormalities were found and 3) prevalence of high-grade lesions.

For cervical disease diagnosis, a biopsy is considered to be the gold standard. In routine and study settings, a biopsy is rarely performed on a woman who is HPV-DNA, cytology, and colposcopy negative. The studies in Table A 9 performed biopsies on all participants. Several routine and study data sources are available where cervical disease status was determined using a Pap smear. This diagnostic test has extremely variable levels of sensitivity and specificity (18,43) and for this reason we did not use cytological data in model calibration.

The likelihood for this calibration step is derived in the same way as in Section A.5.1. Parameter combinations (PCs) from the prior distributions are randomly sampled 100,000 times. Due to small numbers in the modelled population, each PC is used to run 3 simulations, and the results are aggregated across the 3 simulations to obtain a model estimate $M_i(\varphi^*)$ of prevalence for each study, for each PC. The total log-likelihood for each PC, over all the studies, is then calculated by summing all log-likelihood values of Equation 2. The 100 PCs that produced the highest total log-likelihoods were chosen as best fitting, and further analyses were performed using these 100 PCs.

Table A 9 - Prevalence of cervical pre-cancer

Study	HIV status	Year	Location	Outcome	Age range	N	Prevalence
McDonald (25)	positive	2000	Khayelitsha	CIN2+	17-30	512	9.6%
McDonald	positive	2000	Khayelitsha	CIN2+	30-40	582	10.3%
McDonald	positive	2000	Khayelitsha	CIN2+	40-65	277	6.5%
McDonald	negative	2000	Khayelitsha	CIN2+	17-25	884	2.5%
McDonald	negative	2000	Khayelitsha	CIN2+	25-30	662	2.1%
McDonald	negative	2000	Khayelitsha	CIN2+	30-35	666	3.6%
McDonald	negative	2000	Khayelitsha	CIN2+	35-40	2272	2.9%
McDonald	negative	2000	Khayelitsha	CIN2+	40-45	1400	3.1%
McDonald	negative	2000	Khayelitsha	CIN2+	45-50	982	3.1%
McDonald	negative	2000	Khayelitsha	CIN2+	50-55	617	2.1%
McDonald	negative	2000	Khayelitsha	CIN2+	55-65	567	1.4%
Cronje (44)	not tested	2001	Free State	CIN1+	21-65	1093	34.9%
Cronje	not tested	2001	Free State	CIN2+|CIN1+*	21-65	382	23.6%
Denny (21)	not tested	1996	Khayelitsha	CIN1+	35-65	2922	6.1%
Denny	not tested	1996	Khayelitsha	CIN2+|CIN1+	35-65	178	46.6%
Kuhn (27)	negative	2015	Cape Town	CIN1+	30-65	378	12.7%
Kuhn	negative	2015	Cape Town	CIN2+|CIN1+	30-65	48	41.7%
Kuhn	not on ART	2015	Cape Town	CIN1+	30-65	67	37.3%
Kuhn	not on ART	2015	Cape Town	CIN2+|CIN1+	30-65	25	56.0%
Kuhn	on ART	2015	Cape Town	CIN1+	30-65	263	28.9%
Kuhn	on ART	2015	Cape Town	CIN2+|CIN1+	30-65	76	55.3%

*CIN2 or worse given any abnormality (CIN1+)

Table A 10 – Type-specific HPV prevalence

			Date		Number	Prevalence												
i	**Study**	**HIV status**	t_i	**Location**	n_i	**16**	**18**	**31**	**33**	**35**	**39**	**45**	**51**	**52**	**56**	**58**	**59**	**68**
	General population data - females																	
1	McDonald (25)	negative	2000*	Khayelitsha	8050	2.7%	1.5%	1.3%	1.3%	2.9%	0.8%	1.8%	1.3%	1.6%	0.9%	1.9%	1.0%	1.3%
2	McDonald (25)	positive	2000*	Khayelitsha	1371	8.2%	6.2%	4.1%	4.3%	8.5%	3.7%	5.7%	5.1%	5.4%	3.7%	7.9%	3.3%	6.2%
3	Giuliano (45)	negative	2012	Kraaifontein	391	14.1%	6.4%	4.9%	1.8%	9.0%	4.4%	6.1%	8.7%	11.3%	3.3%	10.0%	7.2%	6.4%
4	Snyman (46)	not tested	2011	Tshwane	253	5.7%	4.9%											
5	Snyman (47)	not tested	2012	Tshwane	160	4.4%	5.7%											
6	Adler (48)	negative	2013	Masiphumelele	50	6.0%	4.0%	0.0%	0.0%	2.0%	4.0%	0.0%	6.0%	4.0%	0.0%	2.0%	2.0%	6.0%
7	Adler (48)	positive	2013	Masiphumelele	35	20.0%	14.3%	2.9%	2.9%	14.3%	5.7%	25.7%	8.6%	11.4%	5.7%	2.9%	2.9%	20.0%
8	Mbulawa (49)	negative	2014	Masiphumelele	148	10.8%	6.8%	6.1%	1.4%	6.1%	4.1%	6.8%	10.8%	6.1%	4.1%	13.5%	6.1%	8.1%
9	Mbulawa (49)	negative	2014	Soweto	143	12.6%	8.4%	2.1%	1.4%	7.7%	4.9%	1.4%	7.0%	6.3%	1.4%	7.0%	7.0%	6.3%
10	Mbulawa (50)	positive	2006	Gugulethu	277	11.2%	8.7%	4.0%	4.3%	7.6%	5.1%	9.7%	7.2%	11.2%	4.7%	10.5%	5.1%	7.2%
11	Mbulawa (50)	negative	2006	Gugulethu	207	3.4%	2.4%	2.9%	1.4%	4.8%	2.9%	0.5%	1.4%	3.9%	0.5%	4.3%	2.4%	2.4%
12	Denny (51)	positive	2002	Cape Town	311	16.3%	8.9%	5.6%	5.6%	11.5%	6.7%	7.0%	7.8%	13.3%	8.1%	10.0%	8.9%	8.5%
13	Liebenberg (52)	negative	2007	KZN	779	10.8%	7.1%	6.2%	6.8%	9.4%	5.4%	5.5%	9.8%	6.0%	3.0%	9.0%	6.2%	5.3%
	General population data - males																	
14	Vardas (53)	negative	2005	Soweto	538	4.4%	4.4%	1.4%	1.6%	3.1%	2.1%	2.9%	4.3%	5.2%	4.1%	3.5%	3.3%	
15	Mbulawa (50)	positive	2006	Gugulethu	277	13.3%	7.0%	3.8%	3.2%	9.5%	7.0%	15.2%	9.5%	7.6%	2.5%	10.1%	11.4%	9.5%
16	Mbulawa (50)	negative	2006	Gugulethu	207	5.8%	3.8%	1.9%	1.3%	1.9%	3.5%	3.5%	5.4%	5.1%	1.3%	2.9%	4.2%	4.8%
17	Chikandiwa (54)	positive	2015	Johannesburg	283	13.0%	7.0%	2.0%	5.0%	13.0%	5.0%	7.0%	10.0%	7.0%	5.0%	7.0%	9.0%	8.0%
	ART clinics (initiating ART)																	
18	Moodley (55)	positive	2007	Cape Town	109	13.8%	15.6%	5.5%	8.3%	4.6%	10.1%	15.6%	12.8%	9.2%	5.5%	17.4%	7.3%	11.0%
19	Firnhaber (56)	positive	2009	Johannesburg	147	29.9%	18.4%	7.5%	8.2%	19.7%	8.8%	16.3%	13.6%	13.6%	15.0%	9.5%	10.9%	8.2%
	Family planning clinics																	
20	Mbulawa (57)	not tested	2015	5 provinces	330	7.0%	6.1%	2.1%	1.2%	4.8%	6.7%	7.6%	6.7%	3.0%	3.0%	7.6%	4.8%	3.0%

A.6.3 Calibration to cervical cancer data

For the third calibration step, we fix the HPV infection parameters for all 13 high risk HPV types and the cervical pre-cancer parameters at the medians of the posterior distributions obtained in the first two steps and vary the parameters that drive cervical cancer incidence and diagnosis (described in Section A.6). For this step, we use two sources of data: 1) Pathology confirmed cervical cancer incidence as reported by the National Cancer Registry, and 2) studies on the proportion of cervical cancer diagnoses in different stages.

A.6.3.1 Diagnosed cervical cancer incidence

Public and private laboratories in South Africa report cytology and histology confirmed cervical cancer cases to the National Cancer registry. Data are cleaned to remove duplicates, and a woman is only counted at first diagnosis of cervical cancer. Although this a nationally representative data source, the data suffer from the limitation that women who only receive a clinical diagnosis (no pathology), or never receive a diagnosis, will not be included in this estimate of cervical cancer incidence. In the Eastern Cape (EC) population-based cancer registry, which includes all cancers diagnosed in health care settings, 14% of diagnosed cervical cancer cases in the 2008-2012 period were only diagnosed clinically (58). This registry covers a rural area with a population of around 1 million people. In the Ekurhuleni population-based cancer registry, 7% of CC cases were only clinically diagnosed in 2018 (59). This registry covers an urban area with a population of around 3 million people.

We calibrate the model to overall crude and age-specific NCR incidence data between 2000 and 2016. Although estimates have been published since 1994, numbers of cases per 5-year age group were only published since 2000. We will fit our model under different assumptions about under-reporting: assuming that on average 7%, 10% or 14% of diagnosed cases are not captured in the pathology-based NCR, or that under-reporting decreases linearly from 25% in 2000 to 7% in 2018 (combining estimates from EC (58) and Ekurhuleni (59)).

Since cervical cancer is a rare disease, with age-standardised incidence rates around 30 per 100,000 women per year, model results are severely affected by stochasticity. To get around this problem, and stay within reasonable computing time, we follow this approach:

1) We run 50 simulations for each of 30,000 parameter combinations and aggregate results across the 50 simulations to have around 750,000 adult women in the aggregated model population in 2016.
2) We smooth the time series of incidence estimates by fitting a Poisson regression model to the model results.

$$log(CC_{ij}) = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \beta_{3j} age_{ij} + \beta_{4j} age_{ij} t_i + \beta_{5j} age_{ij} t_i^2$$

   For each of the 30,000 parameter combinations, CC_{ij} is the model estimate of diagnosed CC incidence per 100,000 women aged j (age groups 20-24 to 70-75) at time i (2000 to 2016). An example of the model estimates (black dots), regression estimates (dashed line) and data (red line) for one parameter combination is shown in Figure A 14.
3) We then calculate likelihood values by assuming that the difference between the log-transformed Poisson regression estimate and the log-transformed observed incidence is normally distributed.

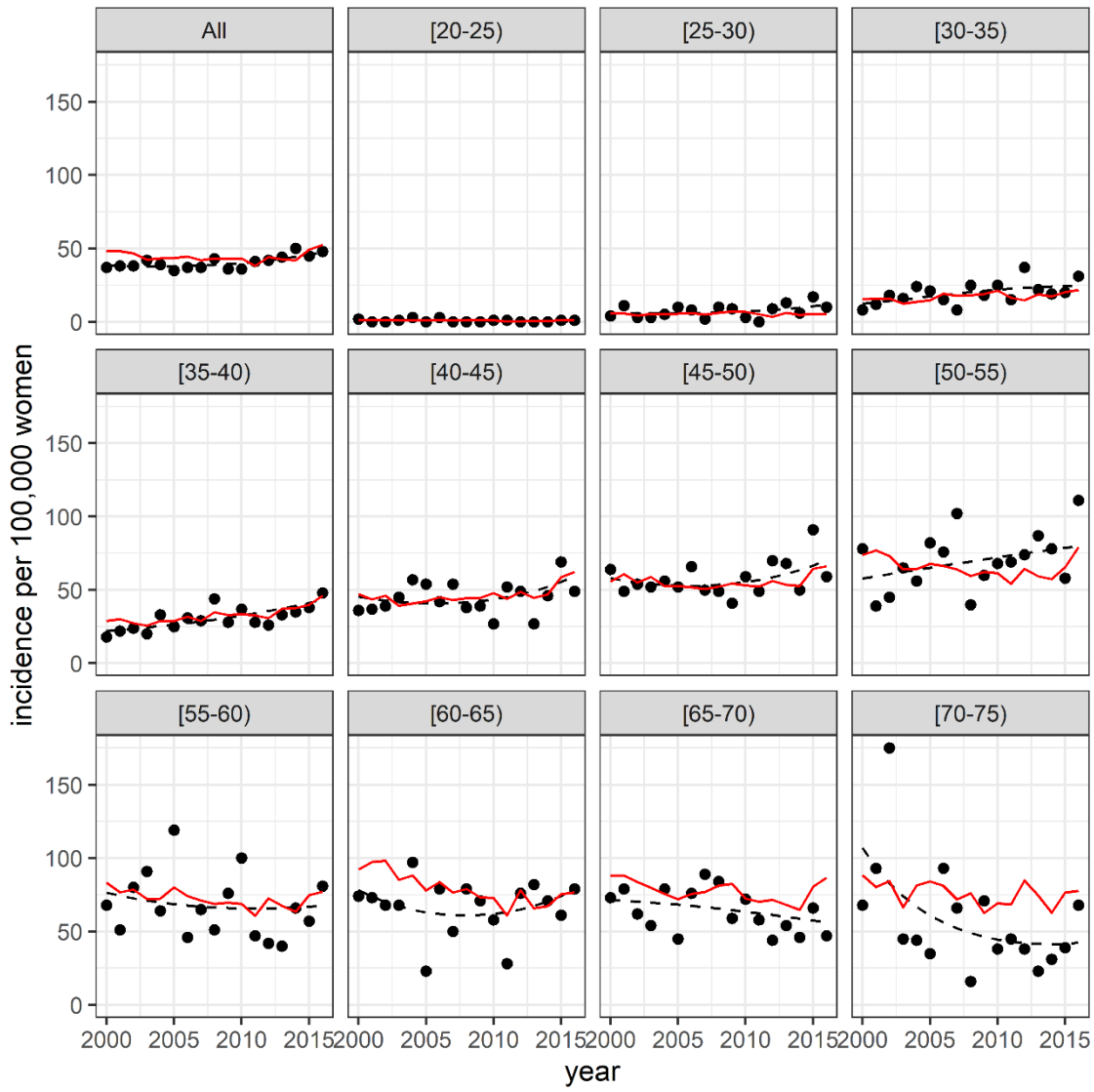


Figure A 14 - Cervical cancer incidence per 100,000 women as estimated by one parameter combination (black dots). The dashed line shows the fit of a Poisson regression through the black dots and the red line shows the crude estimate calculated using NCR data, adjusted by assuming that 10% of diagnosed CC only receives a clinical diagnosis (no pathology).

A.6.3.2 Stage at diagnosis

Cervical cancer is diagnosed as one of four stages of severity, according to the International Federation of Gynecology and Obstetrics (FIGO) classification (60). The disease progression model we are assuming is shown in Figure A 16 in Section A.6.2.7. In order to inform diagnosis rates, we performed a review of South African reports on the stages of CC diagnosis, shown in Table A 11. In all of these studies, the disease stage at the time of treatment initiation is shown, and we make the simplifying assumption that the stage of diagnosis is the same as the stage at which cervical cancer treatment is started. In Groote Schuur Hospital (GSH), the proportion of women treated in each stage has stayed fairly stable over the period 1984-2013, with slight increases in early-stage diagnosis and decreases in late-stage diagnosis. It is notable that a larger fraction of women is diagnosed in early stages of disease than in other hospitals, likely due to higher levels of screening in the Western Cape.

Table A 11 - The proportion of women diagnosed in each stage of cervical cancer

	Year	**N**	**Stage I**	**Stage II**	**Stage III**	**Stage IV**	**Sample**
GSH	2000-2004	741	18.5%	23.5%	44.5%	13.5%	All women treated in GSH in 2000-04
GSH	2005-2009	839	19.2%	19.3%	48.6%	12.9%	All women treated in GSH in 2005-09
GSH	2010-2013	747	24.8%	24.2%	39.2%	11.8%	All women treated in GSH in 2010-13
Lomalisa (61)	2000	836	8.4%	35.6%	41.0%	15.0%	All women treated in Johannesburg hospital 1997-8
Mbodi (62)	2013	104	8.4%	22.1%	59.7%	9.6%	Sample of women treated in Chris Hani Baragwanath hospital 2013
Snyman (63)	2011	85	7.0%	17.6%	58.9%	16.5%	Sample of women treated in Kalafong hospital in 2011
Sabulei (64)	2015	153	5.2%	44.4%	49.7%	0.7%	All women treated in an academic hospital in Gauteng in 2017

We calculate the likelihood of observing the proportions of women in each stage found in the model, given the data in Table A 11, by using a Dirichlet-multinomial distribution. For each parameter combination $\boldsymbol{\varphi}$, the model produces the fraction of women diagnosed in each stage which we will express as $\pi_i(t_j, \boldsymbol{\varphi})$, where i represents one of four stages and t_j is the year of the j^{th} study. The number of women diagnosed in each stage i in the j^{th} study in Table A 11 can be expressed as n_{ij}, and we can assume that these values are multinomially distributed with likelihood function

$$\binom{n_j}{n_{1j}\ n_{2j}\ n_{3j}\ n_{4j}} \pi_1(t_j, \boldsymbol{\varphi})^{n_{1j}} \pi_2(t_j, \boldsymbol{\varphi})^{n_{2j}} \pi_3(t_j, \boldsymbol{\varphi})^{n_{3j}} \pi_4(t_j, \boldsymbol{\varphi})^{n_{4j}}$$

where n_j is the total number of CC cases in study j. This model assumes that the fractions of women diagnosed in each stage is the same everywhere in South Africa. We know this is not the case, due to differences in knowledge, and access to healthcare and screening. To account for possible variation in

fractions diagnosed in the four stages in different settings, we define the true fraction of individuals in stage i from population j as ρ_{ij} and assume that these terms are Dirichlet distributed:

$$p(\boldsymbol{\rho}_j|\boldsymbol{\pi}(t_j,\boldsymbol{\varphi}),\theta) = \Gamma(\theta)\prod_{i=1}^{4}\left(\rho_{ij}^{\theta\pi_i(t_j,\boldsymbol{\varphi})-1}\right)/\Gamma\left(\theta\pi_i(t_j,\boldsymbol{\varphi})\right)$$

From the properties of the Dirichlet distribution, we know that

$$E[\rho_{ij}] = \pi_i(t_j,\boldsymbol{\varphi})$$

$$Var[\rho_{ij}] = \frac{\pi_i(t_j,\boldsymbol{\varphi})\left(1-\pi_i(t_j,\boldsymbol{\varphi})\right)}{\theta+1}$$

and therefore the θ parameter controls the variability in proportions diagnosed in each stage, between studies. We estimate this parameter by making the simplifying assumption that the proportions are constant over time and using the data to calculate maximum likelihood estimates of θ and the constant proportions. We use this maximum likelihood estimate ($\hat{\theta} = 30.0$) as a fixed value in the likelihood function:

$$p(\boldsymbol{n}_j|\boldsymbol{\pi}(t_j,\varphi),\theta) = \int_{\boldsymbol{\rho}_j} p(\boldsymbol{n}_j|\boldsymbol{\rho}_j)p(\boldsymbol{\rho}_j|\boldsymbol{\pi}(t_j,\boldsymbol{\varphi}),\theta)d\boldsymbol{\rho}_j,$$

a Dirichlet-multinomial likelihood function which reduces to (65):

$$p(\boldsymbol{n}_j|\boldsymbol{\pi}(t_j,\varphi),\theta) = \frac{\Gamma(\theta)}{\Gamma(n_j+\theta)}\binom{n_j}{n_{1j}\ n_{2j}\ n_{3j}\ n_{4j}}\prod_{i=1}^{4}\frac{\Gamma(n_{ij}+\theta\boldsymbol{\pi}(t_j,\boldsymbol{\varphi}))}{\Gamma(\theta\boldsymbol{\pi}(t_j,\boldsymbol{\varphi}))}$$

In order to find the parameter combinations $\boldsymbol{\varphi}$ that maximises this likelihood function, we can maximise the log-likelihood and leave out the terms that are independent of $\boldsymbol{\varphi}$. Summing over all studies to obtain the total log-likelihood, we therefore aim to maximise:

$$\sum_{j=1}^{7}\sum_{i=1}^{4}\left\{\log\left(\Gamma\left(n_{ij}+\theta\boldsymbol{\pi}(t_j,\boldsymbol{\varphi})\right)\right)-\log\left(\Gamma\left(\theta\boldsymbol{\pi}(t_j,\boldsymbol{\varphi})\right)\right)\right\}$$

Finally, for each parameter combination (PC), the log-likelihood value from the stage-at-diagnosis data is added to the log-likelihood value from the diagnosed incidence data, and we choose the 100 PCs with the highest total log-likelihoods as a sample from the posterior distributions of the parameters. We then run simulations for each of these 100 PCs and aggregate results to show fits to the data and validation data in Section A.8.

7. Prior distributions of parameters

A.7.1 HPV infection parameters

Prior distributions assigned to the HPV natural history model parameters are given in Table A 12. Each parameter is discussed in Sections A.6.1.1 to 3. Except in the case of the HPV infection duration parameter, the prior distribution assigned to each parameter is the same for all HPV types. However, parameter combinations are sampled from these prior distributions for each HPV type separately.

Table A 12 - Prior distributions of the HPV parameters in the model

		HPV 16	**HPV 18**	**Other types**
		Uniform prior distributions		
Section	**Parameter**	**range**	**range**	**range**
A.6.1.1	Transmission probability (per sex act)			
	Male to Female	0 - 1	0 - 1	0 - 1
	Female to Male	0 - 1	0 - 1	0 - 1
A.6.1.2	Relative HPV duration in HIV infection (66,67)			
	Latent HIV vs HIV-negative	1 - 2	1 - 2	1 - 2
	Acute HIV/late HIV/recent ART* vs HIV-negative	1 - 3	1 - 3	1 - 3
A.6.1.3	Time to reactivation (in years) if HIV-negative			
	Males	0 - 30	0 - 30	0 - 30
	Females	0 - 30	0 - 30	0 - 30
A.6.1.3	Proportion who become latently infected after clearance	0 - 1	0 - 1	0 - 1
A.6.1.3	Relative HPV reactivation rate (66,68)			
	Latent HIV vs HIV-negative	1 - 5	1 - 5	1 - 5
	Acute HIV/late HIV/recent ART* vs latent HIV	1 - 3	1 - 3	1 - 3
A.6.1.3	Duration of immunity (in years)			
	Males	0 - 30	0 - 30	0 - 30
	Females	0 - 30	0 - 30	0 - 30
		Gamma prior distributions		
A.6.1.2	Duration of HPV infection (in months) if HIV-negative	**mean (sd)**	**mean (sd)**	**mean (sd)**
	Males	18 (9)	9 (9)	9 (9)
	Females	18 (9)	18 (9)	18 (9)
	Standard deviation of study effect	0.55 (0.24)	0.55 (0.24)	0.55 (0.24)

*Recent ART is defined as ART initiation within last 2 years. People who have been on ART for longer than two years are assumed to be the same as HIV-negative people.

A.7.1.1 Transmission probabilities

Susceptible individuals in the model acquire HPV from infected sexual partners through a per sex act transmission probability. This value has not been empirically estimated and estimates from other modelling studies vary widely, for example for HPV-16 the value ranges from 6.9% per sexual contact (both sexes) in Matthijsse *et al.* (STDSIM) (69) to almost 100% for male to female transmission in Van de Velde *et al.* (HPV-ADVISE) (70). For these reasons, uniform prior distributions ranging between zero and one were assumed. HIV status does not influence per sex act HPV transmission probabilities and vice versa.

A.7.1.2 Duration of infection

Prior distributions for the duration of infection are based on a review of studies that estimated the type specific median duration of incident HPV infections in females (Table A 13). Studies that used baseline prevalent HPV infections in the calculation of duration were excluded. In the model, infections clear at a constant rate, therefore the median durations are assumed to be generated from exponential distributions. The corresponding means of type 16 and 18 infection duration were calculated and an overall mean, weighted according to overall sample size, was calculated. The mean of an exponential distribution is 1/ln(2) (roughly 1.5) times the value of the median. The standard deviation of the prior distribution was assumed large enough to include the upper and lower 95% confidence limits for each study. The median durations of the eleven other high risk (HR) HPV types were assumed to be one year and the confidence limits include the confidence limits of all the study estimates. This assumption was made due to low prevalence of other types in the studies - sample sizes to calculate the type-specific durations were small and confidence intervals were very wide.

Table A 13 - Studies included in estimating the average HPV infection duration in women.

Paper	Duration of study (years)	Interval between visits (months)	Sample size	Age	Definition of clearance	Median duration (months) (95% CI)	
						HPV-16	HPV-18
Trottier (71)	5	4 - 6	2500	all ages	One negative	7.3 (6.3-10.7)	6.9 (6-12)
Richardson (72)	2	6	621	university	One negative	19.4 (11.4-27.5)	9.4 (4.8-14)
Ho (73)	3	6	608	university	One negative	11 (7-12)	12 (6-17)
Woodman (74)	4	6	1075	15-19	One negative	10.3 (6.8-17.3)	7.8 (6-12.6)
Goodman (75)	5	4	972	all ages	One negative	9.7 (4.5-24.2)	14.3 (4.9-)
Munoz (76)	2.5	6	1610	13-85	One negative	13.7 (8.4-18.8)	11.9 (9.1-16.6)
Insinga (77)	4	6	1203	16-23	Two negatives	13.2 (12-17.6)	13.2 (11.7-17.7)
Insinga (78)	4	6	1788	16-23	Two negatives	17.1 (15.1-20.2)	12.4 (11-17.7)
Jaisamrarn (79)	2	6	4825	15-25	Two negatives	17.1 (7.8-30.3)	11.8 (6.2-23.1)
Weighted average of corresponding means						18.9	15.5

Only one study estimated duration of type specific HPV infection in men (80). The median duration estimates for types 16 (12.2 months (95% CI 7.4-20.2)) and 18 (6.3 months (6.0-12.7)) were used to calculate prior means using the assumption of exponentially distributed durations, i.e. means=medians/ln(2). A median of half a year for the other oncogenic HPV types (9 of the 11 other types had median duration of ~6 months) was assumed, or roughly a mean of 9 months. The standard deviations of the prior distributions were the same as for the female prior distributions.

In the model, durations of HPV infections depend on the stage of an individual's HIV infection. Individuals in the latent stage of HIV clear HPV infections at a lower rate than HIV negative individuals (prior range for relative duration: 1 to 2), and those in the acute phase, late phases (pre-AIDS or AIDS) or on ART for less than two years could clear HPV at a lower rate (prior range, relative to HIV negative: 1 to 3) (66,67). People who have been on ART for longer than 2 years are assumed to clear HPV infection at the same rate as HIV-negative people. The average duration of HPV is assumed to be the same for all age groups, since results from studies estimating age differences in durations are inconsistent (71,73,75,76,81).

A.7.1.3 Natural immunity to re-infection and reactivation of latent infections

Individuals in the latent stage of HIV are assumed to reactivate HPV infections at a higher rate than HIV negative individuals (prior range for relative reactivation rate: 1 to 2), and those in the acute phase, late phases (pre-AIDS or AIDS) or on ART for less than two years clear HPV at a rate that is a multiple of the rate in the latent phase, with this multiple being between 1 and 2 (66,68). People who have been on ART for longer than 2 years are assumed to reactivate HPV infections at the same rate as HIV-negative people.

Uniform prior distributions are assigned to represent the uncertainty around the average durations of immunity and viral latency. Since no data on these durations exist, the range was chosen to be between zero and thirty years. Prior distributions for males and females are the same, but parameters for males and females are sampled separately.

It is uncertain whether reactivated infections have sufficient viral loads to contribute to transmission and lead to persisting infections. We assume that reactivated infections are as infectious and persistent as new infections.

A.7.2 Cervical pre-cancer and cancer parameters

Prior distributions assigned to the parameters of the cervical disease components of the model are given in *Table A 1*Table A14. Each parameter is discussed in Sections A.6.2.1 to 6.2.7. During cervical pre-cancer calibration, the HPV infection parameters were kept fixed at their posterior medians, and the cervical cancer parameters were kept fixed at the prior means. During the cervical cancer calibration, the HPV infection and cervical pre-cancer parameters were kept fixed at the posterior medians. Other parameters that have fixed values in our model are the proportion that clears HPV infection during CIN1 regression (discussed in Section 6.4), parameters of progression through cancer stages and cancer mortality by stage of diagnosis (discussed in Section 6.7).

Table A14 - Prior distributions of cervical pre-cancer and cancer parameters. Means (standard deviations) are shown for beta, normal and gamma distributions and ranges are shown for uniform distributions. In some cases, the parameter for HPV-16 is gamma/beta distributed and the parameters for HPV-18/other HR are uniformly distributed multipliers of the HPV-16 parameter (indicated with an asterisk).

Section	Cervical pre-cancer parameters	Distribution	Type 16	Type 18	Other HR-HPV
A.6.2.1	Multiplier for duration of HPV among women	Uniform	0.2 - 1	0.2 - 1	0.2 - 1
A.6.2.2	Proportion that will progress from HPV infected to CIN1	Beta	0.26 (0.1)	0.14 (0.1)	0.5 – 1*
A.6.2.3	Annual progression rate from CIN1 to CIN2	Gamma	0.09 (0.05)	0 – 1*	0 – 1*
A.6.2.3	Annual regression rate from CIN1 to Normal	Gamma	0.43 (0.2)	1 – 2*	1 – 2*
A.6.2.3	Annual regression rate from CIN2 to CIN1 (aged <30)	Gamma	0.458 (0.04)	1 – 2*	1 – 2*
A.6.2.3	Regression multiplier for women aged 30 or older	Uniform	0.45 – 0.75		
A.6.2.4	HIV multiplier for CIN1 progression	Uniform	2 – 5.32		
A.6.2.4	ART multiplier for HIV progression multiplier	Uniform	0.55 – 0.9		
A.6.2.4	HIV multiplier for CIN1/2 regression	Uniform	0.56 – 0.82		
A.6.2.4	ART multiplier for HIV multiplier for CIN1/2 regression	Uniform	1.3 – 2		
	Cervical cancer parameters				
A.6.2.3	Annual progression rate from CIN2 to CIN3 (aged <30)	Gamma	0.058 (0.03)	0 – 1*	0 – 1*
A.6.2.3	Multiplier for women aged 30 or older (a)	Uniform	2 – 3		
A.6.2.3	Multiplier for (a) for women aged 50 or older	Uniform	1 – 2		
A.6.2.4	HIV: multiplier for CIN2 progression	Uniform	1.1 – 1.6		
A.6.2.6	CIN3 duration: scale (years)	Uniform	5 – 20		
A.6.2.6	CIN3 duration: shape	Uniform	1.5 – 3		
A.6.2.7	Annual probability of getting diagnosed in Stage I **	Uniform	0 – 0.03		
A.6.2.7	Annual probability of getting diagnosed in Stage II	Uniform	0 – 0.2		
A.6.2.7	Annual probability of getting diagnosed in Stage III	Uniform	0.4 – 0.8		
A.6.2.7	Annual probability of getting diagnosed in Stage IV	Uniform	0.7 – 1.0		

*Multiplier for HPV-16

** In a process separate from routine screening.

A.7.2.1 Multiplier for duration of HPV among women

In the first stage of model calibration, the duration of type-specific HPV was estimated by calibrating to type-specific HPV prevalence data and "HPV prevalent" included all cervical disease stages. In the second stage of calibration, we need to estimate the duration of disease-free HPV infection for women. We will use the median value of type-specific HPV duration estimated in the first stage of calibration, multiply this value with a fraction to estimate the disease-free type-specific infection duration. We set the Uniform prior distribution of this multiplier between 0.2 and 1.

A.7.2.2 Proportion moving from HPV infected to CIN1

To estimate these parameters, we use data from two studies performed by Insinga *et al.* in the United States in 2007 and 2011 (77,82). Participants in the 2007 study were 2400 women in the placebo arm of an HPV-16 vaccine trial, and participants in the 2011 study were 1800 women in the placebo arm of a quadrivalent vaccine trial. These studies showed the proportions of those HPV positive at baseline that cleared, persisted, or progressed at 12, 24 and 36 months. We grouped all three stages CIN stages as "disease" and therefore have three health states: Cleared, persistent infection and diseased. We fitted exponential distributions to the 3 data points of the fractions in cleared and diseased to estimate the rate of clearance and rate of progression and use these rates to estimate the proportion that ever progress (Table A 15). By using the lower and upper estimates of the 95% confidence intervals in the same fitting procedure, we estimate lower and upper estimates of the proportion that progress. To use this method, we assume that women only clear, persist or progress during the three-year period, and do not for example fluctuate between infected and diseased or fluctuate between uninfected and infected.

The two studies found different, but overlapping proportions of HPV-16 progression, and we will set the mean of our beta prior distribution to the weighted average of the two studies (26%), with enough uncertainty to include the lower and upper estimates of both studies. The mean of our beta prior distribution for HPV-18 progression will be set to 14%, with standard deviation large enough to include the lower and upper estimates. The proportion of people progressing from infection with other HPV types to CIN1 will be expressed as a fraction of the proportion of HPV-16 infections who progress. Other studies that estimated the proportion of women who progressed during follow-up (due to any HPV types) found mean values ranging from 11-25% (73,79,83,84).

Table A 15 - Proportions that progress from HPV to CIN1

Type	N	Clearance rate	Progression rate	Total rate	Proportion progress	Lower estimate	Upper estimate
Insinga 2011							
16	273	0.38	0.11	0.50	23%	11%	38%
18	113	0.55	0.09	0.64	14%	4%	31%
31	157	0.38	0.11	0.50	23%	7%	44%
33	57	0.55	0.06	0.61	10%	1%	34%
35	52	0.63	0.08	0.71	11%	1%	39%
45	77	0.61	0.05	0.66	8%	1%	23%
52	173	0.44	0.06	0.50	13%	3%	27%
58	109	0.40	0.08	0.49	17%	3%	43%
59	172	0.87	0.03	0.90	4%	1%	9%
Insinga 2007							
16	142	0.37	0.17	0.54	32%	12%	55%
18	62	0.48	0.08	0.56	14%	1%	42%

A.7.2.3 Progression and regression from CIN1 and CIN2

To our knowledge, no South African study has estimated these values. We rely on meta-analyses of studies to estimate these rates (85,86). The studies included in the meta-analyses typically enrolled women who, at baseline, had either CIN1 or CIN2. The women were followed over time, and the fractions who have regressed, persisted, or progressed are reported at different time points. These values are not cumulative fractions, and may be biased because women may move in and out of states multiple times. In addition, left-censoring at baseline and interval-censoring at each subsequent visit may bias estimates. We use the fractions (p) who progressed/regressed at 24 months and assume that these processes happen at constant yearly rates (r). We approximate the rate using the formula

$$p = 1 - \exp(-2 * r).$$

In Liu *et al.* (86), studies of different duration were included in the meta-analyses, and we did an analysis using only the studies with 24 month follow-up. Overall, 16.4% of women progressed from CIN1 to CIN2 after 2 years, and 57.4% regressed from CIN1. This translates to rates of 0.09 and 0.427 per year, respectively. Uncertainty was estimated using values from the individual studies.

Tainio *et al.* performed a meta-analysis of studies on regression and progression from CIN2 (85). In this study, they showed that there are differences in rates by age. Although there were no studies in the meta-analysis that *only* included women older than 30, the authors performed the analysis by first only using studies where all the participants were younger than 30, and then including studies where participants could be older than 30 (proportions p in

Table A *16*). We used the estimates of rates for women younger than 30 to inform our priors for progression and regression of CIN2. The rates in the two analyses in

Table A *16* are not independent, but for the prior of the multiplier of regression among women older than 30 we are guided by the ratio 0.29/0.458=0.63 and the ratios of the intervals. For the prior of the

multiplier for progression among women aged 30 and older we were again guided by the ratio of 0.131/0.058=2.3. Rates of progression of CIN2 in women older than 30 were further disaggregated in our model into rates for women aged 30-50 and women older than 50 (87–90).

Table A 16 – Regression and progression of CIN2 by 24 months from Tainio et al. (85)

	Regression of CIN2 to CIN1 by 24 months		Progression of CIN2 to CIN3 by 24 months	
	p	r	p	r
Only studies where max age <=30	0.6 (0.57-0.63)	0.458 (0.422-0.497)	0.11 (0.05-0.19)	0.058 (0.026-0.105)
Including studies where max age >30	0.44 (0.36-0.52)	0.29 (0.223-0.367)	0.23 (0.12-0.37)	0.131 (0.064-0.231)

A.7.2.4 HIV multipliers for progression and regression of disease

Since both HIV and HPV are sexually transmitted infections, we expect incidence of cervical abnormalities to be higher among HIV-positive than HIV-negative women and relative incidence rates have been estimated in many studies (91). However, it has also been shown that the level of immunosuppression plays a role in this incidence ratio and that abnormalities are less likely to regress and more likely to progress among HIV-positive women than among HIV-negative women. These associations are summarised in a meta-analysis by Liu *et al.* (92), which we will draw from in this study. A meta-analysis of studies comparing progression and regression of cervical disease in HIV-positive women by ART use showed that, although beneficial effects from individual studies are not always significant, the pooled estimates show stronger beneficial effects of ART in terms of both progression and regression (93). However, no study to our knowledge directly compared rates for women on ART to HIV-negative women. As shown in Rohner *et al.* (94), cervical cancer incidence among women on ART in South Africa is almost 10-fold higher than among their European and North-American counterparts and in South Africa, incidence does not depend on duration of ART use. The authors propose that this may be due to ART initiation at late stages of HIV disease in South Africa and that women have already progressed to non-reversible stages of cervical pre-cancer by the time of ART initiation. The meta-analysis by Kelly *et al.* also highlights the CD4 count at ART initiation as a modifier of impact of ART (93).

HPV to CIN1: The relative risk for progression for untreated HIV+ vs HIV- women estimated in Liu *et al.* (92) of 3.73 (95% CI 2.62-5.32) were derived from two studies in the late 1990s. In Kelly *et al.* (93) it was estimated that women on ART had 0.7 (95% CI 0.55-0.9) times the risk of progression in women not on ART. We will assume that women who start ART in late stages of disease will have the same risk of progression in the first two years of ART use as women not on ART, and thereafter the same risk as women who started ART early. The prior ranges for the multiplier for women not on long-term ART will be between 2 and 5.32, and the range for women on long-term ART will be between 0.55 and 0.9 times the multiplier for women not on ART.

Progression after CIN1: For later stages of CIN and progression between cervical cancer stages, we use the relative risk of progression to HSIL from Liu et al of 1.32 (95% CI 1.1-1.58). In Kelly *et al.* it was estimated that women on ART had 0.74 (95% CI 0.61-0.9) times the risk of progression from CIN1 to higher stages than women not on ART. The prior ranges for the multiplier for women not on long-term ART will be between 1.1 and 1.6. Since the ART relative risk for this parameter is very

similar to that of the previous parameter, the same value drawn from 0.55 to 0.9 above will be multiplied by the value drawn from 1.1 to 1.6. This product will have a minimum of 1 so that women on ART are not less likely to progress than HIV-negative women.

CIN1 and CIN2 regression: The relative risk for regression for HIV+ vs HIV- women estimated in Liu is 0.67 (95% CI 0.56-0.82) and the relative risk of regression for women on ART vs not on ART estimated in Kelly is 1.62 (95% CI 1.32–1.99). The prior ranges for the multiplier for women not on ART will be between 0.56 and 0.82, and the range for women on ART will be between 1.3 and 2 times the multiplier for women not on ART. This number will have a maximum of one, so that women on ART are not more likely to regress than HIV-negative women.

A.7.2.5 Proportion that clears HPV infection during CIN1 regression

In a study by Nobbenhuis *et al.* (95), 79 of 87 women (90.8%) who regressed from lower grade lesions also cleared their HPV infections during the 5 years of follow-up. In Schiffman *et al.* (96), this fraction was 447/534 (83.7%) during 2 years of follow-up. Two other studies showed short mean time differences of less than three months between regression of lower grade lesions and HPV clearance (97,98). In the model, 90% of women are assumed to clear the HPV infection at the same time that CIN1 regression takes place.

A.7.2.6 Duration of CIN3

In the model, women who develop CIN3 cannot naturally regress from this state. When a woman progresses to CIN3, we draw a time of progression to cancer from Weibull distributions. The prior distribution for the scale parameter of this Weibull distribution is chosen as uniform between 5 and 20 years. The prior distribution of the shape parameter is uniform between 1.5 and 3. HIV-positive women will have a shorter duration of CIN3, determined by multiplying the scale parameter by the inverse of the multiplier for rate of progression of higher-grade pre-cancer as discussed in Section 6.4.

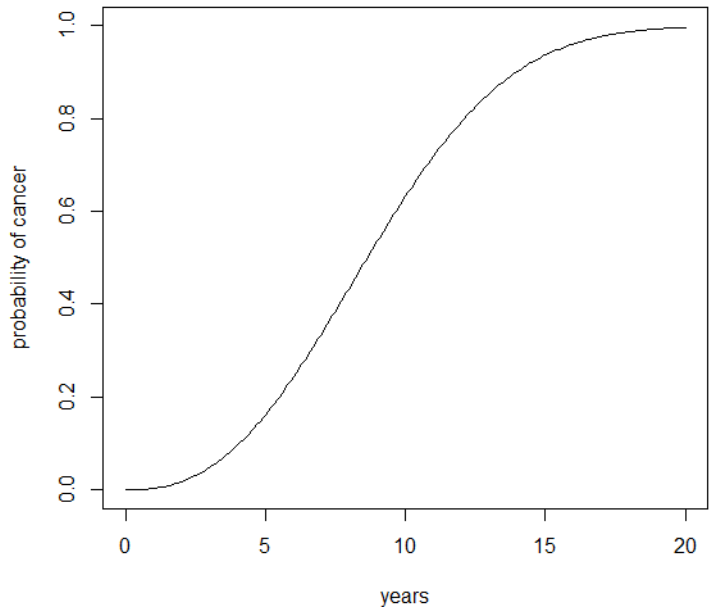


Figure A 15 – Cumulative probability function of a Weibull distribution with scale of 10 years and shape of 2.5. This curve represents the cumulative probability that an HIV-negative woman who has CIN3 will progress to cancer over time.

A.7.2.7 Progression, diagnosis and mortality of cervical cancer

Since we aim to calibrate our model to the incidence of diagnosed cancer, we simulate the progression through stages of cancer severity, and diagnosis at each stage. Figure A 16 illustrates this process in the model.

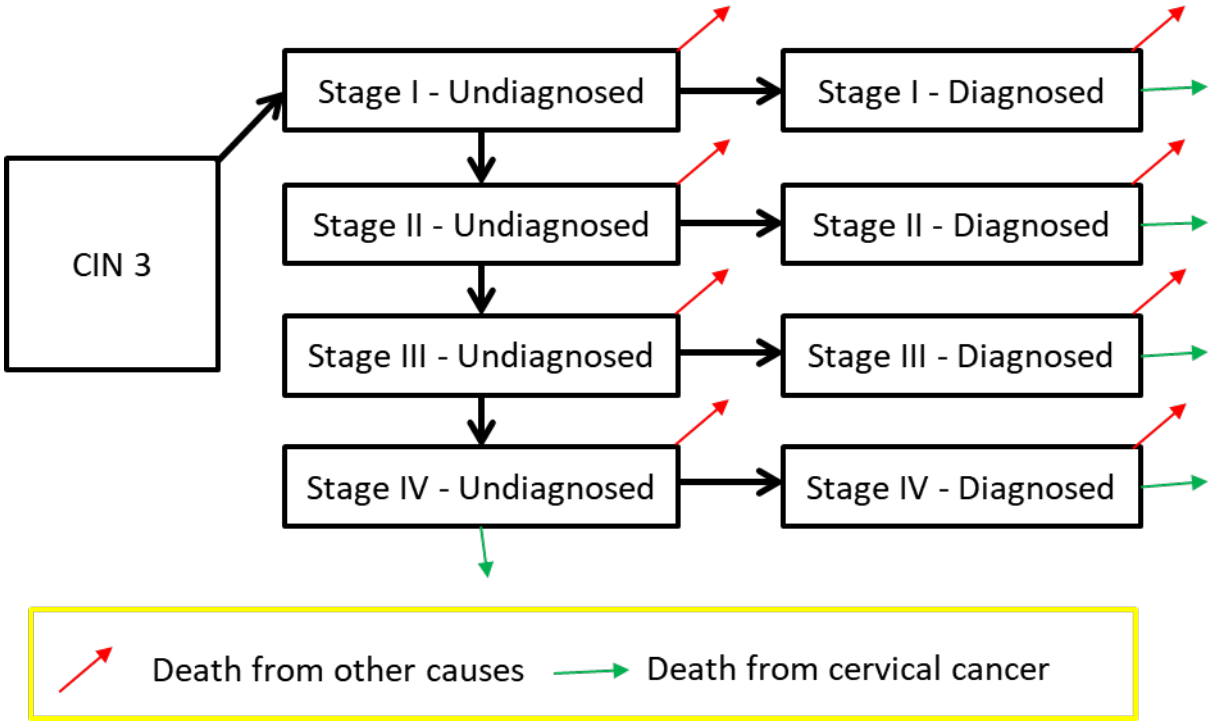


Figure A 16 – Progression and diagnosis of cervical cancer in our model. Undiagnosed women progress through all the stages before dying of cancer, while mortality in diagnosed women depends on the stage at diagnosis.

Data on natural progression through cancer stages are not available, since women are treated upon diagnosis. We will use (for HIV-negative women) the same progression rates that the majority of cervical cancer modelling studies use, as shown in

Table A *17* (99). A more recent analysis of the American National Cancer Institute's data found very similar estimates (100). We will not include these parameters in uncertainty analyses, but will vary them in sensitivity analyses. HIV-positive women will progress at increased rates, using the same multiplier for progression of higher grade lesions as discussed in Section 6.4. Although there are no studies that investigated increased rates of cancer progression among HIV-positive vs negative women, two studies showed that in the pre-ART era in South Africa, HIV-positive women were diagnosed at more advanced stages of disease than HIV-negative women (101,102).

In the model, a woman can be diagnosed through two separate processes: 1) Via the screening algorithm through either routine Pap-smear screening or follow-up colposcopy visit or 2) Diagnosis after seeking health care due to cancer symptoms. Prior distributions for the yearly probabilities of diagnosis with cancer symptoms were originally informed by Myers *et al.* (99), but the point estimates from this study resulted in distribution of stage at diagnosis that is inconsistent with South African data (Table A 11). In particular, the fraction of cases diagnosed in the early stages was too high. We therefore used prior ranges for diagnosis in Stages I and II that were lower than the estimates in Myers *et al.* (99).

To obtain mortality rates by stage of cancer diagnosis, cause of death information from the Groote Scguur Hospital (GSH) database was analysed. We make the simplifying assumption that the stage of diagnosis is the same as the stage at which cervical cancer treatment is started. We use the Weibull accelarated failure time model in the *survival* package in R to estimate survival probabilities, and these values (for the first five years after treatment was started) are shown in Figure A 17 (the black lines). In the model, we randomly draw time to death from these Weibull distributions, by stage at diagnosis (parameters in

Table A *17*). After five years, the probability of cancer death is very small, and in the model we assume that a woman dies of CC in the first five years after diagnosis, or dies of other causes.

A woman who does not get diagnosed with CC will progress through the stages and die from Stage IV. On average, women who were diagnosed at GSH with stage IV cancer and received only palliative care, lived 3 months. We will assume that women in stage IV will either get diagnosed and experience mortality as shown in

Table A *17*, or live on average 6 months.

Table A 17 – Rates of cancer progression, and parameters for Weibull survival distributions.

Stage	Progression rate per year	Shape	Scale (year)
I	0.225	0.61	126.5
II	0.3	0.67	16.28
III	0.45	0.56	3.91
IV	NA	0.78	0.53

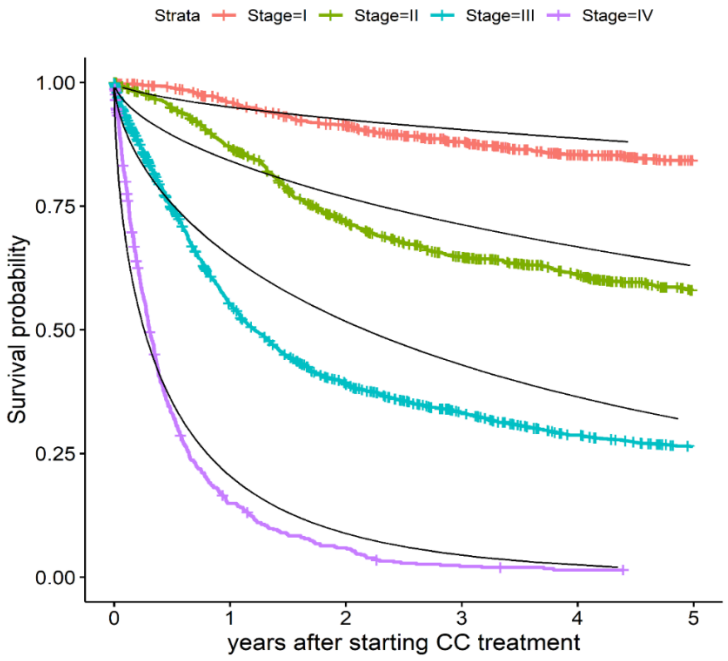


Figure A 17 – Cervical cancer survival probailities over time after treatment was started, by stage of diagnosis. The crosses represent the Groote Schuur Hospital data, while the black lines represent Weibull survival distributions.

8. Posterior distributions of parameters

Table A 18 - Type specific medians (and interquartile range) of the 500 samples from the posterior distributions of the HPV infection and transmission parameters in the model

	16	18	31	33	35	39	45	51	52	56	58	59	68
Transmission probability (per sex act)													
Male to Female	0.6 (0.31-0.77)	0.4 (0.18-0.7)	0.54 (0.28-0.8)	0.18 (0.07-0.48)	0.48 (0.24-0.74)	0.44 (0.21-0.68)	0.19 (0.06-0.5)	0.53 (0.26-0.77)	0.46 (0.21-0.75)	0.15 (0.05-0.48)	0.54 (0.27-0.77)	0.46 (0.23-0.73)	0.37 (0.17-0.64)
Female to Male	0.09 (0.04-0.25)	0.38 (0.18-0.68)	0.28 (0.14-0.54)	0.31 (0.1-0.54)	0.2 (0.07-0.51)	0.23 (0.08-0.56)	0.24 (0.08-0.56)	0.33 (0.12-0.59)	0.41 (0.21-0.69)	0.56 (0.3-0.77)	0.32 (0.11-0.64)	0.28 (0.12-0.58)	0.57 (0.32-0.8)
Relative HPV duration in HIV infection													
Latent HIV vs HIV-negative	1.3 (1.1-1.6)	1.3 (1.1-1.5)	1.3 (1.1-1.5)	1.3 (1.2-1.6)	1.4 (1.2-1.6)	1.3 (1.1-1.6)	1.4 (1.2-1.7)	1.3 (1.1-1.6)	1.4 (1.2-1.7)	1.4 (1.2-1.6)	1.3 (1.1-1.5)	1.4 (1.2-1.6)	1.4 (1.2-1.7)
Acute HIV/late HIV/recent ART vs HIV-negative*	2.2 (1.8-2.6)	2.1 (1.6-2.6)	1.6 (1.2-2)	2 (1.5-2.4)	2 (1.6-2.5)	1.8 (1.4-2.3)	2.1 (1.6-2.5)	1.9 (1.5-2.4)	2.1 (1.6-2.5)	2.3 (1.7-2.6)	1.8 (1.4-2.4)	1.8 (1.4-2.4)	2.1 (1.7-2.6)
Average time to reactivation (in years) if HIV-negative													
Males	17.5 (9.4-23.2)	16.5 (10.8-23.7)	18.8 (10-23)	16.2 (9.3-22.5)	12.4 (7-20.5)	13.7 (8.7-20.7)	13.9 (8-21.5)	14.4 (7.4-21.7)	15.7 (7.7-23.6)	20.1 (12.6-24.5)	13.9 (8.1-19.6)	13.8 (5.8-21.9)	17.2 (9.4-24.9)
Females	19.4 (13.9-24.1)	18.5 (13.6-24.5)	20.8 (13-26)	17.7 (10.5-25.4)	19.8 (11.3-25.6)	17.2 (12.4-24.3)	17.9 (12.5-23.9)	17.6 (11.4-23.8)	18.8 (13-24.4)	18.5 (12.8-24.6)	19.1 (11.6-24.9)	18.2 (11.7-23.6)	19.5 (12.7-24.5)
Proportion who become latently infected after clearance	0.56 (0.42-0.74)	0.65 (0.48-0.81)	0.3 (0.16-0.52)	0.51 (0.28-0.72)	0.5 (0.32-0.72)	0.66 (0.39-0.82)	0.73 (0.58-0.85)	0.58 (0.37-0.77)	0.44 (0.27-0.62)	0.54 (0.34-0.69)	0.62 (0.4-0.78)	0.51 (0.34-0.71)	0.36 (0.2-0.53)
Relative HPV reactivation rate													
Latent HIV vs HIV-negative	2.4 (1.8-3.4)	2.1 (1.5-3)	2.3 (1.5-3.1)	2.2 (1.5-3.5)	2.4 (1.7-3.6)	1.8 (1.3-2.6)	2.5 (1.7-3.6)	2 (1.4-3)	2.5 (1.7-3.6)	2.7 (1.8-3.6)	2 (1.4-2.9)	1.9 (1.4-3.2)	2.7 (1.7-3.8)
Acute HIV/late HIV/recent ART vs latent HIV*	2.1 (1.6-2.5)	1.9 (1.6-2.5)	1.7 (1.4-2.3)	1.9 (1.5-2.5)	2 (1.4-2.4)	1.8 (1.4-2.4)	2 (1.5-2.5)	1.9 (1.4-2.5)	2 (1.5-2.5)	1.9 (1.5-2.3)	2 (1.4-2.4)	1.9 (1.4-2.5)	2 (1.5-2.4)
Average duration of immunity (in years)													
Males	16.6 (9.3-22.9)	10.5 (5.8-15.6)	9.1 (4-12.7)	9 (4.6-13.3)	6.3 (2.1-11.3)	9.5 (4.3-14.2)	10.1 (4.6-14.9)	9.9 (5.3-14.8)	11.5 (6.5-15.4)	11.2 (6.7-15.2)	8 (3.6-13.2)	8.4 (3.9-14.6)	11.9 (7.2-15.6)
Females	15.7 (9.4-22.3)	17.5 (9.5-24.2)	16.5 (11.5-22.4)	16.1 (9.1-24.3)	17.3 (9.8-23.6)	18.8 (11.1-26)	17.7 (10.3-25.1)	18.1 (11.3-25.1)	16.2 (9.2-23.2)	16.5 (8.5-22.7)	17.4 (10.8-23.6)	19.6 (11.7-24.8)	19 (11.2-24.7)
Average duration of HPV infection (in months) if HIV-negative													
Males	12.1 (8.9-16.1)	5.6 (4.1-7)	3.3 (2.4-4.7)	3.3 (2-4.8)	5.2 (3.5-8)	4.6 (3.1-6.7)	5.3 (3.6-7.6)	8 (5.3-10.8)	6.8 (5.1-9.5)	3.2 (2.5-4.3)	4.7 (3.1-6.8)	7.6 (5.2-11)	9.6 (7.2-12.2)
Females	11.4 (9.9-13.4)	9.7 (8.2-11.7)	7.8 (6.6-9.9)	7 (4.9-10.4)	10 (8.4-11.9)	8.5 (6.5-10.9)	10.5 (7.9-14.3)	9.3 (7.6-11.1)	8.5 (7.1-10.2)	7.3 (5.3-11.7)	11.5 (9.5-13.8)	6.8 (6-8.3)	7.5 (6.3-9)

*Recent ART is defined as ART initiation within last 2 years. The same parameters are used for people who have been on ART for longer than two years and HIV-negative people.

Table A 19 – Medians and 95% percentile intervals for 100 best fitting parameter combinations

Cervical pre-cancer parameters	**Type 16**	**Type 18**	**Other HR-HPV**
Multiplier for duration of HPV among women	0.53 (0.23-0.94)	0.64 (0.28-1.0)	0.65 (0.43-0.95)
Proportion that will progress from HPV infected to CIN1	0.23 (0.11-0.34)	0.17 (0.08-0.3)	0.15 (0.08-0.27)
Progression rate from CIN1 to CIN2	0.15 (0.1-0.24)	0.07 (0-0.17)	0.11 (0.05-0.16)
Regression rate from CIN1 to Normal	0.66 (0.31-1.01)	0.91 (0.43-1.64)	1.02 (0.51-1.66)
Regression rate from CIN2 to CIN1 (<=30)	0.44 (0.38-0.52)	0.67 (0.46-0.89)	0.61 (0.43-0.9)
Multiplier for regression in women older than 30	0.59 (0.46-0.73)		
HIV: multiplier for CIN1 progression	2.54 (2.01-4.79)		
ART: multiplier for HIV progression multiplier	0.72 (0.57-0.88)		
HIV multiplier for rate of CIN1/2 regression in HIV-negative women	0.76 (0.61-0.82)		
ART multiplier for rate of CIN1/2 regression in HIV-negative women	1.0 (0.97-1.0)		
Cervical cancer parameters*			
Progression rate from CIN2 to CIN3 (<=30)	0.041 (0.024-0.072)	0.015 (0.001-0.047)	0.008 (0.005-0.013))
Multiplier for women aged 30-50	2.48 (2.04-2.93)		
Multiplier for women aged 50+	3.77 (2.5-5.38)		
HIV: multiplier for CIN2 progression	1.21 (1.12-1.42)		
CIN3 duration: scale (years)	16.45 (11.8-19.5)		
CIN3 duration: shape	2.55 (2.1-2.9)		
Yearly probability of getting diagnosed in Stage I	0.023 (0.002-0.049)		
Yearly probability of getting diagnosed in Stage II	0.12 (0.06-0.19)		
Yearly probability of getting diagnosed in Stage III	0.61 (0.41-0.79)		
Yearly probability of getting diagnosed in Stage IV	0.93 (0.86-1)		

*The assumption that 10% of cervical cancer cases do not receive pathological diagnosis led to the best fits to data, and the values of the cervical cancer parameters shown here are for this scenario.

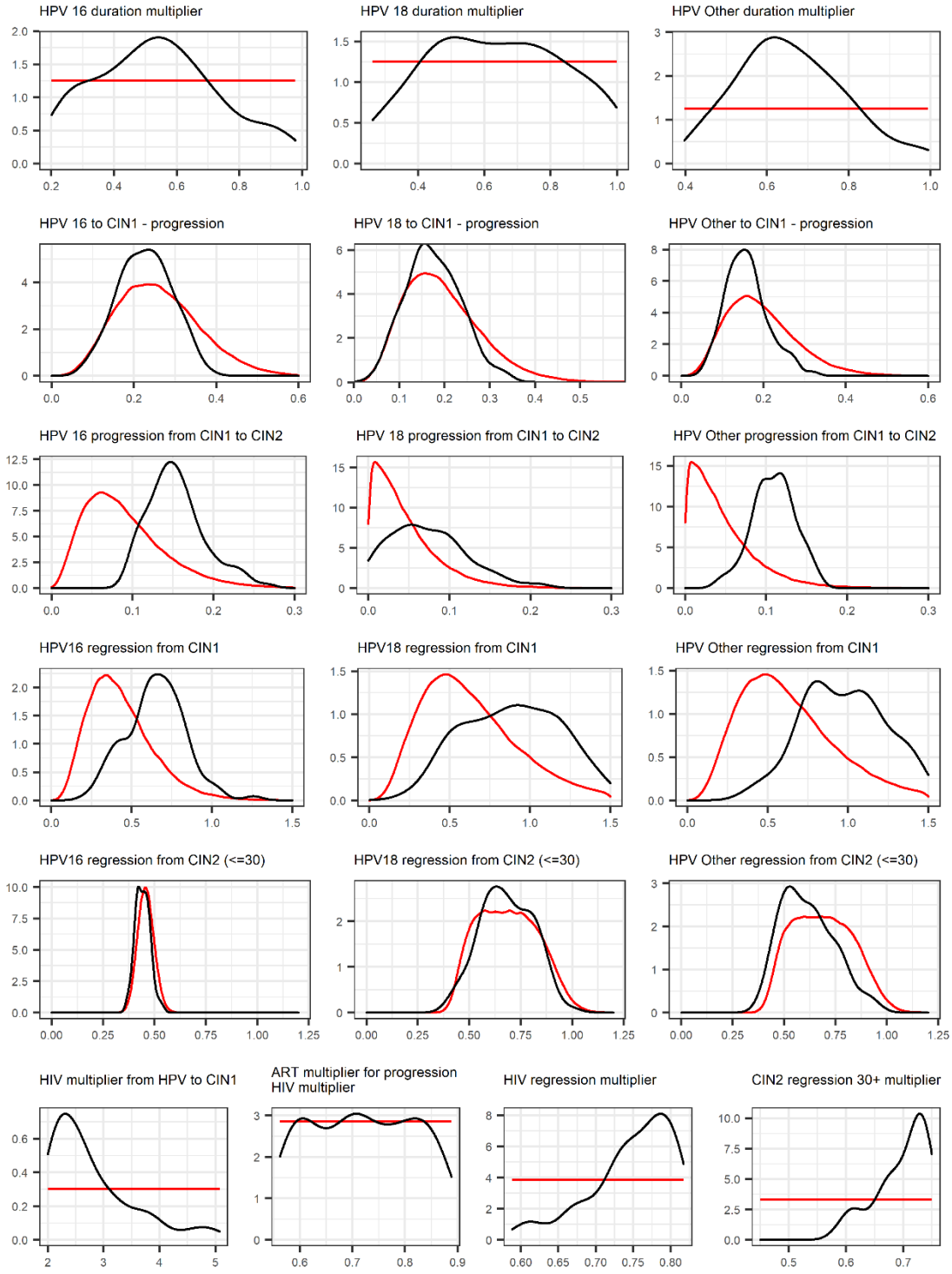


Figure A 18 – Prior and posterior distributions of cervical pre-cancer parameters. Red lines represent prior and black lines represent posterior distributions.

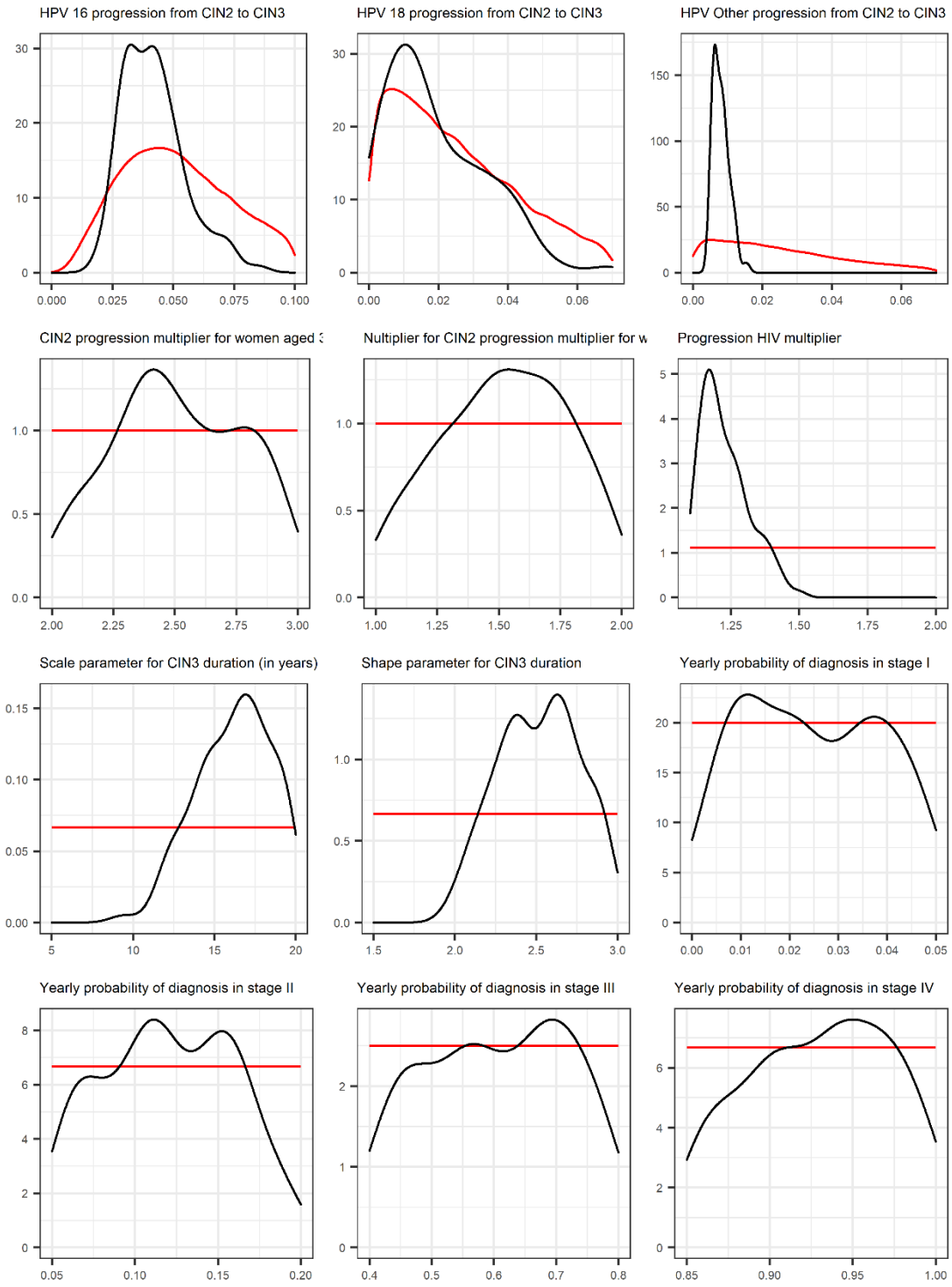


Figure A 19 - Prior and posterior distributions of cervical cancer parameters. Red lines represent prior and black lines represent posterior distributions.

9. Model fits to data

A.9.1 HPV prevalence

In Figures A 20 to 21, the model fits to type specific HPV prevalence data are shown. The red dots and error bars represent the data in Table A10 (number on the x-axis matches study number in the table) and the black dots and error bars represent the mean prevalence estimate produced by the sample of 500 parameter combinations from the posterior distributions.

Figure A 20 - Model fits to HPV types 16, 18, 31, 33, 45, 52, 58 (vaccine types)

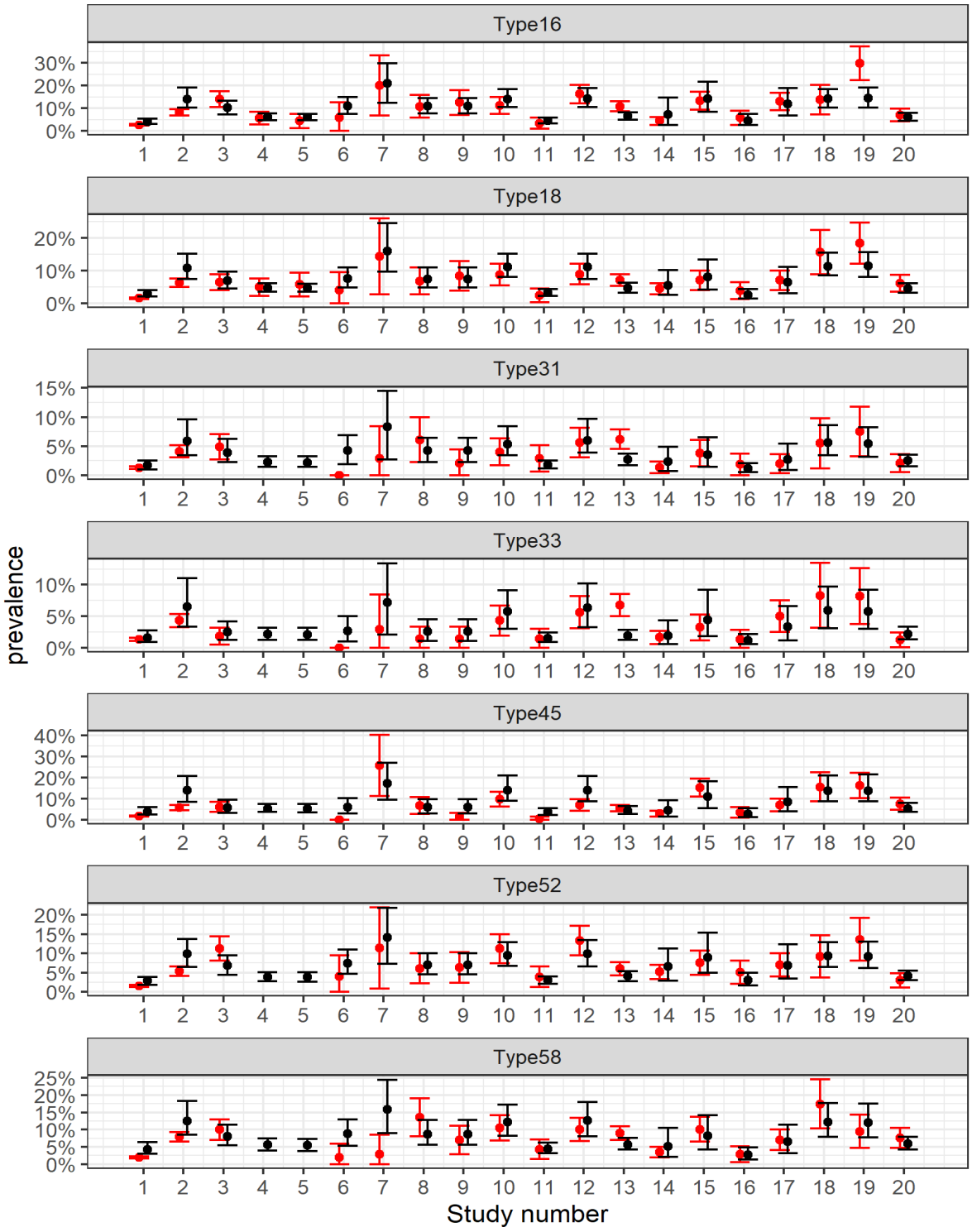
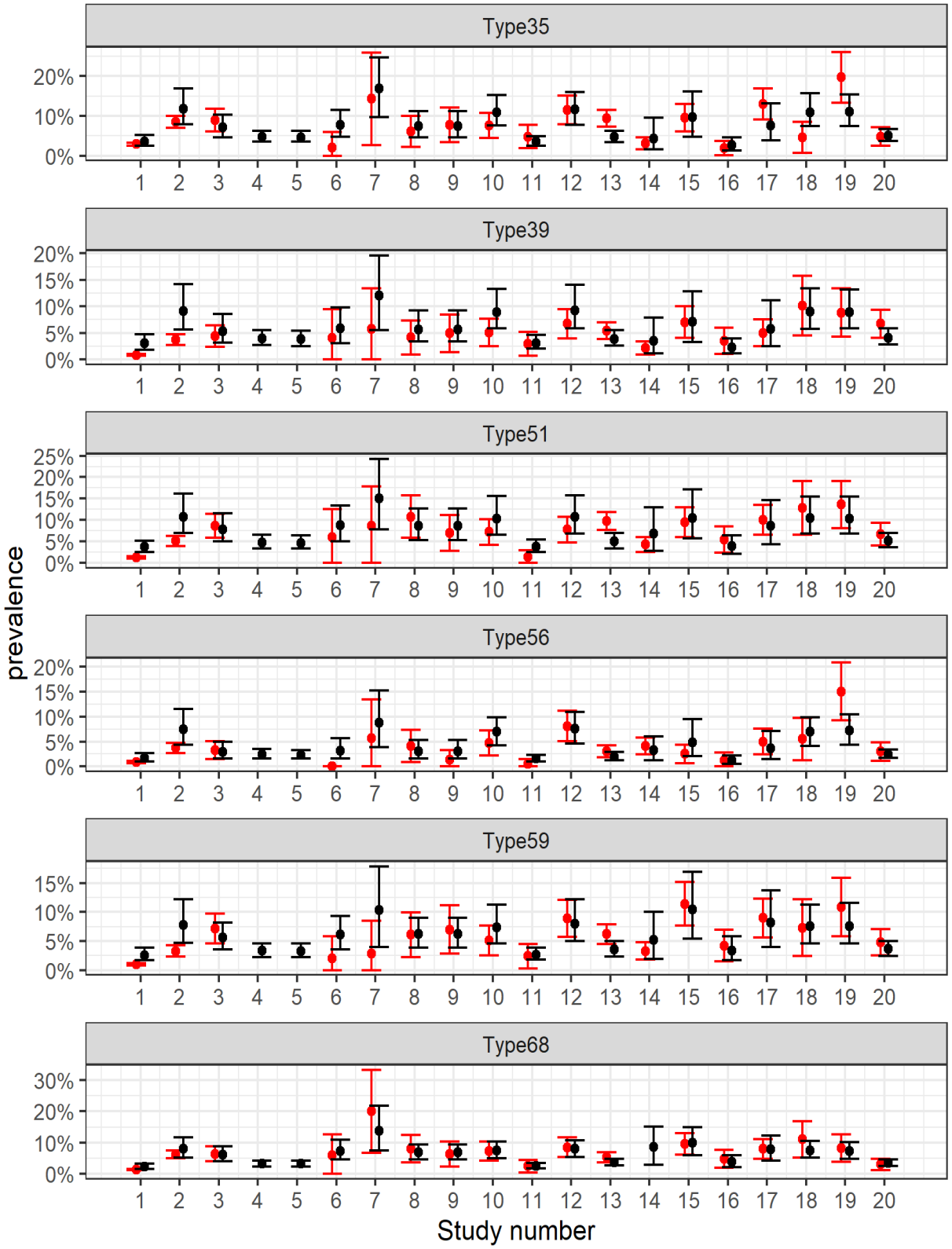


Figure A 21 - Model fits to HPV types 35, 39, 51, 56, 59, 68 (non-vaccine types)



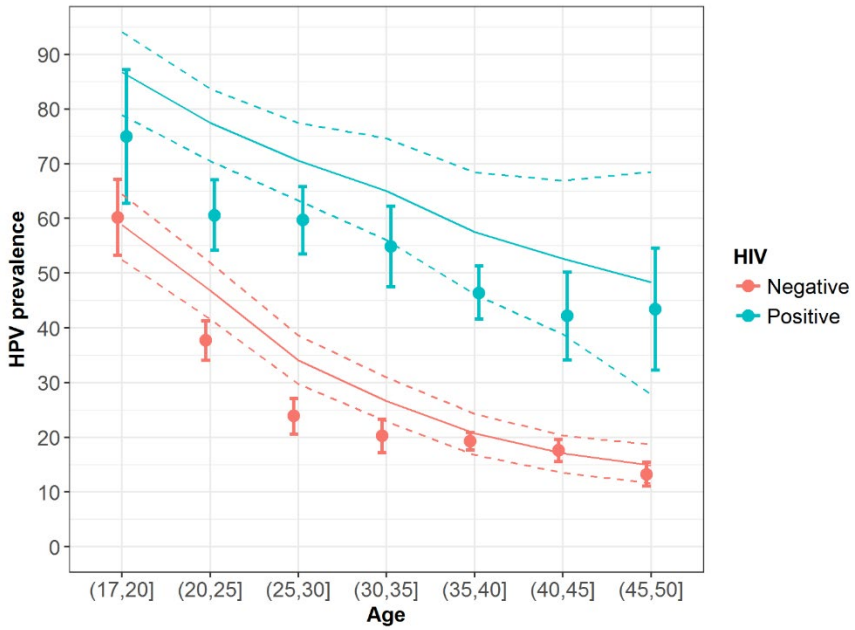


Figure A 22 - The mean overall high risk HPV prevalence in 2000 from 500 populations simulated using the sample of parameter combinations from the posterior distributions (solid lines). The confidence bands (2.5th and 97.5th percentiles) around these estimates are shown in dashed lines. Data points (in closed circles) represent HPV prevalence results from a population level study in Khayelitsha, Cape Town (103).

A.9.2 Cervical pre-cancer

Table A 20 and Figure A show the model fits to the data in Table A 9. Although the data on HIV-positive women in the McDonald study (25) were grouped during calibration, we show disaggregated estimates here.

Table A 20 – Model fits to cervical pre-cancer data

Study	HIV	Measure	Ages	Sample size	Observed prevalence	Model prevalence
Cronje (44)	Not tested	CIN1+	21-65	1093	34.9 (32.1-37.8)	15.2 (12.2-18.2)
Cronje	Not tested	CIN2+|CIN1+*	21-65	382	23.6 (19.3-27.8)	27.6 (23.2-35.5)
Denny (21)	Not tested	CIN1+	35-65	2922	6.1 (5.2-7)	11.2 (8.9-13.4)
Denny	Not tested	CIN2+|CIN1+	35-65	178	46.6 (39.3-54)	24.3 (19.5-32.1)
Kuhn (27)	Negative	CIN1+	30-65	378	12.7 (9.3-16.1)	8.5 (6.5-10.6)
Kuhn	Negative	CIN2+|CIN1+	30-65	48	41.7 (27.7-55.6)	34.1 (27.2-40.7)
Kuhn	No ART	CIN1+	30-65	67	37.3 (25.7-48.9)	41.1 (32.7-49)
Kuhn	No ART	CIN2+|CIN1+	30-65	25	56 (36.5-75.5)	39.8 (33.5-47)
Kuhn	On ART	CIN1+	30-65	263	29 (23.4-34.4)	34.5 (26.7-43.8)
Kuhn	On ART	CIN2+|CIN1+	30-65	76	55 (44.1-66.5)	43.7 (35.7-51.6)

*CIN2 or worse given any abnormality (CIN1+)

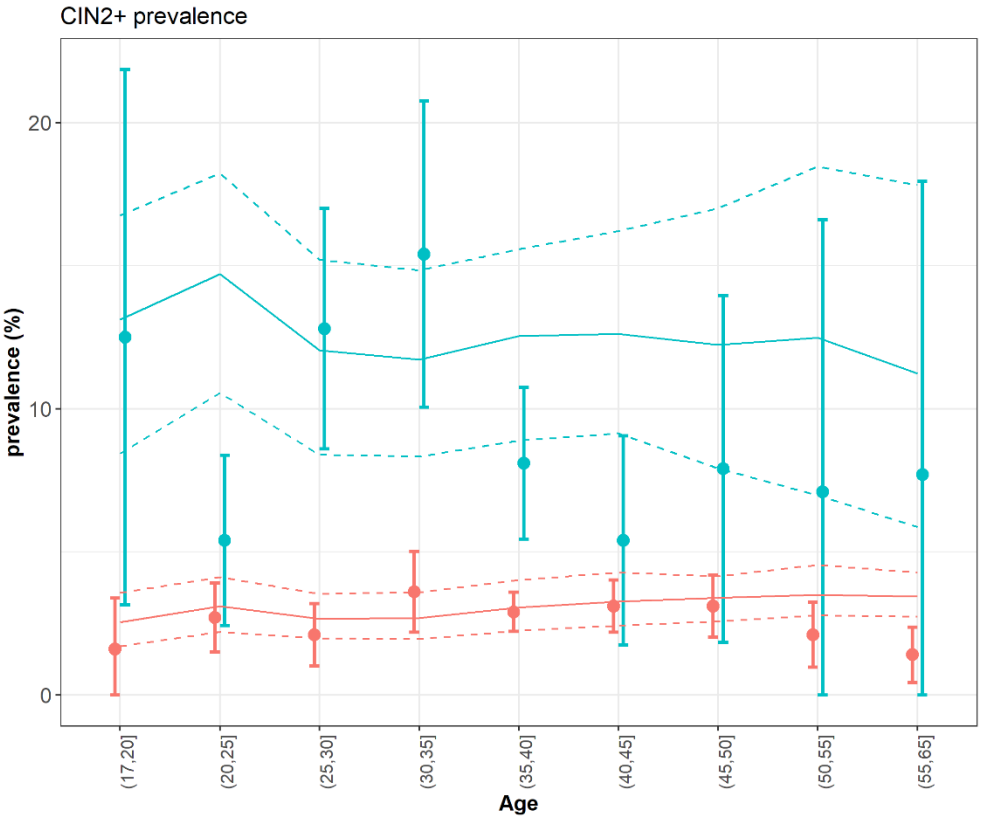


Figure A 23 – CIN2+ prevalence by age and HIV status (red=HIV-negative and blue=HIV-positive) from McDonald et al. (25)

A.9.3 Cervical cancer

As described in Section A.5.3, to fit to cervical cancer data, we used the medians of the posterior distributions of the HPV and cervical pre-cancer parameters and sampled 30,000 parameter combinations from the prior distributions of the cervical cancer parameters. The total population size was around 750,000 women in each simulation in 2016. We calibrated the model using 4 different assumptions regarding the fraction of cervical cancer cases that only receive a clinical diagnosis (no pathology), as described in Section A.5.3. Of the four, the assumption about a linear decrease in the fraction of cases who receive only a clinical diagnosis did not fit well to data. The other three scenarios fit equally well to the data, and the 100 best fitting parameter combinations for the 7% and 14% scenarios overlapped by 89% and 82% with the 10% scenario. The scenario where 10% of cases do not receive a pathology diagnosis had a marginally higher mean total log-likelihood value than the 7% and 14% scenarios, and for this reason we use this scenario in further analyses. Figure A shows the model fit to the age specific cervical cancer data that were used in calibration, using the 100 best fitting parameter combinations and assuming that 10% of CC cases only receive a clinical diagnosis (at all ages).

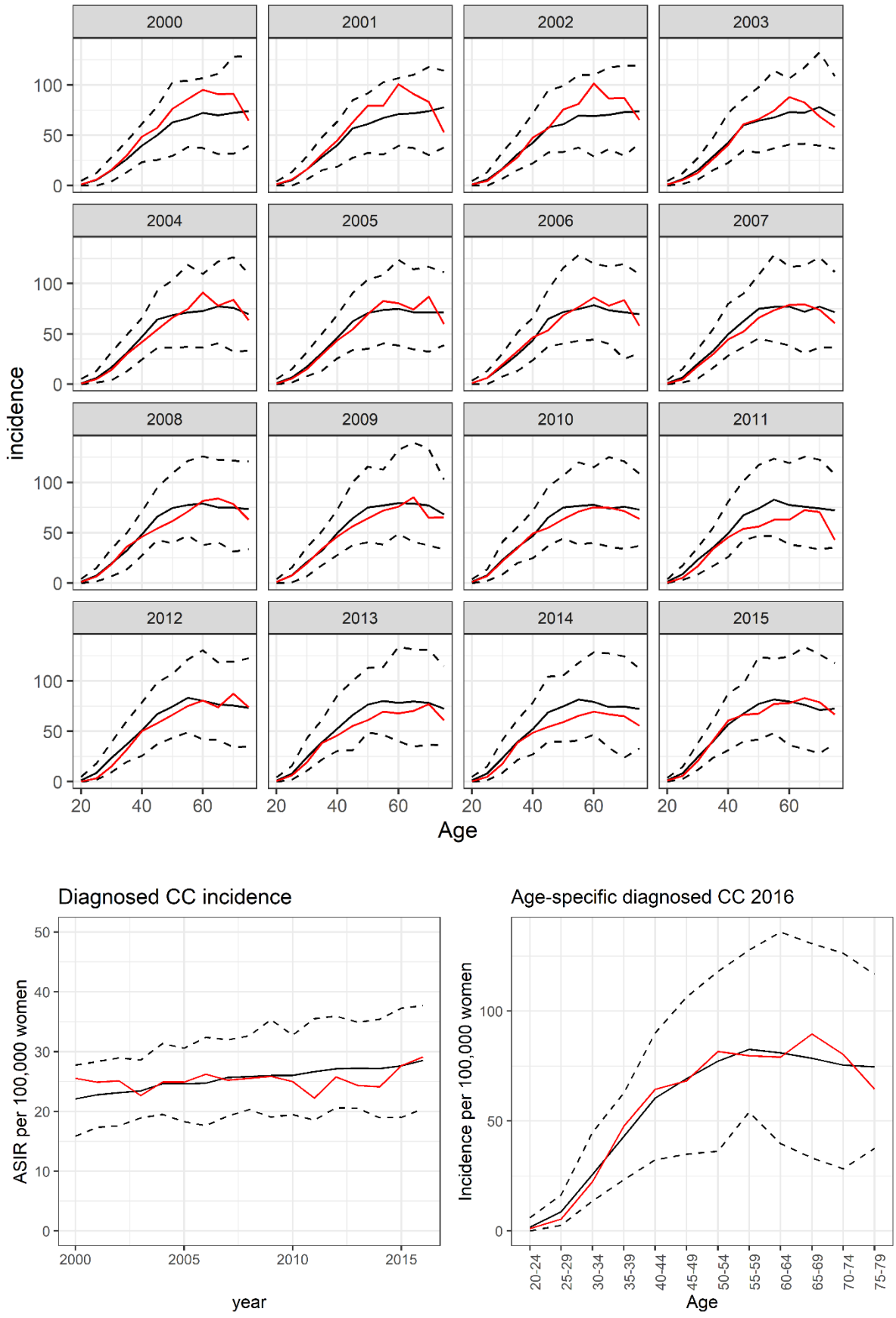


Figure A 24 – Age specific diagnosed cervical cancer incidence per 100,000 women as calculated from NCR data (red lines) and the 100 best fitting parameter combinations (black lines show mean of 100 estimates, dashed lines show 95% percentile intervals).

Figure A 25 shows the model fit to the other sources of data that were used in calibration – the fraction of cases that get diagnosed in each of four stages of cancer (Table A 11). In this figure, model estimates are plotted against the fraction of cases diagnosed in each stage in Groote Schuur Hospital (solid red lines) and the other data points included in the calibration (red dots). The model did not in all cases fit well to the three studies in 2011, 2013 and 2015 with the smallest sample sizes (Table A 11), but this is not a major cause for concern.

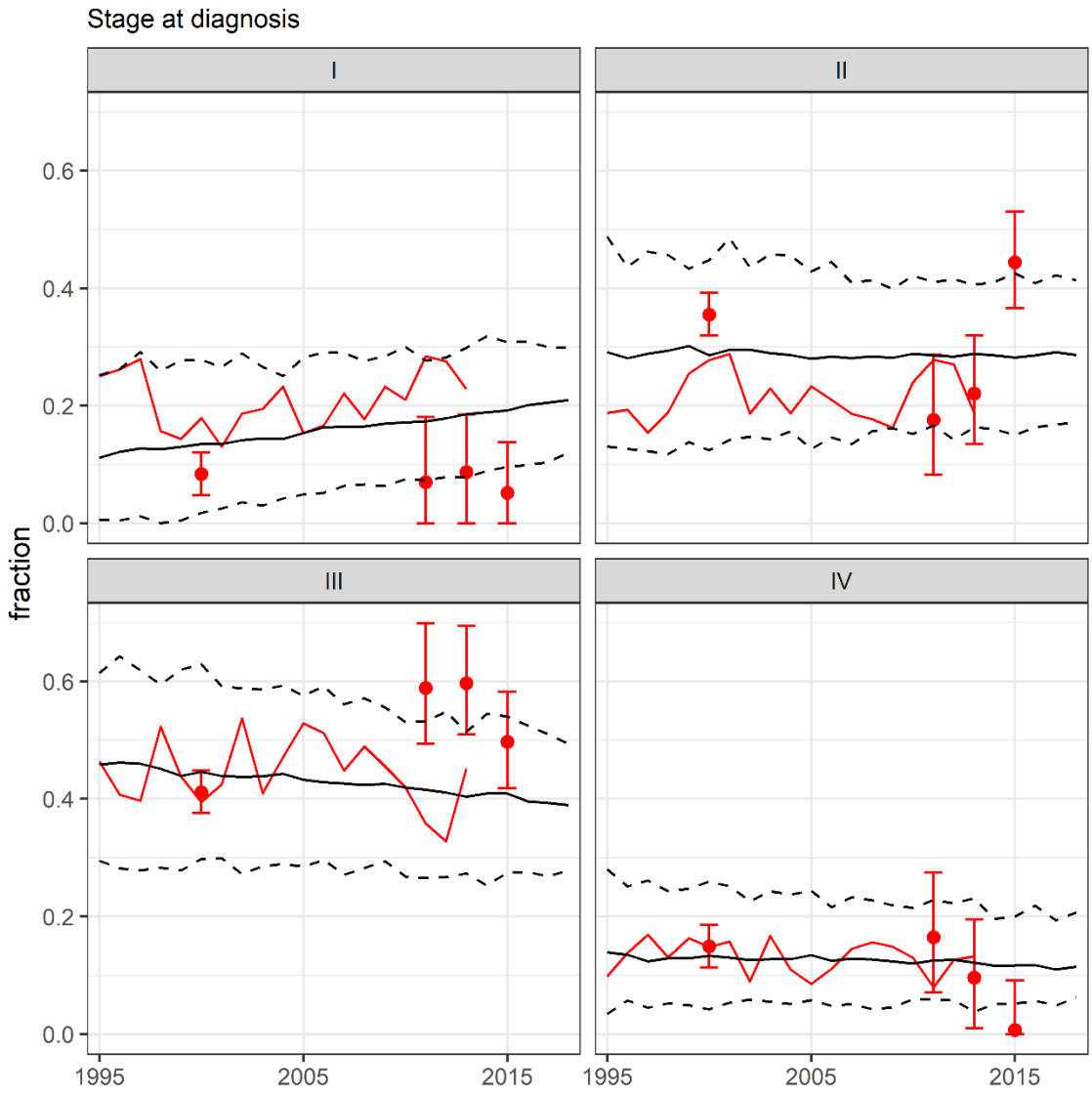


Figure A 25 – The fraction of cervical cancer cases diagnosed in each stage of the disease. The black lines show mean model estimates and the dashed lines show 95% percentile intervals. The red lines represent GSH data and red dots represent data as shown in Table A11 (Lomalisa 2000, Snyman 2011, Mbodi 2013, Sabulei 2015).

In Figure A 26 (A), we show the model fit to the age-standardised incidence rate (ASIR) that we calculated using crude and age-specific incidence data from NCR (the same data that we calibrated the model to) between 2000 and 2016, as well as age-standardised incidence published in Olorumfemi *et al.* (104) between 1994 and 1999. Age-specific incidence was not available for this period and was therefore not included in the calibration. All NCR data were inflated by 10% corresponding to the best

fitting assumption about under-reporting. Our model under-estimates diagnosed cervical cancer incidence in the early years, and in Figure A 24 it seems that this under-estimation is concentrated among older women. This may be because our model's starting conditions (Section A.3) make the implicit assumption that sexual behaviour have always remained constant, while it is likely that people behaved differently in the pre-HIV era and that fractions of women in the pre-cancer stages might have been higher in 1985. Unfortunately, we have no data to validate these claims.

Figure A 22 (B) shows the CC mortality to incidence ratio. We calculated CC mortality in our model as the number of women who die of diagnosed cervical cancer. We divide the age-standardised mortality rate with the age-standardised diagnosed cancer incidence to calculate the model's mortality to incidence ratio (MIR). We compare this to the MIR as calculated from the data in the appendix of Olorumfemi *et al.*(105). They show the pathology diagnosed ASIR from NCR, as well as the age-standardised mortality rate obtained from cause of death data from Statistics South Africa between 2004 and 2013.

Our mean model estimates (black line) are slightly lower than the MIR calculated from the data shown in Olorumfemi (red line). Since CC survival probabilities in the model at this stage depends solely on data from GSH, a hospital with above average resources in South Africa, it is possible that our model under-estimates mortality. On the other hand, by taking the ratio of data in Olorunfemi *et al.* to reflect the true MIR of CC in South Africa, we are essentially assuming that under-reporting in the NCR and Stats SA is similar. It is mandatory to register all deaths in South Africa and therefore under-reporting of CC deaths may happen to a lesser extent than the under-estimation of CC by the pathology-only NCR, and this ratio may be an over-estimate of mortality to incidence.

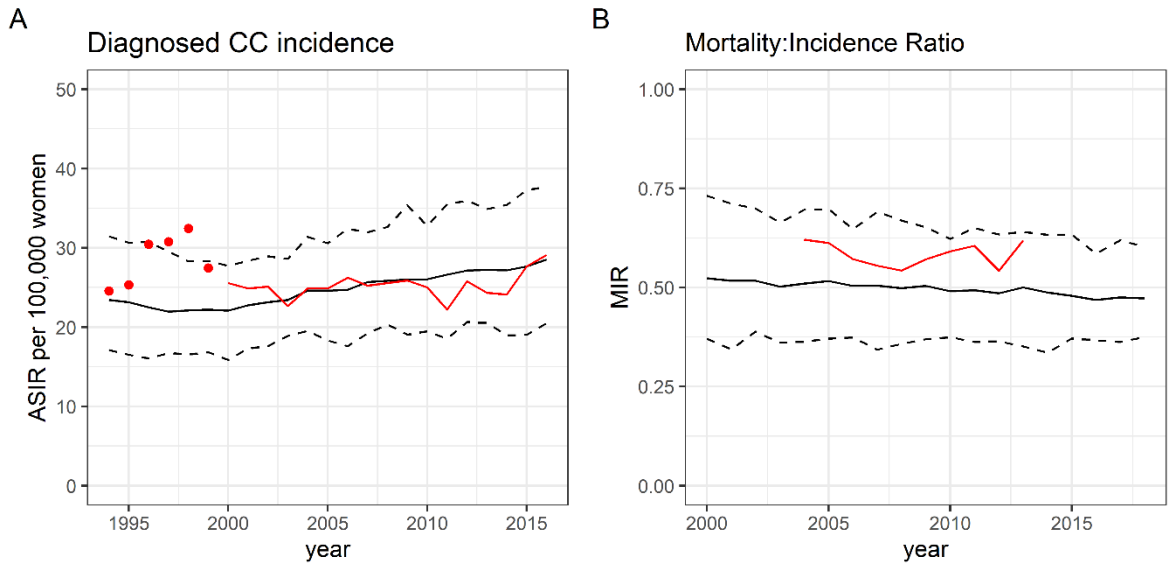


Figure A 226 – Diagnosed ASIR and mortality to incidence ratio using 100 best fitting parameter combinations. The black line represents the mean and the dashed lines represent the 95% percentile intervals. Red lines and points represent data from Olorumfemi et al.(105).

In Denny *et al.* (106), 65.7% of 300 cervical cancer cases had HPV 16/18 infections, while 83.4% had HPV16/18/31/33/45/52/58 infections (red dots in Figure A 27). In a study by Van Aardt *et al.* (107), prevalence of HPV16/18 among cancer patients was 63.2% (red dot in Figure A). Our average model estimates for the same time-period (2008-9) are 63.9% (95% CI 56.2-71.1%) and 79.8% (95% CI 74.7-84.4%) respectively. In the model, we measure the fraction of cancer cases that was *caused* by each type, while in the studies they cannot determine the causal HPV type. Since women in the model

can be infected with more than one HPV type, the fraction of cancer cases *infected* with e.g. HPV16/18 will be higher than the fraction *caused* by HPV 16/18.

On average in 2018, 54.8% (95% CI 46.9-65.6%) of women with cervical cancer in the model were co-infected with HIV. In a recent analysis, Stelzle *et al.* (108) performed a meta-analysis of the relative risk of cervical cancer among women living with HIV and used the pooled relative risk, the GLOBOCAN estimates of cervical cancer incidence and UNAIDS estimates of HIV prevalence to estimate the fraction of cervical cancer cases who are living with HIV. Their estimate for South Africa is 63.5% (red dot in Figure A 27).

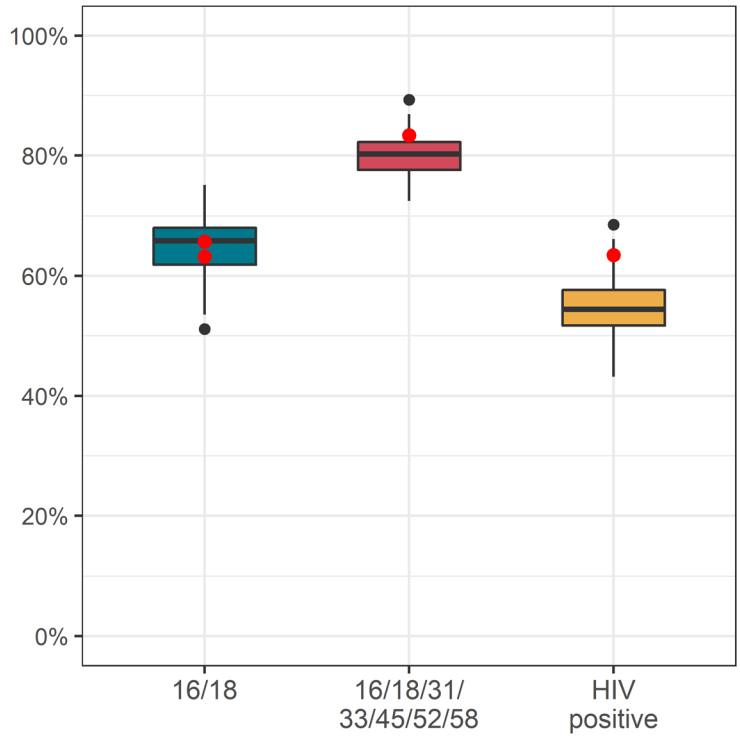


Figure A 27 – Fractions of CC cases that was caused by HPV16/18, HPV16/18/31/33/45/52/58, and fraction of CC cases who were co-infected with HIV. Boxplots represent model estimates and red points represent data. Data are from Denny et al. (106) and Van Aardt et al. (107) for cancer caused by HPV16/18, from Denny et al. (106) for cancer caused by other vaccine types, and from Stelzle et al. (108) for the fraction of cases who are HIV positive.

As another final model validation, we show the mean age of cancer diagnosis by HIV status in Table A 21. Model results are in line with those from studies.

Table A 21 – The average age at cancer diagnosis, by HIV stage

		Study sample size		Study average		Model average (95% CI)	
Study	Year	HIV-negative	HIV-positive	HIV-negative	HIV-positive	HIV-negative	HIV-positive
Lomalisa (61)	1997	776	60	53	44	56 (54-59)	38 (30-48)
Moodley (101)	1999	522	138	55.2	39.8	56 (54-60)	39 (33-43)
Moodley (109)	2000	457	29	46	40	56 (53-58)	40 (35-45)

Van Aardt (107)	2008	154	77	55.8	41.3	57 (55-59)	43 (40-46)
Van Bogaert (110)	2009	905	143	59.1	41.3	57 (55-60)	43 (40-47)

Figure A28 shows model estimates (in black) compared to estimates from the WHO's International Agency for Research on Cancer (IARC), also known as the GLOBOCAN estimates (111) (in red). Although these estimates are widely relied on as a credible source, they are calculated using assumptions that are not always context specific. For example, the 2018 estimates were calculated using the overarching assumption that the mortality to incidence ratio of cervical cancer in South Africa is the same as the mortality to incidence ratio among black Americans. In addition, the shape of the GLOBOCAN age-specific curve was not informed by South African data. Instead, the shape of age-specific incidence curves of low HIV burden countries was equally adjusted at all ages, to reflect the higher cervical cancer burden in South Africa [personal communication: Jacques Ferlay]. This does not take into account that a large proportion of cervical cancer cases in South Africa are among HIV-positive women, who develop cancer at much earlier ages than HIV-negative women.

Figure A 28 – Age-standardised and age-specific cervical cancer incidence as calculated by GLOBOCAN (111) *(red) and the 100 best fitting parameter combinations (black line show mean of 100 estimates, dashed lines show 95% percentile intervals).*

10. The Thembisa model

The Thembisa model is a deterministic compartmental model that simulates the population demographics and HIV epidemic of South Africa. The model provides estimates at the national and provincial level, and the methodology of the model at both scales is described in detail in technical documents available for download at www.thembisa.org. Its contributions to the South African HIV Investment Case helped to identify the most cost-effective HIV interventions, which has been the basis for government's decision to provide lifelong ART to all HIV-positive South Africans from September of 2016 (112). South African HIV estimates for the official UNAIDS reports have been based on Thembisa output since 2017 (113).

The model population is stratified by sex and 5-year age groups and stratification for sexual behaviour is similar to that of MicroCOSM, except that men can have sexual relationships with other men. All adults are classified according to their HIV testing history (never tested, ever tested, and ever diagnosed positive) and according to CD4 count, initiation of ART and ART duration if they are HIV positive. HIV prevention methods simulated include condom use, male circumcision, pre-exposure prophylaxis and ART (to prevent mother-to-child transmission and general transmission). The use of a wide range of different data sources in the model calibration, and the extensive validation of the model, make it the most reliable and informative model for assessing the impact of HIV in South Africa.

MicroCOSM v1 is an individual-based model that simulates a representative sample of the South African population. In addition, it does not simulate international migration and HIV prevention is limited to increases in ART coverage and changes in condom use. For these reasons, we reweight the population totals in MicroCOSM-HPV using the projected population demographics (age, sex, HIV and ART status) of the Thembisa model, on the assumption that the Thembisa model estimates future HIV and demographic trends more realistically.

11. Age standardisation

We use two world- standard populations in analyses of our model results. To be consistent with the method followed by the National Cancer Registry and IARC, we age-standardise cancer incidence according to the SEGI world population (105,111). To be consistent with the Brisson *et al.* paper and the WHO cervical cancer elimination strategy, we age-standardise according to the United Nations Development Programme's 2015 (UNDP 2015) world population (114). These standard populations are shown in Table E 1.

Table E 1 – World standard populations used in this study.

Age Group	SEGI	UNDP 2015
00-04	12000	8895
05-09	10000	8508
10-14	9000	8082
15-19	9000	7850
20-24	8000	7974
25-29	8000	8191
30-34	6000	7444
35-39	6000	6756
40-44	6000	6565
45-49	6000	6198
50-54	5000	5510
55-59	4000	4701
60-64	4000	4115
65-69	3000	3092
70-74	2000	2249
75-79	1000	1763
80-84	500	1154
85+	500	954
Total	100000	100000

Figure E 1 shows three different estimates of cervical cancer incidence. The red and black lines show mean model estimates of cervical cancer incidence calculated using the two world standard populations given in Table E1. Since the UNDP 2015 population gives more weight to women in older age groups who experience higher rates of CC incidence, this estimate is consistently higher than the estimate using the SEGI world population.

The blue line shows the mean model estimates of *diagnosed* cervical cancer incidence. Since the majority of cervical cancer cases in South Africa are diagnosed in advanced stages (Table A 11), there is a long delay between cancer incidence and diagnosis. However, the curves of incidence and diagnosed incidence are not similar in shape or scale, with a much slower increase in diagnosed incidence and lower overall levels. This happens because women are diagnosed in an age-category that carries less weight in the standard population, leading to a lower age-standardised estimate. In addition, a small fraction (~5%) of cervical cancer cases in the model die without receiving a diagnosis (from other causes or from undiagnosed cervical cancer).

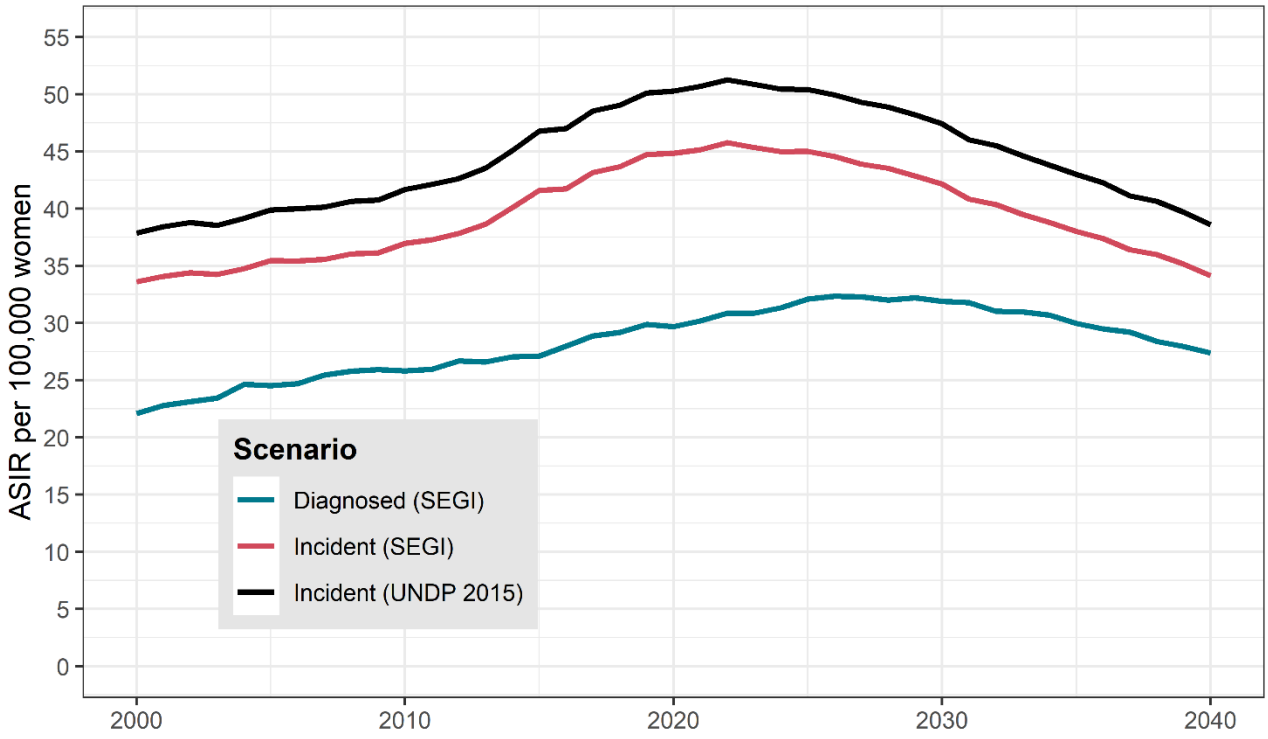


Figure E 1 – Diagnosed and incident cervical cancer calculated using different world standard populations.

Model results are consistent with the stable diagnosed cervical cancer incidence over the 2000-2016 period (section A.9.3). It is important to note here that estimated incidence of cervical cancer is higher and shows a different trend to the *diagnosed* cervical cancer incidence. Several factors contribute to this difference: 1) There is a delay between incidence of cancer and the time of diagnosis (the majority of cancer cases are diagnosed in advanced stages (Table A11)). Women may be in an age-category that carries less weight in the standard population by the time they receive a diagnosis. 2) The population was age-standardised using two different world populations and 3) a small fraction (~5%) of cancer cases in the model die without receiving a diagnosis. The differences are illustrated in Figure E 1.

12. References

1. van Schalkwyk C, Moodley J, Welte A, Johnson LF. Estimated impact of human papillomavirus vaccines on infection burden: The effect of structural assumptions. Vaccine. 2019;37(36).

2. van Schalkwyk C, Moodley J, Welte A, Johnson LF. Are associations between HIV and human papillomavirus transmission due to behavioural confounding or biological effects? Sexually transmitted infections. 2019;95(2):122–8.

3. Johnson LF, Geffen N. A Comparison of Two Mathematical Modeling Frameworks for Evaluating Sexually Transmitted Infection Epidemiology. Sexually transmitted diseases. 2016 Mar;43(3):139–46.

4. Johnson L, Dorrington R. Thembisa version 4.2: A model for evaluating the impact of HIV / AIDS in South Africa [Internet]. 2019 [cited 2020 Jul 20]. Available from: www.thembisa.org

5. 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 Research. 2009;21:289–340.

6. Massyn N, Pillay Y, Padarath A, Editors. District Health Barometer 2017/18. Health Systems Trust; 2019.

7. South African Department of Health. Clinical guidelines for the management of HIV and AIDS in adults and adolescents. 2010;42.

8. Boulle A, Heekes A, Tiffin N, Smith M, Mutemaringa T, Zinyakatira N, et al. Data Centre Profile: The Provincial Health Data Centre of the Western Cape Province, South Africa. International Journal of Population Data Science. 2019 Nov 20;4(2).

9. Johnson LF, Dorrington RE, Moolla H. Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa. Southern African Journal of HIV Medicine. 2017 Jul 27;18(1).

10. Nayar R, Wilbur DC. The Pap Test and Bethesda 2014. Acta Cytologica. 2015 May 19;59(2):121–32.

11. Mamahlodi M, Kuonza L, Candy S. Cervical cancer screening programme in Limpopo province : January 2007 to December 2010. South African Journal of Gynaecological Oncology. 2013;5(1):4–10.

12. Makura C, Schnippel K, Michelow P, Chibwesha C, Goeieman B, Jordaan S, et al. Choropleth Mapping of Cervical Cancer Screening in South Africa Using Healthcare Facility-level Data from the National Laboratory Network. AIMS Public Health. 2016;3(4):849–62.

13. Schnippel K, Michelow P, Chibwesha CJ, Makura C, Lince-Deroche N, Goeieman B, et al. Cost-effectiveness of using the Cervex-Brush (broom) compared to the elongated spatula for collection of conventional cervical cytology samples within a high-burden HIV setting: a model-based analysis. BMC Health Services Research. 2015 Jun 6;15(1):499.

14. Moodley J, Kawonga M, Bradley J, Hoffman M. Challenges in implementing a cervical screening program in South Africa. Cancer Detection and Prevention. 2006;30(4):361–8.

15. Fonn S, Bloch B, Mabina M, Carpenter S, Cronje H, Maise C, et al. Prevalence of pre-cancerous lesions and cervical cancer in South Africa - multicentre study. South African Medical Journal. 2002;92(2).

16. South African Department of Health. National Guideline for Cervical Cancer Screening Programme [Internet]. 2000 [cited 2019 Nov 27]. Available from: http://www.kznhealth.gov.za/cervicalcancer.pdf

17. Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J. Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. Obstetrics and gynecology. 2008 Jan;111(1):167–77.

18. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, et al. Accuracy of the Papanicolaou Test in Screening for and Follow-up of Cervical Cytologic Abnormalities. Journal of Lower Genital Tract Disease. 2001;5(1):60.

19. Wright TC, Denny L, Kuhn L. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. Journal of the American Medical Association. 2000;283(1):81–6.

20. Taylor S, Kuhn L, Dupree W, Denny L, De Souza M, Wright TC. Direct comparison of liquid-based and conventional cytology in a South African screening trial. International Journal of Cancer. 2006;962(118):957–62.

21. Denny L, Kuhn L, Pollack A, Wainwright H, Wright TC. Evaluation of alternative methods of cervical cancer screening for resource-poor settings. Cancer. 2000;89(4):826–33.

22. Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PP, Mustafa RA, et al. Cytology versus HPV testing for cervical cancer screening in the general population. Cochrane Database of Systematic Reviews. 2017 Aug 10;(8).

23. Kitchener H, Nelson L, Adams J, Mesher D, Sasieni P, Cubie H, et al. Colposcopy is not necessary to assess the risk to the cervix in HIV-positive women: An international cohort study of cervical pathology in HIV-1 positive women. International Journal of Cancer. 2007;121(11):2484–91.

24. Firnhaber C, Mayisela N, Mao L, Williams S, Swarts A, Faesen M, et al. Validation of Cervical Cancer Screening Methods in HIV Positive Women from Johannesburg South Africa. Samimi G, editor. PLoS ONE. 2013 Jan 11;8(1):e53494.

25. McDonald AC, Tergas AI, Kuhn L, Denny L, Wright TC. Distribution of Human Papillomavirus Genotypes among HIV-Positive and HIV-Negative Women in Cape Town, South Africa. Frontiers in oncology. 2014;4(March):48.

26. Segondy M, Kelly H, Magooa MP, Djigma F, Ngou J, Gilham C, et al. Performance of careHPV for detecting high-grade cervical intraepithelial neoplasia among women living with HIV-1 in Burkina Faso and South Africa: HARP study. British Journal of Cancer. 2016;115(4):425–30.

27. Kuhn L, Saidu R, Boa R, Tergas A, Moodley J, Persing D, et al. Clinical evaluation of modifications to a human papillomavirus assay to optimise its utility for cervical cancer screening in low-resource settings: a diagnostic accuracy study. The Lancet Global Health. 2020;8(2):e296–304.

28. Arbyn M, Rezhake R, Yuill S, Canfell K. Meta-analysis on the accuracy of methods to triage hrHPV-positive women. In: 33rd International Conference of the Papillomavirus Society: Barcelona (Spain), 20-24 July, 2020.

29. Adam Y, van Gelderen CJ, de Bruyn G, McIntyre J a, Turton D a, Martinson N a. Predictors of persistent cytologic abnormalities after treatment of cervical intraepithelial neoplasia in Soweto, South Africa: a cohort study in a HIV high prevalence population. BMC cancer. 2008;8:211.

30. Mitchell MF, Schottenfeld D, Tortolero-Luna G, Cantor SB, Richards-Kortum R. Colposcopy for the diagnosis of squamous intraepithelial lesions: A meta-analysis. Vol. 91, Obstetrics and Gynecology. 1998. p. 626–31.

31. Cantor SB, Cárdenas-Turanzas M, Cox DD, Atkinson EN, Nogueras-Gonzalez GM, Beck JR, et al. Accuracy of colposcopy in the diagnostic setting compared with the screening setting. Obstetrics and Gynecology. 2008;111(1):7–14.

32. Martin-Hirsch PP, Paraskevaidis E, Bryant A, Dickinson HO, Keep SL. Surgery for cervical intraepithelial neoplasia. In: Martin-Hirsch PP, editor. Cochrane Database of Systematic Reviews. Chichester, UK: John Wiley & Sons, Ltd; 2010.

33. Zeier MD, Nachega JB, Van Der Merwe FH, Eshun-Wilson I, Van Schalkwyk M, La Grange M, et al. Impact of timing of antiretroviral therapy initiation on survival of cervical squamous intraepithelial lesions: a cohort analysis from South Africa. International journal of STD & AIDS. 2012;23(12):890–6.

34. Batra P, Kuhn L, Denny L. Utilisation and outcomes of cervical cancer prevention services among HIV-infected women in Cape Town. South African Medical Journal. 2010;100(1).

35. Noël CJ. Excision margins in Human Immunodeficiency Virus seropositive women undergoing Large Loop Excision of the Transformation Zone for cervical dysplasia. University of the Witwatersrand; 2015.

36. Smith JS, Sanusi B, Swarts A, Faesen M, Levin S, Goeieman B, et al. A randomized clinical trial comparing cervical dysplasia treatment with cryotherapy vs loop electrosurgical excision precedure in HIV-seropositive women from Johannesburg, South Africa. The American Journal of Obstetrics & Gynecology. 2017;217(2):183.e1-183.e11.

37. Kabir F, Gelderen C Van, McIntyre J, Michelow P, Turton D, Adam Y. Cervical intra-epithelial neoplasia in HIV-positive women after excision of the transformation zone – does the grade change ? South African Medical Journal. 2012;102(9):757–60.

38. Kocken M, Helmerhorst TJM, Berkhof J, Louwers JA, Bais AG, Hogewoning CJA, et al. Risk of recurrent high-grade cervical intraepithelial neoplasia after successful treatment: a long-term multi-cohort study. Lancet Oncology. 2011;12:441–50.

39. Kreimer AR, Katki H a., Schiffman M, Wheeler CM, Castle PE. Viral determinants of human papillomavirus persistence following loop electrical excision procedure treatment for cervical intraepithelial neoplasia grade 2 or 3. Cancer Epidemiology Biomarkers and Prevention. 2007;16(January):11–6.

40. Paraskevaidis E, Arbyn M, Sotiriadis A, Diakomanolis E, Martin-hirsch P, Koliopoulos G, et al. The role of HPV DNA testing in the follow-up period after treatment for CIN: a systematic review of the literature. Cancer Treatment Reviews. 2004;30:205–11.

41. Johnson LF, Alkema L, Dorrington RE. A Bayesian approach to uncertainty analysis of sexually transmitted infection models. Sexually transmitted infections. 2010;86(3):169–74.

42. Smith AFM, Gelfand AE. Bayesian statistics without tears: a sampling–resampling perspective. The American Statistician. 1992;46(2):84–8.

43. Mapanga W, Girdler-Brown B, Feresu SA, Chipato T, Singh E. Prevention of cervical cancer in HIV-seropositive women from developing countries through cervical cancer screening: a systematic review. Systematic reviews. 2018;7(1):198.

44. Cronjé HS, Parham GP, Cooreman BF, de Beer A, Divall P, Bam RH. A comparison of four screening methods for cervical neoplasia in a developing country. American journal of obstetrics and gynecology. 2003;188(2):395–400.

45. Giuliano AR, Botha MH, Zeier M, Abrahamsen ME, Glashoff RH, van der Laan LE, et al. High HIV, HPV, and STI prevalence among young Western Cape, South African women: EVRI HIV prevention preparedness trial. Journal of acquired immune deficiency syndromes (1999). 2015;68(2):227–35.

46. Snyman LC, Dreyer G, Botha MH, van der Merwe FH, Becker PJ. The Vaccine and Cervical Cancer Screen (VACCS) project: Linking cervical cancer screening to HPV vaccination in the South-West District of Tshwane, Gauteng, South Africa. South African medical journal. 2015 Jan 6;105(2):115–20.

47. Snyman LC, Dreyer G, Visser C, Botha MH, Van der Merwe FH. The Vaccine and Cervical Cancer Screen project 2 (VACCS 2): Linking cervical cancer screening to a two-dose HPV vaccination schedule in the South-West District of Tshwane, Gauteng, South Africa. South African Medical Journal. 2015;105(3):191.

48. Adler DH, Wallace M, Bennie T, Mrubata M, Abar B, Meiring TL, et al. Cervical dysplasia and high-risk human papillomavirus infections among HIV-infected and HIV-uninfected adolescent females in South Africa. Infectious diseases in obstetrics and gynecology. 2014;2014:498048.

49. Mbulawa ZZA, van Schalkwyk C, Hu NC, Meiring TL, Barnabas S, Dabee S, et al. High human papillomavirus (HPV) prevalence in South African adolescents and young women encourages expanded HPV vaccination campaigns. PloS one. 2018;13(1):e0190166.

50. Mbulawa ZZA, Marais DJ, Johnson LF, Boulle A, Coetzee D, Williamson AL. Influence of human immunodeficiency virus and CD4 count on the prevalence of human papillomavirus in heterosexual couples. The Journal of general virology. 2010 Dec;91(Pt 12):3023–31.

51. Denny L, Boa R, Williamson AL, Allan B, Hardie D, Stan R, et al. Human papillomavirus infection and cervical disease in human immunodeficiency virus-1-infected women. Obstetrics and gynecology. 2008 Jun;111(6):1380–7.

52. Liebenberg LJP, McKinnon LR, Yende-Zuma N, Garrett N, Baxter C, Kharsany ABM, et al. HPV infection and the genital cytokine milieu in women at high risk of HIV acquisition. Nature Communications. 2019;10(1):1–12.

53. Vardas E, Giuliano AR, Goldstone S, Palefsky JM, Moreira ED, Penny ME, et al. External genital human papillomavirus prevalence and associated factors among heterosexual men on 5 continents. The Journal of infectious diseases. 2011 Jan 1;203(1):58–65.

54. Chikandiwa A, Chimoyi L, Pisa PT, Chersich MF, Muller EE, Michelow P, et al. Prevalence of anogenital HPV infection, related disease and risk factors among HIV-infected men in inner-

city Johannesburg, South Africa: baseline findings from a cohort study. BMC public health. 2017 Jul 4;17(Suppl 3):425.

55. Moodley JR, Constant D, Hoffman M, Salimo A, Allan B, Rybicki E, et al. Human papillomavirus prevalence, viral load and pre-cancerous lesions of the cervix in women initiating highly active antiretroviral therapy in South Africa: a cross-sectional study. BMC cancer. 2009 Jan;9:275.

56. Firnhaber C, Zungu K, Levin S, Michelow P, Montaner LJ, McPhail P, et al. Diverse and high prevalence of human papillomavirus associated with a significant high rate of cervical dysplasia in human immunodeficiency virus-infected women in Johannesburg, South Africa. Acta cytologica. 2009;53(1):10–7.

57. Mbulawa ZZA, Hu NC, Kufa-Chakezha T, Kularatne R, Williamson AL. Sentinel surveillance of human papillomavirus genotypes among patients attending public healthcare facilities in South Africa, 2014-2016 133. Vol. 14, Communicable Diseases Surveillance Bulletin. 2016. 133–136 p.

58. Somdyala NIM, Bradshaw D, Dhansay MA, Stefan DC. Increasing Cervical Cancer Incidence in Rural Eastern Cape Province of South Africa From 1998 to 2012: A Population-Based Cancer Registry Study. JCO Global Oncology. 2020;1–8.

59. National Cancer Registry South Africa. Ekurhuleni population-based cancer registry Annual Report 2018. 2020.

60. Creasman W. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. International journal of gynaecology and obstetrics. 2009 May;105(2):109.

61. Lomalisa P, Smith T, Guidozzi F. Human Immunodeficiency Virus Infection and Invasive Cervical Cancer in South Africa. Gynecologic oncology. 2000;77:460–3.

62. Mbodi L, Adam Y. Reasons Why Women present with late stages of Cervical Cancer at Chris Hani Baragwanath Academic Hospital. University of the Witwatersrand; 2016.

63. Snyman L, Herbst U. Reasons why unscreened patients with cervical cancer present with advanced stage disease. South African Journal of Gynaecological Oncology. 2013;5(1):16–20.

64. Sabulei C, Maree J. An exploration into the quality of life of women treated for cervical cancer. Curationis. 2019;42(1):1–9.

65. Johnson LF. Access to antiretroviral treatment in South Africa, 2004 - 2011. Southern African Journal of HIV Medicine. 2012 Mar 13;13(1):22–7.

66. Strickler HD, Burk RD, Fazzari M, Anastos K, Minkoff H, Massad LS, et al. Natural history and possible reactivation of human papillomavirus in human immunodeficiency virus-positive women. Journal of the National Cancer Institute. 2005 Apr 20;97(8):577–86.

67. Ahdieh L, Klein RS, Burk R, Cu-Uvin S, Schuman P, Duerr A, et al. Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. The Journal of infectious diseases. 2001 Sep 15;184(6):682–90.

68. Theiler RN, Farr SL, Karon JM, Paramsothy P, Viscidi R, Duerr A, et al. High-risk human papillomavirus reactivation in human immunodeficiency virus-infected women: risk factors for cervical viral shedding. Obstetrics and gynecology. 2010 Jun;115(6):1150–8.

69. Matthijsse SM, Van Rosmalen J, Hontelez JAC, Bakker R, De Kok IMCM, Van Ballegooijen M, et al. The role of acquired immunity in the spread of human papillomavirus (HPV): Explorations with a microsimulation model. PLoS ONE. 2015;10(2):1–14.

70. Van de Velde N, Boily MC, Drolet M, Franco EL, Mayrand MH, Kliewer E V., et al. Population-level impact of the bivalent, quadrivalent, and nonavalent human papillomavirus vaccines: a model-based analysis. Journal of the National Cancer Institute. 2012 Nov 21;104(22):1712–23.

71. Trottier H, Mahmud S, Prado JCM, Sobrinho JS, Costa MC, Rohan TE, et al. Type-specific duration of human papillomavirus infection: implications for human papillomavirus screening and vaccination. The Journal of infectious diseases. 2008 May 15;197(10):1436–47.

72. Richardson H, Kelsall G, Tellier P, Voyer H, Abrahamowicz M, Ferenczy A, et al. The natural history of type-specific human papillomavirus infections in female university students. Cancer epidemiology, biomarkers & prevention. 2003 Jun;12(6):485–90.

73. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. The New England journal of medicine. 1998 Feb 12;338(7):423–8.

74. Woodman CB, Collins S, Winter H, Bailey A, Ellis J, Prior P, et al. Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. Lancet (London, England). 2001 Jun 9;357(9271):1831–6.

75. Goodman MT, Shvetsov YB, McDuffie K, Wilkens LR, Zhu X, Thompson PJ, et al. Prevalence, acquisition, and clearance of cervical human papillomavirus infection among women with normal cytology: Hawaii Human Papillomavirus Cohort Study. Cancer research. 2008 Nov 1;68(21):8813–24.

76. Muñoz N, Méndez F, Posso H, Molano M, van den Brule AJC, Ronderos M, et al. Incidence, duration, and determinants of cervical human papillomavirus infection in a cohort of Colombian women with normal cytological results. The Journal of infectious diseases. 2004 Dec 15;190(12):2077–87.

77. Insinga RP, Dasbach EJ, Elbasha EH, Liaw KL, Barr E. Incidence and duration of cervical human papillomavirus 6, 11, 16, and 18 infections in young women: an evaluation from multiple analytic perspectives. Cancer epidemiology, biomarkers & prevention. 2007 Apr;16(4):709–15.

78. Insinga RP, Perez G, Wheeler CM, Koutsky LA, Garland SM, Leodolter S, et al. Incidence, duration, and reappearance of type-specific cervical human papillomavirus infections in young women. Cancer epidemiology, biomarkers & prevention. 2010 Jun;19(6):1585–94.

79. Jaisamrarn U, Castellsagué X, Garland SM, Naud P, Palmroth J, Del Rosario-Raymundo MR, et al. Natural history of progression of HPV infection to cervical lesion or clearance: analysis of the control arm of the large, randomised PATRICIA study. PloS ONE. 2013;8(11):e79260.

80. Giuliano AR, Lee JH, Fulp W, Villa LL, Lazcano E, Papenfuss MR, et al. Incidence and clearance of genital human papillomavirus infection in men (HIM): a cohort study. Lancet (London, England). 2011 Mar 12;377(9769):932–40.

81. Muñoz N, Hernandez-Suarez G, Méndez F, Molano M, Posso H, Moreno V, et al. Persistence of HPV infection and risk of high-grade cervical intraepithelial neoplasia in a cohort of Colombian women. British journal of cancer. 2009 Apr 7;100(7):1184–90.

82. Insinga RP, Perez G, Wheeler CM, Koutsky LA, Garland SM, Leodolter S, et al. Incident Cervical HPV Infections in Young Women: Transition Probabilities for CIN and Infection Clearance. Cancer epidemiology, biomarkers & prevention. 2011;(11):287–97.

83. Skinner SR, Wheeler CM, Romanowski B, Castellsagu X, Rosario-raymundo MR Del, Vallejos C, et al. Progression of HPV infection to detectable cervical lesions or clearance in adult women : Analysis of the control arm of the VIVIANE study. International Journal of Cancer. 2016;2438:2428–38.

84. Moscicki A barbara, Hills N, Shiboski S, Powell K, Jay N, Hanson E, et al. Risks for Incident Human Papillomavirus Infection and Low-Grade Squamous Intraepithelial Lesion Development in Young Females. Journal of the American Medical Association. 2001;285(23):2995–3002.

85. Tainio K, Athanasiou A, Tikkinen KAO, Aaltonen R, Cárdenas J, Hernándes, et al. Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis. BMJ. 2018 Feb 27;360:k499.

86. Liu M, Yan X, Zhang M, Li X, Li S, Jing M. Influence of Human Papillomavirus Infection on the Natural History of Cervical Intraepithelial Neoplasia 1: A Meta-Analysis. BioMed Research International. 2017;2017:1–9.

87. Roura E, Travier N, Waterboer T, de Sanjosé S, Bosch FX, Pawlita M, et al. The Influence of Hormonal Factors on the Risk of Developing Cervical Cancer and Pre-Cancer: Results from the EPIC Cohort. PloS one. 2016;11(1):e0147029.

88. McCredie MRE, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. The Lancet Oncology. 2008 May;9(5):425–34.

89. Canfell K, Barnabas R, Patnick J, Beral V. The predicted effect of changes in cervical screening practice in the UK: Results from a modelling study. British Journal of Cancer. 2004;91(3):530–6.

90. Tan N, Sharma M, Winer R, Galloway D, Rees H, Barnabas R V. Model-estimated effectiveness of single dose 9-valent HPV vaccination for HIV-positive and HIV-negative females in South Africa. Vaccine. 2018;36(32):4830–6.

91. Denslow S, Rositch A, Firnhaber C, Ting J, Smith J. Incidence and progression of cervical lesions in women with HIV: a systematic global review. International journal of STD & AIDS. 2014 Mar;25(3):163–77.

92. Liu G, Sharma M, Tan N, Barnabas R V. HIV-positive women have higher risk of human papilloma virus infection, precancerous lesions, and cervical cancer. AIDS (London, England). 2018 Mar 27;32(6):795–808.

93. Kelly H, Weiss HA, Benavente Y, de Sanjose S, Mayaud P, Qiao Y lin, et al. Association of antiretroviral therapy with high-risk human papillomavirus, cervical intraepithelial neoplasia, and invasive cervical cancer in women living with HIV: a systematic review and meta-analysis. The Lancet HIV. 2018;5(1):e45–58.

94. Rohner E, Bütikofer L, Schmidlin K, Sengayi M, Maskew M, Giddy J, et al. Cervical cancer risk in women living with HIV across four continents: A multicohort study. International Journal of Cancer. 2019 Jun 19;ijc.32260.

95. Nobbenhuis MAE, Helmerhorst TJM, Van Den Brule AJC, Rozendaal L, Voorhorst FJ, Bezemer PD, et al. Cytological regression and clearance of high-risk human papillomavirus in women with an abnormal cervical smear. Lancet. 2001;358(9295):1782–3.

96. Schiffman M, Wheeler CM, Castle PE. Human Papillomavirus DNA Remains Detectable Longer than Related Cervical Cytologic Abnormalities. The Journal of Infectious Diseases. 2002;186(8):1169–72.

97. Zielinski DG, Snijders PJF, Rozendaal L, Voorhorst FJ, Runsink AP, De Schipper FA, et al. High-risk HPV testing in women with borderline and mild dyskaryosis: Long-term follow-up data and clinical relevance. Journal of Pathology. 2001;195(3):300–6.

98. Syrjänen S, Shabalova IP, Petrovichev N, Kozachenko VP, Zakharova T, Pajanidi A, et al. Clearance of high-risk human papillomavirus (HPV) DNA and PAP smear abnormalities in a cohort of women subjected to HPV screening in the New Independent States of the former Soviet Union (the NIS cohort study). European Journal of Obstetrics and Gynecology and Reproductive Biology. 2005;119(2):219–27.

99. Myers ER, McCrory DC, Nanda K, Bastian L, Matchar DB. Mathematical model for the natural history of human papillomavirus infection and cervical carcinogenesis. American journal of epidemiology. 2000 Jun 15;151(12):1158–71.

100. Brisson M, Laprise, Jean-François Drolet M, Van de Velde, Nicolas Boily MC. HPV-Advise: Technical Appendix [Internet]. [cited 2019 Jan 18]. Available from: http://www.marc-brisson.net/HPVadvise.pdf

101. Moodley M, Moodley J, Kleinschmidt I. Invasive cervical cancer and human immunodeficiency virus (HIV) infection: a South African perspective. International Journal of Gynecologic Cancer. 2001 May 1;11(3):194–7.

102. Lomalisa P, Smith T, Guidozzi F. Human immunodeficiency virus infection and invasive cervical cancer in South Africa. Gynecologic Oncology. 2000;77(3):460–3.

103. McDonald AC, Denny L, Wang C, Tsai WY, Wright TC, Kuhn L. Distribution of high-risk human papillomavirus genotypes among HIV-negative women with and without cervical intraepithelial neoplasia in South Africa. PloS one. 2012 Jan;7(9):e44332.

104. Olorunfemi G, Ndlovu N, Masukume G, Chikandiwa A, Pisa PT, Singh E. Temporal trends in the epidemiology of cervical cancer in South Africa (1994-2012). International Journal of Cancer. 2018 Nov 1;143(9):2238–49.

105. Olorunfemi G, Ndlovu N, Masukume G, Chikandiwa A, Pisa PT, Singh E. Temporal trends in the epidemiology of cervical cancer in South Africa (1994–2012). International Journal of Cancer. 2018;143(9):2238–49.

106. Denny L, Adewole I, Anorlu R, Dreyer G, Moodley M, Smith T, et al. Human papillomavirus prevalence and type distribution in invasive cervical cancer in sub-Saharan Africa. International journal of cancer. 2014;134(6):1389–98.

107. van Aardt MC, Dreyer G, Pienaar HF, Karlsen F, Hovland S, Richter KL, et al. Unique human papillomavirus-type distribution in South African women with invasive cervical cancer and the effect of human immunodeficiency virus infection. International journal of gynecological cancer. 2015;25(5):919–25.

108. Stelzle D, Tanaka LF, Lee KK, Khalil AI, Baussano I, Mcallister DA, et al. Estimates of the global burden of cervical cancer associated with HIV. Lancet Global Health. 2020;(In press).

109. Moodley J, Hoffman M, Carrara H, Allan B, Cooper D, Rosenberg L, et al. HIV and pre-neoplastic and neoplastic lesions of the cervix in South Africa: a case-control study. BMC Cancer. 2006;6:1–6.

110. Van Bogaert LJJ. Age at Diagnosis of Preinvasive and Invasive Cervical Neoplasia in South Africa: HIV-positive versus HIV-negative women. International Journal of Gynecological Cancer. 2011;21(2):363–6.

111. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. International Journal of Cancer. 2019;144(8):1941–53.

112. South African Department of Health, South African National AIDS Council. South African HIV and TB Investment Case - Summary Report Phase 1 [Internet]. 2016 [cited 2020 Jul 2]. Available from: http://www.heroza.org/wp-content/uploads/2016/03/SA-HIV_TB-Investment-Case-Full-Report-Low-Res.pdf

113. Marsh K, Eaton JW, Mahy M, Sabin K, Autenrieth CS, Wanyeki I, et al. Global, regional and country-level 90-90-90 estimates for 2018: Assessing progress towards the 2020 target. AIDS. 2019;33(April 2019):S213–26.

114. Brisson M, Kim JJ, Canfell K, Drolet M, Gingras G, Burger EA, et al. Impact of HPV vaccination and cervical screening on cervical cancer elimination: a comparative modelling analysis in 78 low-income and lower-middle-income countries. The Lancet. 2020;395(10224):575–90.