Modelling the impact of prevention strategies on cervical cancer incidence in South Africa

by
Catherina van Schalkwyk
Thesis presented for the degree of
DOCTOR OF PHILOSOPHY
in the School of Public Health and Family Medicine
Faculty of Health Sciences, University of Cape Town
December 2020

Supervisors:
Dr Leigh Johnson
Prof Jennifer Moodley

ABSTRACT

Background: In 2019, the World Health Organisation called for the elimination of cervical cancer as a public health concern. In South Africa, despite having a national screening policy in place since 2000, diagnosed cervical cancer incidence has shown no signs of decline. Since 2014, girls aged 9 have been vaccinated against HPV infection using the bivalent vaccine, with high coverage. However, due to the long delay between HPV infection and progression to cancer, the impact that vaccination will have on cervical cancer incidence will be unobservable in the near future. This thesis sets out to quantify this impact using a mathematical model, and will estimate the impact of scaling up current cancer prevention strategies, as well as proposed alternative strategies.

Methods: This research extends a previously developed individual-based model for HIV to include infection with 13 high-risk HPV types and progression to cervical cancer. HPV infection and cervical disease parameters were calibrated to a wide range of South African data sources using a likelihood based approach. In the process of developing an appropriate model for cervical cancer incidence in South Africa, important aspects related to HIV/HPV co-infection dynamics, the natural history of HPV and the current and historic levels of cervical cancer prevention in the Western Cape were investigated. The calibrated and validated model was used to estimate the impact of current and proposed alternative prevention strategies on cervical cancer incidence in the next century.

Findings: Using a model structure that does not include a biological transmission co-factor, we show that simulated associations between HIV and HPV transmission are similar to corresponding empirical estimates and therefore these associations may result from residual confounding by sexual behavioural factors and network-level effects. Using simulated vaccine trials, we show that viral latency and reactivation of latent infections is necessary in the natural history of HPV to match results from empirical trials. The model's screening algorithm reflects findings from the Western Cape's public health sector – low levels of screening coverage and linkage to treatment facilities, and poor adherence to screening schedules. The model matches stable trends in *diagnosed* cervical cancer incidence in South Africa, but it estimates increases in cervical cancer incidence over the last number of years (due to increased life expectancy of women on ART), which will result in sharp increases in diagnoses. While decreasing HIV prevalence and HPV vaccination will substantially reduce cervical cancer incidence in the long term, improvements in South Africa's current screening strategy, as well as switching to new screening technologies, will have significant impact in the short term.

Conclusions: This thesis presents an epidemiological model of cervical cancer in South Africa – the first to dynamically simulate infection with both HIV and HPV at national level. It allows for estimation of the impact of both HIV and cervical cancer prevention on cancer incidence, and provides the opportunity to identify the vaccination and screening strategies with the greatest public health significance.

ACKNOWLEDGEMENTS

This has been a challenging and fulfilling journey and I have many people to thank for helping me along the way.

- I really hit the supervisor jackpot. Leigh Johnson leads by example, does not miss a thing, is incredibly patient and inspires me to be the best I can be. I have learned more from him than I ever could have imagined, and I know that we have not even scratched the surface. Jennifer Moodley was amazingly supportive throughout the process, and truly someone to look up to. Her contributions to this work are invaluable.
- My SACEMA family. I am extremely grateful to everyone who has come and gone at SACEMA through the years. I am sure I have learned a skill from every single one. John Hargrove, for the incredible cheerleading, life lessons and Basic breakfasts; Alex Welte, for (always!) challenging me to think; Juliet Pulliam, for making me feel valued and being an amazing mentor; Gavin Hitchcock, for always inspiring us to be better people, in all aspects of life; Faikah Bruce, for being awesome. Special thanks to Reshma, Roxy, Wim and Brian for your friendship and support.
- My ICI3D family. Attending a clinic instructed by Steve Bellan, Juliet Pulliam, Jonathan Dushoff, Brian Williams, John Hargrove, Travis Porco, Jim Scott and Carl Pearson will inevitably suck you into the magical world of infectious disease modelling. They are a group of remarkable instructors, and it is an honour to learn from them, and now also teach with them.
- The team at the Provincial Health Data Centre at the Western Cape Department of Health has done incredible work to make the routine data tell a story. I am extremely grateful to Mariette Smith, Mieke Willemse, Florence Phelanyane, Rory Pienaar and Andrew Boulle for allowing me to access to this data and for fielding my endless questions.
- The group I led at the 2019 MMED clinic developed the first version of the simulation model to estimate age-specific probabilities of entering cervical cancer screening. Special thanks to Gui Liu, Joseph Sempa, Sylvie Djiomba and Kossi Bekou for figuring it out with me.
- I am extremely grateful to many others who have given important advice to help understand the natural history of HPV, what the data mean and what actually happens in the clinics and labs of our public health sector. Profs Hennie Botha and Patti Gravitt deserve a special mention.
- For the last few months I have worked with two other teams in the WHO's Cervical Cancer Elimination Modelling Consortium and their input has vastly improved the model that I present in this thesis. Thanks to Shona Dalal, Marie-Claude Boily, Ruanne Barnabas, Minttu Ronn, Darcy Rao, Cara Bayer and Gui Liu.

- I gratefully acknowledge the Centre for High Performance Computing (CHPC) for providing computational resources to this research project. The study received financial support from the Cancer Association of South Africa, the South African Department of Science and Technology and National Research Foundation, and the World Health Organization.
- A big thank you to my extended family and friends who have cheered me on through this journey. Thank you to my parents for cooking frozen dinners, doing much more babysitting than most grandparents sign up for, and for always checking in and having my back. Thank you to Dirk, who has patiently waited to have his wife back, and my kids who haven't met the PhD free me!

PREFACE

This thesis is presented in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) in the School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town. The work included in this thesis is original research and has not, in whole or in part, been submitted for another degree at this or any other university.

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publication(s) in my PhD thesis, and where co-authorships are involved, my co-authors have agreed that I may include the publication(s):

1. Van Schalkwyk C, Moodley J, Welte A, Johnson LF. (2019) Are associations between HIV and HPV transmission due to behavioural confounding factors or biological effects? Sexually Transmitted Infections 95(2):122-128
2. Van Schalkwyk C, Moodley J, Welte A, Johnson LF. (2019) Estimated impact of human papillomavirus vaccines on infection burden: the effect of structural assumptions. Vaccine 37 (36), 5460-5465
3. Van Schalkwyk C, Phelanyane F, Smith M, Pienaar C, Willems M, Moodley J, Welte A, Boulle A, Johnson LF. (draft) Cervical cancer surveillance in South Africa: lessons from the public sector of the Western Cape.
4. Van Schalkwyk C, Moodley J, Welte A, Johnson LF. (In press) Modelling the impact of prevention strategies on cervical cancer incidence in South Africa. International Journal of Cancer

My contribution is described in the introduction to each paper. I was lead and corresponding author on all papers, developed all of the HPV and cervical disease components of the model, conducted the analyses for all articles and wrote all the first drafts of the articles. I circulated the manuscripts to co-authors, reviewed their suggestions and comments and made the final decisions regarding further revisions to the manuscripts. All co-authors reviewed and approved the submitted manuscripts. I was primarily responsible for responding to reviewer comments and circulating these to co-authors. My supervisor has confirmed to the University of Cape Town Doctoral Degrees Board that the included papers reflect overwhelmingly my own scientific work.

Signed by candidate

Catherina van Schalkwyk

LIST OF ABBREVIATIONS

AIDS	Acquired immune deficiency syndrome
ART	Antiretroviral therapy
ASIR	Age standardised incidence rate
CC	Cervical cancer
CCEMC	Cervical cancer elimination modelling consortium
CI	Confidence interval
CIN	Cervical intraepithelial neoplasia
DHB	District Health Barometer
DoH	Department of Health
FIGO	International Federation of Gynecology and Obstetrics
HIV	Human immunodeficiency virus
H/LSIL	High/low grade intraepithelial lesion
HPV	Human papillomavirus
HR-HPV	High risk human papillomavirus
HR	Hazard ratio
IARC	International Agency for Research on Cancer
IeDEA	International epidemiologic Databases to Evaluate AIDS
IQR	interquartile range
LLETZ	Large loop excision of the transformation zone
LMIC	Lower- and middle-income country
MicroCOSM	Microsimulation for the Control of South African Morbidity and Mortality
MMC	Medical Male Circumcision
NCR	National Cancer Registry
NHLS	National Health Laboratory Service
OECD	Organization for Economic Co-operation and Development
Pap	Papanicolaou
PHDC	Provincial Health Data Centre
PMI	Patient master index
RCT	Randomised control trial
SA	South Africa
STI	Sexually transmitted infection
VIA	Visual inspection with acetic acid
WC	Western Cape
WHO	World Health Organization

TABLE OF CONTENTS

Chapter 1 – Introduction

1.1 Background

1.1.1 Global context

Cervical cancer (CC) is the fourth most common cancer among females, with 570 000 new cases reported globally in 2018 (1). Around half of these cases result in death and nearly 90% of deaths occur in lower- and middle-income countries (LMICs). Across Sub-Saharan Africa, CC is the second most common type of cancer, but the leading cause of cancer death (2). CC is caused by persistent infection with oncogenic (or 'high risk') types of the sexually transmitted human papillomavirus (HPV) (3), with types 16 and 18 accounting for around 70% of cases (4).

Cervical cancer is a preventable disease – either through primary prevention against HPV by vaccination, condom use or male circumcision, or by secondary prevention through early detection and treatment of pre-cancerous lesions. Current official World Health Organisation (WHO) guidelines recommend that every woman should undergo screening at least once between the ages of 30 and 50 and that screenings should be around 5 years apart (5). In high income countries, widespread screening with Papanicolaou (Pap) smear has led to significant declines in CC incidence, despite varying levels of sensitivity (30-90%) (6,7). This method has high specificity (86-100%), which limits unnecessary treatment. HPV-DNA testing is a screening method with higher sensitivity (over 90% to detect higher grade lesions in women with atypical cells on cytology), but much lower specificity (~60%) and is used as a screening/triage method in many high income countries (8,9).

Primary prevention through two HPV vaccines that protect against types 16 and 18 – the two types that cause around 70% of cervical cancer worldwide (4) - has been available since 2006. The bivalent vaccine (Cervarix, GlaxoSmithKline) protects against only types 16 and 18, while the quadrivalent vaccine (Gardasil, Merck & Co.) also protects against types 6 and 11, which are associated with genital warts. Some evidence of cross-protection has been shown against the oncogenic types 31, 33, 45 and 52 (10). A decade after the approval of these two vaccines, Gardasil9 was approved (another vaccine manufactured by Merck & Co.) that protects against the same types as the quadrivalent vaccine, as well as types 31, 33, 45, 52 and 58. These five HPV types, along with type 16 and 18, cause around 90% of cervical cancer worldwide (4).

The current WHO recommendation states that countries should include HPV vaccination in national vaccination programmes if CC is seen as a public health priority and feasibility and affordability have been established. Vaccination strategies should focus on girls aged 9-13 and at least 2 doses of either the bivalent or quadrivalent vaccines should be administered, at least 6 months apart. Since

vaccination will not have an impact on cervical cancer incidence in the short term, these guidelines stress that screening programmes should stay in place for the foreseeable future (5).

In January 2019, the WHO called for the elimination of cervical cancer as a public health threat, defined as below 4 incident cases per 100,000 women (1). It was shown, through a collaborative modelling approach, that elimination can be achieved in LMICs by the end of the century if, by 2030, the following targets can be achieved and maintained (1,11):

1. 90% vaccination coverage among pre-adolescent girls – using the nonavalent vaccine,
2. 70% coverage of cervical cancer screening at ages 35 and 45 using a highly sensitive diagnostic test, and
3. 90% linkage to appropriate treatment for those who screen positive.

Throughout this thesis the strategy will be referred to as the **90-70-90 targets**.

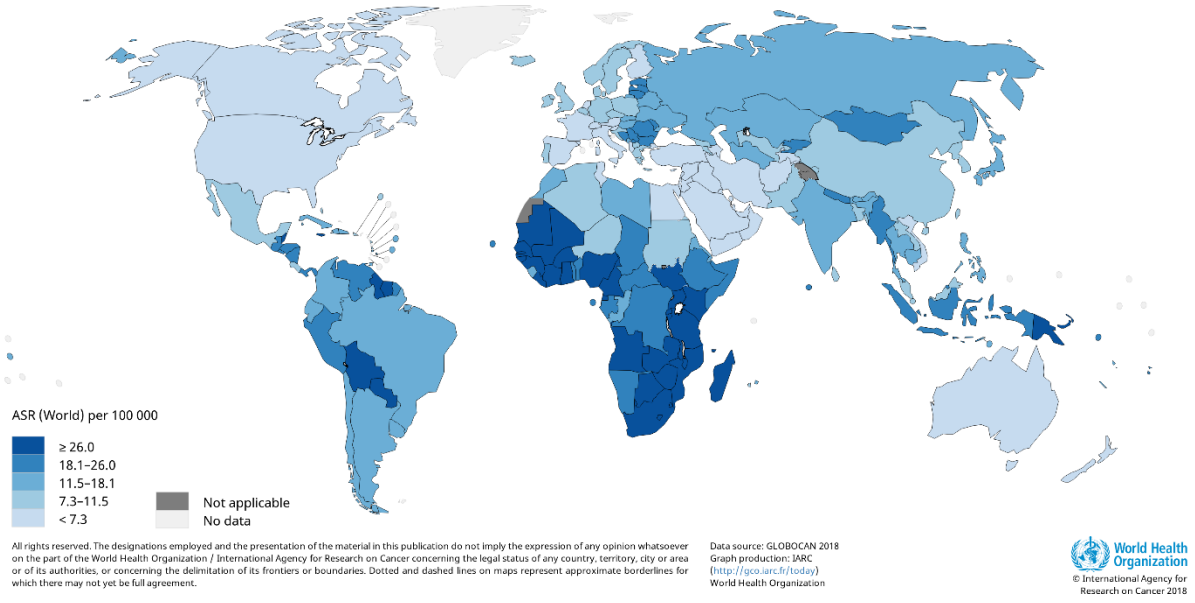


Figure 1 – Estimated age-standardised cervical cancer incidence rates in 2018 (2).

LMICs are disproportionately affected by cervical cancer, and it is clear from Figure 1 that Sub-Saharan Africa carries the highest burden. The link between HPV, cervical cancer and the human immunodeficiency virus (HIV) has been well established. Since 1993, invasive CC has been classified as an acquired immune deficiency syndrome (AIDS) defining cancer, along with Non-Hodgkin's Lymphoma and Kaposi's Sarcoma (12). HIV-positive women have higher HPV prevalence, have higher risk of persistent HPV infection and are more likely to progress to cervical disease than HIV-negative women (13–18). Studies from several countries have shown that HIV-positive women are more commonly infected with non-vaccine type high risk HPV infections, such as types 52 and 58

(19). Although the biological mechanisms involved have not been well described, meta-analyses have also suggested that prevalent HPV infection doubles the risk of HIV acquisition (20,21). These co-infection dynamics might play an important role in estimating the impact of HPV and cervical cancer intervention strategies in countries with high HIV prevalence.

Antiretroviral therapy (ART) has a clear impact on the reconstitution of an HIV-positive person's immune system and the effect of ART on the natural history of HPV is also beneficial. A meta-analysis of several studies that estimated the relative rates of incidence, regression and progression of pre-cancerous lesions between HIV-positive women on ART and ART naïve found that benefits of ART are significant (22). However, due to higher HPV prevalence in HIV-positive women in general, women on ART are still at high risk of cervical cancer. A recent multi-national analysis of women on ART showed that CC incidence is much higher in these women than in the general population of their countries (23).

1.1.2 South African context

Cervical cancer is the second most common cancer and the leading cause of cancer mortality among women in South Africa. The International Agency for Research on Cancer (IARC) estimates that almost 13,000 new cases of cervical cancer occurred in 2018 and that 5600 women died as a result (2). The main source of cancer incidence data in the country is the pathology-based National Cancer Registry (NCR). Although this registry only represents cancer cases that were diagnosed by biopsy or cytology and does not count cases that only received a clinical diagnosis (no pathological confirmation), it is useful for identifying time trends and age patterns. In 2016, cervical cancer reported to the NCR was the most common cancer among young women aged 15-44 and age-standardised incidence for black women was almost double the rate of white women (24). This reflects differences in socio economic status and access to screening services.

In addition to the pathology-based NCR, two small population-based cancer registries exist in the country. In a rural, low-income area of the Eastern Cape with a catchment of around one million people, diagnosed age-standardised cervical cancer incidence has increased from 22 per 100,000 women in 1998-2002 to 29.2 per 100,000 women in 2008-2012. In this registry, 14% of CC cases only received a clinical diagnosis and did not get a laboratory confirmed diagnosis. The Ekurhuleni population-based cancer registry was established in 2018, and the first estimate of CC incidence from this registry is 24.3 per 100,000 women. This highly urbanised health district serves more than 3 million people of varying socio-economic status. In this area, only 7% of CC cases had no laboratory confirmation of cancer.

In South Africa, as in many parts of Sub-Saharan Africa, cervical cancer incidence is fuelled by a generalised HIV epidemic, with prevalence of 26.3% in women of reproductive age (15-49) and overall ART coverage of 62.3% in 2017 (25). Local studies have confirmed the strong association

between HIV and HPV. Prevalence of high-risk HPV infections and of high grade lesions are 2-3 times higher among HIV-positive women than among HIV-negative women (14,26–28) and studies in Gauteng, KwaZulu-Natal and Limpopo have shown that HIV-positive women presented with CC 10 to 15 years earlier than HIV-negative women (29–31). Recently, cervical cancer cases recorded in the National Health Laboratory Service (NHLS) database (i.e. pathology confirmed cancers) were linked to HIV data in the same database (HIV diagnostic tests, CD4 counts and viral loads) using names, surnames, dates of birth and ID numbers (where available) (32). The study showed that the odds ratio for cervical cancer versus other cancers among HIV-positive compared to HIV-negative women increased from 0.6 in 2004 to 3.2 in 2014, indicating that increased life expectancy due to ART increases HIV-positive women's risk of developing and being diagnosed with cervical cancer.

The South African National Department of Health (NDoH) established a CC screening programme in 2000 and the policy recommended three free Papanicolaou (Pap) smears in a woman's lifetime, starting at 30 years of age and at 10 yearly intervals (33). In 2017, the policy was updated and contains detailed guidelines on screening according to HIV status, and algorithms for follow-up and treatment (34). The recommendation on age and interval in the 2000 policy now applies to HIV-negative women, while HIV-positive women should be screened at HIV diagnosis, and at three-yearly intervals thereafter. National coverage of Pap smear screening, defined as the total number of routine smears divided by a 10^{th} of the eligible population, was estimated to be 65.1% in 2018. This fraction varies widely across provinces and districts (35) and the estimate is biased in several ways. Firstly, due to the lack of unique health identifier in South Africa, HIV-negative women who do not adhere to the ten yearly schedule may be counted more than once. HIV status is not consistently captured at the time of the screen, and therefore coverage cannot be disaggregated according to HIV status. This may also lead to over-counting in the coverage numerator, since HIV-positive women *should* be screened more than once in 10 years. Lastly, a tenth of the entire female population older than 30 is used as the denominator, even though a fraction of females are screened in the private sector. This biased estimate is the only element in the cervical cancer prevention cascade that is routinely reported. Data on the element that will have an impact on cervical cancer incidence – the linkage of women with indicative smears to treatment facilities – is not routinely reported, but research studies have shown very low levels of follow-up (36–39).

In high income countries, where screening programmes were initiated a decade or more before South Africa, CC incidence trends show significant reductions (7), while the NCR has demonstrated no decline between 2001 and 2016 (24,40). In addition, CC related mortality has shown no significant change during the same period (41). Some reasons for the lack of decline in CC incidence may be 1) increased lab testing of cancer cases over time; or 2) increased life expectancy of HIV-positive women because of ART; or 3) failure of the screening programme to screen, refer and treat women who have high grade pre-cancer timeously.

Following a number of feasibility and acceptability studies (42), and after significant price reductions from GlaxoSmithKline, the NDoH rolled out an HPV vaccination programme in April 2014. The bivalent vaccine, Cervarix, is administered according to a two-dose schedule to Grade 4 girls, aged 9 and older, in primary schools across the country. Uptake of the first round of vaccination was very high at 86.6% coverage in 2014 (42) and consistently high coverage of the first dose has been maintained. However, latest estimates show that only around 60% of girls aged 9 received both doses of the vaccine in recent years (43). The vaccination programme should have a substantial impact on cervical cancer incidence in the long term, as will be investigated in this thesis.

1.1.3 Epidemiological modelling

Due to the long time between acquiring an HPV infection and developing cervical cancer, the only short-term feasible way to estimate the long-term impact of prevention strategies on CC incidence is through mathematical modelling. Globally, many mathematical models have been developed to illustrate this and most have stressed the importance of implementing and maintaining screening programmes as HPV vaccination programmes are rolled out. In Sub-Saharan African countries, with the highest burden of cervical cancer, it has been shown that elimination is unlikely to be achieved this century with 90% coverage of 9-valent vaccination alone, but that it can be achieved with vaccination and at least one screen (11). The main limitation of this study – which combined outputs from three different transmission dynamic models – was that none of the models considered HIV as an associated factor in the natural history of HPV. High HIV prevalence may alter the short-term trajectory of predicted CC incidence trends, and such settings may require different strategies for elimination.

In Chapter 2, the literature on epidemiological modelling of cervical cancer will be reviewed, with a focus on the modelling of intervention strategies, modelling in high HIV burden settings and a comprehensive review of HPV and cervical cancer modelling for the South African context in particular.

1.2 Problem statement and rationale

Cervical cancer incidence has substantially reduced in most regions of the world in the last 40 years (7). South Africa has had sustained high levels of laboratory reported CC incidence, even after screening was introduced at a national level 20 years ago. This study aims to develop a South African model that can provide estimates of HPV prevalence and cervical cancer incidence in this high HIV prevalence setting. The model will be used to estimate the impact that the screening programme has had over the last 20 years, what impact it will have in future, and what the impact of the vaccination programme will be on cervical cancer incidence. National level data on the performance of the

screening programme is sparse, but such information is crucial to the predictive potential of the model. In the public health sector in the Western Cape, a unique patient identifier is captured in the record of each interaction of a patient with a health facility. This makes it possible to produce individual-level data to describe the cervical cancer care cascade, which will be used to evaluate the success of the programme and inform the screening algorithm in the model.

Enough evidence regarding the impact of co-infection exists to suggest that in a country with severe HIV disease burden such as South Africa, it is naïve to ignore the dynamic relationship between HPV and HIV in developing a model to estimate the long-term impact of interventions on cervical cancer risk. To date, only one published study has attempted to dynamically simulate the transmission of both HIV and HPV in South Africa, focusing on KwaZulu-Natal and only considering the potential impact of rolling out the nonavalent vaccine (44). No transmission dynamic model has explored the mechanisms that are potentially important in the natural history of HPV in high HIV burden settings.

1.3 Aim and objectives

The aim of this thesis is to assess the long-term impact of existing prevention methods on the epidemiology of cervical cancer in the high HIV burden context of South Africa. We will expand on MicroCOSM version 1 (Microsimulation for the Control of South African Morbidity and Mortality), an individual-based model that was developed to simulate HIV, other sexually transmitted infections and processes that drive these epidemics in South Africa.

The main objectives of the study are to:

1) Review the literature on the epidemiological modelling of HPV infection and cervical cancer, globally and in South Africa, and to identify gaps in the literature that need to be addressed.

2) Add stages for HPV infection to MicroCOSM and calibrate the model to HPV prevalence data obtained from a review of South African data sources. Use this model to answer the following research questions:

    a) Are observed associations between HIV and HPV transmission due to behavioural confounding factors or biological effects?
    b) Which structure of the natural history of HPV infection (excluding cervical disease stages) fits best with observed HPV prevalence data, considering different mechanisms of naturally acquired immunity against reinfection, and reactivation of latent infections?

3) Use individual-level data on the cervical cancer prevention cascade in the Western Cape province to:

    a) Estimate levels of Pap smear coverage, screening intervals and linkage to treatment facilities in the public sector of the Western Cape. Estimates will be derived according to HIV status.
    b) Develop the screening algorithm that will be applied in the model. The data is used to estimate (in a separate simulation model) yearly rates of entering screening by age and HIV status; distributions of time to next screen or treatment; and rates of access to treatment.

4) Further develop the individual-based model to add stages of cervical disease. Use this model to evaluate the impact that the HIV epidemic has had on cervical cancer incidence, the impact that screening has had since 2000 and that different strategies and methods of screening and vaccination will have in future.

MicroCOSM (version 1) simulates an open and growing heterosexual population of all ages representative of the South African population, from 1985. Sexual behaviour is simulated at the individual level, and from 1990 people can become infected with HIV. For the purposes of addressing the aims of this thesis, the model will be developed to also simulate HPV infection and its progression to cervical cancer. The individual-based modelling framework is the most appropriate framework to address the objectives, since these involve simulating cohort studies, randomised clinical trials and a detailed screening algorithm where individuals are rescreened based on screening results. The model is developed in C++, and all analyses and visualisation of the data are performed in R.

1.4 Data Sources and ethics

No primary data were collected for the purpose of this thesis. The model is calibrated to HPV and cervical disease prevalence data obtained for reviews of published research studies. Parameters are also informed by published data.

The National Health Laboratory Service (NHLS) screening database that contain cytology data of women screened for cervical cancer in public facilities in the Western Cape (WC) is curated by the Provincial Health Data Centre (PHDC). Using the WC's patient master index (PMI) - a unique patient

identifier - the cytology database has been linked to HIV testing, management and treatment data to infer HIV and ART status per individual. The cytology data has also been linked to data that identifies colposcopy attendance. This data will be used to inform the assumptions of the model's screening algorithm.

Clinical data on all women receiving treatment for cervical cancer at Groote Schuur Hospital, and information about long-term outcomes, are recorded in a database (HREC REF NO: R016/2013). This data were crucial to estimate parameters determining rates of diagnosis and mortality for women with cancer in the model.

Both of these databases were accessed as anonymised datasets with only the fields necessary to obtain aggregated summary statistics.

This modelling study has ethical approval from the UCT Health Sciences Faculty Research Ethics Committee (HREC REF: 260/2016).

1.5 Overview and structure of thesis

This thesis proceeds with a literature review, four chapters that addresses the objectives, a discussion chapter that summarises the findings of the work and supporting appendices in which additional information on methodology, data sources, assumptions and sensitivity analyses are shown. Appendix A can be regarded as an in-depth methodology chapter.

Chapter 2 provides a review of the literature on cervical cancer modelling. It is not intended to be an exhaustive review but rather to present an overview of the methods and research questions typically addressed in these studies. The chapter provides an overview of cervical cancer models in general populations and in HIV-positive women, discusses modelling work that pertains specifically to South Africa and identifies some limitations of existing models.

Chapters 3-7 present the studies included in the thesis. These studies broadly map to the objectives, exploring and discussing the key research questions and findings. Chapters 3 and 4 relates to the second objective and involve the development of MicroCOSM to include stages for HPV infection. Chapter 3 aims to investigate HIV and HPV transmission dynamics. Empirical cohort studies have shown that women and men who are HIV-positive are more likely to subsequently have newly detectable HPV infection. Similarly, studies have shown an association between detectable HPV and subsequent HIV acquisition. Both infections are sexually transmitted and therefore share similar risk factors, but after controlling for these factors, meta-analyses have found a two-fold increased risk in both directions, and it has been suggested that this may be due to biological factors that increase risk of transmission. This chapter shows, by simulating cohort studies using the calibrated individual-based

model, that the empirical estimates of association can be matched without having to invoke biological co-factors. The chapter concludes that the observed transmission associations may result only from residual confounding by behavioural factors and network-level effects that are difficult to measure in studies.

In Chapter 4 different structures of the natural history model of HPV infection are explored. Latent HPV infection has not been definitively proven in humans, but studies suggest that reactivation of latent infections may explain newly detected HPV infections in sexually abstinent women. This chapter shows that, although model structures with and without a latent stage fit equally well to data, a model that does not include reactivation of latent infection cannot match the difference in vaccine efficacy between more/less HPV-naïve groups from randomised clinical trials (RCTs) that estimated vaccine efficacy. The different model structures are used to predict the long-term impact of HPV vaccination on infection burden, and the chapter concludes that models without latency may overestimate the impact.

In Chapter 5, the cervical cancer case cascade in the Western Cape (WC) is described. The provincial department of health in the WC has assigned a patient master index (PMI) for each person accessing public health services in 2007. The Provincial Health Data Centre (PHDC) was established to curate and link different data sources in order to provide improved patient-level care. This chapter uses data curated by the PHDC to describe coverage of Pap smear overall, and by HIV status – an indicator that has not been estimated using routine screening data before. It is also illustrated in this chapter that the recommended screening intervals are not adhered to and that linkage to colposcopy clinics is very low.

Chapter 6 addresses the fourth and final objective – assessing the impact of past and future prevention strategies on incidence of cervical cancer. This chapter illustrates the massive impact that the HIV epidemic has had on cervical cancer incidence and that HIV prevention efforts will lead to substantial declines in cervical cancer. We show that the national screening programme has already prevented thousands of cervical cancer cases, and if current levels of HPV vaccination, screening and linkage to pre-cancer treatment are maintained, cervical cancer incidence will reduce by 75% in the next century. To achieve greater impact in the short term it will be crucial to either substantially increase screening coverage and linkage to pre-cancer treatment, or to switch to more sensitive screen-and-treat technologies.

Chapter 7 summarizes and discusses the findings of the thesis as a combined body of work. It makes recommendations for cervical cancer prevention efforts and lays out future research that will result from this thesis. This chapter is followed by a list of supplementary materials, including a technical appendix describing the methods of calibration and the screening algorithm, and appendices specific to each of Chapters 3 to 6.

Chapter 2 – Literature Review

2.1 Background

Mathematical models of infectious diseases have provided many insights into hypotheses regarding long-term epidemic trends and the impact that potential interventions might have on transmission dynamics. These models prove to be most useful when validated against good epidemiological data and when the parameters of the model are informed by good evidence in the literature. Over the years, several models of the natural history of HPV and cervical cancer (CC) have been published, with models questioning the biological beliefs about disease progression and regression as early as the mid-seventies (45). The early models were focused on estimating the potential impact of cervical screening with cytology and optimizing the timing and frequency of screening, and a first review of such models were published in 1985 (46). During recent years, with major advances in technology, both computing and biomedical, adding complexity to natural history models of HPV has become feasible and the development of new screening techniques and vaccines has sparked interest in building more comprehensive mathematical models.

Transmission dynamics between HIV and HPV have not been explored in a modelling study and a comparison of different natural history structures of HPV in a context of high HIV burden has not been performed. Only one recent dynamical model that simulates both HIV and HPV has been used to estimate the impact that HIV has had on cervical cancer incidence and the impact that ART may have on cervical cancer incidence in the future (47).

This review provides a picture of the types of models that have been used to answer questions regarding the natural history of HPV, the impact of interventions and the cost-effectiveness of proposed interventions. The review will also highlight the main gaps in the literature.

2.2 Different types of models used

This section provides a brief overview of the types of model frameworks that have been used for cervical cancer modelling. The aim is to define terminology that will be frequently referred to in subsequent sections.

Similar to models of other infectious diseases, modelling frameworks can be divided into the following broad categories (48,49):

1) *Static vs dynamic models.* In static models, the incidence of infection has predetermined values and no interaction between individuals is considered. In dynamic models, the time dependent prevalence of the infection and the sexual contact patterns of individuals play a role in transmission of the infection. Static models are often used in cost-effectiveness analyses and are used to estimate the cost-effectiveness of interventions such as screening: a cohort of women are followed through time, and they have fixed rates of acquiring HPV infection and progressing to cervical cancer. These rates may depend on factors such as age or sexual behaviour. Effectiveness of interventions can be calculated, and the effect of screening in one woman will not affect the infection or disease status of another woman. When we want to estimate the cost-effectiveness – and epidemiological impact – of interventions such as vaccines, a dynamical model is more appropriate. In a dynamical model, the indirect protection that unvaccinated individuals obtain by interacting with vaccinated or immune individuals – also known as herd immunity – is considered.

2) *Deterministic vs stochastic models.* In deterministic models, we calculate expected numbers of events (not allowing for random variation) like infection or recovery, while in stochastic models, events can occur by chance, governed by some distributional function. Deterministic models are typically easier to calibrate, since the output is entirely determined by the parameters that drive the static or dynamic processes, without random fluctuations.

3) *Compartmental vs individual-based models (IBMs).* In a compartmental model, individuals move between compartments or health states according to rates determined by parameter values at an aggregate level, and within each compartment all individuals are assumed to be homogeneous. An individual-based (microsimulation) model keeps track of individuals in the population and events are randomly simulated at the individual level, allowing for heterogeneity between individuals, instead of groups of individuals. Using an IBM that has been calibrated to data, it is possible to simulate epidemiological studies such as randomised clinical trials (RCTs), a method that will be applied in this thesis. Compartmental models may be less computer intensive, but sometimes it is more intuitive to use IBMs to simulate different screening algorithms, where the result of a particular test may determine the type of and time to the subsequent test or treatment. Both compartmental and individual-based models can be dynamic or static. Although IBMs allow for more granular analyses on simulated data, they are typically much harder to calibrate. In a stochastic IBM, it is hard to separate uncertainty due to variation in parameters or from random variation. IBMs usually have many more parameters than compartmental models and typically require more steps and more data to calibrate all parameters.

The choice of modelling framework a study will implement depends on the research question, the available computing capacity and the availability of data and information on parameters. Early studies that aimed to estimate the impact of screening on a woman's risk of cervical cancer were static: it could be argued that treatment and the resulting clearance of infection will not substantially influence new infections in other women. On the other hand, studies that aim to estimate the impact of vaccines on cervical cancer risk should be dynamical, since removing a large fraction of women from the susceptible pool at an early age will substantially influence the risk of HPV infection in the unvaccinated group.

A "Markov cohort model" is a phrase frequently used by health economists in studies that aim to estimate cost-effectiveness of interventions. These models are typically static, compartmental models where single or multiple birth cohorts are followed through time. The name Markov suggests that these models follow the principle of Markov chains, where the health state in the next time step is only dependent on the individual's current health state.

The following sections will describe some existing models that aimed to estimate the epidemiological and economic impact of HPV and cervical cancer prevention strategies.

2.3 Modelling the natural history of HPV

The earliest mathematical models attempted to answer the same broad questions that we are interested in today. What are the economic implications of novel interventions against cervical cancer? What epidemiological impact will these interventions have? In attempting to answer these questions, modellers realised that uncertainty about the natural history of HPV infection and its progression to CC influences the validity of their models and several models have attempted to address this uncertainty.

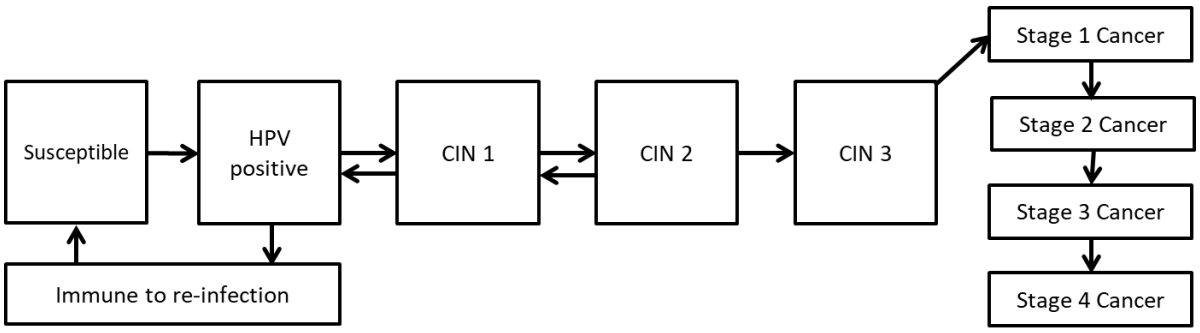


Figure 2 – An example of the natural history diagram used in modelling studies.

Most models agree on the stages of progression from HPV infection to cervical cancer: three stages of pre-cancer that correspond to histological diagnosis of cervical intraepithelial neoplasia (CIN) of grades 1 to 3 (Figure 2); or, two stages that correspond to cytological diagnosis of lower- or higher-grade squamous intraepithelial lesions (L/HSIL). Models that are interested in CC mortality usually includes four stages of cancer progression according to the FIGO classification system (50), since these stages have different mortality rates. Early models used estimates of progression and regression rates obtained from cohort studies, and manually varied these parameters to obtain better fits to data (51–53). In recent years, estimates from studies are used as initial values or prior ranges in more sophisticated, computing intensive fitting methods. Jit *et al.* (54) estimated parameters by using a search algorithm to minimise sums of squared residuals between model output and HPV prevalence data. Bogaards *et al.* (55) used Bayesian methods to fit models to observed data. Posterior distributions were calculated by weighting sampled parameter combinations according to a likelihood, which represented the degree of consistency between the model outputs and surveillance data. Kim *et al.* (56,57) followed a two-step approach to calibrate their model. Firstly, plausible ranges for each separate parameter were the values that produced model output within the 95% CIs of one dataset. Secondly, parameter sets were searched that produced model output consistent with multiple data sources and likelihood-based methods were used to choose samples of parameter sets that best fit the data. Van de Velde *et al.* (58) follows an approach similar to the first step in Kim *et al.* These more sophisticated, statistical fitting methods have started to reveal the extent of uncertainty in natural history parameters – an insight crucial in interpreting predictions of impact of intervention strategies on cervical cancer incidence.

Different modelling frameworks have been used to investigate natural history parameters. Studies that were particularly interested in estimating the progression and regression of HPV used static, deterministic models and all women in the cohorts had fixed probabilities of becoming HPV infected (51–54). Others were stochastic, individual-based models, either static (56,57) or dynamic (58).

One aspect of the natural history of HPV that is not consistent among different mathematical models is the mechanism of acquired or natural immunity against re-infection with HPV. A meta-analysis of studies concluded that women who were HPV seropositive and DNA negative at baseline were at significantly lower risk of new HPV detection during follow-up than seronegative women (pooled risk ratio of 0.65; 95% CI 0.5 – 0.8) (59). As reviewed in Franceschi and Baussano (12), models have to make strong assumptions due to paucity of good long-term follow-up data, and mechanisms after clearance of DNA-detectable HPV infection assumed in models include no natural immunity; partial immunity with lifelong duration; or a proportion of people obtain full immunity that wanes over time according to some distribution. Recent studies have shown that some form of immunity is required to improve fit of models to data (61), and that the extent of this immunity influences vaccine effectiveness (58). In a recent review of transmission dynamic models that estimated vaccine impact, only one in 16 included models assumed no form of natural immunity (62).

A mechanism that is largely ignored in models of the natural history of HPV is latent HPV infection and the reactivation of such infections. Latent infection implies that the HPV viral load is lower than the thresholds detected by current HPV-DNA tests. Latency and reactivation of oral papilloma virus infections has been proven in rabbits in a laboratory setting (63). Although such a study cannot be performed in humans, studies of HIV-positive women have shown newly detected HPV infections among women who report sexual abstinence, and that these rates depend on levels of immune suppression (64,65). In a meta-analysis of two of the largest vaccine efficacy RCTs (66), it was shown that for all of the important endpoints, estimated vaccine protection was lower in the modified total vaccinated cohort (those testing HPV-DNA positive for the vaccine type at baseline are excluded) than the HPV-naïve cohort (those testing seropositive or HPV-DNA positive for any HPV types are excluded). This difference may be attributed to infections present before the trials were started that were reactivated during follow-up. Despite these pieces of evidence, most HPV modellers have not included a stage for latent infection in their models. Recently, two mechanistic modelling studies showed that for both men (67) and women (68), models that included reactivation of previous infections significantly improved fits to longitudinal cohort data.

2.4 Modelling of prevention strategies

It is possible to prevent cervical cancer through primary prevention of HPV transmission or through secondary prevention by screening for and treating pre-cancerous abnormalities. As biomedical primary prevention, three different vaccines are available. Several methods of secondary prevention interventions are available and the most commonly used methods will be briefly discussed here. Then a brief overview of modelling studies that have investigated the epidemiological impact and cost-effectiveness of these strategies will follow.

2.4.1 Primary prevention methods

Primary prevention of two oncogenic types of HPV has been possible since 2006, when the American Food and Drug Administration (FDA) approved the first HPV vaccines. Currently there are three different vaccines available worldwide: the bivalent vaccine (Cervarix, GlaxoSmithKline) protects against HPV types 16 and 18, which causes around 70% of cervical cancer worldwide (4). The quadrivalent vaccine (Gardasil4, Merck & Co.) protects against the same high-risk types, as well as types 6 and 11, which are associated with genital warts. The nonavalent Gardasil9 has been available since 2014 and protects against the same four types as Gardasil4, as well as against the 5 additional oncogenic types 31, 33, 45, 52 and 58. These seven HPV types cause around 90% of cervical cancer worldwide (4).

In randomised clinical trials, effectiveness against persistent infections with these types was shown to be in excess of 95% in per-protocol populations for all these vaccines, among women aged 15 to 25 (69–71). The bivalent vaccine also showed significant cross-protection against types 31, 33 and 45 (72) and high levels of effectiveness against persistent infection (~80%) were demonstrated in women aged 26 to 45, when considering HPV-naïve groups (73,74). Women vaccinated in these trials are still followed up in some settings, and stable immunogenicity has been shown for at least 12 years (75). Very few studies have been performed to estimate efficacy of HPV vaccines in HIV-positive people. While vaccinated HIV-positive people do show lower levels of immune response than HIV-negative people, it is still orders of magnitude higher than the immune response of natural infection (76). A few small studies among HIV-positive people have been performed to measure effectiveness of the vaccine against HPV infection, but results are inconclusive (77). An important gap in the literature is the lack of randomised clinical trials that estimate the immunogenicity and effectiveness of HPV vaccines among HIV-positive women, and women who acquire HIV after HPV vaccination.

Although early efficacy trials of both the bivalent and quadrivalent vaccines used the full three-dose schedule (at zero, one month and six months), the efficacy of one dose and two dose schedules has also been studied. Following a systematic review (78), the WHO now recommends a two dose schedule for girls younger than 15, at least six months apart (79). The main recommendation is to target girls aged 9-13, in order to reach girls before sexual debut. Women older than 15 and immune-comprised women (such as HIV-positive women) should receive the three-dose schedule at zero, one and six months. In addition, the WHO recommends that HPV vaccination should be included in national immunisation programmes if 1) prevention of CC is a public health priority; 2) the introduction is programmatically feasible and economically sustainable, and where 3) cost-effectiveness aspects have been duly considered. Vaccination campaigns should include education about sexual behaviour and stress the importance of continued screening for CC.

The protective effect of consistent condom use on HIV transmission is well established (80). Although data from previous studies were inconsistent in showing a protective effect of condom use on HPV infection (81), the first longitudinal study specifically designed to study this relationship found that consistent condom use reduced the risk of HPV infection by ~70% among female university students in the United States of America (82). Another effective HIV transmission prevention method is medical male circumcision (MMC) (83,84). In both of these RCTs, HPV prevalence was measured at the end of the trial and the risk ratio of HPV prevalence was around two thirds in both (85,86). However, a meta-analysis of cohort studies showed no impact of MMC on HPV acquisition (87). This indicates that although MMC may not protect men from HPV transmission, it may shorten duration of infection due to the presence of fewer skin cells.

2.4.2 Secondary prevention methods

Globally, the Papanicolaou (Pap) smear is still the most common method of screening for cervical cancer (88). A sample of cells is obtained from a woman's cervix at a health facility and conventional cytology is performed at a laboratory. Results are classified according to the Bethesda system (89) with the broad categories of normal cervix, presence of CC precursor lesions ranging from mild to severe, and cancer. Cytological results do not determine the presence of an HPV infection, only the presence of abnormal cells. The major limitations of this method in lower resource settings are the lack of decentralised laboratories and qualified technicians, and that delayed communication of results may lead to loss to follow-up. Pap smears have varying levels of sensitivity (30-90%) (6), but high specificity (86-100%), which limits unnecessary treatment. Liquid based cytology was developed as a refinement of conventional cytology – the sample is transferred to a tube with preservative solution instead of being smeared on a glass slide. This method is not superior to conventional cytology in terms of sensitivity or specificity (90).

Molecular tests are available that are able to detect DNA from high-risk HPV types in cells from cervical samples. Although these tests are more expensive than Pap smears, fewer of these tests per lifetime are needed due to higher sensitivity (80-95%) than cytology (8). However, they are less specific (80-95%) and since HPV testing only determines the presence of an HPV infection and do not determine the presence of abnormal cells, risk of unnecessary treatment is high. This is especially true in a high HIV prevalence setting such as South Africa – in a recent study performed in Cape Town, specificity of HPV-DNA testing as a screening method for high grade pre-cancer of as low as 60% was shown among HIV-positive women (28).

In most developed countries, a screen, diagnose and treat approach is followed – after a positive screen result, the diagnosis will be confirmed and the appropriate treatment approach chosen (79). This confirmation is done through colposcopy (magnified visual inspection of cervix) and taking a biopsy for histological confirmation.

Ablative therapy is a treatment that superficially eliminates lesions by cryotherapy (freezing abnormal areas), laser therapy or electrical coagulation. The treated area will in time regenerate to normal epithelial cells. This method is not recommended for extensive lesions (>75% of ectocervix) and in this case, a large loop excision of the transformation zone (LLETZ) is typically used. With this method, the removed tissue can be sent to the laboratory for histology. A Cochrane review of all available treatment methods have found that LLETZ has a success rate of 90-98%, with greater variability in the success rate of cryotherapy (77-93%) (91). Treatment, however, does not necessarily remove the HPV infection – studies have shown persistence of carcinogenic HPV types of ~15% following LLETZ (92–94).

2.4.3 Modelling the impact of prevention

The impact and cost-effectiveness of screening strategies were the focus of modelling studies in the 1990s and early 2000s, as reviewed in Goldie *et al.* (95). These were mostly static, compartmental models but some were static individual-based models that could keep track of and more easily account for a woman's previous screening results. Studies agreed that any screening is more cost-effective than no screening, but that the optimal screening strategies (tests used, frequency of visits, and age of screening) varies according to countries' resources.

With the introduction of HPV vaccines, the focus of both cost-effectiveness analyses and epidemiological modelling studies has shifted towards the use of dynamic transmission models to fully incorporate the effect of herd immunity. One of the earliest dynamic, deterministic compartmental transmission models by Hughes *et al.* in 2002 (96) investigated the potential impact of a theoretical HPV vaccine on cervical cancer incidence. The model divided women into sexual activity classes and assigned different levels of vaccine effectiveness and duration of effectiveness. In subsequent years, as more information about the type-specific novel vaccines became available, static and dynamical models explored the potential economic and epidemiological impact of these vaccines using different assumptions about sexual activity, duration of vaccine induced immunity, age at vaccination and including boys in vaccination programmes. In a review of modelling studies in high income countries, Kim *et al.* (97) concluded that vaccines, along with continued screening programmes, will reduce disease burden in the long term and will be cost-effective when compared to only screening programmes, as long as vaccine protection is lifelong and high coverage is maintained. In lower- and middle-income countries (LMICs), a review of studies showed that cost-effectiveness is very dependent on the price of the vaccine, and that it is more cost-effective in high-burden countries without widespread screening (98).

In an effort to consolidate findings regarding the impact of HPV vaccinations, a collaborative meta-analysis of model findings was published by Brisson *et al.* (62). In this study, 16 different dynamical transmission models developed for high-income countries were used to estimate the same outputs under the same rules: what is the relative reduction in HPV6/11/16/18 prevalence after 70 years, if girls, or girls and boys are vaccinated and if coverage is 40% or 80%. The main findings were consistent across models that had different natural history structures, different modelling frameworks, were calibrated to different data sets in different settings, and used different calibration methods. The study found that even at moderate coverage of vaccination among pre-adolescent girls, herd immunity leads to similar reductions in HPV prevalence among men and women, and that increasing vaccination among girls has greater impact for decreased prevalence in men and women, than vaccinating boys.

In January 2019, the WHO called for the elimination of cervical cancer as a public health threat, defined as below 4 incident cases per 100,000 women (1). A cervical cancer elimination modelling consortium (CCEMC) was founded, with the aim to assess strategies that will lead to elimination by

the end of the century. The focus was on LMICs, since the vast majority of cervical cancer occurs in these regions. Three independent dynamical models were included in the final analyses of the consortium, and the main finding was that with 90% of girls vaccinated every year, elimination can be achieved in most LMICs by the end of the century, except Sub-Saharan Africa (SSA) which suffers from the greatest current burden (11). Widespread screening will need to be implemented to reach elimination in SSA. This analysis informed the WHO strategy of achieving, by 2030, the 90-70-90 targets (1). The modelling consortium found that by achieving and maintaining these targets, 74 million cervical cancer cases and 62 million related deaths will be averted in 78 LMICs by the end of the century (99). As mentioned in Section 1.1.3, the main limitation of this study was that none of the models considered the influence of HIV in the natural history of HPV. The next section will investigate the modelling studies that have been performed for HIV-positive populations, and general populations with high HIV prevalence.

2.5 Modelling HIV and HPV

Infection with HIV increases the risk of developing cervical cancer, which has been defined as an AIDS defining cancer since 1993 (12). Although cross-sectional studies provide evidence that prevalence of HPV and pre-cancerous lesions is higher among HIV-positive women (14,15,100), it is impossible to infer causality at one time point. Indeed, the sexually transmitted nature of both infections implies that the observed association between HPV and HIV could be due to confounding by behavioural factors. The interaction between HIV and HPV transmission has been studied longitudinally, and meta-analyses of studies found around two-fold risk of newly detecting HPV infection among women with HIV and two-fold risk of HIV acquisition among women with HPV (20,21,101). The studies included in these meta-analyses adjusted for behavioural confounders when assessing associations between HIV and HPV transmission risk.

Apart from the apparent increased transmission risk, there is also evidence of increased persistence of HPV infection and progression to cervical disease among HIV-positive women (18). HIV-positive women have reduced regression rates and these reduced rates are associated with low CD4 counts and high HIV viral loads (102,103). The beneficial effect of ART use on regression and progression were shown to be significant in a meta-analysis of studies that compared HIV-negative to HIV-positive women (22).

WHO recommends that all women who live in high HIV prevalence countries who attend HPV screening services should receive HIV testing and counselling and that all women who test HIV-positive should receive HPV screening, regardless of age (79). This HPV screening should be repeated every 3 years, instead of every 10 years which is the minimum recommendation for HIV-negative

women. Recommended screening and treatment methods are the same for HIV-positive and negative women.

Despite these well-established associations, and the fact that the region in the world with the highest cervical cancer incidence (Figure 1) is also the region with the highest HIV prevalence, very few modelling studies have considered HIV status as a modifier in estimating the cost or epidemiological impact of prevention strategies in Sub-Saharan Africa.

The first model to estimate cost-effectiveness of screening for CC in HIV-positive women was a static, Markov model for women in the United States by Goldie *et al.* in 1999 (53). This model had 5 health states: Normal cervix, LSIL, HSIL, cancer, and death, all stratified by CD4 count categories (>500, 200-500 and <200 cells/mm^3). Infection with HPV was not included in the natural history. The authors compared the strategies of no screening, different intervals of Pap smear screening, and different intervals of colposcopy as a screening method. Parameters on incidence, progression and regression of cervical disease by HIV stage were obtained from literature and individually changed in sensitivity analyses. The study found that different screening strategies for women with CD4 counts above 200 cells/mm^3 are cost-effective, but only under certain assumptions is screening cost-effective for women at late stage HIV disease, who gain very little life-years if screened for cancer. If these women are started on treatment, cost-effectiveness was dependent on assumptions about duration of survival on ART.

In a follow-up study, the authors added a stage for HPV infection in the natural history of HPV in order to estimate cost-effectiveness of DNA testing as a screening method (104). The model was still static, with predetermined probabilities of becoming infected. This study found that, for women in all stages of HIV disease, the most effective strategy would be 6-monthly Pap smears for women who test HPV-DNA positive at HIV diagnosis and annual Pap smears for those who do not.

The framework of the Goldie *et al.* studies have been used in other settings. Atashili *et al.* (105) showed that in Cameroon, ART initiation doubled the lifetime risk of cervical cancer mortality if no screening was done and that one cervical cancer death would be avoided for every 262 women screened at ART initiation. For Brazil, Vanni *et al.* (106) used the same framework, but applied calibration methods to estimate a best-fitting parameter set by choosing from over 100,000 parameter sets the combination that gave the smallest sum of residuals between target statistics and model outputs. This study estimated that annual HPV-DNA testing with Pap smear triage would be the most cost-effective strategy for HIV-positive women in Brazil. Goldie *et al.*'s framework was also used in a study to determine cost-effectiveness of screening and preventative cryotherapy for HIV-positive women in Kenya (107). This study found that performing preventative cryotherapy on an HIV-positive woman presenting for screening had lowest cost and led to the greatest addition to life expectancy, but that same day screen and treat strategies with combinations of VIA, Pap and HPV could lead to similar gains.

Another study that followed the natural history model of Goldie *et al.*, but used a different modelling framework, was a study by Smit *et al.* (108). This study estimated current and future non-communicable disease burden in Kenya, by HIV status. Cervical cancer was one of the cancers included in this analysis. In this static, individual-based model, people become HIV infected, initiate ART and progress to AIDS at yearly rates as estimated by other models. Women have a fixed lifetime risk of becoming HPV infected, and this risk is increased when becoming HIV infected. Progression and regression parameters were taken from literature – some were used as estimated in studies, and others were calibrated. The study estimated that 51% of Kenyans suffer from at least one non-communicable disease, and that cervical cancer will be the most common of all cancers between 2018 and 2035 at 20% of 550,000 cases. Perez-Guzman *et al.* (109) used the same model framework to estimate the epidemiological impact of four different screening strategies compared to the status quo of screening in Kenya, focusing only on HIV-positive females.

A dynamical, compartmental model of both HIV and HPV transmission was recently published for Tanzania (47). The aim of this study was to measure the impact that the HIV epidemic has had on CC incidence, and what impact HIV infection and HIV prevention strategies (MMC, ART and pre-exposure prophylaxis (PrEP)) will have on CC incidence in the future. The model has compartments for people with higher and lower risk sexual behaviour, and the female population has a sub-category of female sex workers. The majority of parameters in the model were taken as fixed values from literature, with only certain sexual behaviour parameters and the relative per-sex-act probability of HIV acquisition for females compared to males estimated in a calibration algorithm. The study estimates that if 80% of men are circumcised, 73% of PLHIV are virally suppressed on ART and 90% of female sex-workers use daily PrEP between 2020 and 2070, cervical cancer incidence will be reduced from 65 per 100,000 women under current levels of HIV interventions, to 36 per 100,000 women in 2070. The authors do not quantify uncertainty in their estimates but acknowledge that results are highly sensitive to changes in several of the parameters. They conclude that HIV interventions alone are not sufficient to reduce CC.

Other cervical cancer modelling studies that considered HIV in the natural history of HPV were for the South African context and will be discussed in the next section.

2.6 HPV and cervical cancer modelling for the South African context

Very few models have been developed to answer HPV and cervical cancer related research questions for South Africa. The search string (HPV[Title/Abstract] OR "Human Papillomavirus"[Title/Abstract]) AND "South Africa"[Title/Abstract] AND ("mathematical model" OR simulation OR "cost-effectiveness" OR "cost effectiveness") in Pubmed yielded 12 results. After reading all the abstracts, 2

were discarded because they were not relevant to this review, and 3 will not be discussed in depth here. Two of the three estimated cost-effectiveness of screening or treatment methods within two clinical trials among HIV-positive women attending an outpatient clinic in Johannesburg (110,111) by simulating treatment algorithms given the measured cervical disease prevalence. Since the studies did not model the natural history of HPV, we did not consider these relevant. The third study was focussed on HIV prevention and estimated the cost-effectiveness of dual vaccination with the HPV vaccine and a hypothetical HIV vaccine, considering only the costs of HIV disease, and not the costs of diagnosing and treating cervical pre-cancer or cancer (112). The other 8 studies will be discussed in this section.

In the first two of the eight relevant modelling studies, Goldie *et al.* used the same model: first in a South Africa-only analysis of the cost-effectiveness of different screening strategies and then adding four other low-resource countries to the analysis to be more generalizable (113,114). This model was a static Markov cohort model and distinguished between HIV-positive and negative women in terms of progression and regression of cervical disease. The model used parameters obtained from literature, prevalence estimates from a screen-and-treat study performed in South Africa (115) and calibrated the model to published estimates of lifetime risk of CC in South Africa (116). These studies concluded that, considering costs to patients and providers, a once-off screening combining direct visual inspection with acetic acid and HPV-DNA testing may be a clinical and cost-effective alternative to a three-visit cytology-based screening strategy. Since these studies were performed at a time before ART was widely available, the impact of ART on these rates was not considered. These studies also assumed that HPV-DNA testing in South Africa costs the same as cytology.

Vijayaraghavan *et al.* aimed to determine the cost-effectiveness of screening strategies including combinations of cytology and HPV-DNA testing (117). The model was a static individual-based model that followed a cohort of women from the age of 13. Similar to Goldie *et al.*, this model distinguished between HIV-positive and negative women in terms of progression and regression of cervical disease, but the effect of ART was also included. HIV incidence and ART coverage were assumed to be constant during the women's lifetimes. Although this study estimated effectiveness of two of the same tests as in Goldie *et al.*, they used a more realistic relative price of HPV-DNA testing to cytology, of about three times. Nevertheless, the study concluded that conventional cytology followed by HPV-DNA testing for triage would be less expensive and more effective than using cytology alone.

In another study that used modelling to estimate the cost-effectiveness of cervical cancer screening, Campos *et al.* focused on the HIV-positive sub-population (118). This model is a static, individual-based model that follows women from the age of 9 years. Each month, women have probabilities of moving between HPV health states. All women become HIV-positive at exactly 20 years of age and initiate ART at exactly 25 years of age. Monthly, age-specific HPV infection, progression and regression probabilities among HIV-negative women, which the authors estimated in previous

modelling studies for other countries, were used as point estimates in this study. Values from ranges of multipliers for these monthly probabilities for HIV-positive women were chosen by fitting to local datasets and choosing the 50 parameter sets with the highest likelihoods. This study assumed 70% coverage of screening and no loss to follow-up between screening and treatment. Biennial HPV-DNA testing followed by treatment was estimated to reduce lifetime risk of cervical cancer by 56% and was shown to be cost-effective at the threshold of South African GDP per capita. At lower thresholds of willingness to pay, the HPV-DNA testing followed by treatment could still be cost-effective at reduced costs of the test.

Sinanovic *et al.* aimed to determine the cost-effectiveness of adding an HPV vaccine to the current screening policy in South Africa (119). In a static Markov model, a cohort of women was followed from the age of 0 to 85, with vaccination at age 12. Parameters were derived from the literature and women had fixed probabilities of becoming HPV infected. The parameters were not calibrated to local data and interactions with HIV were not considered. The study found that adding an HPV vaccine to the national vaccination programme could be cost effective, but price reductions of up to 60% would be necessary to make it affordable.

Kim *et al.*, 2013 estimated cost-effectiveness of HPV vaccination strategies for 48 countries in Sub-Saharan Africa (120). Since very little epidemiological data exist for most of these countries, a simple static Markov model with parameters obtained from literature was used. For two countries, South Africa and Uganda, enough data were available to calibrate a more comprehensive individual-based microsimulation model using likelihood-based methods. This model was also static and interactions with HIV were not considered. The study also concluded that, at a substantial reduction in the vaccine price, a strategy of vaccination and screening could be cost effective and affordable in South Africa.

In a third static Markov model that estimated the cost-effectiveness of adding the HPV vaccine to the cervical cancer prevention strategy in South Africa, Li *et al.* also showed that vaccination will lead to substantial reductions in cancer incidence and is cost-effective (121). This study followed one cohort of girls from age 12 for life, and girls could become HIV and HPV infected at fixed rates though their lifetime. Parameters for progression and regression for HIV-negative and positive women were obtained from literature, with manual calibration of the rate of progression of CIN3 to cancer to match cervical cancer incidence as estimated by the pathology based national cancer registry.

Although all three studies showed cost-effectiveness of HPV vaccination, the assumptions in the models were very different. All three models were static and therefore did not consider the beneficial impact of herd immunity. Kim *et al.* and Li *et al.* considered the impact of HIV, while Sinanovic *et al.* did not. Kim *et al.* assumed a three-dose vaccine schedule at 70% coverage and 100% lifelong efficacy. Sinanovic *et al.* assumed a three-dose schedule plus a booster shot, 80% coverage and lifelong protection of 90%. Li *et al.* assumed the two-dose schedule that was approved by WHO at the time, but assumed efficacy levels as low as 50% against CIN1 and 65% against CIN2+. Li *et al.* used a

discount rate of 5%, while Sinanovic *et al.* and Kim *et al.* used a rate of 3%. Sinanovic *et al.* did not calibrate or validate the model against any local data, while Kim *et al.* and Li *et al.* calibrated the model to some local data sources.

The final modelling study is the first dynamical model to be developed for a South African setting. Tan *et al.* developed a deterministic compartmental model that simulates both HIV and HPV transmission dynamics to estimate the potential reduction in cervical cancer incidence after introduction of the 9-valent HPV vaccine (44). The HIV and sexual activity aspects of the model were developed and validated in a previous study, for Kwazulu-Natal specifically. In this study, compartments for HPV infection and cervical disease were added. Heterosexual HPV transmission depends on sexual risk group, age, and HIV disease stage, implying higher susceptibility when HIV-positive. Rates of progression and regression also depend on age and HIV stage. The authors assume that only females develop natural immunity (at low levels) and that this immunity wanes over time. Initial values for the parameters were obtained from literature and an algorithm searched for the parameter combination that maximised the likelihood of observing a set of data points. These data points were age-specific prevalence of CIN2/3 in a population-based study in Cape Town (by HIV status), age-specific overall HPV prevalence from the same study and age-specific HPV prevalence from another study in Cape Town. The authors estimate that in the base case of 90% coverage of vaccination and 80% lifelong efficacy, vaccination with the 9-valent vaccine will lead to a reduction of 74% in cervical cancer incidence in 70 years. The study assumed that no cervical cancer screening has happened to date, and that none will happen in future.

Another study estimated the impact of vaccination and HPV-DNA based cervical cancer screening on cervical cancer trends in 181 countries, including South Africa (122). In this study, Simms *et al.* use 1) estimates of herd effects from the Australian *Policy1-Cervix* transmission dynamic model, 2) age-specific South African cervical cancer incidence in 2012 estimated by the WHO's International Agency for Research in Cancer (IARC) (known as the GLOBOCAN estimates), and 3) trends in reported cervical cancer incidence in Uganda and Zimbabwe's population-based registries (1993 to 2012) to estimate cervical cancer incidence for South Africa up to 2099. The authors estimate that age-standardised incidence will be 19.3/100,000 women in 2099 under the 2014 levels of screening (no vaccination assumed), but that this value could reduce to 4.9/100,000 women if women were screened twice per lifetime with an HPV-DNA based test (70% coverage) and if 100% of women are vaccinated. In this study, no country-specific data were used to inform the transmission dynamic model, and herd effects are specific to the Australian context, where sexual behaviour is different than in South Africa. The impact of HIV was not considered in the calculation of the GLOBOCAN estimates, or the trends in cervical cancer incidence over time.

2.7 Limitations of existing models and modelling studies

Each model described in previous sections has its own inherent limitations, as acknowledged by the authors, but all models have some overarching limitations. In many settings, local data to inform models are sparse (e.g., HPV prevalence) or biased (e.g., only cytology diagnosed cervical disease data or only pathology confirmed cervical cancer incidence data). Rates of disease progression and regression are often measured in a biased way (e.g., double interval censoring is inherent to the data collection process); studies are not large enough to inform age-specific rates; or parameters cannot be measured directly because it is unethical to measure certain rates (such as progression of CIN3 or early stages of cancer). Cohort studies are typically large and expensive, and not many are performed in LMICs. Many models overcome these limitations by fitting uncertain parameters to data using a calibration algorithm, or use rates obtained from empirical studies as fixed values and perform sensitivity analyses to confirm robustness of conclusions.

There are two states in the natural history of HPV for which very little data to inform models exist: acquired or natural immunity against reinfection, and viral latency. Although most models assume some form of immunity against reinfection after clearance of an HPV infection (62), the lack of definitive data leads to a wide variety of assumptions about the level and duration of protection that natural immunity provides. Viral latency and the possibility of reactivation of latent infections have largely been ignored by the modelling community. Two recent mechanistic models showed that including a latent state in the natural history model could improve fits to HPV prevalence data observed in cohorts (67,68). It is yet to be explored what the impact of reactivation of latent infections will have on population-level cervical cancer incidence – in general, but especially in settings with high HIV prevalence. Although evidence of reactivation of latency exists (64,65), data that models require such as the fraction of cases that become latently infected after clearance, or the duration of latency, is not available.

As discussed in Section 2.5, very few models for HIV-positive populations have been developed, and only two have been published that dynamically simulate both HIV and HPV infection, and progression to cervical cancer (44,47). However, several models have estimated the impact and cost-effectiveness of cervical cancer prevention in Sub-Saharan Africa (120), East Africa (123) and LMICs, many of which are in Sub-Saharan Africa (11). Without considering the effect of HIV, results from these models should be interpreted with caution and the question of impact and cost-effectiveness of interventions in high HIV prevalence settings remain unanswered.

A general limitation of the South African modelling studies discussed in Section 2.6 is that they have not been calibrated using national data. No nationally representative HPV or cervical disease prevalence survey has been performed, and the National Cancer Registry reports only on pathology confirmed cervical cancer in South Africa. Modellers have to rely on prevalence estimates from

research studies and many of these studies are not performed among populations representative of the general population. The majority of South African studies used Pap smear to determine cervical disease, which has uncertain levels of diagnostic accuracy.

Both the models by Goldie *et al.* (113,114) and Vijayaraghavan *et al.* (117) have the advantage of using different rates according to HIV status, but parameters were not calibrated to local data and results were not stratified by HIV status. The study by Campos *et al.* (118) calibrated parameters to local data, but this study was focused only on the HIV-positive population and made very crude assumptions about time of HIV acquisition and ART initiation. Therefore, no study to date has estimated the impact of screening on cervical cancer or the cost effectiveness of different strategies in South Africa by using a model that is calibrated to local data, that formally estimates uncertainty and that stratifies transition rates and results by HIV status.

The vaccination studies by Sinanovic *et al.* (119), Kim *et al.* (120) and Li *et al.* (121) shared the major limitation that the static design of their models did not allow for indirect protection through herd immunity. Although this limitation leads to underestimates of vaccine impact, the models all found that HPV vaccination will be cost-effective. The population-level impact of the bivalent vaccine rollout in public schools since 2014 on cervical cancer incidence in the long term therefore still needs to be estimated in a transmission dynamic model. Tan *et al.* (44) took the first step in this direction by developing a dynamic model to estimate the impact that the 9-valent vaccine will have on cervical cancer incidence in 70 years' time. This analysis applies only to KwaZulu-Natal, although all the data on HPV and cervical disease that was used to calibrate the model are from the Western Cape, a province with much lower HIV prevalence than KwaZulu-Natal. Since this model applies to only one province, estimates only the impact of the 9-valent vaccine, and does not consider current levels of screening, the impact of current intervention strategies on cervical cancer incidence in South Africa remains to be estimated.

Mathematical modelling studies of HPV and cervical cancer mostly focus on the potential impact of new prevention strategies, and while this is necessary – and commendable that WHO is using dynamical modelling in the development of new guidelines (1) – there are several types of research questions that are not commonly investigated using HPV and cervical cancer models. For example, our search of the literature has not resulted in any modelling studies that compared the impact of historical screening with the counterfactual of no screening. The statistical method called age-period-cohort models have been used to estimate the number of cervical cancer cases that were prevented due to screening since the 1960/70s in some high-income countries (124,125). This type of model allows for age effects (aging of the population), period effects (e.g., scale up of Pap smear screening over time) and cohort effects (e.g., changes in the sexual behaviour of cohorts over time). The method suffers from identifiability problems (126), and it would be complicated to separate the opposing period effects of HIV and screening on the incidence of cervical cancer. Mathematical modelling

studies can be used to evaluate the impact of historical screening. In future, it will also be necessary for models to evaluate progress towards reaching the 90-70-90 targets and elimination of cervical cancer, given the prevention methods adopted by the country studied.

Another use of dynamical models is to design empirical studies, but there have been relatively few such applications in the field of HPV and cervical cancer epidemiology. Using appropriate models, RCTs or cohort studies can be simulated to estimate potential effect sizes, optimal study design and sample sizes. Examples include such study design simulations for Ebola (127), HIV (128) and malaria (129). Research questions that involve sexual behaviour and the sexual network of participants – aspects that are notoriously hard to measure in empirical studies – can also be explored by simulating studies using individual-based sexual network models. An example question is the estimated increased risk of HIV acquisition among those infected with herpes simplex virus type 2 – two infections that share the same mode of sexual transmission. A recent simulation study showed that this association can be explained by behavioural factors (130). Similar questions that involve sexual behaviour and HPV transmission, such as the synergistic relationship between HIV and HPV transmission, or the fraction of new HPV detections that are due to new infections versus reactivation of latent infections can be explored using simulation models of HPV transmission.

Chapter 3 – Are associations between HIV and HPV transmission due to behavioural confounding factors or biological effects?

Van Schalkwyk C, Moodley J, Welte A, Johnson LF

Publication status

Published: *Sexually Transmitted Infections* 2019; 95(2):122-128

Author contributions

LJ developed the HIV and sexual network components of MicroCOSM. CvS developed the HPV component and conceptualised and conducted the analysis, led data interpretation and drafted the manuscript for this study. LJ provided technical guidance on the modelling, conceptualisation of the study and overall leadership. AW was responsible for partial funding. LJ, JM and AW assisted with data interpretation and writing. All authors read and approved the final manuscript.

3.1 Abstract

Background: Cohort studies have shown significant increased risk of HIV acquisition following human papillomavirus (HPV) detection and increased risk of new HPV detection in HIV-positive individuals, after adjusting for behavioural risk factors. This study uses an individual-based model to assess whether confounding sexual behaviour factors and network level effects can explain these associations between HIV and HPV infection status, without biological interactions.

Methods: The model simulates infection with 13 oncogenic HPV types and HIV. It allows for different relationship types, with heterogeneity in probabilities of concurrency and rates of partner change. No effect of prevalent HPV infection on HIV acquisition is assumed and vice versa. The model is calibrated to South African HIV and type-specific HPV prevalence data using a Bayesian approach. The model is used to simulate cohorts with quarterly HIV and HPV testing from 2000 to 2002. These simulated data are analysed using proportional hazard models.

Findings: The mean of the unadjusted hazard ratios of HIV acquisition following detection of an oncogenic HPV type calculated for each simulated cohort is 2.6 (95% CI 2.2–3.1). The mean of the unadjusted hazard ratios for the effect of HIV on newly detected HPV is 2.5 (95% CI 2.2–2.8). Simulated associations between HIV and HPV infection status are similar to corresponding empirical estimates. In sensitivity analyses in which HIV and HPV were assumed to increase each other's transmission risk, simulated associations were stronger but not inconsistent with empirical estimates.

Conclusions: Although we cannot rule out the possibility that associations between HIV and HPV transmission may be due in part to biological interactions, these results suggest that observed associations could be explained entirely by residual confounding by behavioural factors and network-level effects that observational studies cannot account for.

3.2 Introduction

Worldwide, more than 35 million people live with HIV and despite prevention efforts and increases in treatment coverage, around 2.1 million people were newly infected in 2015 (131). Human papillomavirus (HPV) is the most common STI globally, and several cohort studies have investigated the role that it plays in the transmission of HIV in recent years. Estimates of the association between HPV infection and HIV acquisition varied, from no effect (132) to a five-fold increased risk (133). Two meta-analyses have both estimated a two-fold increased risk of HIV acquisition in HPV infected individuals (20,21). Since an association between the two STIs would be expected, due to their common mode of transmission, the studies included in these meta-analyses controlled for individual level sexual behavioural risk factors, defined differently in each study. Several biological mechanisms for increased HIV risk in HPV infected individuals have been postulated, as reviewed in Houlihan *et al.* (21) and Lissouba *et al.* (20). For example, HPV infected epithelial cells have increased density of HIV target cells, such as CD4+ T cells (134).

There is also evidence of increased risk of new HPV detection in the presence of HIV infection, with relative risk estimates obtained from observational cohort studies ranging between 2 and 5 (16,17,135–139)[a]. Biological mechanisms of increased susceptibility to HPV or increased infectiousness of HPV in the presence of HIV are not well established, but reactivation of latent infections in immune suppressed individuals may play a role in this association (64,65).

The apparent epidemiological synergy between HIV and HPV could serve as additional motivation for integrated STI prevention programmes. However, in the absence of randomized controlled trial data, observational evidence of effects of HPV on HIV acquisition and HIV on HPV acquisition has to be treated with caution. Although the studies control for self-reported individual behaviour, risk behaviour at the sexual network level is difficult to control for accurately and the observed associations might be due to residual confounding.

Although observational studies have similarly shown significant associations between other STIs and HIV (140), clinical trials of STI control for HIV prevention have failed to show significant impact (141), and it has been noted that the associations reported in observational studies could be explained to some extent by confounding behavioural and network factors (142). An epidemiological model that explicitly simulate sexual networks in a population has been used to show how the precision of estimated associations between diseases in simulated cohorts can be improved (143) and recently such a model was used to show that observed associations between herpes simplex virus 2 (HSV-2) and HIV can be explained by behavioural factors (130). In this study, we simulate cohort studies using an individual-based model to assess whether confounding behavioural factors and network effects are sufficient to explain the observed associations between HIV and HPV infection status, in the absence of biological transmission effects.

[a] A meta-analysis by Looker *et al.* was published after our analysis was completed (101).

3.3 Methods

3.3.1 The model

MicroCOSM, an individual-based network model, represents the South African population by age and sex and simulates population growth since 1985 (144). The model simulates infection with human immunodeficiency virus (HIV) and was extended to include the transmission dynamics of 13 oncogenic (high-risk or HR) HPV types for this study.

The sexual behaviour assumptions of MicroCOSM are described in detail elsewhere (144) and in Appendix A. Briefly, the adult population is divided into low risk and high risk groups. Individuals in the low-risk group never have concurrent sexual partners, while individuals in the high-risk group can have up to two partners at any given time. Partnerships can be short-term, long-term (marital) or client/sex worker interactions. All long-term relationships start off as short-term relationships. Only heterosexual relationships are modelled, and all sex workers are female. Coital frequency and rates of relationship formation and dissolution depend on factors such as age, sex, risk group, relationship type and stage of HIV infection. Condom usage increases over time and depends on age, sex, and relationship type (Appendix A.2). Sexual behaviour assumptions and parameter estimates are based on South African data (144,145).

The current focus of the extension of MicroCOSM is on the aspects that drive HR-HPV prevalence: transmission and clearance of infection, natural immunity against re-infection, viral latency (defined as infections not detectable by nucleic acid tests) and reactivation of latent infections to detectable virus levels. Figure 3 represents the structure of the model of HPV natural history. The 13 oncogenic or high-risk (HR) HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) are modelled individually and independently from each other. No effect of HPV on HIV transmission is assumed, and vice versa. HIV infection is however assumed to increase the duration of an HPV infection and increase the rate of reactivation from latency. HIV natural history assumptions are described elsewhere (144) and in Appendix A.

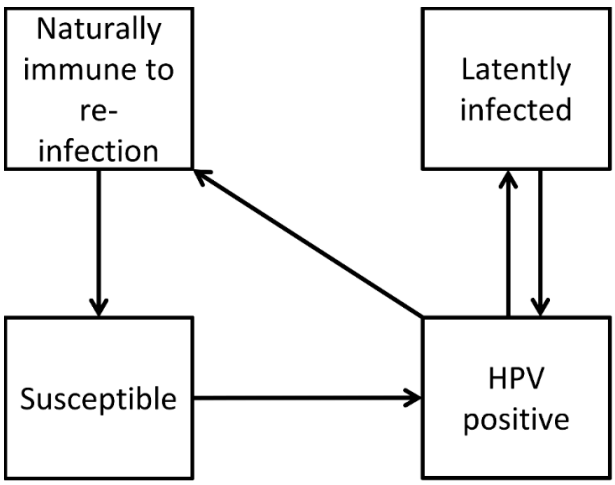


Figure 3 - Model of the natural history of HPV

The model is fitted to South African type-specific HPV prevalence data using a Bayesian approach, allowing for uncertainty in key HPV parameters. We performed an extensive literature review and identified 20 prevalence estimates for each of the 13 HR-HPV types from studies (*Table A 10*). Prior distributions are specified for all of these parameters (*Table A 12*). For each HPV type, 500,000 combinations of parameters are sampled from the prior distributions, and for each parameter combination a likelihood value is calculated to represent how well the model estimates of type-specific HPV prevalence match the empirical data. Using these likelihood values as weights, 500 parameter combinations are resampled to represent a sample from the posterior distribution (146). More details on the calibration method and the prevalence studies included in the calibration procedure are shown in Appendix A.5.1. Although the appendix describes calibration to cervical disease stages in Sections A.5.2 and A.5.3, only the methods described in Section A.5.1 apply to this chapter.

In this analysis, we use the 500 parameter combinations from the posterior distributions for each HPV type to simulate 500 cohort studies. These cohorts consist of sexually active men and women, aged 15 to 49, who undergo quarterly HIV and HPV testing. The cohorts are enrolled in 2000 (a time of high HIV incidence in SA) and followed for 3 years, corresponding to observational studies. Ethical approval for this study was obtained from the Human Research Ethics Committee at the University of Cape Town (260/2016).

3.3.2 Statistical Analysis

The Cox proportional hazard model was used to calculate hazard ratios to assess association. For all estimates, the mean of the 500 point estimates is shown and uncertainty is expressed by the 2.5^{th} and 97.5^{th} percentiles of the estimates.

Observational studies of the association between HIV and HPV were compared to the model if they met criteria listed in Appendix B.1. Data from the simulated cohorts are analysed using the same statistical methods used in the observational studies. Results of four of the studies are compared to high risk females in the simulated data to match sexual behaviour (132,133,147,148). Crude estimates are compared, since studies adjusted for different demographic and STI information and sexual behaviour indicators were defined differently. Although the primary focus is on the crude analyses, adjusted estimates were also calculated (controlling for sex, age, the number of new sexual partners in the preceding 6 months, marital status and index of concurrency (149)). All statistical analyses are performed in R (150).

In the following sections, the first visit of new detection of HIV or HPV will be denoted as v_d, while the previous visit is denoted as $v_d - 1$ and baseline visit as v_0.

3.3.3 Effect of baseline HPV on HIV acquisition

In the analysis of the effect of HPV status on HIV acquisition, individuals who were HIV-positive at study enrolment (v_0) were excluded from the simulated cohorts. To estimate the effect of HPV status at v_0 on HIV acquisition later in the study, two analysis methods were used: 1) a Cox proportional hazards model with baseline HPV status (151) or number of HPV types detected (147) as a fixed covariate to produce hazards ratios and 2) logistic regression to produce odds ratios (133,152).

3.3.4 Effect of HPV status at visit prior to HIV acquisition

The effect of HPV status at visit $v_d - 1$ on HIV acquisition between $v_d - 1$ and v_d are estimated in two ways in observational studies. The first is HPV positivity (any HPV type) at $v_d - 1$, regardless of the HPV status at v_d. The second is HPV clearance, i.e., any HPV type that was detected at $v_d - 1$ cleared before the visit of first HIV detection (v_d), while other HPV types may have persisted. The reference group in the majority of observational studies is those who were HPV uninfected at visit $v_d - 1$. Two analysis methods were used to estimate these effects: 1) a Cox proportional hazards model with HPV status (153) or HPV clearance (151,153) as a time-varying covariate to produce hazards ratios and 2) conditional logistic regression to produce odds ratios (132,152). For the conditional logistic regression, cases (HIV seroconverters) were matched to three controls (those who remained HIV-negative) using time in the study as matching criteria.

3.3.5 Effect of HIV status at study enrolment on new HPV detection

In the analysis of the effect of HIV status at time of study enrolment (v_0) on new HPV detection later in the study, individuals who became newly HPV infected or experienced a reactivation of a latent infection in the model contributed to new detections. Two analysis methods were used: 1) a Cox proportional hazards model with baseline HIV status as a fixed covariate to produce hazard ratios (65,136,148) and 2) Poisson regression to produce incidence rate ratios (17,135,137). Although the exact stages in the HPV natural history, and exact times of transition between states, are known in the simulated data, two assumptions were made to correspond to observational studies. First, time of HPV detection and viral clearance were assumed to occur at the midpoint between two visits and second, individuals in the immune, latent and susceptible stages all contribute to exposure time.

3.3.6 Effect of HIV incidence on new HPV detection

Two methods were used to estimate the effect of HIV incidence on new HPV detection. In the first analysis, all individuals who were HIV-negative at enrolment were followed and a Cox model was used to estimate hazards ratios, using HIV status as a time-varying covariate (16). In the second analysis, HIV seroconverters were matched to individuals who remained HIV-negative on time in

study. Multinomial logistic regression was used to estimate odds ratios for new HPV detections (zero, one or ≥2) in visits subsequent to first HIV detection (138).

3.3.7 Sensitivity analyses

Model results were investigated for sensitivity to two sources of uncertainty: 1) the structure of the HPV natural history model and 2) the parameter estimates given the base model structure (Figure 3). To explore the first source of uncertainty, the model was calibrated using 6 alternative HPV natural history structures (which differed in terms of natural immunity and reactivation of latent infections) and cohorts were simulated to calculate associations. To explore the second source, cohorts were simulated using the means of the posterior distributions of the main parameters in the base model structure. Then each parameter in turn was decreased and increased by 50% and cohorts were simulated for each parameter change.

In addition, cohorts were simulated using the 500 parameter combinations from the posterior distributions for each HPV type with 1) infectivity of HIV doubled in the presence of HPV infection; 2) susceptibility to HIV doubled in the presence of HPV; 3) infectivity of HPV doubled in the presence of HIV and 4) susceptibility to HPV doubled in the presence of HIV.

3.4 Results

3.4.1 Model fit

The mean simulated prevalence of HIV and oncogenic HPV in the population aged 15-49 in 2000 are 11.0% (95% CI 6.8–15.1%) and 29.5% (95% CI 26.4-32.3%), respectively. Mean type specific HPV prevalence estimates, split by HIV status, are shown in Figure 4. Model fits to the calibration data are shown for each HPV type in Figures A 20-32. The mean overall HPV prevalence in 2000 calculated using the sample of 500 parameter combinations from the posterior distributions compares well to the HPV prevalence pattern by age and HIV status found in a population level study in Khayelitsha, Cape Town (14) (Figure A22). Consistent with the calibration data, the model estimates a substantially higher HPV prevalence in HIV-positive individuals compared to HIV-negative individuals, for all HR-HPV types.

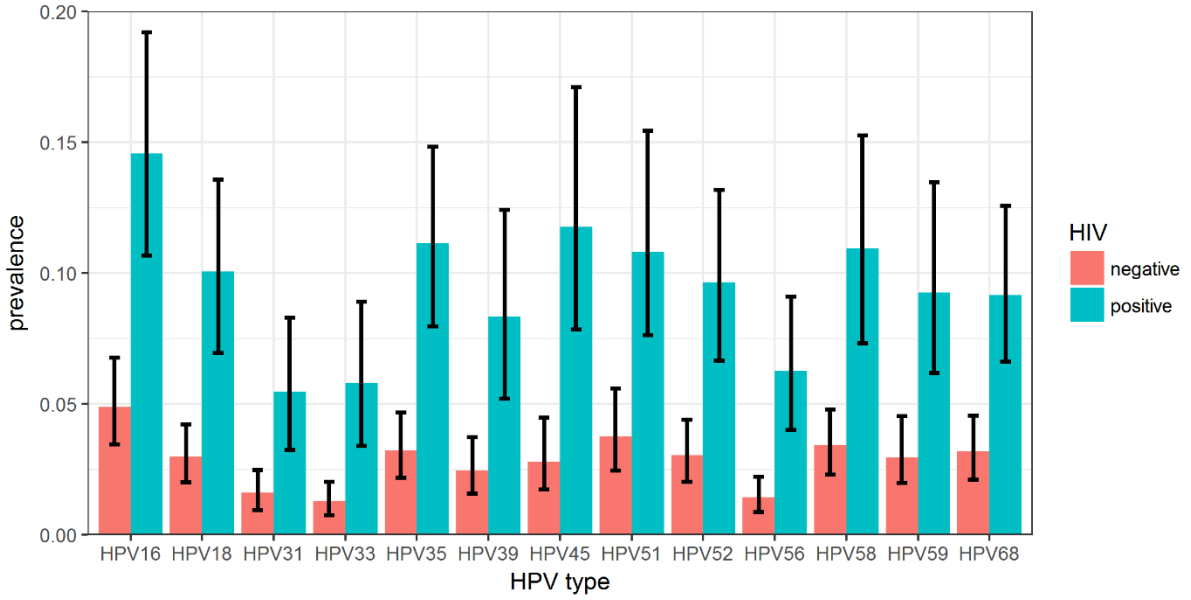


Figure 4 - Model estimates of type specific HPV prevalence by HIV status. Results are the mean HPV prevalence levels for individuals aged 15-49 in 2000, simulated using the sample of 500 parameter combinations from the posterior distributions. Error bars represent the 2.5th and 97.5th percentiles of the 500 prevalence estimates.

3.4.2 Associations between HPV status and HIV acquisition

Amongst the 500 simulated cohorts, the mean unadjusted hazard ratio for the association between an oncogenic HPV type at $v_d - 1$ and HIV acquisition at v_d, in the general population aged 15-49, is 2.6 (95% CI 2.2–3.1). The corresponding mean hazard ratio for the association between HPV status at v_0 and subsequent HIV acquisition is 2.4 (95% CI 2.0-3.0). The association between HPV clearance and HIV acquisition between $v_d - 1$ and v_d is 1.9 (95% CI 1.4-2.4). For this estimate, the reference group is those who were HPV negative at visit $v_d - 1$, though the hazard ratio is similar when the reference group comprises everyone who was HPV negative at both $v_d - 1$ and v_d or everyone who did not clear an infection. The corresponding estimates adjusted for sex, age, number of new sexual partners in the preceding 6 months and marital status are 1.7 (95% CI 1.4-2.1), 1.6 (95% CI 1.3-2.0) and 1.3 (95% CI 1.0-1.7). Figure 5 compares the unadjusted estimates of association from observational studies and the simulated data. In general, the simulated associations between HPV and HIV acquisition are roughly consistent with the corresponding empirical estimates.

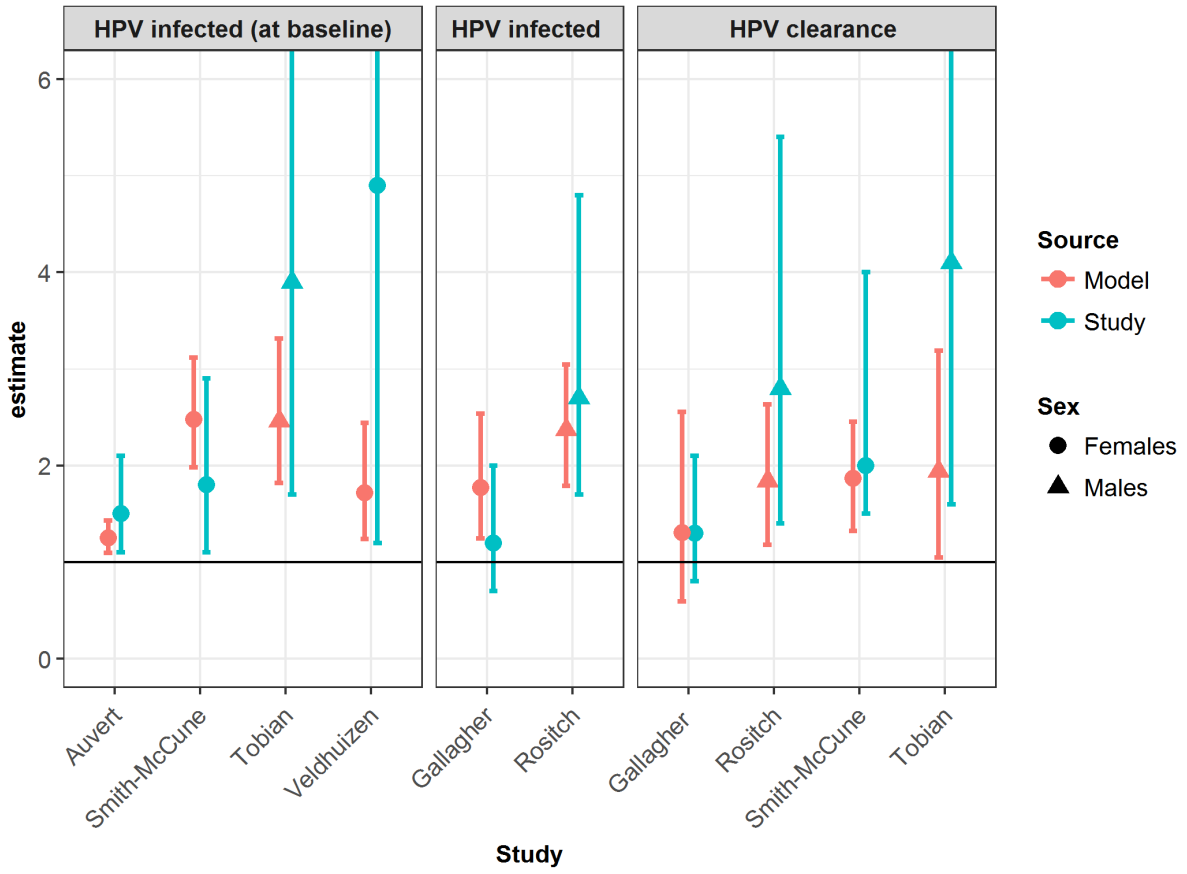


Figure 5 - Associations between HPV status and HIV acquisition. Point estimates represent hazard ratios and odds ratios, as shown in Table B 1. Intervals around point estimates reflects uncertainty due to sample size for study estimates, and parameter uncertainty and stochastic variation for model estimates.

3.4.3 Associations between HIV status and new HPV detection

Amongst the 500 simulated cohorts, the mean unadjusted hazard ratio for the association between HIV status at v_0 and newly detected HPV, in the general population aged 15-49, is 2.5 (95% CI 2.2–2.8). The corresponding mean hazard ratio for the association between new HIV detection at v_d - *1* and new HPV detection at v_d is 2.9 (95% CI 2.5-3.3). The corresponding adjusted estimates are 2.1 (95% CI 1.8-2.3) and 2.2 (95% CI 1.8-2.5). Figure 6 compares the unadjusted estimates of association between observational studies and the simulated data. Simulated associations between HIV and new HPV detection were again consistent with corresponding empirical estimates.

Further adjustment of associations for the index of concurrency in the sexual network brought the model results closer to the null, but only the association between HPV clearance and HIV acquisition did not remain significantly greater than one (Table B 3).

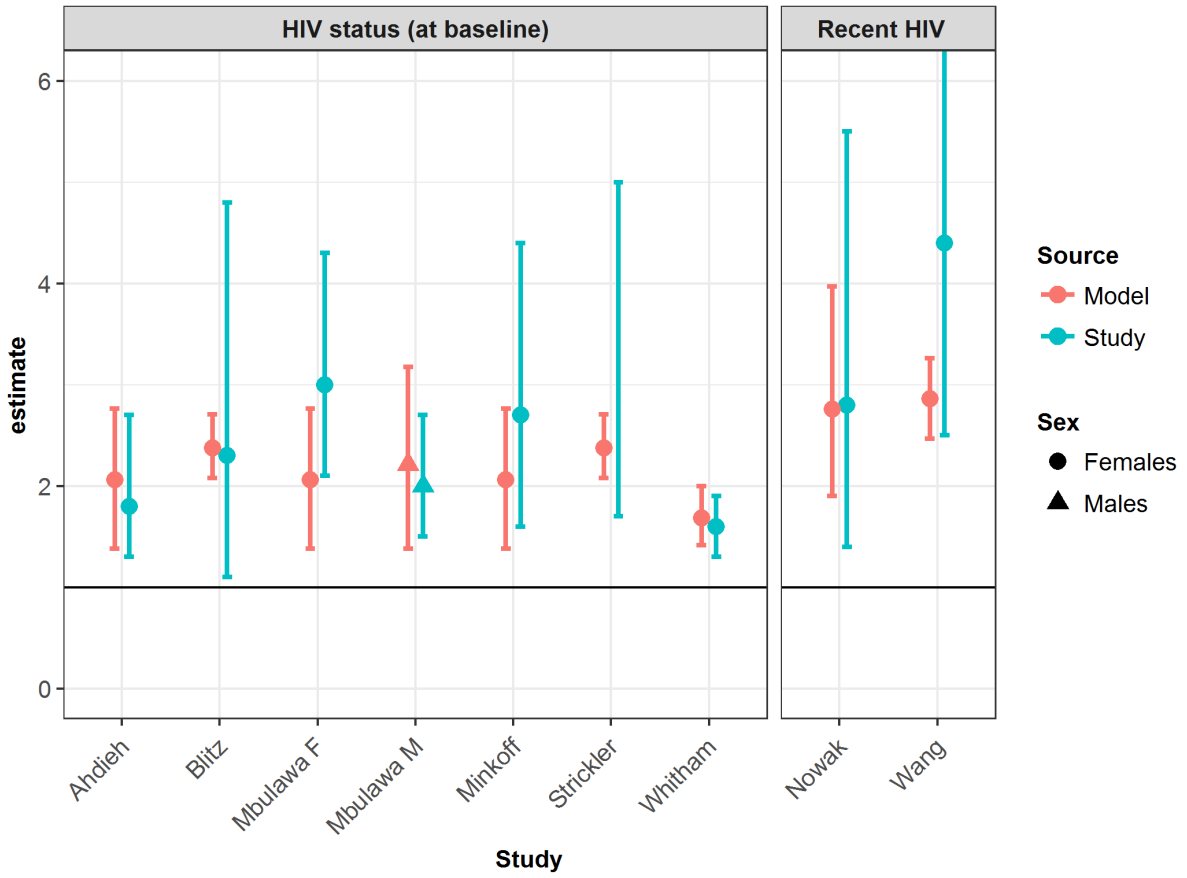


Figure 6 - Associations between HIV status and new HPV detection. Point estimates variously represent hazard ratios, odds ratios and incidence rate ratios, as shown in Table B 2. Intervals around point estimates reflect uncertainty due to sample size for study estimates, and parameter uncertainty and stochastic variation for model estimates. The interval for the Strickler study (65) *is the smallest and largest point estimate for different stages of HIV infection.*

3.4.4 Sensitivity analyses

The mean hazard ratio for the association between HPV infection at v_d - *1* and new HIV detection at v_d ranged from 2.2 to 3.6 for the 6 alternative model structures (Table B 4). The mean hazard ratio for the association between HIV status at baseline and new HPV detection ranged from 2.2 to 3.3. The mean hazard ratio for the association between HPV infection at v_d - *1* and new HIV detection at v_d was 2.4 when using the means of the posterior distributions of the HPV parameters and ranged from 2.1 (increasing the proportion of individuals who acquire viral latency by 50%) to 2.7 (decreasing the duration of HPV infection or the transmission probabilities by 50%) (Table B 5). The mean hazard ratio was at a minimum of 2.1 and at a maximum of 2.6 when changing the most important HIV and sexual behaviour parameters.

The mean hazard ratio for the association between HIV status at baseline and new HPV detection was 2.5 when using the means of the posterior distributions of the HPV parameters and ranged from a

minimum of 2.1 (decreasing the proportion of individuals who acquire viral latency by 50%) to a maximum of 3.2 (decreasing the duration of HPV infection by 50%). The hazard ratio ranged from 2.4 to 2.7 for 50% changes in the most important HIV and sexual behaviour parameters.

Measures of association compared favourably with study estimates when infectivity of or susceptibility to the one infection was doubled in presence of the other (Figure B 1-2). Full results of the sensitivity analyses performed are shown in Appendix B.2.

3.5 Discussion

In this study, unadjusted measures of association between HIV and HPV transmission in simulated cohorts compare well to unadjusted results from observational studies, even though the model assumes no direct effect of HPV on HIV transmission and no direct effect of HIV on HPV transmission. It is worth noting that the model has not been calibrated to match the observed associations between HPV and HIV transmission, although we attempted to standardize the comparisons by using consistent statistical methods and sub-populations. This suggests that most, if not all, of the observed association between HIV and HPV transmission could be due to confounding factors, in particular heterogeneity in risk behaviour. In sensitivity analyses, findings were consistent for different assumptions about the natural history of HPV in terms of acquired immunity and viral latency and were robust to changes in the main parameter values.

Our simulated measures of association remained significantly different from 1 even after controlling for sex, age and sexual behaviour. Although some studies have found that associations between HPV and HIV acquisition remain significant even after controlling for reported risk behaviours (20,21), sexual behaviour data are frequently distorted by misreporting, which leaves the possibility for residual confounding. Even when individual behaviour is accurately reported and controlled for, there is still potential for confounding due to network level effects. These network level effects are extremely difficult to adjust for in observational studies, since participants can mostly only report on their own behaviour, and not on the behaviour of the sexual network of their partners. A recent mathematical modelling study has shown that the observed associations between HSV-2 and HIV infection may be attributable mostly to confounding factors, even when individual behaviours are controlled for in multivariable analysis (130). We found that network size accounted for much of the residual association between HIV and HPV after controlling for individual-level risk factors, thus confirming the importance of network-level effects.

Some of the studies included in the comparison of associations between HIV status and new HPV detection have shown a dose-response relationship between HIV viral load or CD4 count and the rate of new HPV detection (17,65,136). This may be attributable to an effect of immune suppression on

HPV reactivation (which we allow for in the model), rather than an effect of immune suppression on susceptibility to new HPV acquisition. In sensitivity analyses, the association between HIV status and new HPV detection was weaker when HPV latency and reactivation were not included in the model, but still significantly non-null. This suggests that behavioural confounding may also partially explain the observed associations between HIV status and new HPV detection.

In sensitivity analyses, we explored the effects of assuming direct effects of HIV on HPV transmission and vice versa. Simulated associations between HIV and HPV in these scenarios were greater than those simulated when assuming no direct effects of HIV and HPV on each other's transmission, though not inconsistent with empirical estimates (Figure B 1-2). Thus, we cannot rule out the possibility that HIV transmission may indeed be inflated in the presence of HPV (and vice versa).

There are some limitations to this study. Fundamental aspects of the natural history of HPV infections are unknown, such as the extent and protectiveness of naturally acquired immunity against reinfection, the extent of viral latency and the contribution of reactivated infections to transmission, and the effects of HIV infection on duration of HPV infection and viral latency. Several sensitivity analyses were performed to check consistency of the main results against changes in model assumptions about the HPV natural history structure and changes in parameter estimates. Other modelling studies suggest that simulated associations between HIV and other STIs may be closer to null association when there is greater uniformity of sexual behaviour (130,143). The true extent of heterogeneity in sexual behaviour is unknown, and sensitivity of the model results to variation in heterogeneity has not been assessed in this analysis. Model results were however closer to null association when only considering the population with high risk behaviour (Figure 5 and 6, (132,133,147,148)). Our model considers only the South African population, and it is possible that different levels of confounding may exist in other settings, where sexual networks are different.

We conclude that observed associations between HPV and HIV transmission may result from confounding by behavioural factors and network-level effects that observational studies cannot measure and account for. This highlights the need for studies that can establish causality more conclusively, and for further modelling studies to assess what biological interactions between HIV and other STIs are plausible in other settings. Hope has been pinned on the possibility that HPV vaccination may have an impact on HIV incidence (154). While HPV vaccination is undoubtedly important in the prevention of HPV and cervical cancer, current observational evidence is insufficient to suggest that this will have a substantial impact on HIV incidence.

Chapter 4 – Estimated impact of human papillomavirus vaccines on infection burden: the effect of structural assumptions

Van Schalkwyk C, Moodley J, Welte A, Johnson LF

Publication status

Published: *Vaccine* 2019; 37 (36), 5460-5465

Author contributions

CvS developed the HPV components of the model and conceptualised and conducted the analysis, led data interpretation and drafted the manuscript for this study. LJ provided technical guidance on the modelling, conceptualisation of the study and overall leadership. AW was responsible for partial funding. LJ, JM and AW assisted with data interpretation and writing. All authors read and approved the final manuscript.

4.1 Abstract

Background: Mathematical models have been used to estimate the impact of human papillomavirus (HPV) vaccines on infection burden and cervical cancer. Models assume different mechanisms of naturally acquired immunity against re-infection, but processes of latency and reactivation of latent infection have not been explored.

Methods: This study uses an individual-based dynamic model to simulate randomised controlled trials (RCTs) for vaccine efficacy, using different assumptions about naturally acquired immunity and viral latency after clearance of HPV infection. Model estimates of vaccine effectiveness are compared to those from published RCTs. We then estimate the impact of the bivalent vaccine on HPV-16 and -18 infection burden in South Africa under these different assumptions.

Findings: When assuming no latency, simulated vaccine effectiveness overestimates results from RCTs and the model cannot match the observed difference in vaccine effectiveness between total vaccinated cohorts and more HPV-naïve cohorts. The reduction in HPV-16 and -18 burden by 2045, following roll-out of vaccination in 2014, does not depend on assumptions about natural immunity, but models that assume no latency predict ~25% greater reduction in HPV-16 and -18 burden than models that include reactivation of latent infection for all men and women.

Conclusions: Mathematical models that do not allow for reactivation of latent HPV infections may therefore overestimate the long-term impact of HPV vaccines.

4.2 Introduction

Mathematical models of infection with human papillomavirus (HPV) and its progression to cervical cancer have been crucial in estimating the epidemiological and economic impact of cervical cancer prevention strategies such as vaccination and screening. There is general consensus in the literature on the progression of disease – persistent HPV infections progress to cervical cancer through two or three stages of pre-cancerous cervical disease (155). On the other hand, there is uncertainty regarding the biological process after HPV is no longer detectable. Three possible events have been proposed in literature: 1) immediate susceptibility to re-infection with the same HPV type, 2) naturally acquired immunity against re-infection with the same HPV type and 3) viral latency that may reactivate (155).

There is evidence of both immunity and latency, but their extent and duration are difficult to estimate from epidemiological studies. It has been shown that only about 40-60% of women seroconvert following HPV infection (156) and this fraction is much lower for men and HIV-positive women (157,158). Beachler *et al.* (59) performed a systematic review and meta-analysis of studies that estimated the impact of HPV seropositivity on HPV re-detection. They estimate that seropositive women have a 30% reduced risk of re-detection compared to seronegative women, but no significant difference in risk of HPV re-detection could be shown in the studies in heterosexual men. The longitudinal studies with shorter follow-up found stronger protection against new HPV detection associated with baseline seropositivity, compared to studies with longer follow-up. This may indicate that natural immunity wanes over time but reactivated latent infections could also account for some of the new HPV detection.

Maglennon *et al.* (63) have shown that inducing immune suppression in rabbits previously infected with papillomaviruses leads to detectable levels of virus. Such a study would be unethical in humans, but studies that included immune suppressed (HIV-positive) women have shown new detection of HPV in women who reported no recent sexual activity, suggesting reactivation of latent infection is possible, particularly in immune-suppressed individuals (64,65). In a review paper, Gravitt (159) concluded that not all recurrence of HPV infections can be explained by new sexual partners and subsequent analyses of longitudinal studies support this conclusion (160–163).

Epidemiological models of HPV natural history have used different assumptions about naturally acquired immunity. For example, some do not include such a state (164,165), while a number of models allow individuals to enter an immune state upon clearance and become susceptible again according to some rate (55,61,166–168). In some models a proportion of individuals clearing HPV infection are assumed to obtain lifelong immunity and the remainder become susceptible to new infection upon clearance (169,170). In another model, all individuals who clear an HPV infection are assumed to enter a lifelong partially immune state, with a constant reduced risk of being re-infected (57).

Only one model explicitly includes viral latency in the natural history of HPV. Korostil and Regan (171) developed a dynamical model of HPV-16 transmission in Australia where both men and women obtain naturally acquired immunity or enter the latently infected stage. In this model only women can experience reactivation of the infection and all the infections of latently infected women reactivate at menopause. Ranjeva et al (67) shows that a model that includes a higher risk of newly detected HPV for individuals with previous infections fits better to longitudinal HPV prevalence data than a memoryless model. The authors suggest that reactivation of latent infection or autoinoculation from other sites could explain this increased risk.

Models with different HPV natural history structures may fit equally well to epidemiological data from the pre-vaccine era but may estimate different levels of vaccine effectiveness. Two highly effective HPV vaccines have been available for the last decade, and modelling studies have estimated the potential impact of these vaccines on HPV burden (62). The current study uses an individual-based dynamic model to investigate how well model estimates of vaccine effectiveness compare to those in randomised controlled trials (RCTs) for different assumptions about naturally acquired immunity and viral latency after infection becomes undetectable. We then estimate the future impact of the bivalent vaccine on HPV-16 and -18 infection burden in South Africa under these different assumptions.

4.3 Methods

The individual-based model, MicroCOSM, was used to estimate the burden of HPV-16 and -18 in this analysis (144,172). The model simulates, at weekly time steps, the HIV epidemic, infection with thirteen oncogenic HPV types and the underlying sexual network in the South African population. HPV types are independently simulated, and HIV co-infection is assumed to increase duration of HPV infection and rates of reactivation of latent infections, where applicable. The HIV and sexual behaviour components are described in the Appendix A and in previous publications (144, Chapter 3).

In this study, four distinct stages of HPV infection are considered: 1) Susceptible to HPV infection, 2) HPV-DNA positive infection, 3) Latent HPV infection i.e. infection not detectable by nucleic acid amplification tests and 4) cleared infection with naturally acquired immunity to HPV re-infection (Figure 3). Six different models of movement between these stages, as described in Table 1, are compared.

Table 1 - Six structures for the natural history of human papillomavirus

Model Name	**Description of possible events after HPV-DNA becomes undetectable**
Model 1	All individuals become either immune to re-infection or latently infected. Natural immunity wanes (individuals become susceptible again) and latent infections can reactivate (Figure 3).
Model 2	All individuals become immune to re-infection and immunity wanes. No one becomes latently infected.
Model 3	All individuals become either immune to re-infection or latently infected. Immunity wanes and infections can only reactivate in HIV-positive individuals.
Model 4	Women become either immune to re-infection or latently infected. Men become immune. Immunity wanes and infections can reactivate in women.
Model 5	All individuals become either immune to re-infection, latently infected or immediately susceptible to re-infection. Those who become immune remain immune forever. All latent infections can reactivate.
Model 6	All individuals become either partially immune to re-infection or latently infected. Latent infections can reactivate and the reduced risk of re-infection in the immune stage is lifelong and constant.

The six models are calibrated to South African type-specific HPV prevalence data (Table A 10) using a likelihood-based approach. Prior distributions are specified for the key parameters driving HPV infection: per sex-act transmission probabilities; durations of detectable and latent infection (both dependent on stage of HIV infection); duration of natural immunity; probability of becoming susceptible, naturally immune, or latently infected after HPV clearance; and a parameter quantifying between study variability (*Table A 12*). Five hundred thousand parameter combinations are randomly sampled from the prior distributions for each HPV type. For each parameter combination, a likelihood value is calculated to quantify how well model estimates of HPV prevalence compare to data collected in observational studies. This likelihood value is used as a weight to resample 500 parameter combinations that represent the posterior distributions of the parameters (146). The appendices contain detailed descriptions of the calibration targets (Table A 10), the calibration method (Section A.5.1), prior distributions (Section A.6.1) and posterior distributions for HPV types 16 and 18, for each model structure (Appendix C.1).

4.3.1 Vaccine effectiveness

The samples of 500 parameter combinations from the posterior distributions of each HPV type and of each model structure are used to simulate RCTs of vaccine efficacy. Trials are simulated to correspond

in design to the RCTs performed for young women (aged 15-25) (66,173) and women older than 25 (74). For both types of trial, HIV-negative women in the simulated population are enrolled in 2014 and randomised to receive the vaccine or not. Women are followed bi-annually for four years. Vaccine effectiveness is assessed at two levels of prophylactic efficacy against incident HPV16/18 infection through sexual contact: 100% or 95% of vaccinated women obtain lifelong protection (from here on called "take efficacy").

In the analysis of the simulated trial data, the outcome of interest is infection with HPV16/18 that persists for at least six months, similar to the primary outcomes in the RCTs (66,74,173). In the analysis of Kreimer *et al.* (66), results were shown for the "modified total vaccinated cohort" (m-TVC), which excluded women who were HPV16/18 DNA positive at enrolment, and the "naïve total vaccinated cohort" (n-TVC), which excluded women DNA positive for any oncogenic HPV or seropositive for HPV16/18. In the study described by Harper (173), only the "naïve" cohort was enrolled in the study i.e. only women seronegative for HPV16/18 and DNA negative for all oncogenic HPV types were randomised to receive the vaccine. Results from the study in women aged 25 and older (74) are compared for the "total vaccinated cohort" (TVC) where women with prevalent infection or seropositive to HPV16/18 were not excluded, and the "according to protocol" cohort, where these women were excluded. To simplify notation in our analysis, we will label the latter cohort similar to the naïve cohort in younger women, i.e., n-TVC.

In the PATRICIA trial, which contributed ~75% of women in the combined analysis in (66), and in the trial described in (173), women who reported more than 6 lifetime partners (LTP) were excluded. Women tend to under-report total numbers of sexual partners (174). We show results for simulated cohorts that only included women with <=6 LTP at baseline, but as a sensitivity analysis we also show results for cohorts without this exclusion criteria. To compare results to (66), we match distribution of the reported LTP in the RCT by ensuring that 75% of women enrolled in the simulated cohorts had fewer than 6 LTP. For older women, the under-reporting of LTP may be subject to not only social desirability bias, but also recall bias and therefore we also show results for simulated cohorts with no exclusion based on LTP. In our setting, with much higher HIV prevalence than in the vaccine RCT settings, the women with the highest numbers of partners are mostly excluded based on HIV-positive status.

We do not explicitly model HPV seropositivity, but consider two scenarios to approximate the effect of excluding women based on positive serostatus: 1) all women in the naturally immune stage are seropositive or 2), in addition, half of women in the latently infected stage are seropositive. There is an association between seropositivity and viral load (156), and since latent infections are undetectable, women in the latently infected stage are less likely to be seropositive.

4.3.2 Long-term impact of vaccination

The sample of 500 parameter combinations from the posterior distributions are also used to simulate HPV infection up to 2045. Since 2014, the bivalent vaccine that protects against HPV types 16 and 18 has been administered in a two-dose schedule to 9-year-old girls at public schools in South Africa. We estimate the reduction in population level HPV-16 and -18 prevalence relative to the counterfactual in which there is no vaccination. Estimates are obtained assuming 100% of vaccinated women obtain lifelong protection against incident infection. Efficacy is assumed to be the same after HIV seroconversion. Vaccination uptake is assumed to have been 90% since 2014 (42,175) and to stay constant until 2045.

4.4 Results

The six model structures all fit well to type specific prevalence data (Figures C 3-4) and the age patterns of overall HPV prevalence are consistent with data (Figure C 5). The standard deviation of the parameter quantifying between-study variability is largest for Model 2, which may indicate worse fit than the other models (Tables C 1-3).

4.4.1 Vaccine effectiveness

Vaccine effectiveness for simulated RCTs of women aged 15-25, using the six model structures described in Table 1 is shown in Table 2. In the models that allowed for latency, vaccine effectiveness is estimated to be higher when excluding HPV-seropositive individuals (n-TVC analysis) than when including seropositive individuals (m-TVC analysis), with the difference being particularly substantial if women with higher numbers of partners are not excluded. This is consistent with the observations of Kreimer et al (66). In contrast, Models 2 and 3 (which assume no or little HPV reactivation) estimate little change in effectiveness associated with different exclusion criteria. Models 2 and 3 are more consistent with observed vaccine effectiveness when assuming 95% prophylactic take efficacy than when assuming 100% prophylactic take efficacy; in the other models, either assumption could be consistent with observed effectiveness, depending on the extent to which high-risk women are excluded. In this analysis, only women who become HIV-positive during the simulated trials contribute to reactivation in Model 3, and therefore results of Models 2 and 3 are very similar. This would not be the case if HIV-positive women were included in the simulated RCTs.

Table 2 - Vaccine effectiveness against persistent HPV 16 or 18 infection among women aged 15-25. Mean effectiveness among the 500 simulated trials is shown, along with the 2.5th and 97.5th percentiles.

	<=6 Lifetime partners			**No limit on number of lifetime partners**		
	m-TVC	**n-TVC***	**n-TVC****	**m-TVC**	**n-TVC***	**n-TVC****
Kreimer (66)	89.1 (86.8;91.0)	93.6 (91.2;95.5)	93.6 (91.2;95.5)	89.1 (86.8;91.0)	93.6 (91.2;95.5)	93.6 (91.2;95.5)
Harper (173)		96.0 (75.2;99.9)	96.0 (75.2;99.9)		96.0 (75.2;99.9)	96.0 (75.2;99.9)
100% prophylactic efficacy against HPV16/18 infection						
Model 1	95.9 (90.1;99.3)	97.2 (91.7;100)	98.5 (94.6;100)	86 (72.5;95)	90 (78.3;97.7)	94.4 (86.2;99.2)
Model 2	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)
Model 3	99.7 (98.4;100)	99.9 (99;100)	100 (99.1;100)	98 (95;100)	99 (96.6;100)	99.5 (97.9;100)
Model 4	95.5 (89.1;99.3)	97 (91.7;100)	98.6 (95;100)	84.9 (71.2;95.8)	89.2 (76.7;97.5)	93.9 (85.6;99.2)
Model 5	95.4 (90.3;98.8)	96.8 (93.1;99.5)	98.4 (95.7;100)	84.3 (73.0;92.0)	88.3 (78.6;94.8)	93.4 (87.1;97.7)
Model 6	95.7 (90;99.2)	97 (92.2;100)	98.4 (95.1;100)	85.6 (75.7;92.9)	88.8 (79.5;95.7)	93.7 (87.2;98.5)
95% prophylactic efficacy against HPV16/18 infection						
Model 1	91.3 (85.4;96.5)	92.6 (86.5;98)	93.9 (88.5;98.3)	81.9 (68.6;91.4)	85.7 (74.8;94.2)	89.9 (82.1;96.3)
Model 2	95.5 (91.1;98.8)	95.5 (90.8;99.2)	95.5 (90.8;99.2)	95.4 (92.5;97.9)	95.4 (90.7;98.9)	95.4 (90.7;98.9)
Model 3	95 (90.8;98.4)	95.2 (90.5;99)	95.3 (90.5;99)	93.5 (89.3;96.7)	94.4 (89.8;98.2)	94.8 (90.3;98.4)
Model 4	91 (83.3;96.7)	92.4 (85.1;97.5)	93.9 (87.7;98.4)	81 (67.5;91.4)	84.9 (72.4;94)	89.4 (79.6;96)
Model 5	90.9 (85.3;95.2)	92.4 (87.2;96.4)	93.8 (89.6;97.0)	80.4 (70.0;88.7)	84.1 (74.7;91.6)	89.0 (82.1;94.2)
Model 6	91.1 (84.5;96.2)	92.4 (85.1;97.2)	93.8 (88.2;98.1)	81.5 (71;89.5)	84.6 (74.4;92.2)	89.3 (82.2;95.6)

m-TVC – In the modified total vaccinated cohort, women who were HPV16/18 DNA positive at enrolment were excluded.

n-TVC – In the naïve TVC, women who were HPV16/18 seropositive or DNA positive with any oncogenic type were excluded.

* Only women in the immune stage are seropositive.

** In addition, 50% of women in the latent stage are seropositive.

Table 3 shows the simulated vaccine effectiveness when the vaccine is provided to women aged 25 and older. On average, estimates from Models 1 and 4-6 compare well to RCT estimates when simulated data match RCT data based on LTP distribution, for both the TVC and n-TVC if it is assumed that there is significant seropositivity during latent HPV infection. Confidence intervals around model estimates are wide due to very few incident cases in this low risk population. When all women are included in simulated cohorts, regardless of LTP, model estimates of effectiveness are lower. Wide confidence intervals around model and RCT estimates make it difficult to identify the model structures that are most consistent with the RCT data. Nevertheless, Model 2 produces higher levels of effectiveness than observed in RCTs (74), and only when an efficacy of 80% is assumed does Model 2 match the observed effectiveness in the RCT (results not shown). A sensitivity analysis of the robustness of the results in Tables 2 and 3 to changes in the natural history – from simulating only HPV infection to simulating both HPV and cervical disease – is shown in Appendix C.6.

Table 3 - Vaccine effectiveness against persistent HPV 16 or 18 infection among women aged 25 and older. Mean effectiveness among the 500 simulated trials is shown, along with the 2.5th and 97.5th percentiles.

	Matching LTP distribution			No limit on number of LTP		
	TVC	**n-TVC***	**n-TVC****	**TVC**	**n-TVC***	**n-TVC****
Skinner (74)	47.0 (25.4;62.7)	82.9 (53.8;95.1)	82.9 (53.8;95.1)	47.0 (25.4;62.7)	82.9 (53.8;95.1)	82.9 (53.8;95.1)
100% prophylactic efficacy against HPV16/18						
1	48.4 (-16.2;88.2)	56.3 (-21.4;100)	72.4 (-5.3;100)	40.2 (13.4;66.3)	46.9 (14.0;78.1)	62.6 (22.6;88.1)
2	80.2 (21.4;100)	100 (100;100)	100 (100;100)	76.3 (57.2;90.3)	100 (100;100)	100 (100;100)
3	77.5 (3.1;100)	95.2 (51.8;100)	98.4 (73.7;100)	70.6 (47.9;87.0)	91.0 (71.1;100)	95.7 (81.2;100)
4	44.6 (-40.8;91.9)	52.1 (-32.9;100)	69.3 (-28.3;100)	34.7 (5.1;65.3)	40.7 (5.0;76.1)	57.2 (20.0;86.4)
5	47 (-17.3;90.7)	55.3 (-19.6;100)	70.7 (-13.6;100)	33.4 (4.5;58.5)	40.0 (8.7;68.0)	58.4 (19.5;86.4)
6	46.6 (-25.5;90.1)	53.8 (-21;100)	71.1 (-18.3;100)	34.9 (6.4;60.8)	35.7 (0.6;63.4)	54.3 (15.7;84.5)
95% prophylactic efficacy against HPV16/18						
1	46.4 (-19.1;87.6)	53.5 (-24.5;100)	68.6 (-16.3;100)	38.4 (12.2;64.7)	44.8 (11.5;74.1)	59.7 (18.5;85.3)
2	76.8 (18.9;100)	95.9 (63.4;100)	95.9 (63.4;100)	73.0 (54.1;87.6)	95.6 (86.2;100)	95.6 (86.2;100)
3	74.7 (-3.9;100)	91.5 (42.5;100)	94.4 (52.5;100)	67.2 (43.1;84.3)	86.6 (65.9;100)	91.0 (71.4;100)
4	42.9 (-41.4;91.6)	50 (-34.2;100)	67.4 (-26;100)	33.2 (4.0;62.4)	39.1 (2.3;73.5)	54.7 (14.4;84.6)
5	44.5 (-21.2;88.3)	52.2 (-22;100)	67.4 (-20.8;100)	31.9 (3.0;56.9)	38.1 (7.6;66.6)	55.7 (16.9;83.6)
6	44.4 (-28.9;90)	51.1 (-23.8;100)	67.8 (-23.9;100)	33.4 (6.4;59.7)	34.1 (0.5;63.2)	51.7 (11.1;83.1)

TVC – In the total vaccinated cohort, women with prevalent infection or seropositive to HPV16/18 were not excluded.

n-TVC – In the naïve TVC, women with prevalent infection or seropositive to HPV16/18 were excluded.

* Only women in the immune stage are seropositive.

** In addition, 50% of women in the latent stage are seropositive.

4.4.2 Long-term impact of vaccination

The six different HPV natural history structures estimate similar mean HPV-16 and -18 prevalence for males and females aged 15 or older in 2014, the year that vaccination is initiated in girls aged 9 (Table C 4). All six model structures predict significant reduction in HPV16/18 prevalence by 2045, but with marked differences between structures with and without latency (Figure 7). Model 1 predicts 66.5% (95% CI 52.5-83.3%) and 63.3% (48.5-86.1%) reduction in HPV-16 and -18 prevalence respectively. The model without any latency (Model 2) predicts much greater reduction in HPV-16 and -18 prevalence by 2045 (88.0% (77.4-96.5%) and 89.9% (73.2-100%) respectively). HPV16/18 prevalence also reduces substantially for men, through herd immunity. Although reductions in men are lower than for females and estimates have more uncertainty, reductions predicted for both men and women by the model without latency are ~25% greater than reductions predicted by the models with latency (Figure 7). For men and women, reductions in prevalence for models 1, 5 and 6 are very similar, indicating that different structures for natural immunity do not play an important role in predicting the long-term impact of vaccines. These results are based on the assumption that all vaccinated women experience

lifelong protection against HPV acquisition. Results for three alternative assumptions are shown in Appendix C (Figures C 6-8).

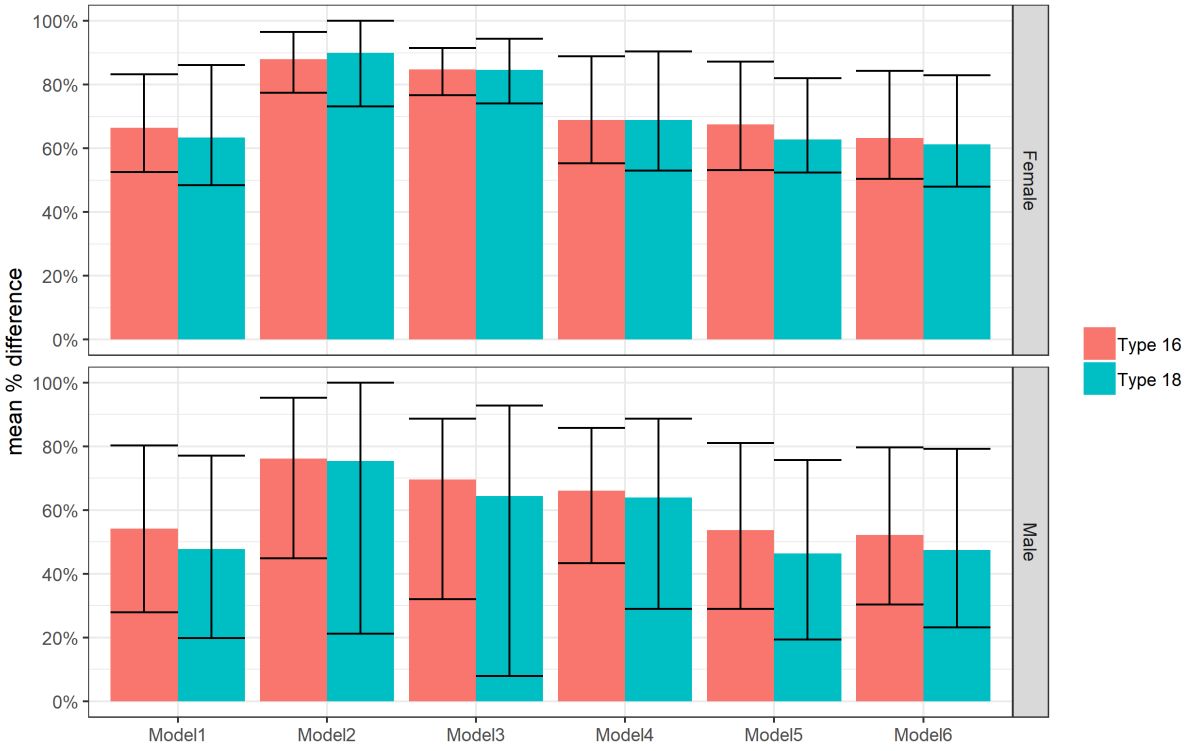


Figure 7 - Mean percentage reduction in HPV16/18 prevalence in 2045 for individuals aged 15+ with prophylactic vaccine efficacy of 100%. Vaccination coverage of girls aged 9 is assumed constant between 2014 and 2045 at 90%.

4.5 Discussion

In this study we consider six different models of HPV natural history that differ in terms of assumptions regarding natural immunity to re-infection and reactivation of latent infection after HPV-DNA is no longer detectable. In simulated RCTs, vaccine effectiveness against persistent HPV16/18 infection is compared to estimates from RCTs. Among young women, the model without latency fails to match the relative difference in effectiveness when applying different DNA- and seropositivity exclusion criteria. In this model structure, all vaccinated women are equally protected against new detection of HPV16/18 and the additional exclusion of seropositive women and women who are DNA positive with other oncogenic HPV types reduces the number of cases and number exposed equally. In the model structures that include reactivation of latent infection, the higher risk women are more likely to have latent HPV16/18 infections. Reactivated infections will lead to vaccine effectiveness that is less than prophylactic vaccine efficacy, but the stricter the trial exclusion criteria are in excluding

higher-risk women and women with prior/current infection, the smaller the difference between effectiveness and efficacy is likely to be.

For older women, the model without latency overestimates vaccine effectiveness. When no latency is assumed, simulated vaccine effectiveness is the same in younger and older women, in contrast to RCTs, which estimate a ~10% difference in effectiveness (66,74,173). This observed age difference in vaccine effectiveness can only be matched by Model 2 if it is assumed that prophylactic vaccine efficacy is lower in older women than in younger women. Although various factors may influence vaccine induced immunity, antibody titre data suggest similar efficacy in vaccinated women in different age groups (66,74,173). In the model structures that allow for reactivation of latent infection of all women, cases resulting from reactivation lead to vaccine effectiveness that better matches data.

In the analyses restricted to lower risk women, assumptions about natural immunity against re-infection did not have a clear impact on vaccine effectiveness. Assumptions about natural immunity also do not seem to play an important role in the long-term impact of vaccination on HPV16/18 prevalence. For both men and women, there is a ~25% difference in HPV prevalence in 2045 between Model 1 and Model 2. Latency and reactivation effectively increase the average duration of detectable HPV infection, which means that it takes longer for the vaccine to reduce the prevalence of HPV in Model 1 than in Model 2. We show impact on HPV prevalence for all ages (15+), as this will be a relevant predictor of cervical disease in 2045. In sensitivity analyses, we change assumptions about duration and degree of vaccine efficacy (Figures C 6-8). Although absolute values of estimates change, the substantial differences between estimates from Models 1 and 2 remain.

The results of this study could be generalizable to other settings. Although we compare results from models calibrated to South African data to results from RCTs performed in very different contexts, we do exclude women at baseline based on positive HIV status. Cost-effectiveness models of HPV-FASTER (strategies involving HPV testing and vaccinating women of all ages as cervical cancer prevention (176)) should consider the potential impact of reactivation of latent infections in the natural history assumptions.

The study has limitations. We do not explicitly include serostatus in the natural history of HPV and make crude assumptions about serostatus in this analysis to illustrate the effects of applying different exclusion criteria. A difference of ~25% reduction in HPV burden between Models 1 and 2 does not directly imply ~25% difference in cervical cancer reduction, but one would expect that model predictions of reduction in cervical cancer would also differ substantially if latency is allowed for or not. For a given model structure, parameter uncertainty leads to wide confidence intervals for model estimates and the confidence intervals around RCT point estimates are wide. This makes it difficult to judge which models are most consistent with the RCT data.

Although there is a growing body of evidence that HPV infections can become undetectable and reactivate to detectable levels, it is unknown whether reactivated infections in immunocompetent

women are likely to persist to be of clinical significance (177). In our model, reactivated infections are assumed to be as likely to persist at detectable levels as new infections. However, our estimated mean durations of latency are more than 15 years and therefore we do not simulate reactivation of intermittently detectable infections, but only reactivations that could be of clinical significance. There is great uncertainty in this duration of latency, since there is no data to inform the parameter. Long-term follow-up of cohorts such as those in (160–163), with viral load and sexual behaviour monitoring at regular intervals, could help inform this parameter.

4.6 Conclusion

This study argues that HPV natural history model structures that do not include reactivation of latent infections may not match the bivalent HPV vaccine effectiveness estimated in RCTs (which included sexually experienced women) as well as model structures that do include reactivation of latent infections. The choice of model structure also influences the predicted impact of HPV vaccination of sexually naïve women on HPV16/18 prevalence. Models that do not include a stage for latent HPV infection, and models in which only the infections of HIV-positive individuals can reactivate, may overestimate the long-term impact of HPV vaccination. Models that allow for latency may predict a slower decline in cervical cancer incidence, which underscores the importance of ongoing screening programmes in addressing comprehensive prevention of cervical cancer.

Chapter 5 – Cervical cancer surveillance in South Africa: lessons from the public sector of the Western Cape

Van Schalkwyk C, Phelanyane F, Smith M, Pienaar RC, Willems M, Moodley J, Welte A, Boulle A, Johnson LF

Publication status:

In preparation

Author contributions:

CvS led the analysis, data interpretation and drafted the manuscript for this study. AB, FP, MS, CP and MW were responsible for data curation, and provided guidance on interpretation. LJ provided technical guidance on the analysis. AW was responsible for partial funding. LJ, JM and AW assisted with data interpretation and writing. All authors read and approved the final manuscript.

5.1 Abstract

Background: South Africa has had a national cervical cancer screening programme since 2000. Monitoring and evaluation of this programme are challenging, since the records of clients accessing public health care do not include a unique patient identifier. The reported coverage of Pap smear screening is biased since both the numerator (counts of screens) and denominator (targeted population) are overestimated. The fraction of those screened who receive treatment is unknown, since there is no routine reporting of linkage of referrals to colposcopy clinics.

Methods: In the Western Cape, a unique health identifier or patient master index (PMI) is used in electronic record keeping in public healthcare facilities. This identifier enables linkage across several surveillance databases and research into aspects of the CC prevention care cascade at the individual level. We demonstrate how this data can be used to estimate screening coverage, screening intervals and colposcopy clinic attendance in the public sector of the Western Cape. Since the identifier also enables linkage of information regarding HIV testing and care, we can estimate these indicators according to HIV status. We calculate screening coverage as the number of unique women screened in the last 10 years (or 3 years for HIV-positive women), divided by the average population of women older than 30 who do not have health insurance in the middle of the given year.

Findings: We find overall screening coverage in the Western Cape to be higher than the estimates based on unlinked, aggregated data reported by the National Department of Health – around 69% in 2018 vs. the reported 56%. Coverage of 3-yearly screening among ART experienced women was 56% in 2018. The screening schedule of once in 10 years for HIV-negative women and once in 3 years for HIV-positive women is not adhered to – 1 in 5 HIV-negative women return for another routine screen within 3 years, but on average return after 15.5 years. The average time between screens for ART experienced women is 8.6 years. 54% of HIV-negative women, 38% of HIV-positive and ART naïve women, and 44.2% of ART experienced women attended colposcopy clinics within two years of receiving an indicative smear in 2017.

Conclusions: Cervical cancer prevention in the public sector of the Western Cape does not meet the 90-70-90 targets as proposed by the WHO's cervical cancer elimination strategy. Although around 90% of pre-adolescent girls who attend public schools get vaccinated and around 70% of women get screened at the appropriate times, only ~50% of women had evidence of accessing colposcopy services from 2016. This study shows the crucial importance of a unique health identifier to improve patient-level care and to monitor and evaluate implementation of prevention strategies and progress toward the 90-70-90 elimination targets.

5.2 Introduction

Despite major advancements in cervical cancer (CC) prevention the last two decades, more than half a million new cases occurred worldwide in 2018 (178). Lower- and middle-income countries (LMICs) bear the brunt of this burden, with 90% of cervical cancer deaths occurring in LMIC. To eliminate CC as a public health problem, WHO has set 90-70-90 targets for cancer prevention: By 2030, 90% of girls should be vaccinated against HPV by age 15; 70% of women should get screened at least twice and 90% of those with cervical disease should receive appropriate treatment. Most LMICs do not have nationwide screening programmes and among those that do, few have surveillance systems in place to monitor and evaluate their programmes. In particular, the use of a unique health identifier is not common, which makes linkage of patient-level data difficult (179).

The South African National Department of Health has identified CC as a major public health concern and introduced primary (HPV vaccination) and secondary (screening) prevention programmes to address the burden (34). Around 90% of pre-adolescent girls who attend public schools get vaccinated every year since 2014 (42), which translates to around 80% of all girls in the target age (43). However, the lack of a unique health identifier makes it difficult at a national surveillance level to quantify uptake of screening and access to appropriate treatment following an indicative screen result.

According to the national screening policy available since 2000, asymptomatic women should be screened by Pap smear at age thirty and at ten-yearly intervals following normal results (180). Cytology data from the National Health Laboratory Service (NHLS) are aggregated and screening coverage gets published in the annual District Health Barometer (DHB) (181). This indicator is defined as the total number of Pap smears performed in the public sector in a given year (excluding repeat or diagnostic smears) divided by a tenth of the population of women aged 30 and older. At national level, this indicator has increased from 3.6% in 2000 to 65.1% in 2018 (35).

The 2000 screening policy provided no guidelines for screening HIV-positive women, but the South African HIV guidelines released in 2010 (182) recommended a Pap smear for all HIV-positive women immediately after diagnosis and at three-yearly intervals thereafter. This policy has not been fully implemented, since increases in Pap smear numbers do not reflect the numbers of women newly diagnosed with HIV (183,184). In addition, HIV status is not well captured on the NHLS cytology form, and the aggregated data do not reflect HIV status, which has made it impossible to disaggregate screening coverage by HIV status.

The cervical cancer screening coverage estimated using aggregated data is therefore biased in the following ways: 1) the same HIV-positive women can be counted 3 times in a ten-year interval and could be younger than 30 years; 2) due to the lack of an unique identifier, it is impossible to assess policy adherence and some women may be screened more than once in 10 years; and 3) only public

sector Pap smears are counted, but the entire population older than 30 (public and private healthcare users) is used as the denominator in calculating coverage.

Two nationally representative studies have estimated screening coverage, defined as the fraction of women older than 30 who reported receiving a Pap smear within the last 10 years (185,186). Although self-reported data may also suffer from bias (187), these studies' coverage estimates of 41.6% in 2012 and 40.9% in 2016 are substantially different from the NHLS estimates for these two periods: 52% in 2012 and 63.6% in 2016 (35).

There is no systematic routine reporting on the third WHO target - linkage of referrals to pre-cancer or cancer treatment facilities - since this requires an information system with capacity to follow individuals. The only published estimates are based on research studies of linkage to pre-cancer treatment facilities. Proportions of women accessing colposcopy clinics following a cytological high-grade pre-cancerous lesion diagnosis ranged from as low as 28% in Johannesburg in 2007 to 50-63% across a range of rural and urban sites in the Eastern and Western Cape between 2003 and 2009 (36–39).

In the Western Cape, a unique health identifier or patient master index (PMI) is used in electronic record keeping in public healthcare facilities. This identifier enables linkage across several surveillance databases and allows healthcare providers to access linked named data of the specific patients who consult them, thereby improving patient-level care. It also allows research into aspects of the CC prevention care cascade at the individual level. In this study, we demonstrate how this data can be used to estimate screening coverage, adherence to screening schedules and colposcopy clinic attendance in the public sector of the Western Cape.

5.3 Methods

5.3.1 Data sources

The establishment, approach and data sources of the Western Cape Provincial Health Data Centre (PHDC) are described in detail elsewhere (188). The main aim of the PHDC is to inform patient-level care, but the platform also provides data at the individual or population level for operational, surveillance and research purposes. For this study, we accessed anonymised individual-level inferred screening and colposcopy episodes generated by the PHDC. We obtained ethical approval to perform this study from the University of Cape Town's Faculty of Health Sciences Human Research Ethics Committee (260/2016).

Data are available from 1 January 2007, when primary care clinics started implementing PMI. To be consistent with other government reports, we report on the financial years 2007/08 up to 2018/2019, i.e., for the period 1 April 2007 to 31 March 2019.

Individual-level screening episodes are defined by Pap smears submitted to the NHLS. The Pap smears are linked to the hospital database using the unique identifier, to obtain evidence that the women who required colposcopy services attended these facilities. Evidence includes diagnosis, procedure, ward, or clinic codes associated with colposcopy services. HIV and ART experience at the time of each Pap smear are inferred by linking the dates of testing, CD4 counts and viral load measurements from NHLS and dates of treatment initiation from the electronic ART register. Data received from NHLS for some screening episodes did not contain enough information to be linked to a unique PMI and as a result, these smears could not be linked to evidence of attending colposcopy or HIV services. The fraction of Pap smear reports that could not be linked to PMI decreased from 12.4% in 2007/08 to 1.3% in 2018/19. The PHDC has not been able to link individuals to the vital registration system, therefore we do not know whether women who were lost to follow-up, have died.

For this study, we use the Thembisa model (189) to obtain estimates of the female population of the Western Cape (the denominators of our coverage calculations). This model is the official source of South African HIV estimates used by UNAIDS, and is also a source of demographic estimates (Appendix D.2). Using this model, we can estimate the number of HIV-negative and positive women at each age, separate the HIV-positive group into women who are and are not yet receiving any care, and further separate the women who are receiving care according to ART experience. The HIV status of each woman receiving a Pap smear is linked on the basis of accessing HIV services, and we assume that a woman will receive care after HIV diagnosis in the form of CD4 or VL measurements. We assign an HIV status at the time of a given Pap smear to a woman as one of three categories: 1) HIV-negative or unknown status – this group corresponds to HIV-negative or HIV-positive, but undiagnosed in the Thembisa model; 2) HIV-positive, but ART-naïve - this group corresponds to those diagnosed, but ART-naïve in Thembisa; and 3) HIV-positive and ART-experienced. We do not consider periods of ART interruption in this analysis.

5.3.2 Screening coverage

We define cervical cancer screening coverage in the same way as in the national HSRC and DHS surveys and as reported in high income countries (88): the fraction of unique women older than 30 who were screened in the last 10 years. However, the Western Cape has experienced substantial growth in this population (on average 2.5% between 2007 and 2018 (189)) and for this reason we use the average population of women older than 30 during the preceding 10 years as the denominator. We only show estimates for 2016/17, 2017/18 and 2018/19, since we do not have historical data before 2007/8. To be consistent with the DHB estimate of coverage, we use the estimate of the entire WC population older than 30, but then also subtract from this denominator the fraction of women who are covered by medical aid to estimate coverage of the population of women using the public sector. Around thirty percent of women aged 30 and older participating in the annual Statistics South Africa General Household Survey reported to be on medical aid since 2010 (190). In the 2016 DHS, this

fraction was also 30% (186). Using the same logic, we calculate age-specific screening coverage and adjust for age-specific medical aid coverage during the last 10 years (Table D.2).

We assume that all the women with smears that could not be mapped to a PMI did not have a smear with a PMI at a later stage, and we count each of them based on the IDs that NHLS assigned to them. As a sensitivity analysis, we calculate coverage for the scenario that all these women received a Pap smear with a PMI at a later time in the given interval and therefore do not contribute to the numerator.

For HIV-positive women, we can estimate coverage – based on 3-yearly intervals – from 2009/10 to 2018/19. For women aged 15 or older, we count the number of unique women who received a Pap smear in the preceding 3 years and divide by the diagnosed HIV-positive population, as estimated in the Thembisa model. The fraction of women who receive HIV care in the private sector (189) is subtracted from the denominator and we separate results according to ART status. With the rapid roll-out of ART during this period, HIV sub-populations changed substantially over a period of three years. Therefore, we use the average of the preceding three years as the denominator.

To illustrate the first two sources of bias in the DHB estimate of cervical cancer screening coverage described earlier, we show the distribution of screening intervals by HIV status. We use the subset of women who received Normal cytological results – these women should ideally return 10 years later if HIV-negative and 3 years later if HIV-positive. We only include women whose subsequent screen is routine (i.e., not symptomatic). The subset of women screened in 2007 or 2008 has the longest follow-up time, so we estimate the median and mean screening interval by performing a parametric survival analysis, fitting a Weibull distribution to the proportion of women who have not returned by x years, where x goes up to twelve years. We also check for time trends in screening intervals by considering the subset of women who received a 'Normal' result in 2012 or 2013, although maximum follow up is seven years for these women. We exclude the screens of women who were not successfully mapped to PMI.

5.3.3 Colposcopy attendance

As described above, evidence of colposcopy clinic attendance was linked to the cytology database. We take the subset of women who received cytological results indicating colposcopy confirmation and show proportions accessing appropriate care by 6 months, 1 year, 2 years or thereafter. These results are separated according to HIV status. Since smears that have not been mapped to PMI could not be linked to other datasets, we exclude these smears from the analysis.

5.4 Results

The total number of Pap smears performed in the public sector of the Western Cape increased from 74 476 in 2007/08 to 128 988 in 2018/19, an increase of 73% (Table D 1). The total estimated number of adult females in the Western Cape increased from 2.1 million to 2.6 million (23%) during the same period. The proportion of smears that were performed for diagnostic reasons decreased from 32.1% to 21.4%, indicating that a greater fraction is performed for routine screening purposes.

5.4.1 Screening coverage

During the 10-year interval 1 April 2007 to 31 March 2017, 692,830 unique women aged older than 30 in 2016/17 received Pap smears in the public sector of the Western Cape (Table 4). If we divide this number by the average total population of females aged 30 and older in the 10 years preceding 2016/17, estimated coverage is 47.5%. However, if we subtract the women believed to access private sector healthcare from the denominator, our estimate of public sector coverage increases to 66.5%. This estimate slightly increases to 68.4% in 2017/18 and 68.9% in 2018/19.

If we exclude smears that could not be linked to PMI from the numerator, our estimates of coverage are 61.5% in 2016/17, 63.9% in 2017/18 and 65.1% in 2018/19.

During the 3-year interval 1 April 2007 to 31 March 2010, 48,137 or 54% of all diagnosed HIV-positive adult women received a Pap smear (Table 5). This coverage increased to 55.7% in 2018/19. Coverage among those diagnosed, but ART-naïve decreased from 64.1% in 2009/10 to 33.6% in 2018/9. Coverage among those ART-experienced increased from 40.3% in 2009/10 to 63.5% in 2018/19.

5.4.2 Screening intervals

When considering all smears, we see that roughly 1 in 5 women return for a routine screen within 3 years (18.6%) and 1 in 3 within five years (29.4%). This finding seems to be consistent over time, as illustrated in Figure 8 A). Screening intervals for HIV-negative women or women with unknown HIV are similar to the overall group. A larger proportion of women who were HIV-positive, ART-naïve at the time of the original screen returned within 3 years (27.3%) and 39.3% of ART-experienced women returned after three years. Screening behaviour of HIV-positive women changed slightly when considering different time-cohorts. The average time between two screens is 15.5 years for women with HIV-negative or unknown HIV status, 16.6 years for ART-naïve women and 8.6 years for ART-experienced women.

5.4.3 Colposcopy attendance

The proportion of smears that require follow-up with colposcopy are roughly 3 times higher among untreated HIV-positive and roughly 4 times higher among treated HIV-positive women than among women who are HIV-negative or of unknown status (Table 6). In general, the proportion of women who attend colposcopy services within 6 months after the indicative Pap smear has increased over time, but these proportions remain low, with only a quarter of HIV-positive women attending colposcopy services within 6 months in 2017. Although the proportion of women with no evidence of any colposcopy attendance has in general decreased over time, these proportions remain high.

Table 4 - Overall screening coverage in the Western Cape

	2016/17	**2017/18**	**2018/19**
A: Number of unique women screened in previous 10 years (NHLS data)	692 830	721 677	740 895
B: Thembisa average population aged 30+ during the last ten years	1459742	1496711	1536023
Coverage (A/B)	47.5%	48.2%	48.2%
Proportion of women aged 30+ on medical aid (190)	30.0%	30.4%	30.3%
C: Thembisa average population aged 30+ not on medical aid during the last ten years	1042182	1055782	1075406
Coverage in public sector (A/C)	66.5%	68.4%	68.9%
SA Health Review estimate (35)	54.9%	57.8%	55.6%
Coverage by age			
30-40	75.7%	77.4%	78.1%
40-50	73.6%	75.3%	74.5%
50-60	64.8%	66.3%	66.5%
60+	41.4%	44.4%	46.9%

Table 5 - Screening coverage among adult diagnosed HIV-positive women, defined as the proportion of these women who received a smear within the last three years.

	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/17	2017/18	2018/19
All diagnosed HIV-positive										
Average Thembisa population aged 15+ during the preceding 3 years, not on medical aid	89076	106051	123798	141497	158370	174347	189609	203840	217487	230275
Number of unique women screened in last three years	48137	63915	77071	87705	95964	104964	112024	118 173	122747	128363
Coverage	54.0%	60.3%	62.3%	62.0%	60.6%	60.2%	59.1%	58.0%	56.4%	55.7%
Diagnosed HIV-positive, not on ART										
Average Thembisa population aged 15+ during the preceding 3 years, not on medical aid	57414	61906	65282	67462	67717	66887	65863	63227	57720	52319
Number of unique women screened in last three years	36805	45876	50814	52035	47960	42451	36739	30961	24 002	17573
Coverage	64.1%	74.1%	77.8%	77.1%	70.8%	63.5%	55.8%	49.0%	41.6%	33.6%
HIV-positive, on ART										
Average Thembisa population aged 15+ during the preceding 3 years, not on medical aid	31662	44145	58517	74034	90653	107460	123746	140613	159767	177956
Number of unique women screened in last three years	12756	20269	29362	39422	52161	66552	79046	90 761	101758	112936
Coverage	40.3%	45.9%	50.2%	53.2%	57.5%	61.9%	63.9%	64.5%	63.7%	63.5%

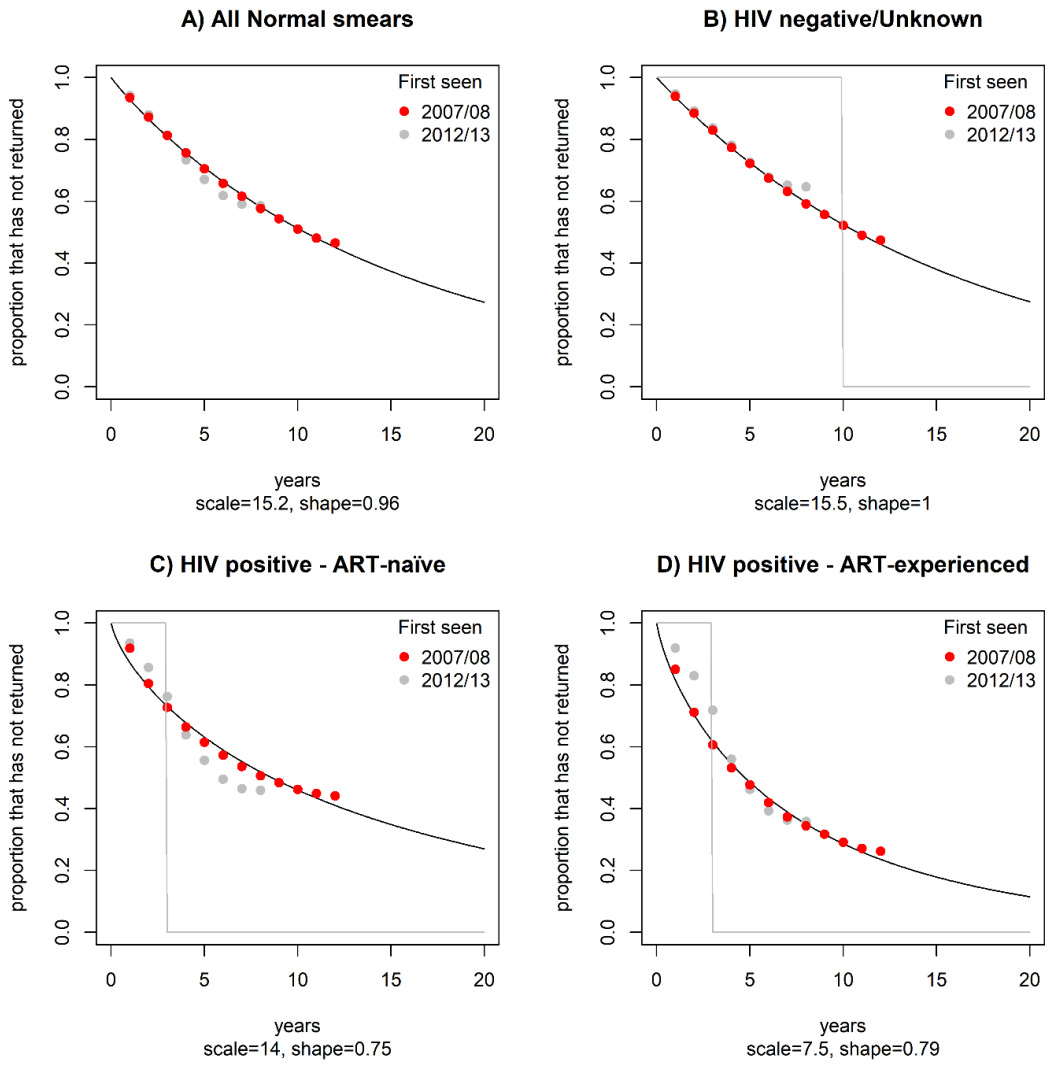


Figure 8 - Screening intervals. Red dots represent the proportion of women with a 'Normal' result in 2007/08 who have not returned for a routine follow up over time. Grey dots represent the same proportions for the 2012/13 cohort. The black line represent the Weibull survival function fit through the red dots, with scale and shape parameters below the plot. ***The grey line represents the survival curve that we would have expected to see if screening policies were perfectly implemented and adhered to.***

Table 6 - Linkage to colposcopy clinics: cumulative proportions of women who attended colposcopy clinics following a cytological result indicating confirmatory diagnosis.

	All smears with results	Number requiring Colposcopy	% Requiring colposcopy	% who attended colposcopy within:			% with no evidence of colposcopy attendance
				6 months	1 year	2 years	
	HIV-negative or unknown status						
2007	46,882	1,886	4	27.1	31.4	33.3	61.2
2008	53,233	2,607	4.9	21.9	26.6	28.3	66.7
2009	59,493	2,908	4.9	21.6	25.9	28.2	66.5
2010	65,551	2,859	4.4	22.9	27.3	29	65.5
2011	68,424	2,632	3.8	23.5	28.1	30.1	63.9
2012	66,458	2,685	4	27.5	32	34.8	60.3
2013	68,762	2,802	4.1	31.4	37.6	40.9	55.3
2014	70,662	2,894	4.1	33.3	42.3	45.1	51.7
2015	71,469	2,849	4	38.7	46.3	49.6	48.4
2016	79,020	2,863	3.6	42.8	51	53.7	44.9
2017	74,507	2,676	3.6	41.4	50.4	54.4	45.1
	HIV-positive, ART-naïve						
2007	8,240	731	8.9	16.5	21	24	68.9
2008	12,366	1,323	10.7	13.1	16.6	18.9	75
2009	15,071	2,049	13.6	13.5	17.5	20.8	73.2
2010	18,883	2,241	11.9	14.3	19	21.7	72.8
2011	18,975	2,199	11.6	13.1	17.7	20.4	74.2
2012	17,013	2,001	11.8	17.4	22.1	25.1	70.7
2013	14,991	1,744	11.6	20.3	25.6	28.6	67.2
2014	12,915	1,479	11.5	18.9	27.3	30.6	66.1
2015	11,194	1,191	10.6	21.7	29.9	34.2	62.8
2016	9,342	896	9.6	29.2	36	40	57.8
2017	6,221	606	9.7	25.5	34.3	38.4	60.5
	HIV-positive – ART-experienced						
2007	2,609	351	13.5	28.8	36.5	41.6	53
2008	4,198	781	18.6	18.2	22.8	24.8	68.8
2009	6,352	1,288	20.3	19.5	24.4	27.7	66.3
2010	9,951	1,932	19.4	18.6	23.6	25.9	68.9
2011	14,286	2,523	17.7	16.9	21.5	24.1	71.1
2012	17,761	3,226	18.2	19.7	25.2	28.2	67.1
2013	23,246	4,002	17.2	21.9	28	31	65.5
2014	30,172	4,700	15.6	18.5	27.6	31.4	65.5
2015	32,854	4,886	14.9	22	30.5	34.9	62.7
2016	39,003	4,874	12.5	26.8	37.1	42	55.7
2017	43,789	5,298	12.1	25.6	38.2	44.2	55.1

5.5 Discussion

Cervical cancer screening coverage among women aged 30 and older in the public sector of the Western Cape is 68.9% in 2018/19, close to the target of 70% coverage set by the WHO (1) and the national screening policy (33). In this study, we define screening coverage in the same way as in many high-income countries – the number of unique women screened in the reporting interval, divided by the average target population size over the interval (88). The 2016 Demographic and Health Survey used the same definition and estimated 62% coverage in the WC (unpublished estimate derived from raw data), similar to our estimate of 66.5%. The estimates of coverage from this study of individual-level data are substantially higher than the estimates from aggregated data (35), driven by the large fraction of women older than 30 who receive private healthcare (30% (190)). An advantage of our data is that it provides information on the distribution of times between cervical cancer screening visits, following a Pap smear with Normal result, which is quite different from the recommended screening interval. The high proportions of women returning for a subsequent screen much sooner than the recommended interval (Figure 8), confirms that the same women are repeatedly screened, which leads to over counting of women in the numerator of the coverage estimate in the DHB.

Coverage of screening among all HIV-positive women has increased for a while, but there was no sharp increase following the release of the HIV guidelines in 2010 that recommended cervical cancer screening at HIV diagnosis and at 3-yearly intervals thereafter. The decrease in coverage in diagnosed, untreated HIV-positive women and the increase in coverage in treated HIV-positive women are striking. This may be explained by the change in ART guidelines – as ART eligibility expanded, the diagnosed women who are untreated become a more 'select' group (i.e., they remain untreated because they prefer to not engage in HIV care, not because they are ineligible for ART). On the other hand, women receiving ART regularly engage with the healthcare system and may have more opportunity for and knowledge of screening. Despite regular interaction with healthcare, screening coverage of women on ART is still low at 63.5% in 2018/19, and the benefits of ART may be offset with cervical cancer morbidity and mortality.

Even at high coverage, a screening programme will have no impact on cervical cancer incidence if women with abnormal cytology results do not attend colposcopy clinics to receive confirmation and relevant treatment. The fractions of women requiring colposcopy services who have evidence of attending these services (based on procedure, diagnosis and ward codes) have slightly increased over time, but remain unacceptably low. Results are consistent with a study in the Overberg district that showed that 51.3% of women who required colposcopy in 2009 accessed this service (37) and a study performed in East London in 2007, in which 51.2% of 864 women accessed colposcopy (38). If strategies are not put in place to increase access to colposcopy and treatment for women with abnormal Pap smears, gains in screening coverage will be in vain.

This study has some limitations. The Western Cape Provincial Health Data Centre was established almost two decades ago, but implementation improved incrementally. From 2007, health care facilities from primary to tertiary level used the unique patient master index (PMI), enabling linkage between sources. However, in 2007, more than a tenth of Pap smears performed could not be linked to a PMI, limiting the interpretation of early results in this study. This fraction has reduced to less than 2% in 2018. The time trends of access to colposcopy service may partly be due to improved data capture and linkage to PMI. Another major limitation of the PHDC is that it is currently not linked to vital registration, and we therefore cannot quantify the fraction of loss to follow-up that is attributable to death and may over-estimate the numerator of our coverage estimates. In this study, the denominators are estimated through modelling and should be interpreted as such. We have assumed that women on medical aid do not access public sector screening and that women not on medical aid do not access private sector screening, since we lack detailed data on screening in the private sector. The analysis of screening intervals may be biased by the fact that we do not account for out-migration (which is estimated to be low in the Western Cape (191)) or mortality. This bias may result in under-estimation of the fraction that has returned at the appropriate time.

These results do not necessarily apply to the other provinces of South Africa. The coverage estimate published by the SA Health Review (35) was lower than our estimate, due to high medical aid coverage. In a province such as Kwazulu-Natal, the sources of bias are very different from the Western Cape: HIV prevalence among females (29.5% vs. 12% (189)), ART coverage among females (71.6% vs. 59.3% (189)), population growth (average 2.1% vs. 2.5% between 2007 and 2018) and medical aid coverage among females older than 30 (15% vs. 30% (190)). Since Kwazulu-Natal has more over-counting in the numerator (more HIV-positive women on ART) and less over-counting in the denominator (fewer women on medical aid), their SA Health Review estimate of 90.7% is likely to be an over-estimate of screening coverage, as opposed to the under-estimate in the Western Cape.

Cervical cancer prevention in the public sector of the Western Cape does not meet the 90-70-90 targets as proposed by WHO (1). Around 90% of pre-adolescent girls who attend public schools get vaccinated (42) and almost 70% of women get screened at the appropriate times. However, the area that requires the greatest level of improvement surrounds the follow-up of women who receives a Pap smear result indicative of colposcopy services. Although the WHO target is 90%, only ~50% of referred women had evidence of accessing colposcopy services from 2016. We have no data to assess how well the Western Cape – or any part of South Africa – is doing in terms of the second pillar of the third target: access to cervical cancer treatment. Access and uptake of screening is better in the Western Cape than in the other provinces (185,186) and we expect even lower levels of successful cervical cancer prevention in the rest of South Africa. Another major lesson learned from this study is that the 10-yearly schedule of routine smears is not adhered to at all, resulting in over-screening of some women and therefore misallocation of scarce resources. This study shows the crucial importance

of a unique health identifier to improve patient-level care and to monitor and evaluate implementation of prevention strategies and progress toward the 90-70-90 elimination targets.

Chapter 6 – Modelling the impact of prevention strategies on cervical cancer incidence in South Africa

Van Schalkwyk C, Moodley J, Welte A, Johnson LF.

Publication status:

In press: *International Journal of Cancer*

Author contributions:

CvS developed the HPV and cervical cancer components of the model and conceptualised and conducted the analysis, led data interpretation and drafted the manuscript for this study. LJ provided technical guidance on the modelling, conceptualisation of the study and overall leadership. AW was responsible for funding. LJ, JM and AW assisted with data interpretation and writing. All authors read and approved the final manuscript.

6.1 Abstract

Background: In 2019, the World Health Organisation called for the elimination of cervical cancer as a public health concern. In South Africa, despite having a national screening policy in place since 2000, diagnosed cervical cancer incidence has shown no signs of decline. This study quantifies the impact of the HPV vaccination programme on cervical cancer incidence using a mathematical model, and estimates the impact of scaling up current cancer prevention strategies, as well as proposed alternative strategies.

Methods: We extend a previously developed individual-based model for HIV and HPV infection to include progression to cervical cancer. The model accounts for future reductions in HIV incidence and prevalence and includes a detailed cervical cancer screening algorithm, based on individual-level data from the public health sector of the Western Cape. The WHO's suggested strategy to achieve elimination involves vaccinating 90% of pre-adolescent girls against HPV, screening 70% of women for pre-cancer and treating 90% of women with pre-cancer by 2030 (the 90-70-90 targets). We use the calibrated and validated model to estimate the impact of current prevention on cervical cancer incidence in the next century, and we estimate the impact of scaling up current prevention to meet the 90-70-90 targets.

Findings: The model matches stable trends in *diagnosed* cervical cancer incidence in South Africa, but it estimates increases in cervical cancer incidence over the last number of years, as a result of the ART programme, which will result in sharp increases in diagnoses. The screening programme prevented 8,600 (95% CI 4,700-12,300) cervical cancer cases between 2000 and 2019. At current levels of screening and vaccination (status quo), age-standardised cervical cancer incidence will reduce from 49.4 per 100,000 women (95% CI 36.6-67.2) in 2020, to 11.7 per 100,000 women (95% CI 7.8 – 16.8) in 2120. Scaling-up our current programme to meet the WHO's 90-70-90 targets by 2030 will prevent around 70,000 additional cancer cases by 2040 and 360,000 cases by 2120, compared to the status quo.

Conclusions: Decreasing HIV prevalence and HPV vaccination will substantially reduce cervical cancer incidence in the long term, but improvements in South Africa's current screening strategy will be required to prevent cases in the short term. Switching to new screening technologies and test-and-treat strategies will have the greatest impact.

6.2 Introduction

In 2019, the World Health Organisation (WHO) called for the elimination of cervical cancer as a public health concern (1). Despite very simple and effective screening methods being available, more than half a million new cases and 300,000 deaths were reported globally in 2018. Lower- and middle-income countries are carrying the highest burden, with 90% of deaths occurring in these settings. In Africa, it was the second most common cancer, and the leading cause of cancer death in 2018 (2). For South Africa, the WHO's International Agency for Research in Cancer (IARC), estimated an age-standardised cervical cancer incidence of 43.5 per 100,000 women in 2018 – more than three times the global estimate of 13.1 per 100,000, but similar to estimates for Southern Africa (43.1) and Eastern Africa (40.1) (2).

In high-income countries, cervical cancer incidence rates have plummeted following implementation and broad coverage of screening programmes, with some countries being close to elimination (less than 4 incident cases per 100,000 women per year is the commonly used definition of elimination) (1). Despite the introduction of a national screening strategy in 2000, the South African National Cancer Registry (NCR) has shown no decline in pathology diagnosed cervical cancer (24,41). Slow uptake of screening and low levels of linkage to colposcopy clinics may explain this slow decline (Chapter 5). Another reason may be that HIV-positive women, who have high rates of HPV and cervical pre-cancer, have longer life expectancy due to increases in anti-retroviral therapy (ART) coverage (41). A recent analysis of women on ART enrolled in South African cohorts, estimated a cervical cancer incidence of 447 per 100,000 person years (23).

Since 2014, girls aged 9 in public schools in South Africa have been vaccinated using the bivalent HPV vaccine, administered in two doses. Initial first-dose coverage was 86.6% of girls in the target group (42), and has remained around 80% since (192). However, due to the long delay between HPV infection and progression to cervical cancer, the impact that vaccination will have on cancer incidence will be negligible in the near future, and mathematical modelling studies are crucial to quantify the expected long-term impact. In the South African context, HPV vaccination modelling studies initially focused on cost-effectiveness, but were not dynamic and therefore did not consider the impact of herd immunity (119,121). One model that dynamically simulated both HIV and HPV has estimated the epidemiological impact of the nonavalent vaccine in Kwazulu-Natal (44). Assuming 90% coverage and 80% lifelong efficacy of one dose, the authors estimated that cervical cancer incidence will reduce by 74% by 2070. No published model has estimated the long-term impact of current intervention strategies on cervical cancer incidence in South Africa at a national level.

This study extends an individual-based model of HIV and HPV infection in South Africa to include progression to cervical cancer (144,172,193). We use this simulation model to estimate cervical cancer incidence over time, and assess the impact that current prevention (vaccination and screening) has had

and will have in the future. The core of the WHO's cervical cancer elimination strategy involves: 1) increasing HPV vaccination coverage to 90%, 2) increasing screening coverage to 70% and 3) increasing treatment of women with cervical disease to 90% by 2030 – the 90-70-90 targets (1). We show projected impact of scaling up our current vaccination, screening and treatment programme to meet these targets, as well as the impact of alternative vaccination and screening scenarios as presented in the South African screening policy (34) and those proposed by the WHO's cervical cancer elimination strategy (1,11).

6.3 Methods

6.3.1 HPV natural history

In a previous study, MicroCOSM (version 1) was developed to simulate individuals who represent the South African population by age and sex over time, their sexual activities, and infection with HIV and other sexually transmitted infections (144). MicroCOSM was then extended to include infection with 13 high-risk HPV types and progression to cervical cancer. The HPV natural history structure in this model is illustrated in Figure 9. The structure of stages representing HPV infection in males and females (shaded in green) 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 (Chapter 4). HPV duration and rates of reactivation of latent infection are dependent on HIV stage, and HPV types are simulated independently of each other.

This study extends the natural history of HPV infection to include cervical disease stages among the female population. Women 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 9). Rates of progression and regression are dependent on age, HIV and ART status, but HIV status does not influence transmission probabilities for HPV infection (172). Details about these parameters are shown in Appendix A.6 and A.7.

The objective of this study is to estimate incidence of cervical cancer in South Africa. The NCR defines incident cancers as first-time diagnosis per anatomical site (194). For this reason, we do not simulate possible recurrence of cervical cancer. Women progress through four stages of cancer, and can be diagnosed during each stage either through routine screening (described in Appendix A.4) or based on cervical cancer symptoms (described in Appendix A.6.2.7). Women with cancer suffer from excess mortality based on the stage of cancer diagnosis (described in Appendix A.6.2.7).

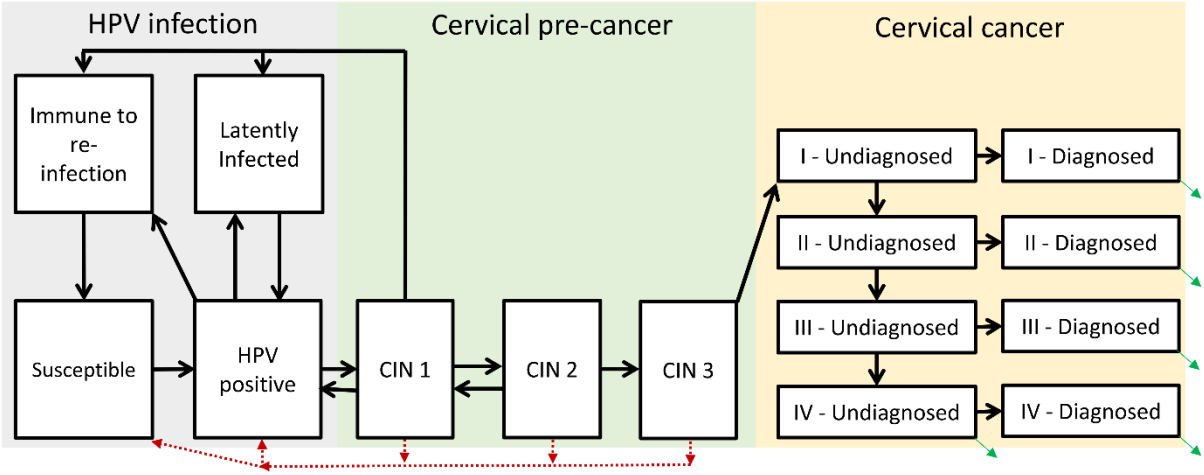


Figure 9 - Model of the natural history of HPV and cervical cancer. Black arrows represent natural movement between states, red arrows represent clearance of disease due to treatment, and green arrows represent excess mortality due to cervical cancer.

6.3.2 Calibration

We adopt a likelihood-based approach to calibration. Prior distributions of the varied parameters (HPV transmission probability per sex-act; durations of HPV infections, natural immunity, and latency; rates of regression/progression; HIV/ART multipliers and probabilities of cervical cancer diagnosis), and values of fixed parameters, were informed by a review of available data and other modelling studies (Appendix A.6). Calibration targets were determined through a review of South African data sources (Appendix A.5). The parameter combinations that were most consistent with both the prior distributions and the calibration targets were used in model simulations. In the first step of the calibration process, the main parameters that determine HPV prevalence among males and females were estimated by calibrating to type-specific prevalence data. This is the calibration process detailed in Chapter 3, and Model 1 in Chapter 4. (In the second and third steps, these parameters are fixed at the medians of the posterior distributions.) In the second step, we calibrate the parameters that determine pre-cancer disease progression (progression to CIN 1, 2 and 3 and regression from CIN 1 and 2) using histological results obtained from biopsies and endocervical curettage performed on all participants during studies in South Africa (14,28,115,195,196). In this step, type-specific HPV prevalence data are also used in calibration. In the third step, we calibrate the parameters that determine progression to cervical cancer (from CIN 3) and diagnosis by using age-specific cervical cancer incidence as reported to the NCR between 2000 and 2016 (24), as well as data on the fractions of women diagnosed in each stage of cancer in South Africa (197–200). Since the NCR only reports on pathologically confirmed cancers, this is an underestimate of the true burden of cervical cancer. We fit our model with different under-reporting assumptions: assuming that on average 7% (201), 10% or 14% (202) 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 (202) and (201)). The assumption that 10% of cervical cancer cases only receive a clinical diagnosis (no pathology

confirmation) resulted in the best fit to data, and this will be our base case assumption. The starting conditions (fractions of women in each of the HPV stages) in 1985 were obtained through an iterative process that is described in detail in Appendix A.3).

6.3.3 Vaccination, screening and pre-cancer treatment

The first national guideline on cervical cancer screening was released in the early 2000s (33), and we will assume that no routine screening happened before 2000 (although women with symptoms of cervical cancer could still be diagnosed). In the model, each woman has an age- and HIV status-specific probability of entering the screening programme. This probability increases over time and stabilises in different years depending on ART status. After the initial Pap smear, time to the next screen or treatment is drawn from Weibull distributions depending on the cytological result. To inform the probabilities of entering screening, we developed a separate, simpler simulation model (Appendix A.4.2) and fitted this model to screening frequencies in the Western Cape, using individual-level Pap smear data from the Provincial Health Data Centre (PHDC) (188). Although the national screening policy states that HIV-negative women should be screened every 10 years at ages 30, 40 and 50, and HIV-positive women 3 yearly after diagnosis, we showed in Chapter 5 that this schedule is not adhered to. Therefore, realistic distributions of time between visits and linkage to colposcopy and pre-cancer treatment were also estimated using the WC PHDC data (Appendix A.4). We assume that 75% of HIV-negative and 40% of HIV-positive women will be cleared of all abnormal lesions following treatment (Appendix A.4.9). We assume that 15% of those treated successfully will not clear the HPV infection (92–94). The status quo screening and pre-cancer treatment algorithm in the model (Figure A 5), and assumptions about diagnostic accuracy of tests, are fully described in Appendix A.4.

Since 2014, around 80% of all South African girls aged nine received one dose of the bivalent vaccine at public schools. Around 60% received both doses (192). In the model, we will assume that all vaccinated women receive 100% life-long protection against infection with HPV types 16 and 18, as well as against types 31/33/45 in 50% of women, regardless of the number of doses (66,203,204). In the model, coverage of vaccination with the bivalent vaccine since 2014 is 80% and we assume that this level will be maintained in the future.

6.3.4 Evaluating the impact of HIV and current programme

We will show the impact of our current screening and vaccination programme (given no changes to coverage, linkage to treatment or screening and treatment methods), on cervical cancer in the next century (“Status Quo”). We simulate three counterfactual scenarios to evaluate the impact that the

HIV epidemic, ART scale-up and the existing CC prevention programme has had on CC incidence in South Africa in the last 20 years. These three counterfactual scenarios are:

1) No cervical cancer interventions
2) No HIV epidemic, no increase in condom use and current levels of screening
3) No ART, increases in condom use (Figure A 4) and current levels of screening

6.3.5 Potential changes to current programme

The impact of improving aspects of our current screening programme to reach the WHO's proposed 90-70-90 targets will be estimated. In addition to the scale-up of coverage to these targets, we show the impact of removing inefficiencies in the scheduling of screens (Chapter 5), to a schedule where HIV-negative and HIV-positive women who are ART naïve are screened once between ages 30 and 40, once between ages 40 and 50, and once between ages 50 and 60, and ART experienced women are screened once in 3 years between the ages of 15 and 60. In addition, we show the impact if follow-up guidelines after inadequate screens, lower grade cytology results and pre-cancer treatment are perfectly followed.

We will consider the following scenarios (Table 7):

A) Increase coverage of vaccination with the bivalent vaccine to 90% by 2030
B) In addition to A), change the schedule of screening and follow-up to closely match policy guidelines, while keeping screening coverage constant
C) In addition to B), linearly increase the screening coverage to 70% by 2030
D) In addition to C), linearly increase linkage to pre-cancer treatment facilities to 70% by 2030
E) In addition to C), linearly increase linkage to pre-cancer treatment facilities to 90% by 2030

We do not simulate the process of HIV diagnosis in our model, and we showed in Chapter 5 that screening intervals are similar for HIV-negative women and HIV-positive women who are ART naïve, therefore we assume that the 3-yearly screening schedule starts after ART initiation.

We also simulate two alternative screening strategies that are being considered by the Department of Health, where Pap smear screening is replaced by HPV-DNA testing (34). Logistically, the most likely HPV-DNA testing platform will be Cepheid's GeneXpert, since the machines are available in many locations in the country for tuberculosis testing. Assuming that this will be the test used, we follow the same algorithm as described in the policy in our first alternative screening strategy: Women who test positive for HPV types 16 or 18 will immediately receive pre-cancer treatment (we assume 10% loss), and women who test positive for other high-risk types will be referred to colposcopy for triage. Rates of linkage to colposcopy are similar to the current strategy – i.e., the implicit assumption is that the

availability of colposcopy will increase with the need. The policy states that HIV-negative women should be screened at the same frequency as with Pap smear (ages 30, 40 and 50), but the schedule for HIV-positive women is not specified. We will assume that the policy for HIV-positive women is 3-yearly screens between the ages of 25 and 60.

This strategy is expected to lead to substantial increases in referral to colposcopy, and therefore we explore a second alternative strategy where the triage test for women testing positive with high-risk types other than types 16 and 18 is a Pap smear instead of colposcopy. Treatment will follow a Pap result with any atypical cells (ASCUS+) and we assume 10% loss between Pap and treatment. We simulate these alternative screening strategies assuming the status quo scheduling, screening coverage and linkage to colposcopy clinics will be maintained, as well as assuming the 90-70-90 targets will be reached by 2030. Assumptions about the diagnostic accuracy of tests are discussed in Appendix A.4.6.

We include in our study a simulation of the strategies implemented by the comparative modelling study by Brisson *et al.* (11), which was the first analysis published by the WHO's cervical cancer elimination modelling consortium (CCEMC). None of the three models included in this study considered HIV and its impact on the natural history of HPV, although 41 of the 78 countries simulated are in Sub-Saharan Africa. The CCEMC initiated a separate comparative modelling study of three models that estimate cervical cancer in South Africa, of which our model is one. The 3 models in the latter CCEMC study simulate HIV and its influence on HPV and cervical disease and consider additional screening and vaccination strategies for HIV-positive women (Table 7). The results of this study have not been published, but we show the results of our model here.

For this analysis, we make similar assumptions to Brisson *et al.*: 1) The HPV-DNA test has sensitivity of 90% to detect CIN2 and 94% to detect CIN3 or cancer. 2) Women are vaccinated with the nonavalent vaccine from 2020 at 90% coverage, 3) the vaccine is assumed to have 100% lifelong prophylactic efficacy, 4) loss to follow-up between screening and treatment is 10% and 5) treatment of pre-cancers is 100% effective.

Brisson *et al.* assumed that there were no existing screening programmes in the 78 countries included in their study. This is however not the case in South Africa, and to avoid initial increases in cancer in those age groups not included in the suggested screening scenarios, we slightly adjust their suggested scenarios: instead of screening once at ages 35 and 45 exactly, we screen once between the ages of 30 and 40, and once between 40 and 50. We phase out Pap smear screening for the age groups not included in these strategies (more details in Appendix E.2).

Table 7 - Scenarios to estimate the impact of changes in the current screening cascade on cervical cancer incidence. Changes will be implemented in 2020.

Scenario	Ages and schedule	Screening	Linkage to pre-cancer treatment	Vaccination coverage*
		Coverage		
Status quo (SQ)	Current levels**			80%
Scale-up of current screening programme				
A) Increase bivalent vaccination coverage	Current levels			90%
B) A + Appropriate schedule	HIV-negative and ART naive: Once aged 30-40, once aged 40-50 and once aged 50-60 ART experienced: 3-yearly, ages 15-60	Current levels		90%
C) B + 70% Screening coverage		70% by 2030	Current levels	90%
D) C + 70% Linked to colposcopy and treated		70% by 2030	70% by 2030	90%
E) C + 90% Linked to colposcopy and treated		70% by 2030	90% by 2030	90%
Replace Pap with HPV-DNA as primary screening method				
SQ1) Treat if positive with HPV16/18, colposcopy triage if positive for other HR-HPV	Current levels HIV-negative and ART naive: ages 30-60 ART experienced: ages 25-60	Current levels		80%
SQ2) Treat if positive with HPV16/18, cytology triage if positive for other HR-HPV		Current levels		80%
E1) Treat if positive with HPV16/18, colposcopy triage if positive for other HR-HPV	HIV-negative and ART naive: Once aged 30-40, once aged 40-50 and once aged 50-60 ART experienced: 3-yearly, ages 25-60	70% by 2030	90% if HPV16/18; 90% by 2030 if colposcopy	90%
E2) Treat if positive with HPV16/18, cytology triage if positive for other HR-HPV		70% by 2030	90%	90%
WHO's Cervical Cancer Elimination Modelling Consortium strategies				
F) Nonavalent vaccination for girls aged 9. Catch-up for girls aged 9-14 in 2020.	Current levels			90%***
G) F + Twice in a lifetime screening with HPV-DNA test	Once aged 30-40, once aged 40-50	70% by 2030, 90% by 2045	90%	90%
H) G + Extra screening for HIV-positive women	HIV-negative: Once aged 30-40, once aged 40-50 HIV-positive: 3-yearly, ages 25-50	70% by 2030, 90% by 2045	90%	90%
I) H + Catch-up vaccination for HIV-positive women aged 15-25		70% by 2030, 90% by 2045	90%	90%

*Coverage among girls aged 9, unless stated otherwise in description of scenario
**As described in Appendix A.4
***To be consistent with Brisson *et al.*, coverage is 90% in 2020, in contrast to Scenario A), where bivalent vaccine coverage linearly increases to 90% in 2030

6.3.6 Analysis

To estimate population-level impact of changes in coverage in South Africa's current prevention methods and the impact of the WHO strategies (Table 7), we will focus on three outcomes: age-standardised CC incidence rate (ASIR); change in ASIR, relative to the status quo; and the cumulative number of cancer cases prevented. We will show the mean estimate resulting from the 100 best fitting parameter combinations, as well as the 2.5^{th} and 97.5^{th} percentiles. The current screening programme in South Africa is struggling with capacity at colposcopy clinics, resulting in low linkage to pre-cancer treatment and long waiting times (37,38) and for this reason we also estimate the increase in referrals for certain scenarios compared to the status quo.

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 our model 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 (204, and a short description in Appendix D.2).

Cancer incidence is age-standardised using two different age weightings: the same standard population that NCR uses, for the purpose of comparing the model to NCR data (206), and the same standard population as in Brisson *et al.* (11,207), for the purpose of comparing elimination targets (Appendix E.1). Similar to Brisson *et al.*, we will consider two proposed thresholds of elimination: 4 or 10 incident cases per 100,000 women.

6.3.7 Sensitivity Analysis

We assess whether our findings are sensitive to 1) the fraction of cervical cancer diagnoses that only receive a clinical diagnosis and are therefore not included in the NCR pathology data, 2) the long-term efficacy of a single dose of the bivalent vaccine, 3) future HIV prevention efforts and 4) assumptions about viral latency and reactivation of latent infection. Results of this sensitivity analysis are shown in Appendix E.

6.4 Results

6.4.1 Calibration

Model fits to type-specific HPV prevalence data, cervical disease data, stage of cancer diagnosis and other model validation data are shown in Appendix A.8. In Figure 10, we show the model fit against overall diagnosed age-standardised CC incidence as reported by the NCR, assuming that 10% of CC

cases receive no pathological confirmation and are therefore not included in the registry. Incidence is standardised using the same world population as NCR and IARC (2). Model results are consistent with the stable diagnosed cervical cancer incidence over the 2000-2016 period. It is important to note here that estimated incidence of cervical cancer (Status Quo in Figures 13-16) is higher and shows a different trend to the *diagnosed* cervical cancer incidence in Figure 10. 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 (Appendix E.1) 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.

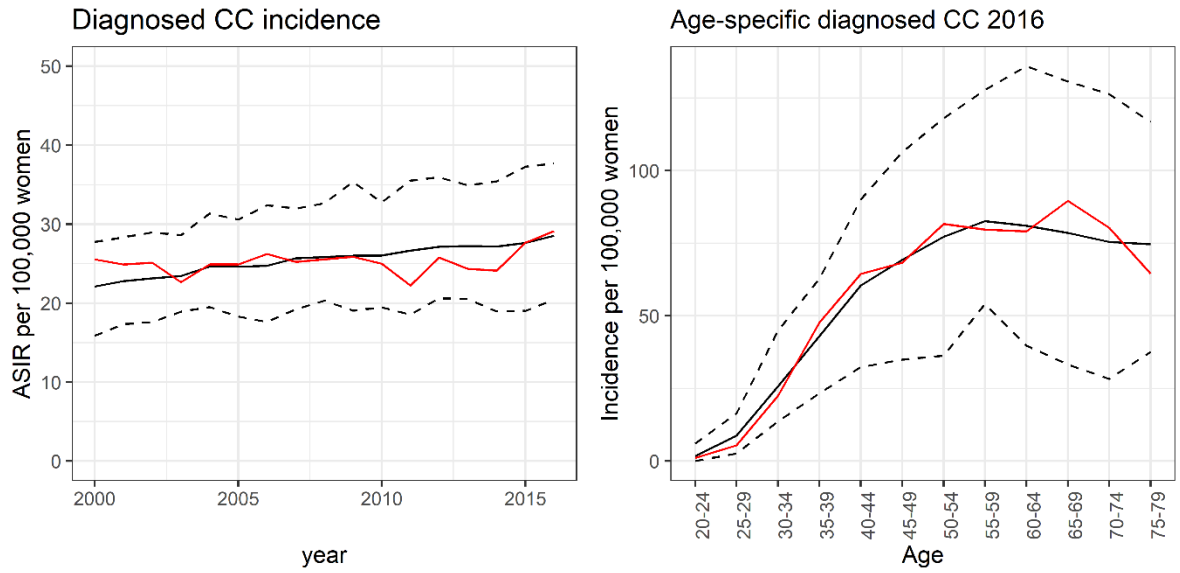


Figure 10 – Diagnosed cervical cancer incidence as calculated from NCR data (red lines) and the 100 best fitting parameter combinations (black lines show mean of 100 estimates, dashed lines show 95% percentile intervals).

6.4.2 Counterfactual scenarios

In Figure 11, we show age-standardised cervical cancer incidence for the status quo scenario, as well as the counterfactual scenarios, between 2000 and 2040. It is clear that the HIV epidemic plays a major role in the CC incidence trend over time. Had the HIV epidemic never occurred in South Africa, CC ASIR would have slowly declined over time due to screening. Had ART never been scaled up, CC ASIR would have slowly increased (as a result of HIV) and then decreased again as the population at high risk of CC died of HIV. We estimate that the ART programme led to an average cumulative increase of 27,800 (95% CI 19,900-39,400) CC cases between 2005 and 2019 due to increased life expectancy. We show in Figure 11 that the current screening programme has already had an impact on

CC ASIR. Under the scenario of no CC screening or pre-cancer treatment, our model estimates that on average 8,600 (95% CI 4,700-12,300) more CC cases would have occurred between 2000 and 2019.

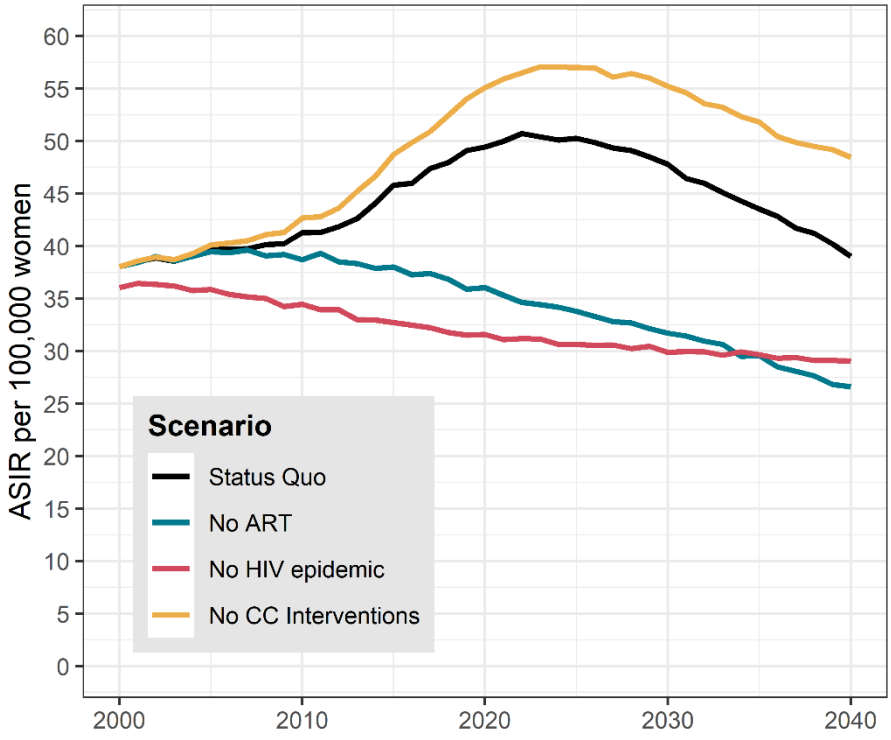


Figure 11 – Mean model estimates of age-standardised CC incidence for the counterfactual scenarios

Figure 12 is another illustration of the impact of the HIV epidemic on cervical cancer incidence. It shows that currently the majority of new cervical cancer cases, 55.6% in 2020 (95% CI 47.8-64.0%), occur among women living with HIV, but that the fraction of cases that occur in women with HIV will decline over time as HIV prevalence among adult women declines.

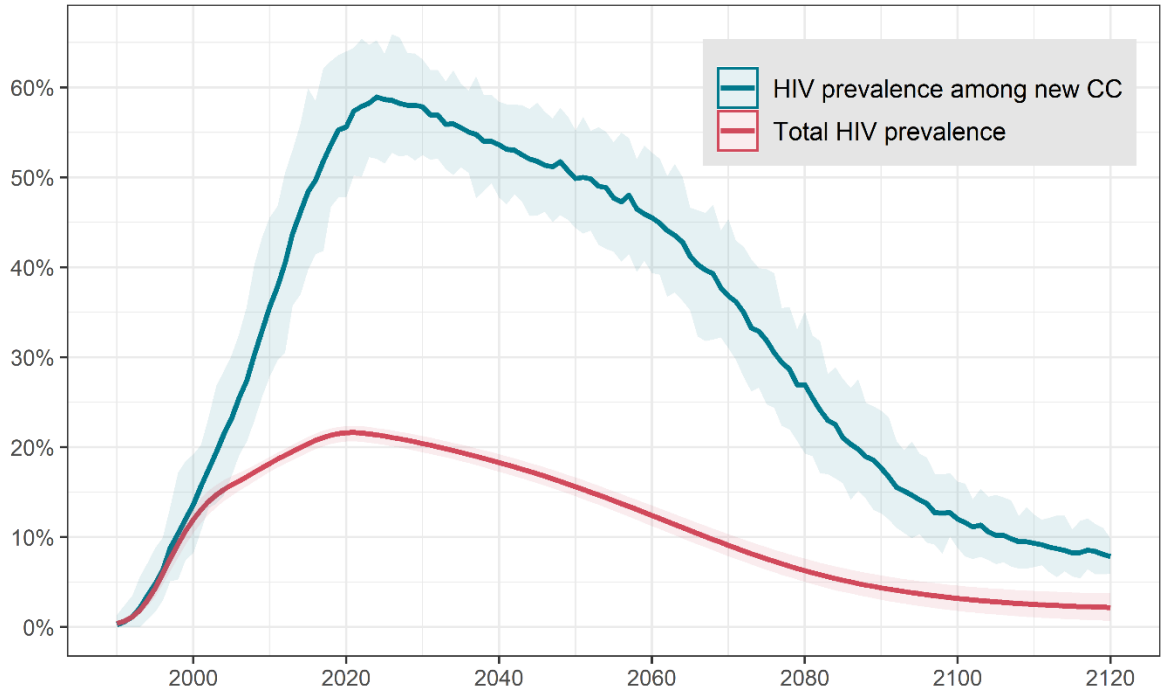


Figure 12 – Mean model estimates (and 95% confidence intervals of total HIV prevalence (among women aged 15 and older) and HIV prevalence among incident cervical cancer cases.

6.4.3 Scale-up of current screening programme

Figures 13-16 show the model's average estimates of the CC ASIR for scenarios A) to I) in Table 7. Estimates and confidence intervals of incidence, reductions in incidence from 2019 and the numbers of cases prevented under each scenario (compared to the status quo) are shown in Table 8.

In Figure 13 we illustrate the impact that the scale-up of our current CC prevention programme to meet the 90-70-90 targets will have on CC ASIR in the next century. Under the current levels of screening, pre-cancer treatment and vaccination (the status quo as described in Appendix A.4), we estimate that CC ASIR was 49.1 per 100,000 women in 2019 (95% CI 34.5-64.8), will peak at 50.9 in 2022 (95% CI 37.2-66.1) and thereafter will slowly reduce and level off at around 11.8 per 100,000 women in 2100 (95% CI 7.7-16.9). Increasing coverage of vaccination to 90% by 2030 will lead to elimination at the 10/100,000 women threshold by 2093 (60% probability of elimination by 2120). In the short term, vaccination will have no impact on cervical cancer incidence, but increases in the coverage of screening and linkage to treatment facilities will lead to substantial reductions. Increasing coverage of Pap smear screening to 70% by 2030 could reduce CC ASIR by 37.3% by 2040 (from the 2019 level) and additionally increasing linkage to treatment to 90% by 2030 could reduce CC ASIR by 52.9% by 2040 (Table 8). Reaching the 90-70-90 targets by 2030 could help South Africa reach elimination of cervical cancer (at the 10/100,000 threshold) by 2078 (90% probability of elimination by 2120). Meeting these targets will increase the number of referrals to colposcopy clinics in 2040 by 2.7-fold compared to the status quo.

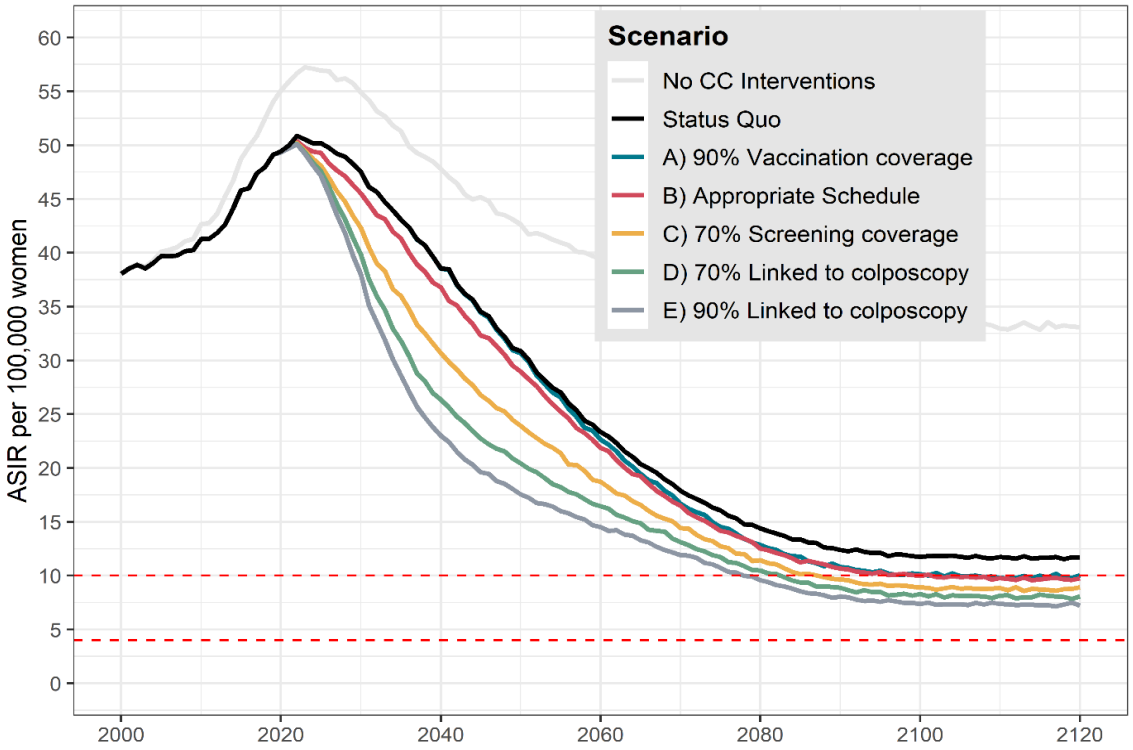


Figure 13 – Mean model estimates of age-standardised CC incidence for the scenarios where the current vaccination, screening and pre-cancer treatment programme is scaled up to meet the 90-70-90 targets by 2030.

6.4.4 Replace Pap screening with HPV-DNA screening

Figures 14 and 15 show the effect of switching our current screening strategy (Pap smear, followed by colposcopy and treatment) to HPV-DNA based approaches. In Figure 14, we assume that vaccination, screening coverage and linkage to colposcopy and treatment stay constant at current levels (as described in Appendix A.4). HPV-types 16 and 18 are the most common high-risk HPV types in South Africa, and will continue to be until the bivalent vaccine starts to significantly reduce prevalence of these types. Since we assume that 90% of women who test positive for these types will immediately receive treatment, the majority of HPV-DNA positive cases will receive immediate pre-cancer treatment leading to a reduction in CC ASIR of 46.3% by 2040 compared to 2019 (95% CI 32.2-57.2%) (Table 8). There is little difference in CC incidence between the colposcopy triage and Pap triage scenarios. However, the approach of HPV-DNA screening, with colposcopy triage of those testing positive for high-risk HPV other than types 16/18, will increase referrals to colposcopy clinics 4.6-fold compared to the status quo.

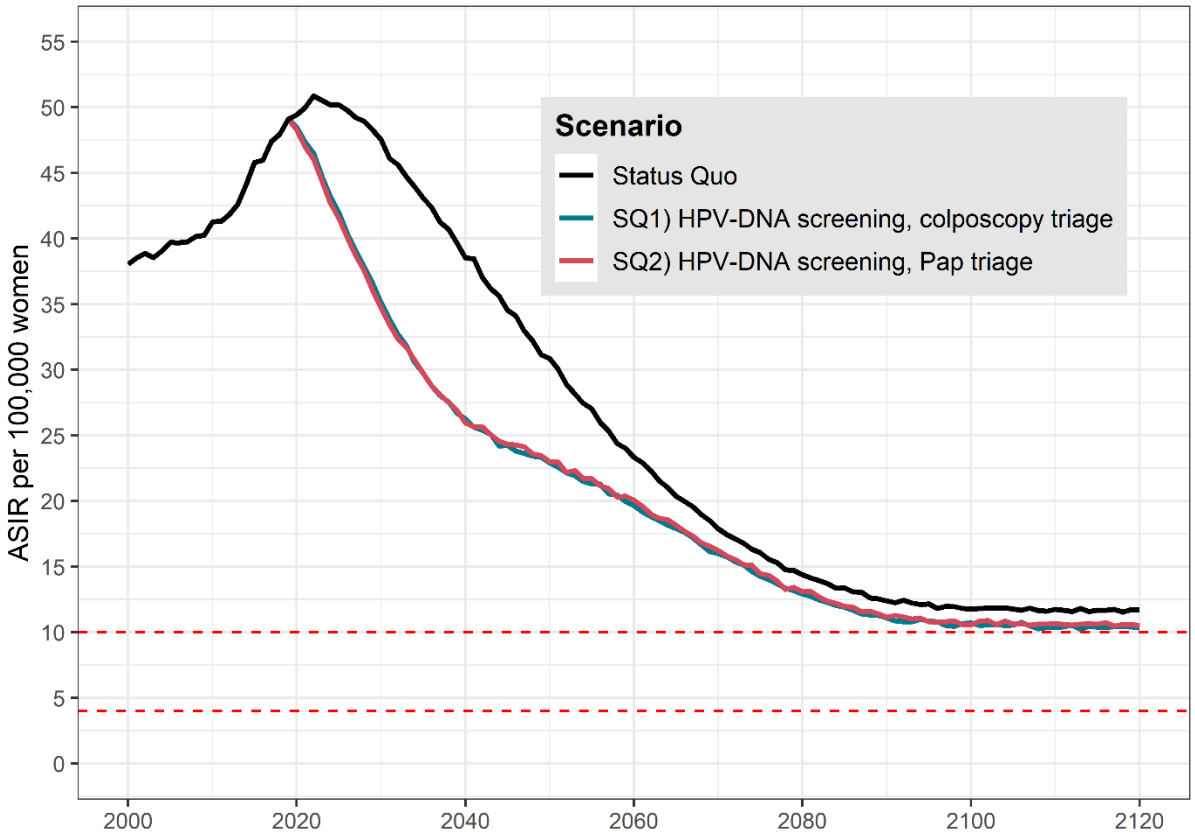


Figure 14 - Mean model estimates of age-standardised CC incidence for the scenarios where the current screening method (Pap smear followed by colposcopy) is replaced by HPV-DNA test based approaches and the current levels of vaccination and screening coverage, and linkage to colposcopy and treatment are maintained.

In Figure 15, we assume that the 90-70-90 targets will be met in 2030, and again compare the effect of switching to HPV-based screening. Again, the HPV-DNA based approaches show an immediate substantial impact on CC ASIR, with a reduction of 70.3% in 2040 compared to 2019 (95% CI 64.4-75.9%) (Table 8). In the scenario with colposcopy triage, referrals to colposcopy clinics will increase 9.6-fold compared to the status quo.

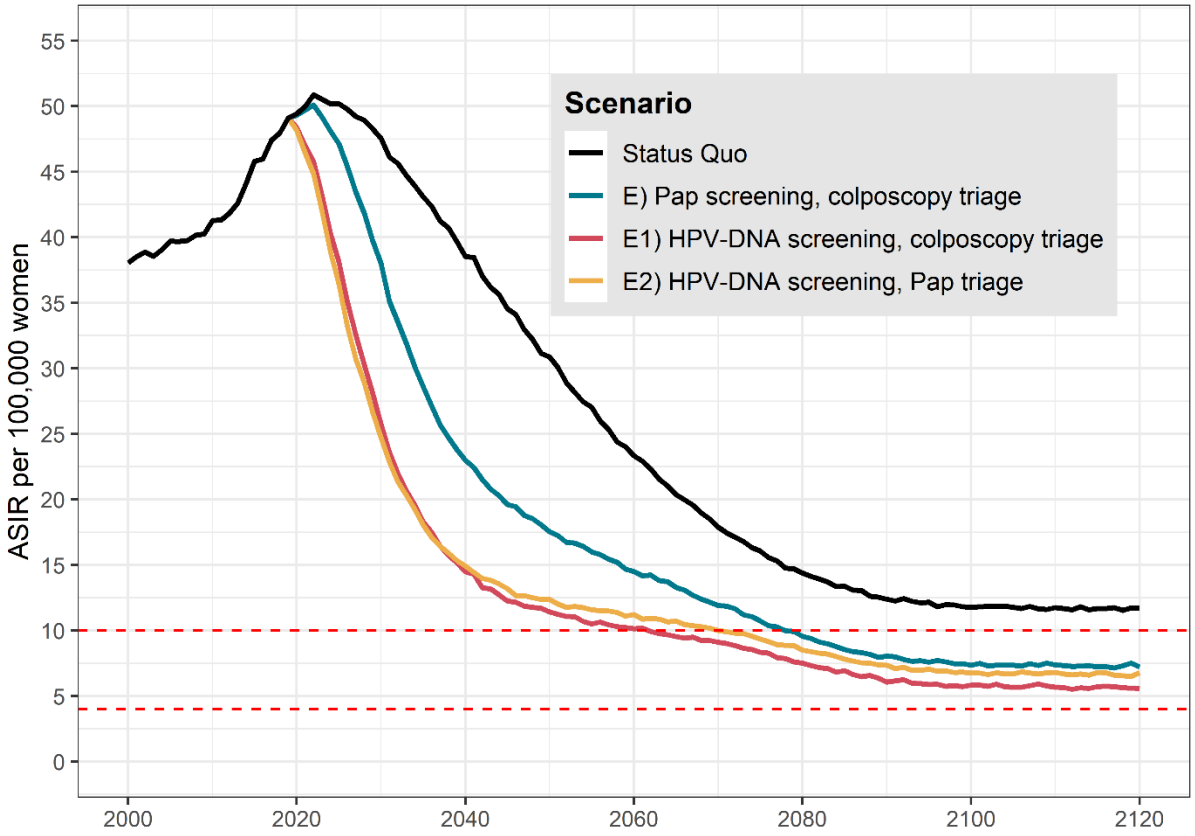


Figure 15 - Mean model estimates of age-standardised CC incidence for the scenarios where the current screening method (Pap smear, followed by colposcopy) is replaced by HPV-DNA test based approaches and the 90-70-90 targets are met by 2030.

6.4.5 WHO's CCEMC strategies

Figure 16 shows the impact of the strategies proposed by the CCEMC on cervical cancer incidence. We show that rolling out the nonavalent vaccine in 2020 (with catch-up vaccination for girls aged 9-14) and maintaining 90% coverage of girls aged 9 thereafter, will reduce CC ASIR by 2090 to around 7 per 100,000 women, compared to around 10 per 100,000 women at 90% coverage of the bivalent vaccine. Twice lifetime HPV-DNA based screens, followed by a 100% effective treatment in 90% of those testing positive, will halve CC ASIR by 2040 (Table 8) and elimination of CC at the threshold of 4 per 100,000 women may be achieved by 2092 (64% probability of elimination by 2120). The addition of 3-yearly screens for HIV-positive women aged 25-50 will further reduce CC ASIR by 2040 and prevent over 100,000 more CC cases between 2020 and 2040 than the status quo strategy (Table 8). Catch-up vaccination of HIV-positive women aged 15-24 will be beneficial for a brief period of time.

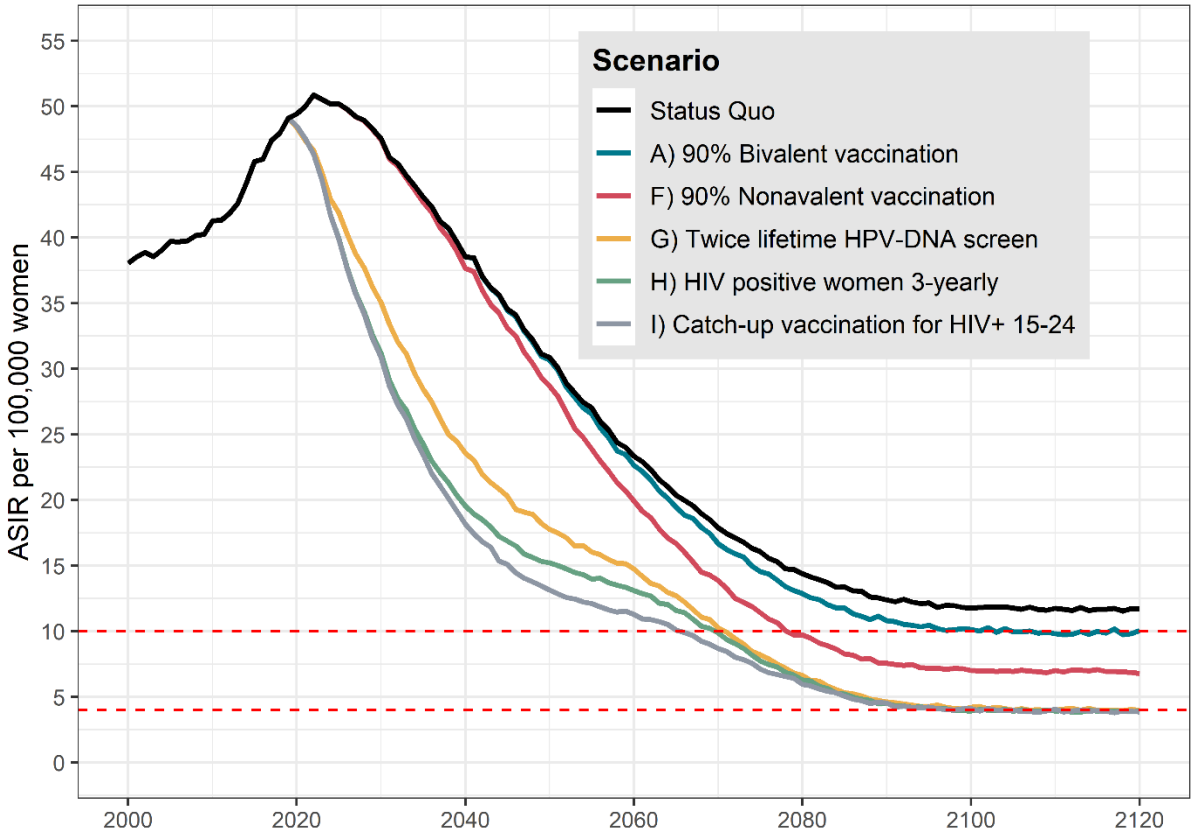


Figure 16 - Mean model estimates of age-standardised CC incidence for the scenarios proposed by the WHO's cervical cancer elimination modelling consortium (CCEMC).

Table 8 – Average model estimates (2.5th and 97.5th percentile) of age-standardised cervical cancer incidence (ASIR), reductions in ASIR from 2019 and the cumulative number of cases prevented (in thousands) for each scenario, compared to the status quo (SQ) by 2040, 2080 and 2120.

Scenario	2040			2080			2120		
	ASIR	**Reduction in ASIR from 2019**	**Cumulative cases prevented vs SQ (1000's)**	**ASIR**	**Reduction in ASIR from 2019**	**Cumulative cases prevented vs SQ (1000's)**	**ASIR**	**Reduction in ASIR from 2019**	**Cumulative cases prevented vs SQ (1000's)**
Status quo (SQ)	38.5 (26.5-51.6)	21.3% (6.0-36.5%)		14.4 (9.6-20.4)	70.6% (63.2-76.9%)		11.7 (7.8-16.8)	76.1% (70.0-81.5%)	
Improvements to current screening programme									
A) Increase vaccination coverage	38.5 (26.5-51.5)	21.3% (6.0-36.5%)	0 (0-0)	12.9 (8.6-18.4)	73.8% (67.0-80.4%)	13 (-17-44)	10.0 (6.3-15)	79.6% (73.4-85.4%)	50 (-45-149)
B) A + Appropriate schedule	36.8 (25.9-51.1)	24.8% (9.5-38.9%)	12 (-14-38)	12.5 (8.1-18.2)	74.5% (67.3-81%)	45 (-65-156)	9.8 (6.1-14.1)	80.2% (75.3-85.7%)	85 (-92-264)
C) B + 70% Screening coverage	30.7 (21.5-40.5)	37.3% (22.4-49.8%)	37 (9-68)	11.4 (7.9-16.4)	76.7% (70.7-82.5%)	139 (24-269)	8.9 (5.7-13.2)	81.8% (75.8-86.8%)	203 (22-405)
D) C + 70% treated	26.4 (18.2-34.8)	46.1% (32.7-56%)	56 (26-90)	10.5 (6.6-14.9)	78.6% (72.5-84.8%)	207 (82-352)	8.1 (5.2-11.4)	83.5% (77.1-88.3%)	288 (98-509)
E) C + 90% treated	23.0 (15.8-30.9)	52.9% (42.3-63%)	70 (37-110)	9.6 (6.2-13.5)	80.4% (75.2-85.7%)	260 (126-425)	7.2 (4.5-11.4)	85.3% (78.9-90.6%)	358 (155-601)
Replace Pap with HPV-DNA as primary screening method									
SQ1) Treat if positive with HPV16/18, colposcopy triage if positive for other HR-HPV	26.3 (18.1-35.3)	46.3% (32.2-57.2%)	79 (42-120)	12.9 (8.4-18.5)	73.7% (67.8-80.1%)	179 (57-323)	10.4 (6.7-15.2)	78.9% (72.3-85.4%)	204 (16-418)
SQ2) Treat if positive with HPV16/18, cytology triage if positive for other HR-HPV	26.0 (17.1-34.6)	46.9% (34.7-57.4%)	80 (44-121)	13.1 (8.5-18.7)	73.3% (64.2-81.4%)	176 (54-317)	10.5 (6.6-15.3)	78.6% (71.3-84.2%)	198 (10-406)
E1) Treat if positive with HPV16/18, colposcopy triage if positive for other HR-HPV	14.5 (10.3-20.4)	70.3% (64.4-75.9%)	138 (87-195)	7.5 (4.5-10.6)	84.7% (79.8-89.3%)	425 (253-638)	5.6 (3.4-8.5)	88.5% (83.7-92.7%)	561 (315-866)
E2) Treat if positive with HPV16/18, cytology triage if positive for other HR-HPV	14.9 (10.1-20.4)	69.4% (61.9-76%)	143 (92-203)	8.5 (5.2-12.1)	82.6% (77.2-88%)	412 (243-621)	6.8 (4.1-10.5)	86.2% (81.5-90.4%)	525 (285-817)
WHO's Cervical Cancer Elimination Modelling Consortium strategies									
F) Nonavalent vaccination for girls aged 9. Catch-up for girls aged 9-14 in 2020.	37.7 (25.7-50.8)	23.1% (7.8-38.1%)	2 (-1-4)	9.7 (6.1-14.5)	80.1% (74.6-85.2%)	62 (1-129)	6.8 (3.7-10.2)	86.2% (82.4-90.6%)	172 (44-319)
G) F + Twice in a lifetime screening with HPV-DNA test	23.6 (15.4-32.7)	51.8% (39.8-61.5%)	86 (47-129)	6.6 (4.5-9.3)	86.4% (82-89.4%)	269 (129-433)	4.0 (2.5-5.9)	91.8% (87.9-94.9%)	432 (215-693)
H) G + Extra screening for HIV-positive women	19.5 (12.8-27.5)	60.1% (50.6-68.3%)	109 (67-158)	6.3 (3.9-9.1)	87.1% (83.0-91.2%)	325 (176-504)	3.9 (2.3-6.1)	92.0% (88.4-94.7%)	492 (266-766)
I) H + Catch-up vaccination for HIV-positive women aged 15-25	18.2 (12.3-25.2)	62.8% (54.2-70%)	113 (70-162)	6.0 (3.4-9.0)	87.9% (83.8-91.7%)	369 (216-555)	3.9 (1.9-6)	92.1% (88.5-95.8%)	538 (309-823)

6.5 Discussion

We estimate that the implementation of the national screening policy has already prevented around 8,500 cervical cancer cases during the last 20 years, and at current levels of screening coverage, linkage to pre-cancer treatment and vaccination coverage, the age-standardised cervical cancer incidence rate is expected to reduce to 11.7 per 100,000 women in the next century.

However, in the absence of changes to current policy and practice, CC incidence is unlikely to decline substantially in the next decade. In the short term, substantial reductions in CC incidence can be achieved by pursuing the 90-70-90 targets set by the WHO, with a reduction of 53% from the 2019 ASIR estimate. Achieving the 90-70-90 targets by 2030, using our current vaccination, screening and treatment strategy could result in elimination of CC at the 10 per 100,000 women threshold and prevent an additional 360,000 cancer cases compared to the status quo in the next century. While increasing vaccination coverage from 80% to 90% among girls aged 9 and increasing screening coverage might be achievable through awareness campaigns and mobile screening services, increasing linkage to colposcopy and treatment facilities will require substantial improvements in health system capacity. In the last 20 years, around 50% of women who required these services accessed the appropriate care, and those who did experienced lengthy delays before getting appointments (Chapter 5, (36–39)). Increasing screening coverage to 70% by 2030 and following the national policy screening schedule will more than double the number of referrals by 2030, and linking those referrals to care will require more colposcopy facilities and doctors who can perform colposcopies.

Due to its higher sensitivity to detect cervical disease, HPV-DNA screening will have a substantial impact on CC incidence in the short term if it replaces Pap smear screening. Without increasing screening coverage or linkage to colposcopy clinics, it will have nearly the same impact by 2040 as maintaining Pap smear screening, but meeting the 90-70-90 targets – 46% vs 53% reduction in cancer. However, the method that requires colposcopy as a triage test will increase referrals to colposcopy clinics by 4.6-fold compared to the status quo. The impact of the two HPV-DNA screening strategies are very similar (Figure 14), since initially the majority of those testing positive will be treated (HPV16/18) and when other HPV types start to dominate, the lower sensitivity of Pap smear to detect pre-cancer is offset by the low linkage to colposcopy clinics in the colposcopy triage scenario. HPV-DNA screening, followed by Pap smear as triage method leads to substantial reductions in CC ASIR, while eliminating the need for colposcopy services, and with the appropriate training of doctors at primary care facilities, the entire cascade of screening to pre-cancer treatment could happen at these decentralised facilities. While this strategy seems like the optimal strategy in the short term, it will not lead to elimination of cervical cancer in the next century, with incidence levelling off at 10.5 per 100,000. This underscores the importance of higher screening coverage to reach elimination targets

and we show that using the same strategies, but meeting the 90-70-90 targets, would reduce cervical cancer incidence to well below 10 per 100,000 women (Figure 15).

The strategies proposed by the WHO's CCEMC are aspirational scenarios that may not be implementable in South Africa in the near future (the price of the nonavalent vaccine is prohibitive), and may not be entirely realistic (treatment of pre-cancer is not 100% effective). In addition, the proposed strategies may lead to substantial overtreatment due to the low specificity of HPV-DNA to detect cervical disease (no triage methods are suggested). Nonetheless, we show the impact that these aspirational strategies might have on cervical cancer incidence in South Africa. The finding that two HPV-DNA screens (once aged 30-40 and once aged 40-50) for all women will lead to elimination of CC at the threshold of 4 per 100,000 women is qualitatively similar to the result for Sub-Saharan Africa in Brisson *et al.* (11). However, they predict that this threshold will be crossed sooner (before 2080 vs after 2090 in our model) and that equilibrium CC ASIR will be 1.4/100,000 women, which is lower than our predicted value of 4.0 (95% CI 2.5-5.9) in 2120. Another modelling study that simulated a similar scenario of vaccination and treatment, and showed results for South Africa, estimated that cervical cancer incidence in 2100 will be 4.9/100,000 (122), similar to our estimate of 4.4/100,000 women (95% CI 2.5-6.3). Additional screening of HIV-positive women will lead to further substantial reductions in the short term, but as HIV prevalence reduces and as larger fractions of the population are vaccinated over time, this additional benefit reduces. Adding one catch-up vaccination campaign of 90% of HIV-positive women aged 15-24 to the strategy of 3-yearly screening of HIV-positive women will prevent around 46,000 additional cumulative cases. Since the majority of HIV-positive women will be diagnosed and on ART (and therefore accessing health services) between the ages of 25 and 50, high coverage of 3-yearly screening seems achievable. However, achieving 90% vaccination coverage of HIV-positive women aged 15-24 may be challenging, since these women may not be diagnosed yet or may not be accessing services.

Our study uses a population-level dynamical model to estimate the impact of HPV vaccination and screening strategies on cervical cancer in South Africa. The model also dynamically simulates infection with HIV and its effects on the natural history of HPV. The model fits well to diagnosed cervical cancer incidence data from the NCR between 2000 and 2016, and the overall age-standardised CC incidence estimate in 2018 of 41.9 per 100,000 women (95% CI 33.3-55.7) compares well to the IARC estimate of 43.5 per 100,000 women (2). We show, by simulating counterfactual scenarios where no HIV epidemic occurred or where no ART programme was implemented, that increases in cervical cancer incidence at a population level are driven by HIV co-infection. If ART was not available, the impact of HIV on cervical cancer would have been less striking, since the majority of women would have died before developing cervical cancer. Instead, the ART programme has allowed women – who have progressed to advanced pre-cancer – to survive long enough to develop cervical cancer and we estimate sharp increases in incidence between 2010 and 2022. Similar to this finding, Hall *et al.* (47) estimated, using a transmission dynamic model of HIV and HPV, that ART roll-out

will lead to short term increases in cervical cancer incidence in Tanzania. HIV prevention will bend the cervical cancer incidence curve in years to come but despite this, age-standardised cervical cancer would have remained at levels above 30 per 100,000 women if no cervical cancer prevention programme was implemented.

This study has several limitations. The majority of the data used in calibration (for estimating HPV infection and cervical pre-cancer parameters), as well as the data used to derive the screening algorithm, are from the Western Cape. The diagnosed cervical cancer incidence data used in calibration is collected at national level, but only includes pathology diagnosed cancer cases, and we had to make assumptions about the fraction of cases that only receive a clinical diagnosis and no pathology confirmation. We show in sensitivity analyses that although the different assumptions would lead to slightly different levels of CC ASIR, our findings are qualitatively similar. Our model estimates have wide confidence intervals, as seen in Table 8. Although this is partly due to parameter uncertainty, it can also be attributed to stochasticity. Due to limits in computing time, we cannot run the individual-based model over a period of 136 years (1985-2120) for a very large population, and since cervical cancer is a relatively rare disease, incidence estimates are influenced by chance. The findings may not be generalizable to other settings since the age-standardised cervical cancer depends on HIV prevalence, ART coverage and the existing prevention programmes. South Africa has the largest HIV epidemic in the world, with exceptionally high HIV prevalence levels, and our findings may therefore be less applicable to other sub-Saharan African regions, in which HIV prevalence is substantially lower. Policy decisions rely heavily on cost-effectiveness analyses, and this is also the case for cervical cancer prevention. Although this study only considers the epidemiological impact of different strategies, the next step in our research is to estimate the most cost-effective approaches by exploring combinations of screening methods, screening intervals, minimum and maximum ages of screening, vaccines (bivalent or nonavalent), catch-up vaccination and vaccinating boys.

In conclusion, South Africa's existing cervical cancer prevention programme has already prevented thousands of cases and – in combination with decreasing HIV prevalence – will substantially reduce cervical cancer incidence in the long term. Improvements in South Africa's current screening strategy could prevent large numbers of new cervical cancer cases, and switching to new screening technologies and immediate linkage to pre-cancer treatment will have the greatest impact in the short term.

Chapter 7 – Discussion

This thesis presents the first individual-based model that dynamically simulates sexually transmitted infection with both HIV and HPV, and the progression of HPV to cervical cancer in the South African population. Important aspects regarding the natural history of HPV in a setting with high HIV prevalence, and transmission dynamics between HIV and HPV are explored. After describing the cervical cancer cascade in the Western Cape province of South Africa and using this information to inform the model's screening algorithm, the model is used to estimate cervical cancer in South Africa, and the impact that existing and suggested prevention strategies will have in the future.

This chapter summarises the findings of the thesis as a combined body of work, discusses strengths and limitations inherent to the methods and findings, and concludes with recommendations to policymakers and highlights areas of future research.

7.1 Summary of key findings

7.1.1 HIV and HPV transmission dynamics

This paper was prompted by the results from the CAPRISA 004 trial, which showed significant associations between genital HPV infection and acquisition of HIV (208). This study found that participants who acquired any HPV infection during follow-up were 9 times more likely to become HIV infected than those participants who did not acquire new HPV infections. This finding led to the obvious question: Would primary prevention of HPV through vaccination have HIV prevention benefits? In South Africa, despite major advances in HIV treatment and prevention, around 240 000 people became infected with HIV in 2016 (209) and therefore the idea that the HPV vaccine could potentially prevent HIV is appealing.

This association between HPV and HIV infection risk, and inversely between HIV and HPV infection risk, has been shown in several studies and summarised in meta-analyses (20,21,101,139). The two sexually transmitted infections share a common mode of transmission, and therefore one would expect this association, but estimates are adjusted by controlling for confounding variables such as self-reported sexual behaviour (e.g., age at sexual debut, marital status, number of sexual partners), age and sex. In some studies, and in the meta-analyses, the risk after controlling for confounders remains significant and authors have suggested several biological reasons for this. However, due to possible biases in self-reported sexual behaviour data and the fact that study participants only report about their own behaviour and not that of their partners, it is plausible that the transmission associations that remain after controlling for confounders can be explained by residual confounding and confounding at the sexual network level.

We use our calibrated, individual-based model to explore these transmission dynamics. In the model, men and women can become infected with HIV and HPV and we can track their status, as well as their sexual behaviour and the behaviour of all their partners over time. We can therefore simulate cohort studies with levels of information on behaviour that are very hard to measure in the real world. For this analysis, we simulated cohorts that matched those of observational studies in design (by age, sex and time of follow-up) and analysed the resulting individual-level data using similar statistical methods. In the model, we assumed that there is no *a priori* increased transmission risk of the one infection in presence of the other due to biological reasons. When analysing our simulated cohorts, we found that we could match the empirical associations even though our model did not include biological factors. After controlling for the same behavioural factors as observational studies, the associations remained significantly greater than one. Controlling for network-level effects, such as the size of an individual's sexual network (the total number of people an individual is connected to through sexual contact, counting their partners' partners etc.), brought the estimated associations closer to the null.

We conclude Chapter 3 by noting that the observed associations can be entirely explained by individual-level and network-level sexual behaviour. Although we cannot rule out that biological mechanisms may explain some of the increased transmission risk, it was unnecessary to include biological co-factors in the model in order to match observed associations. In the subsequent modelling studies included in this thesis, no biological mechanisms to increase the risk of HIV *transmission* in the presence of HPV infection, and vice versa, were included. One biological mechanism that may increase new HPV *detection* in the presence of HIV infection - and therefore exacerbate the apparent transmission association - is the reactivation of latent infections, which we accounted for in this analysis.

7.1.2 HPV natural history structures

The next component of the natural history of HPV that is important to explore in a context with a high burden of HIV infection, is reactivation of latent infections. Two cohort studies found high levels of new HPV detection among women who reported sexual abstinence, and among HIV-positive women incidence increased with decreasing CD4 counts (64,65). Additional evidence of reactivation of latent infection include experimental proof of reactivation of rabbit oral papillomavirus upon induced immunosuppression (63) and the fact that vaccine effectiveness is lower among women with evidence of more sexual experience (66). Models that have estimated the long-term impact of vaccines have not considered latency and reactivation in their natural history model structures.

We compare different model structures by considering their fits to South African HPV prevalence data, assessing how well simulated vaccine RCTs match empirical studies, and predicting the long-term impact of vaccines on HPV prevalence. Different structures include and exclude reactivation of

latent infections and have different assumptions about natural immunity after clearance of infection. We show that models with different natural history structures fit equally well to data, with those models that do include reactivation of latency having slightly higher total likelihood values. What sets the model structures apart is their ability to reproduce the results of vaccine RCTs.

Some of the first phase III clinical trials performed to measure the efficacy of vaccines against HPV infection started in the early 2000s. The first trials enrolled sexually active women aged 15 to 25, followed by trials among women older than 25. All the trials showed very high levels of effectiveness against persistent HPV infection and pre-cancerous lesions. The striking finding of relevance to the issue of latency is that vaccine effectiveness was highest among the youngest groups of women who had no evidence of any previous HPV exposure (DNA- and sero-negative for all types), and that effectiveness was lower for women who had some evidence of previous HPV exposure (but DNA-negative for the vaccine types at baseline). In the latter group, effectiveness was also lower at older ages, although antibody titres of the vaccine do not differ by age. In a meta-analysis of some of the vaccine RCTs, the authors state: "…some outcomes might have resulted from undetected infections present before vaccination" (66). We use our model, with different natural history structures, to simulate vaccine RCTs using the same exclusion criteria, duration of follow-up and analysis methods. The main finding from this analysis is that models that do not include viral latency and reactivation of latent infections cannot match the difference in effectiveness of the vaccines in women with no evidence versus some evidence of previous HPV exposure. A model that assumes a possible latent state with reactivation of infection for everyone and 100% prophylactic efficacy of the vaccine performs best at matching estimates from RCTs.

The implication of this finding when estimating the future impact of vaccination programmes is two-fold:

1) Although RCTs do not always show perfect efficacy of the vaccines in studies of sexually active women, the vaccine itself may actually have 100% prophylactic efficacy when administered to sexually naïve women. If this efficacy does not wane, cohorts of sexually naïve vaccinated girls (and boys) may be fully protected against disease from HPV types included in the vaccines.

2) The impact of the vaccine on reducing HPV infection as predicted by models that do not consider latency may be less marked in the shorter term since individuals may become DNA-positive without sexual contact – i.e., the effect of herd immunity may take longer to manifest.

We show that for the South African context with high HIV prevalence (and therefore high rates of reactivation of latent HPV infections), models that include a latent state of HPV infection predict a ~25% lower vaccine attributable reduction in HPV16 and -18 prevalence by 2045, than a model without latency. This implies that reductions in cervical cancer may be slower than anticipated. In the subsequent modelling study included in this thesis, we assume that males and females can become

latently infected and that these infections can reactivate and progress to cervical disease and cancer among females.

7.1.3 Cervical cancer prevention in the Western Cape

In South Africa as a whole, the record of a patient's visit to a public health facility does not include a unique patient identifier. In general, this complicates individual-level care since information about the same person's visits to different care facilities cannot be linked. For cervical cancer prevention in particular, this makes it difficult to track the screening behaviour of women and their linkage to treatment after an indicative screen result. Since the ultimate aim of this thesis' modelling work is to estimate the impact that prevention strategies have had and will have on cervical cancer incidence in the future, it is important to quantify women's screening behaviour and access to treatment. Although this is not currently possible at the national level, it is possible for women living in the Western Cape.

In this province, public sector facilities offering all levels of care (primary to tertiary) have implemented the use of a 'patient master index' (PMI) since 2007 and electronic data for each individual accessing healthcare are stored at a centralised data centre. Using this PMI, it is possible to count the number of Pap smears per woman, the interval between screens, and whether the woman has been accessing HIV care. By linking individuals with indicative cytology results to information from facilities with colposcopy services or oncology departments, it is possible to quantify linkage to follow-up care.

We use this database to quantify screening coverage, screening intervals and linkage to treatment in the Western Cape, all stratified by HIV and ART status. Public sector screening coverage in South Africa as reported yearly using aggregated data suffer from over-counting in the numerator (women who had more than one screen in the reporting period are counted more than once) and in the denominator (women who only access private healthcare are not excluded). We find that in the WC, these two sources of bias cancel out, and that overall screening coverage (defined as the fraction of unique women screened in the last 10 years) is around 55% between 2016 and 2018 – similar to the estimates calculated using aggregated surveillance data (with no unique person ID) reported in the South African health review (35). We find that coverage among women on ART (defined as the fraction of unique women screened in the last 3 years) who regularly interact with healthcare has also been stable around 55% since 2014. Linkage to treatment was found to be dismally low – with only 40% of HIV-negative women and around a quarter of HIV-positive women accessing colposcopy services within 6 months of an indicative Pap smear result. In addition, we show that guidelines for screening intervals are not adhered to, with almost a fifth of HIV-negative women returning for a routine screen following a normal Pap within 3 years.

We use the information obtained from this analysis to inform screening in our model. Individuals in the model are assigned age-, ART- and time-specific probabilities of entering screening and time to next screen are drawn from distributions – all informed by WC screening behaviour. Probabilities of accessing colposcopy services are also informed by this data. Although the public healthcare in the Western Cape does not necessarily compare well with healthcare in the rest of the country, aggregated coverage estimates for the WC are similar to national estimates in recent years and we use screening behaviour of women in the public sector of the WC as a proxy for that of all women in South Africa.

7.1.4 The impact of prevention on cervical cancer incidence

The main objective of this thesis is to develop an appropriate model for cervical cancer incidence in South Africa. Such a model can serve a number of purposes of public health significance. Firstly, it can provide estimates of current and historic levels of incidence. The only national level estimates of cervical cancer incidence are those provided by the pathology-based National Cancer Registry (NCR). These numbers do not reflect cases that only received a clinical diagnosis and no pathology confirmation, and - of course - cases that never received a diagnosis at all. More importantly, diagnosis of cervical cancer lags several years behind the first occurrence of cancer in a woman's body and therefore an appropriate model can predict changes in diagnosis trends in the short term. Secondly, an appropriate model can be used as a tool to measure and evaluate the success of the cervical cancer prevention programme. Thirdly, an appropriate model can assist policy makers with difficult decisions on the optimal strategies for the prevention of cervical cancer.

An appropriate cervical cancer model for the South African context is one that dynamically simulates HIV infection and prevention efforts in the population. In Chapters 3 and 4 we establish an appropriate model structure for the natural history of HPV infection, and its interaction with HIV. In Chapter 6, this model is further developed to add stages of cervical disease and cancer and a detailed screening algorithm based on the data analysed in Chapter 5. We fit this model to data on prevalence of type-specific HPV infection and pre-cancer by HIV and ART status, stage at cervical cancer diagnosis and age-specific diagnosed cervical cancer incidence, as reported to the pathology based National Cancer Registry.

By simulating counterfactual scenarios, we illustrate the massive impact that the HIV epidemic has had on cervical cancer incidence in South Africa. Had there been no HIV epidemic (but the screening programme was still scaled-up to current coverage levels among HIV-negative women), we estimate that age-standardised cervical cancer incidence rate (CC ASIR) would have been 27.9 per 100,000 women (95%CI 18.5-38.3) in 2020, and consistently declining due to screening. In contrast, we estimate that the actual CC ASIR for 2020 is 44.9 per 100,000 women (95%CI 33.2-60.7) and that incidence, after gradually increasing over the years, will peak in 2022 and then start to decline. The scale-up of ART has increased cervical cancer by extending the life expectancy of the women at

highest risk of cervical cancer and because of this we estimate sharp increases in incidence between 2010 and 2022. Due to the predicted long-term declines in HIV prevalence, we expect CC ASIR to decline after the peak in 2022 – with or without cervical cancer prevention.

We estimate that – despite low coverage and low linkage to treatment – the screening programme has prevented around 8,600 cases between 2000 and 2019 and that – mainly due to reducing HIV prevalence and the sustained coverage of HPV vaccination – cervical cancer incidence will reduce to a quarter of its current levels in the next century. However, in the short term, we estimate that on average 70,000 cases of cervical cancer can be prevented by 2040 if our current screening strategy can be scaled up to meet the WHO's 90-70-90 targets. This may double the number of referrals to colposcopy clinics – the component of the current system that is severely under strain. Currently, the most common strategy followed in South Africa is Pap smear screening followed by colposcopy and immediate treatment if colposcopy positive. A more sensitive but less specific proposed alternative strategy is HPV-DNA based screening, followed by immediate treatment for those testing positive for HPV types 16 or 18 and Pap smear triage for those testing positive for other high risk HPV types, followed by treatment for those with positive cytology. We estimate that even if current levels of screening coverage are maintained, this strategy will have nearly the same impact by 2040 as scaling up our current system to meet the 90-70-90 targets – almost halving cervical cancer incidence in 20 years while eliminating the need for colposcopy clinics.

7.2 Strengths and limitations

The specific strengths and limitations of each analysis in this thesis were presented in the discussion sections of chapters 3 to 6. We will now discuss the overarching strengths and limitations of the complete body of work.

The fact that this work built on an existing individual-based model to investigate HPV and HIV transmission dynamics, HPV natural history and cervical cancer incidence was both a strength and a limitation. The population demographics, sexual network and HIV infection components of the model were already well developed, and could be applied with little change for this project. The individual-based model was the appropriate framework for answering the research questions of Chapters 3 and 4, since it involved simulation of cohorts of individuals. The individual-based approach is also ideal for simulating detailed screening algorithms – although this body of work did not delve into optimizing screening algorithms for maximum impact on cervical cancer incidence, this will be investigated in future work using the foundation that was established here. The limitation of the framework is that the individual-based approach is extremely computationally intensive and time-consuming when simulating incidence of cancer. Since cancer is such a rare event, the number of individuals simulated

over time needs to be large enough to generate sufficient numbers of cancer cases, while also considering the impact of stochasticity.

The parameters that determine the components of the model – HPV infection, pre-cancer and cancer – were calibrated using data obtained by performing extensive reviews of South African data sources. However, the majority of the data are from the Western Cape (WC), which may raise concerns about the generalisability of our results to the entire country. In particular, the model's screening algorithm and cancer mortality rates are based almost entirely on WC data, and only one study outside the WC estimated prevalence of pre-cancerous lesions. Of the 20 study estimates of type-specific HPV prevalence, 12 were from the WC.

Prior distributions of parameters were determined by literature reviews, and where no data exist to inform distributions, we opted for relatively non-informative prior distributions. Uncertainty in our model estimates are driven by those parameters that are difficult to measure in empirical studies, such as per-sex act transmission probabilities, durations of natural immunity and latency, and progression and regression of advanced stages of pre-cancer.

The modelling of HIV infection and progression in version 1 of MicroCOSM, which this thesis built on, has some limitations. Medical male circumcision is not simulated, although this has increased in South Africa in recent years and reduces risk of both HIV and HPV transmission. This limits projection of HIV prevalence in the long term using this model. In addition, this version of MicroCOSM is based on a clinical HIV staging system, rather than a CD4-based staging system, and it doesn't explicitly represent HIV testing and diagnosis, or viral suppression on ART. This means that the model might not represent changes in the cascade of HIV care over time as realistically as we would have liked. The model also does not take international migration into account. To overcome these limitations, we reweight estimates from our model using population estimates (by age, sex, HIV and ART status) from the Thembisa model when projecting after 2020.

In order to estimate the impact of HPV vaccination on cervical cancer, one has to project well into the future. The first cohort of girls aged 9 were vaccinated in 2014 and won't reach the peak age of cervical cancer incidence before 2050. To show the cumulative, increasing impact of vaccination, we project cervical cancer incidence for the next century. In order to do this, we have to make assumptions about long-term population demographics and sexual behaviour. We also assume that drugs, diagnostics and treatments will have the same efficacy over time. All of these assumptions lead to uncertainty in our findings.

The calibration of our main cervical cancer parameters depended on the assumed fraction of diagnosed cervical cancers that do not receive a pathological diagnosis, and that are therefore not included in the NCR database. Estimates of this fraction are very different for the two population-based cancer registries in South Africa and we had to make assumptions about this value at the national level. We

are currently working with colleagues at the WC PHDC to develop an algorithm that will identify women with cervical cancer using the unique patient identifier and laboratory, diagnosis, pharmacy and procedure codes. This will allow us to estimate the fraction of cases that did not receive a pathological diagnosis in the Western Cape since 2007 which, in combination with the values from the two small population registries, will provide a more robust estimate.

7.3 Future research

This study has laid the foundation for future work to determine optimal cervical cancer prevention strategies in South Africa. While this study considered the impact of different strategies on cervical cancer incidence, health system capacity was only superficially considered, and no cost-effectiveness analyses were performed. In our next analysis we will work closely with health economists to determine the optimal screening strategy in terms of cost and impact by varying vaccine choice and screening technologies, as well as the start and stop ages of screening and intervals between screens. In addition, we will perform an analysis of the balance between benefits (preventing more cases) and harms (unnecessary treatment) when comparing the strategies with Pap smear screening (low sensitivity, high specificity) to the strategies with HPV-DNA screening (high sensitivity, low specificity).

MicroCOSM has been extended to include a more detailed HIV natural history, more options for HIV prevention (e.g. medical male circumcision) and will be extended to include smoking and obesity (risk factors for CC) and socio-economic factors (210). We will implement the HPV and cervical cancer components in this new version of MicroCOSM and investigate the impact of medical male circumcision on HPV transmission, as well as the impact of smoking and obesity interventions on cervical cancer incidence.

It is important to identify which sections of the population are being missed by current screening and treatment programmes. This thesis did not distinguish between women screened in the public and private healthcare sectors of South Africa, but it is likely that there are differences in screening coverage and access to treatment between these populations. Differences in levels of access to screening and treatment may also exist for women with different levels of educational attainment and women in urban versus rural locations (211). The extension of MicroCOSM to include socio-economic factors will allow us to model the impact of differences in access on cervical cancer incidence.

In this thesis, we modelled HPV types as if they are independent of each other. However, it is possible that HPV types not covered by the existing vaccines may take the place of the vaccine types by filling an ecological niche (212). There is increasing evidence of this theory as vaccinated cohorts age. In a

recent updated meta-analysis, Drolet *et al.* show that 5-8 years after vaccination, the prevalence of non-vaccine types increased by 12-17% in studies that enrolled women from the general population in pre- and post-vaccination periods (213). Since type replacement could potentially change HPV dynamics substantially, we will estimate the impact of these findings on our model results.

In future work, we will revisit the model's assumptions about viral latency and reactivation of latent infection to consider the possibility that reactivated infections may be less likely to progress to cancer. We may potentially update our assumption that there is no biological interaction between HIV and HPV transmission if new compelling evidence is generated.

7.4 Policy recommendations

Three main recommendations result from this work. Firstly, to have the greatest impact on cervical cancer in the next twenty years, we recommend that Pap smear screening should be replaced by HPV-DNA based screening. Since these tests can be processed by the GeneXpert machines that are available at nearly 200 decentralised clinics across South Africa (214), it is possible that a woman with HPV types 16 or 18 can be screened and treated at the same facility visit. This may require additional training of healthcare workers at primary care facilities to perform pre-cancer treatment procedures, but may significantly reduce loss to follow up. Due to the low specificity of the test to detect high-grade pre-cancer, referring all women testing positive for other high-risk HPV types to colposcopy clinics may overburden the system, and we recommend Pap smear as a triage method. The loss in sensitivity of this method may be offset by the low linkage to colposcopy clinics, which may become worse with an increased number of referrals. This recommendation is based on the strategy that will have the greatest impact on cervical cancer incidence, and future research will focus on identifying the strategy that is most cost-effective.

Secondly, if the current strategy of screening with Pap smear and referral to colposcopy is maintained, urgent increased access to colposcopy clinics and treatment, and decreased waiting times is required. This would require investment in colposcopy equipment and training of healthcare workers. Thirdly, we recommend better adherence to the existing policy of screening HIV-positive women every three years. We estimate that at current levels of screening and linkage to treatment, more than half of CC cases will be HIV-positive for the next 30 years. Although women who are HIV-positive but not yet accessing HIV care may be hard to reach, those who are on ART regularly interact with the health system and should be easy to reach for CC screening.

An important topic that this thesis reinforces is the utility of a unique patient identifier in a public health information system. The ability to link health data from all the levels of healthcare facilities, laboratories and pharmacies at the individual level not only has the potential to improve patient care,

but enables thorough monitoring and evaluation of prevention and treatment programmes that require substantial government investment. This thesis illustrated the benefits of the unique identifier to quantify the cervical cancer screening cascade in the Western Cape but benefits of such a system will extend to all healthcare. A health information system with a unique identifier may also eliminate the need for population-based cancer registries, by enabling identification of cancer cases from laboratory data (contained in the current pathology-based National Cancer Registry) as well as from hospital (diagnosis and procedure codes) and pharmacy data (chemotherapy drugs). The Western Cape's Provincial Health Data Centre has proven that the establishment of a health information system with a unique patient identifier is possible, and this thesis has illustrated the level of programme evaluation that the system enables. We therefore recommend that such a system be implemented at the national level.

7.5 Conclusion

This thesis presents an epidemiological model of HIV, HPV, and cervical cancer in South Africa. The research presented here advances our understanding of HIV and HPV transmission dynamics, as well as the natural history of HPV infection. The analysis of the Western Cape's public sector cervical cancer screening and treatment data shows low coverage of screening, low linkage to treatment and poor adherence to screening intervals, which highlights the need for unique patient identifiers in routine surveillance data to monitor and evaluate the success of the prevention programme. Our model allows for estimation of the impact of both HIV and cervical cancer prevention on cancer incidence, and provides the opportunity to identify the vaccination and screening strategies with the greatest public health significance.

References

1. World Health Organization. Draft: Global strategy towards the elimination of cervical cancer as a public health problem [Internet]. 2019. Available from: https://www.who.int/docs/default-source/documents/cervical-cancer-elimination-draft-strategy.pdf

2. 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. Int J Cancer [Internet]. 2019;144(8):1941–53. Available from: https://gco.iarc.fr/

3. Bosch FX, Lorincz A, Munoz N, Meijer CJLM, Shah K V. The causal relation between human papillomavirus and cervical cancer. J Clin Pathol. 2002 Apr 1;55(4):244–65.

4. Serrano B, Alemany L, Tous S, Bruni L, Clifford GM, Weiss T, et al. Potential impact of a nine-valent vaccine in human papillomavirus related cervical disease. Infect Agent Cancer. 2012;7(1):1–13.

5. WHO. Comprehensive Cervical Cancer Control. Geneva [Internet]. 2014;366–78. Available from: http://apps.who.int/iris/bitstream/10665/144785/1/9789241548953_eng.pdf?ua=1

6. 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. J Low Genit Tract Dis. 2001;5(1):60.

7. Vaccarella S, Lortet-Tieulent J, Plummer M, Franceschi S, Bray F. Worldwide trends in cervical cancer incidence: Impact of screening against changes in disease risk factors. Eur J Cancer. 2013 Oct;49(15):3262–73.

8. Arbyn M, Sasieni P, Meijer CJLM, Clavel C, Koliopoulos G, Dillner J. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. Vaccine. 2006 Aug 31;24 Suppl 3:S3/78-89.

9. Wild C, Weiderpass E, Stewart B, Editors. World Cancer Report: Cancer Research for Cancer Prevention [Internet]. 2020. Available from: http://publications.iarc.fr/586

10. Malagón T, Drolet M, Boily M-C, Franco EL, Jit M, Brisson J, et al. Cross-protective efficacy of two human papillomavirus vaccines: a systematic review and meta-analysis. Lancet Infect Dis. 2012 Oct;12(10):781–9.

11. 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. Lancet. 2020;395(10224):575–90.

12. Castro KG, Ward JW, Slutsker L, Buehler JW, Jaffe HW, Berkelman RL, et al. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. Morb Mortal Wkly Report, Natl Cent Infect Dis Div HIV/AIDS, Centers Dis Control Prev. 1992;41(RR-17):1–19.

13. De Vuyst H, Lillo F, Broutet N, Smith JS. HIV, human papillomavirus, and cervical neoplasia and cancer in the era of highly active antiretroviral therapy. Eur J Cancer Prev. 2008 Nov;17(6):545–54.

14. 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. Front Oncol. 2014;4(March):48.

15. Clifford GM, Gonçalves MAG, Franceschi S. Human papillomavirus types among women infected with HIV: a meta-analysis. AIDS. 2006 Nov 28;20(18):2337–44.

16. Wang C, Wright TC, Denny L, Kuhn L. Rapid rise in detection of human papillomavirus (HPV) infection soon after incident HIV infection among South African women. J Infect Dis. 2011 Feb 15;203(4):479–86.

17. Mbulawa ZZA, Marais DJ, Johnson LF, Coetzee D, Williamson A-L. Impact of human immunodeficiency virus on the natural history of human papillomavirus genital infection in South African men and women. J Infect Dis. 2012 Jul 1;206(1):15–27.

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

19. McKenzie ND, Kobetz EN, Hnatyszyn J, Twiggs LB, Lucci J a. Women with HIV are more commonly infected with non-16 and -18 high-risk HPV types. Gynecol Oncol. 2010 Mar;116(3):572–7.

20. Lissouba P, Van de Perre P, Auvert B. Association of genital human papillomavirus infection with HIV acquisition: a systematic review and meta-analysis. Sex Transm Infect. 2013 Aug;89(5):350–6.

21. Houlihan CF, Larke NL, Watson-Jones D, Smith-McCune KK, Shiboski S, Gravitt PE, et al. Human papillomavirus infection and increased risk of HIV acquisition. A systematic review and meta-analysis. AIDS. 2012 Nov 13;26(17):2211–22.

22. 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. Lancet HIV. 2018;5(1):e45–58.

23. 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. Int J Cancer. 2019 Jun 19;ijc.32260.

24. National Cancer Registry National Institute for Occupational Health National Health Laboratory Service. Cancer in South Africa 2016 Full Report [Internet]. 2020 [cited 2020 Jul 16]. Available from: https://www.nicd.ac.za/centres/national-cancer-registry/

25. Simbayi L, Zuma K, Zungu N, Moyo S, Marinda E, Jooste S, et al. South African national HIV prevalence, incidence, behaviour and communication survey, 2017 [Internet]. Nelson Mandela Foundation. 2017. Available from: http://www.hsrc.ac.za/uploads/pageContent/10779/SABSSM V.pdf

26. 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. Infect Dis Obstet Gynecol. 2014;2014.

27. Mbulawa ZZA, Coetzee D, Williamson A-L. Human papillomavirus prevalence in South African women and men according to age and human immunodeficiency virus status. BMC Infect Dis. 2015;15(1):1–11.

28. 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. Lancet Glob Heal. 2020;8(2):e296–304.

29. Moodley M, Mould S. Invasive cervical cancer and human immunodeficiency virus (HIV) infection in KwaZulu-Natal, South Africa. J Obstet Gynaecol (Lahore). 2005 Jan 2;25(7):706–10.

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

31. Van Bogaert L-JJ. Age at Diagnosis of Preinvasive and Invasive Cervical Neoplasia in South Africa: HIV-positive versus HIV-negative women. Int J Gynecol Cancer. 2011;21(2):363–6.

32. Dhokotera T, Bohlius J, Spoerri A, Egger M, Ncayiyana J, Olago V, et al. The burden of cancers associated with HIV in the South African public health sector, 2004-2014: A record linkage study. Infect Agent Cancer. 2019;14(1):1–12.

33. 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

34. South African Department of Health. Cervical Cancer Prevention and Control [Internet]. 2017 [cited 2020 Oct 8]. Available from: http://www.health.gov.za/index.php/2014-08-15-12-53-24?download=1393:cervical-cancer-policy-pdf

35. Day C, Gray A, Ndlovu N, Cois A. South African Health Review: Health and related indicators. 2019.

36. Jassat W. An evaluation of the cervical screening programme in Johannesburg metro district, Gauteng province. University of the Witwatersrand; 2010.

37. Blanckenberg N, Oettle C, Conradie H, Krige F. Impact of the introduction of a colposcopy service in a rural South African sub-district on uptake of colposcopy. S Afr J Obstet Gynaecol. 2013;19(3):81–5.

38. Knegt Y. Audit of cervical cancer screening and colposcopy attendance in rural South Africa. Afr J Reprod Health. 2014;18(4):70–8.

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

40. The National Cancer Registry. Cancer in South Africa 2001: Full report [Internet]. 2001. Available from: http://www.nicd.ac.za/centres/national-cancer-registry/

41. 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). Int J Cancer. 2018;143(9):2238–49.

42. Delany-Moretlwe S, Kelley KF, James S, Scorgie F, Subedar H, Dlamini NR, et al. Human Papillomavirus Vaccine Introduction in South Africa: Implementation Lessons From an Evaluation of the National School-Based Vaccination Campaign. Glob Heal Sci Pract. 2018;GGHSP-D-18-00090.

43. World Health Organization. Immunisation, Vaccines and Biologicals [Internet]. [cited 2020 Sep 15]. Available from: https://www.who.int/immunization/monitoring_surveillance/data/en/

44. 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.

45. Coppleson LW, Brown B. Observations on a model of the biology of carcinoma of the cervix: a poor fit between observation and theory. Am J Obstet Gynecol. 1975;122(1):127–36.

46. Prorok PC. Mathematical models and natural history in cervical cancer screening. In: IARC Sci Publ 76. 1986. p. 185–98.

47. Hall MT, Smith MA, Simms KT, Barnabas R V., Canfell K, Murray JM. The past, present and future impact of HIV prevention and control on HPV and cervical disease in Tanzania: A modelling study. PLoS One. 2020;15(5):1–23.

48. Vynnycky E, White R. An introduction to infectious disease modelling. Oxford University Press; 2010.

49. Bellan S, Borchering R, Bruce F, Dushoff J, Grebe E, Hargrove J, et al. International Clinics on Infectious Disease Dynamics and Data [Internet]. 2017 [cited 2020 Oct 7]. Available from: https://doi.org/10.6084/m9.figshare.c.3788224.v13

50. Creasman W. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. Int J Gynaecol Obstet. 2009 May;105(2):109.

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

52. 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. Br J Cancer. 2004;91(3):530–6.

53. Goldie SJ, Weinstein MC, Kuntz KM, Freedberg KA. The costs, clinical benefits, and cost-effectiveness of screening for cervical cancer in HIV-infected women. Ann Intern Med. 1999;130(2):97–107.

54. Jit M, Gay N, Soldan K, Choi YH, Edmunds WJ. Estimating Progression Rates for Human Papillomavirus Infection From Epidemiological Data. Med Decis Mak. 2010;30(84).

55. Bogaards JA, Xiridou M, Coupé VMH, Meijer CJLM, Wallinga J, Berkhof J. Model-based estimation of viral transmissibility and infection-induced resistance from the age-dependent prevalence of infection for 14 high-risk types of human papillomavirus. Am J Epidemiol. 2010;171(7):817–25.

56. Kim JJ, Kuntz KM, Stout NK, Mahmud S, Villa LL, Franco EL, et al. Multiparameter calibration of a natural history model of cervical cancer. Am J Epidemiol. 2007;166(2):137–50.

57. Campos NG, Burger EA, Sy S, Sharma M, Schiffman M, Rodriguez AC, et al. An updated natural history model of cervical cancer: derivation of model parameters. Am J Epidemiol. 2014 Sep 1;180(5):545–55.

58. Van de Velde N, Brisson M, Boily MC. Understanding differences in predictions of HPV vaccine effectiveness: A comparative model-based analysis. Vaccine. 2010;28(33):5473–84.

59. Beachler DC, Jenkins G, Safaeian M, Kreimer AR, Wentzensen N. Natural Acquired Immunity Against Subsequent Genital Human Papillomavirus Infection: A Systematic Review and Meta-analysis. J Infect Dis. 2016;213.

60. Franceschi S, Baussano I. Naturally acquired immunity against human papillomavirus (HPV): why it matters in the HPV vaccine era. J Infect Dis. 2014 Aug 15;210(4):507–9.

61. 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.

62. Brisson M, Bénard É, Drolet M, Bogaards JA, Baussano I, Vänskä S, et al. Population-level impact, herd immunity, and elimination after human papillomavirus vaccination: a systematic review and meta-analysis of predictions from transmission-dynamic models. Lancet Public Heal. 2016;1(1):e8–17.

63. Maglennon GA, McIntosh PB, Doorbar J. Immunosuppression Facilitates the Reactivation of Latent Papillomavirus Infections. J Virol. 2014;88(1):710–6.

64. 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. Obstet Gynecol. 2010 Jun;115(6):1150–8.

65. 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. J Natl Cancer Inst. 2005 Apr 20;97(8):577–86.

66. Kreimer AR, Struyf F, Del Rosario-Raymundo MR, Hildesheim A, Skinner SR, Wacholder S, et al. Efficacy of fewer than three doses of an HPV-16/18 AS04-adjuvanted vaccine: combined analysis of data from the Costa Rica Vaccine and PATRICIA trials. Lancet Oncol. 2015;16:775–86.

67. Ranjeva SL, Baskerville EB, Dukic V, Villa LL, Lazcano-ponce E, Giuliano AR, et al. Recurring infection with ecologically distinct HPV types can explain high prevalence and diversity. Proc Natl Acad Sci. 2017;

68. Brouwer AF, Meza R, Eisenberg MC. Integrating measures of viral prevalence and seroprevalence: A mechanistic modelling approach to explaining cohort patterns of human papillomavirus in women in the USA. Philos Trans R Soc B Biol Sci. 2019;374(1773).

69. Paavonen J, Naud P, Salmerón J, Wheeler CM, Chow SN, Apter D, et al. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. Lancet. 2009;374:301–14.

70. The FUTURE II Study Group. Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. N Engl J Med. 2007;356(19):1915–27.

71. Joura EA, Giuliano AR, Iversen OE, Bouchard C, Mao C, Mehlsen J, et al. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med. 2014;372(8):711–23.

72. Wheeler CM, Castellsagué X, Garland SM, Szarewski A, Paavonen J, Naud P, et al. Cross-protective efficacy of HPV-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by non-vaccine oncogenic HPV types: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. Lancet Oncol. 2012;13(1):100–10.

73. Castellsagué X, Muñoz N, Pitisuttithum P, Ferris D, Monsonego J, Ault K, et al. End-of-study safety, immunogenicity, and efficacy of quadrivalent HPV (types 6, 11, 16, 18) recombinant vaccine in adult women 24-45 years of age. Br J Cancer. 2011;105(May):28–37.

74. Skinner SR, Szarewski A, Romanowski B, Garland SM, Lazcano-Ponce E, Salmerón J, et al. Efficacy, safety, and immunogenicity of the human papillomavirus 16/18 AS04-adjuvanted vaccine in women older than 25 years: 4-year interim follow-up of the phase 3, double-blind, randomised controlled VIVIANE study. Lancet. 2014;384:2213–27.

75. Artemchuk H, Eriksson T, Poljak M, Surcel HM, Dillner J, Lehtinen M, et al. Long-term Antibody Response to Human Papillomavirus Vaccines: Up to 12 Years of Follow-up in the Finnish Maternity Cohort. J Infect Dis. 2019;219(4):582–9.

76. Denny L, Hendricks B, Gordon C, Thomas F, Hezareh M, Dobbelaere K, et al. Safety and immunogenicity of the HPV-16/18 AS04-adjuvanted vaccine in HIV-positive women in South Africa: A partially-blind randomised placebo-controlled study. Vaccine. 2013;31(48):5745–53.

77. Lacey CJ. HPV vaccination in HIV infection. Papillomavirus Res. 2019;8(June):100174.

78. D'Addario M, Redmond S, Scott P, Egli-Gany D, Riveros-Balta AX, Henao Restrepo AM, et al. Two-dose schedules for human papillomavirus vaccine: Systematic review and meta-analysis. Vaccine. 2017 May;35(22):2892–901.

79. World Health Organization. WHO guidance note: Comprehensive cervical cancer prevention and control: A healthier future for girls and women [Internet]. 2013 [cited 2014 Nov 12]. Available from: http://apps.who.int/iris/handle/10665/78128

80. Pinkerton SD, Abramson PR. Effectiveness of condoms in preventing HIV transmission. Soc Sci Med. 1997;44:1303–12.

81. Manhart LE, Koutsky LA. Do condoms prevent genital HPV infection, external genital warts, or cervical neoplasia? A meta-analysis. Sex Transm Dis. 2002;29:725–35.

82. Winer RL, Hughes JP, Feng Q, O'Reilly S, Kiviat NB, Holmes KK, et al. Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med. 2006;354:2645–54.

83. Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Puren A. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 trial. PLoS Med. 2005;2(11):1112–22.

84. Gray R, Kigozi G, Serwadda D, Makumbi F, Watya S, Nalugoda F, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet. 2007;369:657–66.

85. Auvert B, Sobngwi-Tambekou J, Cutler E, Nieuwoudt M, Lissouba P, Puren A, et al. Effect of male circumcision on the prevalence of high-risk human papillomavirus in young men: results of a randomized controlled trial conducted in Orange Farm, South Africa. J Infect Dis. 2009 Jan 1;199(1):14–9.

86. Tobian AR, Serwadda D, Quinn TC, Kigozi G, Gravitt PE, Laeyendecker O, et al. Male circumcision for the prevention of HSV-2 and HPV infections and syphilis. N Engl J Med. 2009;360:1298–309.

87. Albero G, Castellsagué X, Giuliano AR, Bosch FX. Male circumcision and genital human papillomavirus: A systematic review and meta-analysis. Sex Transm Dis. 2012;39(2):104–13.

88. Organisation for Economic Co-operation and Development. Health Statistics 2019: Definitions, Sources and Methods [Internet]. 2019 [cited 2019 Oct 4]. Available from: https://stats.oecd.org/

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

90. 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. Obstet Gynecol. 2008 Jan;111(1):167–77.

91. 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.

92. 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 Treat Rev. 2004;30:205–11.

93. 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 Epidemiol Biomarkers Prev. 2007;16(January):11–6.

94. 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 Oncol. 2011;12:441–50.

95. Goldie SJ, Kim JJ, Myers E. Chapter 19: Cost-effectiveness of cervical cancer screening. Vaccine. 2006;24(SUPPL. 3):164–70.

96. Hughes JP, Garnett GP, Koutsky LA. The theoretical population-level impact of a prophylactic human papilloma virus vaccine. Epidemiology. 2002 Nov;13(6):631–9.

97. Kim JJ, Brisson M, Edmunds WJ, Goldie SJ. Modeling Cervical Cancer Prevention in Developed Countries. Vaccine. 2008;26(Suppl 10).

98. Fesenfeld M, Hutubessy R, Jit M. Cost-effectiveness of human papillomavirus vaccination in low and middle income countries: A systematic review. Vaccine. 2013;31(37):3786–804.

99. Canfell K, Kim JJ, Brisson M, Keane A, Simms KT, Caruana M, et al. Mortality impact of achieving WHO cervical cancer elimination targets: a comparative modelling analysis in 78 low-income and lower-middle-income countries. Lancet. 2020;395(10224):591–603.

100. Bruni L, Diaz M, Castellsagué X, Ferrer E, Bosch FX, de Sanjosé S. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J Infect Dis. 2010 Dec 15;202(12):1789–99.

101. Looker KJ, Rönn MM, Brock PM, Brisson M, Drolet M, Mayaud P, et al. Evidence of synergistic relationships between HIV and Human Papillomavirus (HPV): systematic reviews and meta-analyses of longitudinal studies of HPV acquisition and clearance by HIV status, and of HIV acquisition by HPV status. J Int AIDS Soc. 2018;21(6).

102. Schuman P, Ohmit SE, Klein RS, Duerr A, Cu-Uvin S, Jamieson DJ, et al. Longitudinal study of cervical squamous intraepithelial lesions in human immunodeficiency virus (HIV)-seropositive and at-risk HIV-seronegative women. J Infect Dis. 2003;188(1):128–36.

103. Massad LS, Ahdieh L, Benning L, Minkoff H, Greenblatt RM, Watts H, et al. Evolution of cervical abnormalities among women with HIV-1: evidence from surveillance cytology in the women's interagency HIV study. J Acquir Immune Defic Syndr. 2001;27(5):432–42.

104. Goldie SJ, Freedberg KA, Weinstein MC, Wright TC, Kuntz KM. Cost effectiveness of human papillomavirus testing to augment cervical cancer screening in women infected with the human immunodeficiency virus. Am J Med. 2001;111(2):140–9.

105. Atashili J, Smith JS, Adimora A a, Eron J, Miller WC, Myers E. Potential impact of antiretroviral therapy and screening on cervical cancer mortality in HIV-positive women in sub-Saharan Africa: a simulation. PLoS One. 2011 Jan;6(4):e18527.

106. Vanni T, Luz PM, Grinsztejn B, Veloso VG, Foss A, Mesa-Frias M, et al. Cervical cancer screening among HIV-infected women: An economic evaluation in a middle-income country. Int J Cancer. 2012;131(2):96–104.

107. Zimmermann MR, Vodicka E, Babigumira JB, Okech T, Mugo N, Sakr S, et al. Cost-effectiveness of cervical cancer screening and preventative cryotherapy at an HIV treatment clinic in Kenya. Cost Eff Resour Alloc. 2017;15(1):1–10.

108. Smit M, Perez-Guzman PN, Mutai KK, Cassidy R, Kibachio J, Kilonzo N, et al. Mapping the Current and Future Noncommunicable Disease Burden in Kenya by Human Immunodeficiency Virus Status: A Modeling Study. Clin Infect Dis. 2019;(Xx Xxxx):1–10.

109. Perez-Guzman PN, Chung MH, De Vuyst H, Dalal S, Mutai KK, Muthoni K, et al. The impact of scaling up cervical cancer screening and treatment services among women living with HIV in Kenya: a modelling study. BMJ Glob Heal. 2020;5(3):e001886.

110. Lince-Deroche N, Phiri J, Michelow P, Smith JS, Firnhaber C. Costs and cost effectiveness of three approaches for cervical cancer screening among HIV-positive women in Johannesburg, South Africa. PLoS One. 2015;10(11):1–16.

111. Lince-Deroche N, van Rensburg C, Roseleur J, Sanusi B, Phiri J, Michelow P, et al. Costs and cost-effectiveness of LEEP versus cryotherapy for treating cervical dysplasia among HIV-positive women in Johannesburg, South Africa. Andrei G, editor. PLoS One. 2018 Oct 11;13(10):e0203921.

112. Moodley N, Gray G, Bertram M. The Price of Prevention: Cost Effectiveness of Biomedical HIV Prevention Strategies in South Africa. Clin Res HIV/AIDS. 2016;3(1):1–31.

113. Goldie SJ, Kuhn L, Denny L, Pollack A, Wright TC. Policy analysis of cervical cancer screening strategies in low-resource settings: clinical benefits and cost-effectiveness. JAMA. 2001;285(24):3107–

15

114. Goldie SJ, Gaffikin L, Goldhaber-Fiebert JD, Gordillo-Tobar A, Levin C, Mahé C, et al. Cost-effectiveness of cervical-cancer screening in five developing countries. N Engl J Med. 2005 Nov 17;353(20):2158–68.

115. 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.

116. Sitas F. Histologically diagnosed cancers in South Africa, 1988. South African Med J. 1994;84:344–8.

117. Vijayaraghavan A, Efrusy M, Lindeque G, Dreyer G, Santas C. Cost effectiveness of high-risk HPV DNA testing for cervical cancer screening in South Africa. Gynecol Oncol. 2009 Feb;112(2):377–83.

118. Campos NG, Lince-Deroche N, Chibwesha CJ, Firnhaber C, Smith JS, Michelow P, et al. Cost-effectiveness of cervical cancer screening in women living with HIV in South Africa: A mathematical modeling study. J Acquir Immune Defic Syndr. 2018;79(2):195–205.

119. Sinanovic E, Moodley J, Barone M a, Mall S, Cleary S, Harries J. The potential cost-effectiveness of adding a human papillomavirus vaccine to the cervical cancer screening programme in South Africa. Vaccine. 2009 Oct 19;27(44):6196–202.

120. Kim JJ, Campos NG, O'Shea M, Diaz M, Mutyaba I. Model-based impact and cost-effectiveness of cervical cancer prevention in sub-Saharan Africa. Vaccine. 2013 Dec 29;31 Suppl 5:F60-72.

121. Li X, Stander MP, Van Kriekinge G, Demarteau N. Cost-effectiveness analysis of human papillomavirus vaccination in South Africa accounting for human immunodeficiency virus prevalence. BMC Infect Dis. 2015;15(1):1–18.

122. Simms KT, Steinberg J, Caruana M, Smith MA, Lew J Bin, Soerjomataram I, et al. Impact of scaled up human papillomavirus vaccination and cervical screening and the potential for global elimination of cervical cancer in 181 countries, 2020–99: a modelling study. Lancet Oncol. 2019;20(3):394–407.

123. Campos NG, Kim JJ, Castle PE, Ortendahl JD, O'Shea M, Diaz M, et al. Health and economic impact of HPV 16/18 vaccination and cervical cancer screening in Eastern Africa. Int J Cancer. 2012;130(11):2672–84.

124. Vaccarella S, Franceschi S, Engholm G, Lönnberg S, Khan S, Bray F. 50 years of screening in the Nordic countries: Quantifying the effects on cervical cancer incidence. Br J Cancer. 2014;111(5):965–9.

125. Pesola F, Sasieni P. Impact of screening on cervical cancer incidence in England: A time trend analysis. BMJ Open. 2019;9(1):1–7.

126. Smith TR, Wakefield J. A review and comparison of age-period-cohort models for cancer incidence. Stat Sci. 2016;31(4):591–610.

127. Bellan SE, Pulliam JRC, Pearson CAB, Champredon D, Fox SJ, Skrip L, et al. Statistical power and validity of Ebola vaccine trials in Sierra Leone: A simulation study of trial design and analysis. Lancet Infect Dis. 2015;15(6):703–10.

128. Cori A, Ayles H, Beyers N, Schaap A, Floyd S, Sabapathy K, et al. HPTN 071 (PopART): A cluster-randomized trial of the population impact of an HIV combination prevention intervention including universal testing and treatment: Mathematical model. PLoS One. 2014;9(1).

129. Silkey M, Homan T, Maire N, Hiscox A, Mukabana R, Takken W, et al. Design of trials for interrupting the transmission of endemic pathogens. Trials. 2016;17(1):1–16.

130. Omori R, Nagelkerke N, Abu-Raddad LJ. HIV and herpes simplex virus type 2 epidemiological synergy: misguided observational evidence? A modelling study. Sex Transm Infect. 2017;sextrans-2017-053336.

131. UNAIDS. Global AIDS update [Internet]. 2016. Available from: http://www.who.int/hiv/pub/arv/global-AIDS-update-2016_en.pdf?ua=1

132. Gallagher KE, Baisley K, Grosskurth H, Vallely A, Kapiga S, Vandepitte J, et al. The Association between Cervical Human Papillomavirus Infection and Subsequent HIV Acquisition in Tanzanian and Ugandan Women: A Nested Case-Control Study. J Infect Dis. 2016;214(1):87–95.

133. Veldhuijzen NJ, Vyankandondera J, van de Wijgert JH. HIV acquisition is associated with prior high-

risk human papillomavirus infection among high-risk women in Rwanda. AIDS. 2010 Sep 10;24(14):2289–92.

134. Stanley M. Immunobiology of HPV and HPV vaccines. Gynecol Oncol. 2008 May;109(2 Suppl):S15-21.

135. 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. J Infect Dis. 2001 Sep 15;184(6):682–90.

136. Blitz S, Baxter J, Raboud J, Walmsley S, Rachlis A, Smaill F, et al. Evaluation of HIV and highly active antiretroviral therapy on the natural history of human papillomavirus infection and cervical cytopathologic findings in HIV-positive and high-risk HIV-negative women. J Infect Dis. 2013 Aug 1;208(3):454–62.

137. Minkoff H, Feldman J, DeHovitz J, Landesman S, Burk R. A longitudinal study of human papillomavirus carriage in human immunodeficiency virus-infected and human immunodeficiency virus-uninfected women. Am J Obstet Gynecol. 1998 May;178(5):982–6.

138. Nowak RG, Gravitt PE, Morrison CS, Gange SJ, Kwok C, Oliver AE, et al. Increases in human papillomavirus detection during early HIV infection among women in Zimbabwe. J Infect Dis. 2011 Apr 15;203(8):1182–91.

139. 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. 2018 Mar 27;32(6):795–808.

140. Wasserheit JN. Epidemiological synergy. Interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. Sex Transm Dis. 19(2):61–77.

141. Gray RH, Wawer MJ. Reassessing the hypothesis on STI control for HIV prevention. Lancet (London, England). 2008 Jun 21;371(9630):2064–5.

142. Korenromp EL, de Vlas SJ, Nagelkerke NJ, Habbema JD. Estimating the magnitude of STD cofactor effects on HIV transmission: how well can it be done? Sex Transm Dis. 2001 Nov;28(11):613–21.

143. Boily MC, Anderson RM. Human immunodeficiency virus transmission and the role of other sexually transmitted diseases. Measures of association and study design. Sex Transm Dis. 1996;23(4):312–32.

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

145. 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. Demogr Res. 2009;21:289–340.

146. Smith AFM, Gelfand AE. Bayesian statistics without tears: a sampling–resampling perspective. Am Stat. 1992;46(2):84–8.

147. Auvert B, Marais D, Lissouba P, Zarca K, Ramjee G, Williamson A-L. High-risk human papillomavirus is associated with HIV acquisition among South African female sex workers. Infect Dis Obstet Gynecol. 2011;2011.

148. Whitham HK, Hawes SE, Chu H, Oakes JM, Lifson AR, Kiviat NB, et al. A Comparison of the Natural History of HPV Infection and Cervical Abnormalities among HIV-Positive and HIV-Negative Women in Senegal, Africa. Cancer Epidemiol biomarkers Prev. 2017 Jun;26(6):886–94.

149. Kretzschmar M, Morris M. Measures of concurrency in networks and the spread of infectious disease. Math Biosci. 1996 Apr 15;133(2):165–95.

150. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [Internet]. Available from: https://www.r-project.org/

151. Smith-McCune KK, Shiboski S, Chirenje MZ, Magure T, Tuveson J, Ma Y, et al. Type-specific cervico-vaginal human papillomavirus infection increases risk of HIV acquisition independent of other sexually transmitted infections. PLoS One. 2010 Apr 8;5(4):e10094.

152. Tobian AAR, Grabowski MK, Kigozi G, Redd AD, Eaton KP, Serwadda D, et al. Human papillomavirus clearance among males is associated with HIV acquisition and increased dendritic cell density in the

foreskin. J Infect Dis. 2013 Jun 1;207(11):1713–22.

153. Rositch AF, Mao L, Hudgens MG, Moses S, Agot K, Backes DM, et al. Risk of HIV acquisition among circumcised and uncircumcised young men with penile human papillomavirus infection. AIDS. 2014 Mar 13;28(5):745–52.

154. Rositch AF, Gravitt PE, Smith JS. Growing evidence that HPV infection is associated with an increase in HIV acquisition: Exploring the issue of HPV vaccination. Sex Transm Infect. 2013;89(5):357.

155. Gravitt PE. The known unknowns of HPV natural history. J Clin Invest. 2011 Dec;121(12):4593–9.

156. Mollers M, Vossen JM, Scherpenisse M, van der Klis FRM, Meijer CJLM, de Melker HE. Review: current knowledge on the role of HPV antibodies after natural infection and vaccination: implications for monitoring an HPV vaccination programme. J Med Virol. 2013 Aug;85(8):1379–85.

157. Giuliano AR, Viscidi R, Torres BN, Ingles DJ, Sudenga SL, Villa LL, et al. Seroconversion following anal and genital HPV infection in men: The HIM study. Papillomavirus Res. 2015;1:109–15.

158. Kelly H, Faust H, Chikandiwa A, Ngou J, Weiss HA, Segondy M, et al. Human Papillomavirus Serology Among Women Living With HIV: Type-Specific Seroprevalence , Seroconversion , and Risk of Cervical Reinfection. J Infect Dis. 2018;218(6):927–36.

159. Gravitt PE. Evidence and Impact of Human Papillomavirus Latency. Open Virol J. 2012;6.

160. Winer RL, Hughes JP, Feng Q, Stern JE, Xi LF, Koutsky LA. Incident Detection of High-Risk Human Papillomavirus Infections in a Cohort of High-Risk Women Aged 25-65 Years. J Infect Dis. 2016 Sep 1;214(5):665–75.

161. Shew ML, Ermel AC, Tong Y, Tu W, Qadadri B, Brown DR. Episodic detection of human papillomavirus within a longitudinal cohort of young women. J Med Virol. 2015 Dec;87(12):2122–9.

162. Rositch AF, Burke AE, Viscidi RP, Silver MI, Chang K, Gravitt PE. Contributions of Recent and Past Sexual Partnerships on Incident Human Papillomavirus Detection: Acquisition and Reactivation in Older Women. Cancer Res. 2012;1–9.

163. Fu T-CJ, Carter JJ, Hughes JP, Feng Q, Hawes SE, Schwartz SM, et al. Re-detection vs. new acquisition of high-risk human papillomavirus in mid-adult women. Int J cancer. 2016 Nov 15;139(10):2201–12.

164. Olsen J, Jepsen MR. Human papillomavirus transmission and cost-effectiveness of introducing quadrivalent HPV vaccination in Denmark. Int J Technol Assess Health Care. 2010;26(02):183–91.

165. Vanni T, Mendes P, Foss A, Mesa-frias M, Legood R. Economic modelling assessment of the HPV quadrivalent vaccine in Brazil: A dynamic individual-based approach. Vaccine. 2012;30(32):4866–71.

166. Smith MA, Lew J, Walker RJ, Brotherton JML, Nickson C, Canfell K. The predicted impact of HPV vaccination on male infections and male HPV-related cancers in Australia. Vaccine. 2011;29:9112–22.

167. Choi YH, Jit M, Gay N, Cox A, Garnett GP, Edmunds WJ. Transmission dynamic modelling of the impact of human papillomavirus vaccination in the United Kingdom. Vaccine. 2010;28(24):4091–102.

168. De Blasio BF, Neilson AR, Klemp M, Skjeldestad FE. Modeling the impact of screening policy and screening compliance on incidence and mortality of cervical cancer in the post-HPV vaccination era. J Public Health (Bangkok). 2012;34(4):539–47.

169. Van de Velde N, Boily M-C, Drolet M, Franco EL, Mayrand M-H, Kliewer E V., et al. Population-level impact of the bivalent, quadrivalent, and nonavalent human papillomavirus vaccines: a model-based analysis. J Natl Cancer Inst. 2012 Nov 21;104(22):1712–23.

170. Baussano I, Lazzarato F, Ronco G, Lehtinen M, Dillner J, Franceschi S. Different Challenges in Eliminating HPV16 Compared to Other Types: A Modeling Study. J Infect Dis. 2017;216.

171. Korostil IA, Regan DG. The potential impact of HPV-16 reactivation on prevalence in older Australians. BMC Infect Dis. 2014;14(312).

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

173. Harper DM, Franco EL, Wheeler CM, Moscicki A-B, Romanowski B, Roteli-Martins CM, et al. Sustained efficacy up to 4.5 years of a bivalent L1 virus-like particle vaccine against human papillomavirus types 16 and 18: follow-up from a randomised control trial. Lancet. 2006 Apr 15;367(9518):1247–55.

174. Wellings K, Collumbien M, Slaymaker E, Singh S, Hodges Z, Patel D, et al. Sexual behaviour in context: a global perspective. Lancet. 2006 Nov 11;368(9548):1706–28.

175. Jordaan S, Michelow P, Richter K, Simoens C, Bogers J. A Review of Cervical Cancer in South Africa: Previous, Current and Future. Heal Care Curr Rev. 2016;04(04).

176. Bosch FX, Robles C, Díaz M, Arbyn M, Baussano I, Clavel C, et al. HPV-FASTER: broadening the scope for prevention of HPV-related cancer. Nat Rev Clin Oncol. 2016;13(2):119–32.

177. Winer RL, Xi LF, Shen Z, Stern JE, Newman L, Feng Q, et al. Viral load and short-term natural history of type-specific oncogenic human papillomavirus infections in a high-risk cohort of midadult women. Int J cancer. 2014 Apr 15;134(8):1889–98.

178. World Health Organization. Cervical Cancer [Internet]. 2018. Available from: https://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/

179. Beck EJ, Shields JM, Tanna G, Henning G, de Vega I, Andrews G, et al. Developing and implementing national health identifiers in resource limited countries: why, what, who, when and how? Glob Health Action. 2018;11(1).

180. Botha H, Cooreman B, Dreyer G, Lindeque G, Mouton A, Guidozzi F, et al. Cervical cancer and human papillomavirus: South African guidelines for screening and testing. South African J Gynaecol Oncol. 2010;2(1):23–6.

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

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

183. Jordaan S, Michelow P, Simoens C, Bogers J. Challenges and Progress of Policies on Cervical Cancer in South Africa. Heal Care Curr Rev. 2017;05(01):1–5.

184. Johnson LF, Rehle TM, Jooste S, Bekker L-G. Rates of HIV testing and diagnosis in South Africa: successes and challenges. AIDS. 2015 Jul;29(11):1401–9.

185. Phaswana-Mafuya N, Peltzer K. Breast and Cervical Cancer Screening Prevalence and Associated Factors among Women in the South African General Population. Asian Pac J Cancer Prev. 2018 Jun 25;19(6):1465–70.

186. National Department of Health, Statistics South Africa, South African Medical Research Council, ICF. South Africa Demographic and Health Survey 2016. Pretoria, South Africa and Rockville, Maryland, USA; 2019.

187. Howard M, Agarwal G, Lytwyn A. Accuracy of self-reports of Pap and mammography screening compared to medical record: A meta-analysis. Cancer Causes Control. 2009;20(1):1–13.

188. 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. Int J Popul Data Sci. 2019 Nov 20;4(2).

189. Johnson LF, Dorrington RE. Modelling the impact of HIV in South Africa's provinces: 2019 update [Internet]. 2019 [cited 2020 Jul 2]. Available from: www.thembisa.org

190. Statistics South Africa. General Household Survey 2017 [Internet]. 2019 [cited 2019 Oct 7]. Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/723/get_microdata

191. Johnson LF, Mutemaringa T, Heekes A, Boulle A. Effect of HIV Infection and Antiretroviral Treatment on Pregnancy Rates in the Western Cape Province of South Africa. J Infect Dis [Internet]. 2020 Jun 11;221(12):1953–62. Available from: https://doi.org/10.1093/infdis/jiz362

192. World Health Organization. Immunization, Vaccines and Biologicals [Internet]. 2020 [cited 2020 Oct 24]. Available from: https://www.who.int/immunization/monitoring_surveillance/data/en/

193. van Schalkwyk C, Moodley J, Welte A, Johnson LF. Estimated impact of human papillomavirus vaccines on infection burden: The effect of structural assumptions. Vaccine [Internet]. 2019;37(36):5460–5. Available from: https://doi.org/10.1016/j.vaccine.2019.06.013

194. The National Cancer Registry. NCR Methodology [Internet]. Available from: https://cansa.org.za/files/2012/05/NCRMethods-2000-2001.pdf

195. 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.

196. 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. Am J Obstet Gynecol. 2003;188(2):395–400.

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

198. Snyman L, Herbst U. Reasons why unscreened patients with cervical cancer present with advanced stage disease. South African J Gynaecol Oncol. 2013;5(1):16–20.

199. 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.

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

201. National Cancer Registry South Africa. Ekurhuleni population-based cancer registry Annual Report 2018 [Internet]. 2020. Available from: www.ncr.ac.za

202. 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 Glob Oncol. 2020;1–8.

203. Brotherton JM, Budd A, Rompotis C, Bartlett N, Malloy MJ, Andersen RL, et al. Is one dose of human papillomavirus vaccine as effective as three?: A national cohort analysis. Papillomavirus Res. 2019;8(July):100177.

204. Kreimer AR, Sampson JN, Porras C, Schiller JT, Kemp T, Herrero R, et al. Evaluation of Durability of a Single Dose of the Bivalent HPV Vaccine: The CVT Trial. J Natl Cancer Inst. 2020;112:1–9.

205. 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

206. 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). Int J Cancer. 2018 Nov 1;143(9):2238–49.

207. UN Department of Economic and Social Affairs. World population prospects: 2017 revision [Internet]. 2017 [cited 2020 Sep 14]. Available from: https://population.un.org/wpp/Download/Standard/Population/

208. 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. Nat Commun. 2019;10(1):1–12.

209. Johnson L, Dorrington R. Thembisa 4.3: A model for evaluating the impact of HIV / AIDS in South Africa. 2020; Available from: www.thembisa.org

210. Johnson LF, van Rensburg C, Govathson C, Meyer-Rath G. Optimal HIV testing strategies for South Africa: a model-based evaluation of population-level impact and cost-effectiveness. Sci Rep. 2019;9(1):1–12.

211. Moodley J, Constant D, Mwaka AD, Scott SE, Walter FM. Mapping awareness of breast and cervical cancer risk factors, symptoms and lay beliefs in Uganda and South Africa. PLoS One. 2020;15(10 October):1–17.

212. Tota JE, Ramanakumar A V., Jiang M, Dillner J, Walter SD, Kaufman JS, et al. Epidemiologic approaches to evaluating the potential for human papillomavirus type replacement postvaccination. Am J Epidemiol. 2013;178(4):625–34.

213. Drolet M, Bénard É, Pérez N, Brisson M, Ali H, Boily MC, et al. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. Lancet. 2019;394(10197):497–509.

214. World Health Organization. Global TB report - laboratory diagnostic services [Internet]. 2019 [cited 2020 Nov 12]. Available from: https://www.who.int/tb/country/data/download/en/

215. Johnson LF, Dorrington RE, Moolla H. Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa. South Afr J HIV Med. 2017 Jul 27;18(1).

216. Mamahlodi M, Kuonza L, Candy S. Cervical cancer screening programme in Limpopo province : January 2007 to December 2010. South African J Gynaecol Oncol. 2013;5(1):4–10.

217. 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 Heal. 2016;3(4):849–62.

218. 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 Serv Res. 2015 Jun 6;15(1):499.

219. 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 Med J. 2002;92(2).

220. Wright TC, Denny L, Kuhn L. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. J Am Med Assoc. 2000;283(1):81–6.

221. 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. Int J Cancer. 2006;962(118):957–62.

222. 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 Syst Rev. 2017 Aug 10;(8).

223. 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. Int J Cancer. 2007;121(11):2484–91.

224. 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. Br J Cancer. 2016;115(4):425–30.

225. 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.

226. 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.

227. 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.

228. 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. Obstet Gynecol. 2008;111(1):7–14.

229. 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. Int J STD AIDS. 2012;23(12):890–6.

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

231. 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.

232. 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. Am J Obstet Gynecol. 2017;217(2):183.e1-183.e11.

233. 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 Med J. 2012;102(9):757–60.

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

235. 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. Syst Rev. 2018;7(1):198.

236. 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. J Acquir Immune Defic Syndr. 2015;68(2):227–35.

237. 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 Med J. 2015 Jan 6;105(2):115–20.

238. 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 Med J. 2015;105(3):191.

239. 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. Infect Dis Obstet Gynecol. 2014;2014:498048.

240. Mbulawa ZZA, van Schalkwyk C, Hu N-C, 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.

241. Mbulawa ZZA, Marais DJ, Johnson LF, Boulle A, Coetzee D, Williamson A-L. Influence of human immunodeficiency virus and CD4 count on the prevalence of human papillomavirus in heterosexual couples. J Gen Virol. 2010 Dec;91(Pt 12):3023–31.

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

243. 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. J Infect Dis. 2011 Jan 1;203(1):58–65.

244. 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.

245. 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.

246. 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 Cytol. 2009;53(1):10–7.

247. Mbulawa ZZA, Hu NC, Kufa-Chakezha T, Kularatne R, Williamson A-L. Sentinel surveillance of human papillomavirus genotypes among patients attending public healthcare facilities in South Africa,

2014-2016 133 [Internet]. Vol. 14, Communicable Diseases Surveillance Bulletin. 2016. 133–136 p. Available from: http://nicd.ac.za/assets/files/Sentinel surveillance of HPV.pdf

248. Johnson LF. Access to antiretroviral treatment in South Africa, 2004 - 2011. South Afr J HIV Med. 2012 Mar 13;13(1):22–7.

249. 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. J Infect Dis. 2008 May 15;197(10):1436–47.

250. 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 Epidemiol biomarkers Prev. 2003 Jun;12(6):485–90.

251. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. N Engl J Med. 1998 Feb 12;338(7):423–8.

252. 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.

253. 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 Res. 2008 Nov 1;68(21):8813–24.

254. 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. J Infect Dis. 2004 Dec 15;190(12):2077–87.

255. Insinga RP, Dasbach EJ, Elbasha EH, Liaw K-L, 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 Epidemiol biomarkers Prev. 2007 Apr;16(4):709–15.

256. 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 Epidemiol biomarkers Prev. 2010 Jun;19(6):1585–94.

257. 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.

258. Giuliano AR, Lee J-H, 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.

259. 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. Br J Cancer. 2009 Apr 7;100(7):1184–90.

260. 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 Epidemiol biomarkers Prev. 2011;(11):287–97.

261. 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. Int J Cancer. 2016;2438:2428–38.

262. Moscicki A, 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. J Am Med Assoc. 2001;285(23):2995–3002.

263. 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.

264. 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 Res Int. 2017;2017:1–

9.

265. 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.

266. 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. Lancet Oncol. 2008 May;9(5):425–34.

267. 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.

268. Schiffman M, Wheeler CM, Castle PE. Human Papillomavirus DNA Remains Detectable Longer than Related Cervical Cytologic Abnormalities. J Infect Dis. 2002;186(8):1169–72.

269. 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. J Pathol. 2001;195(3):300–6.

270. 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). Eur J Obstet Gynecol Reprod Biol. 2005;119(2):219–27.

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

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

273. McDonald AC, Denny L, Wang C, Tsai W-Y, 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.

274. 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. Int J cancer. 2014;134(6):1389–98.

275. 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. Int J Gynecol cancer. 2015;25(5):919–25.

276. 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 Glob Heal. 2020;(In press).

277. 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.

278. 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

279. 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.

Appendix A – Technical Appendix

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

A.1 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 (144). 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)

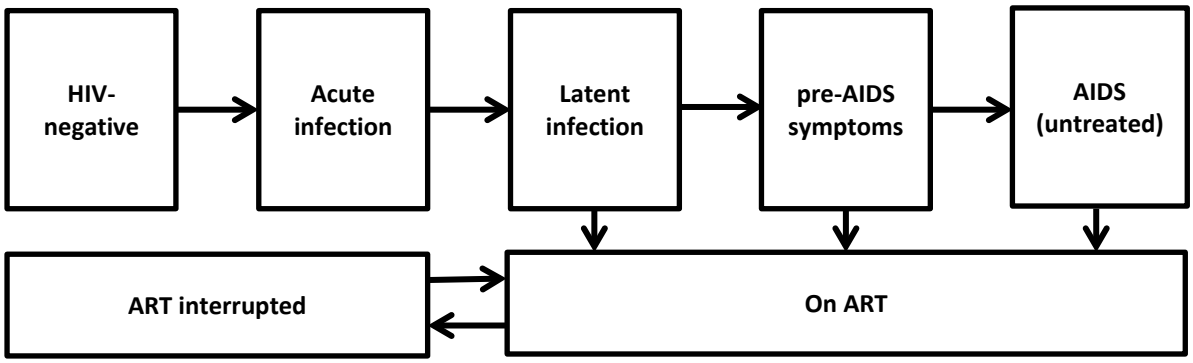


Figure A 1- Model of the natural history of HIV

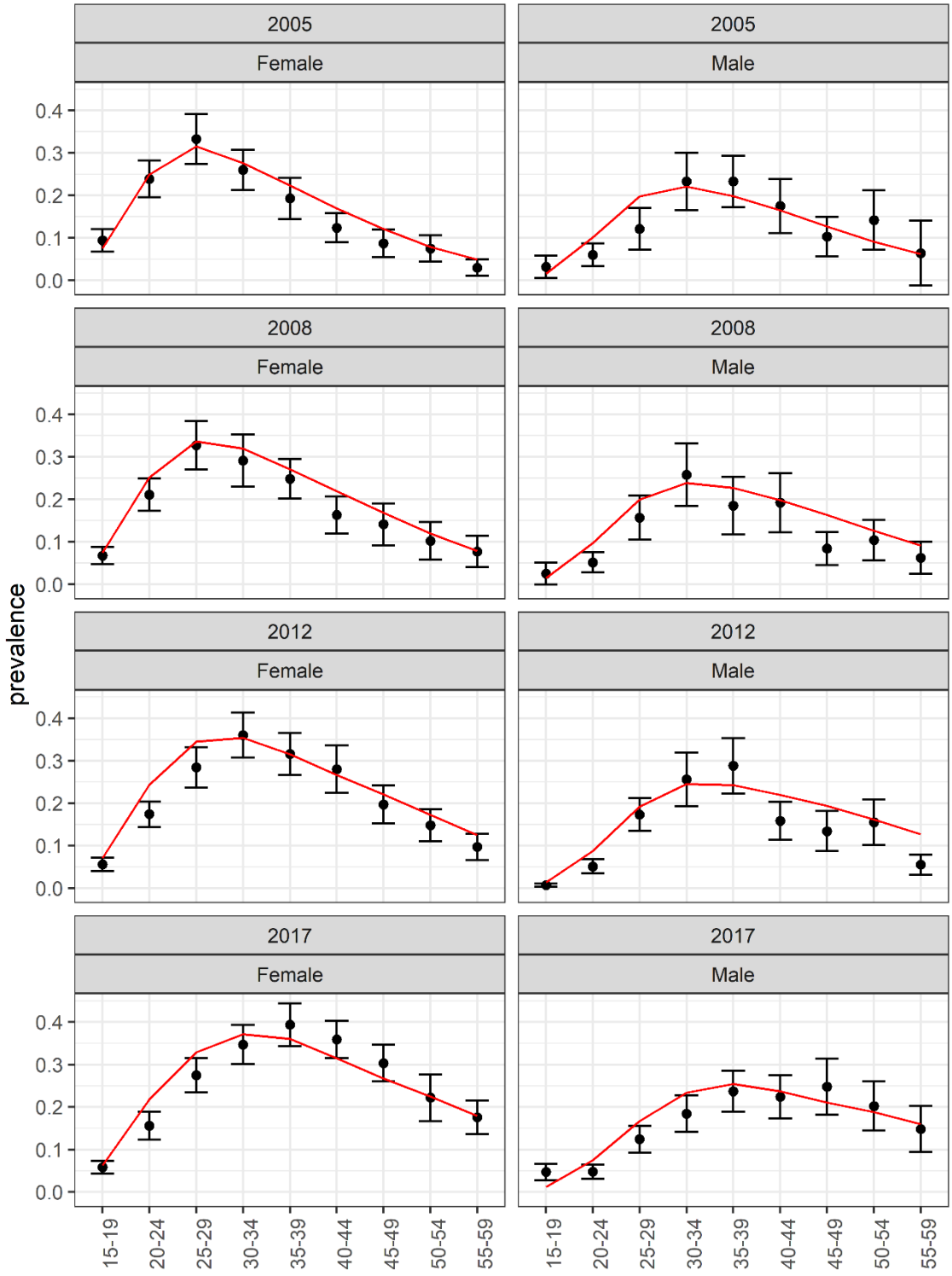


Figure A 2 – HIV prevalence by age, sex and time. Black dots and error bars are data from the HIV prevalence surveys performed by the Human Sciences Research Council (HSRC), and the red lines are estimates from the model.

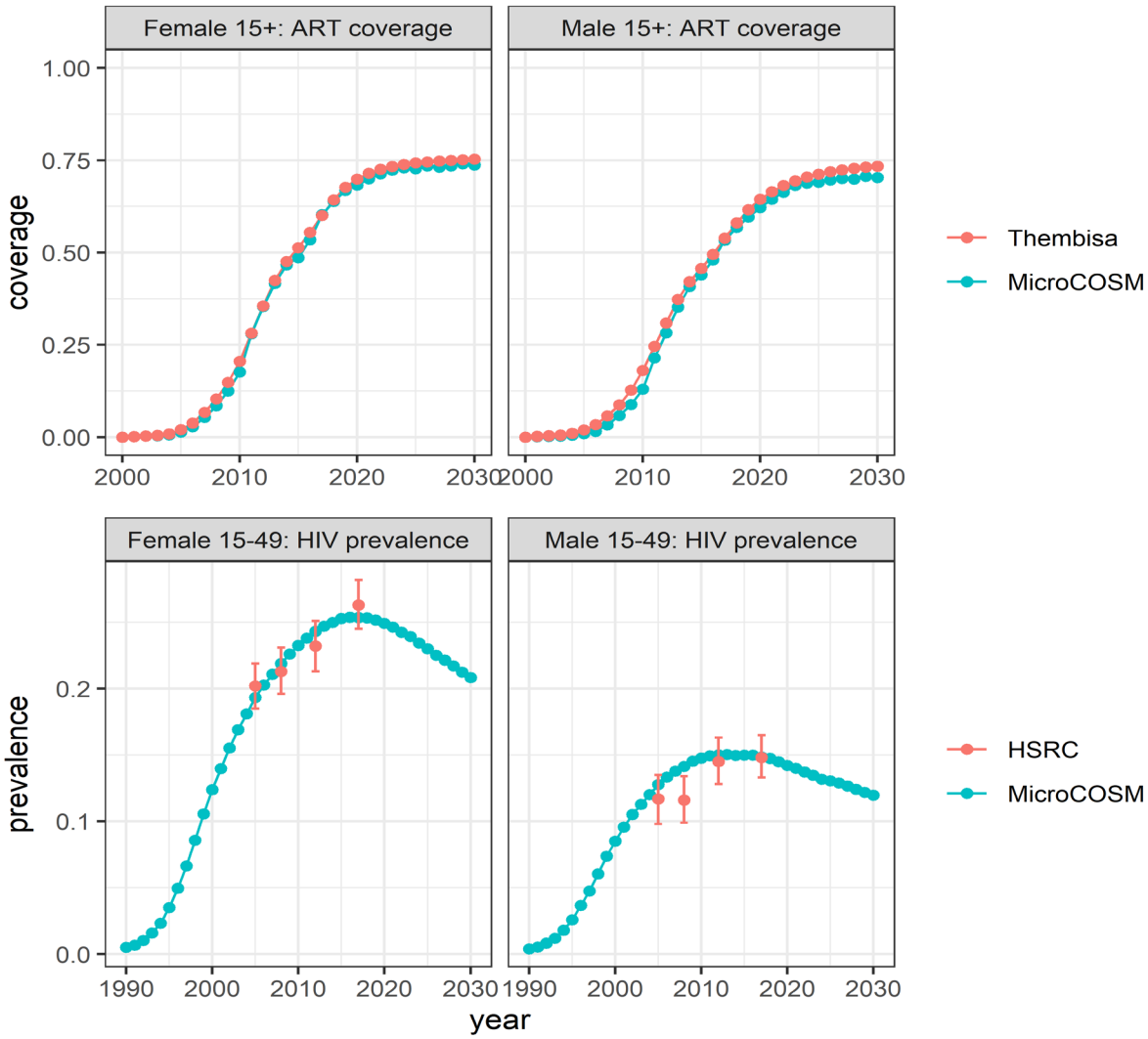


Figure A 3 – ART coverage and HIV prevalence over time.

A.2 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 (205), 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 (144,145) and details regarding the wide variety of data sources and the calibration methods to estimate the rates are documented in these publications.

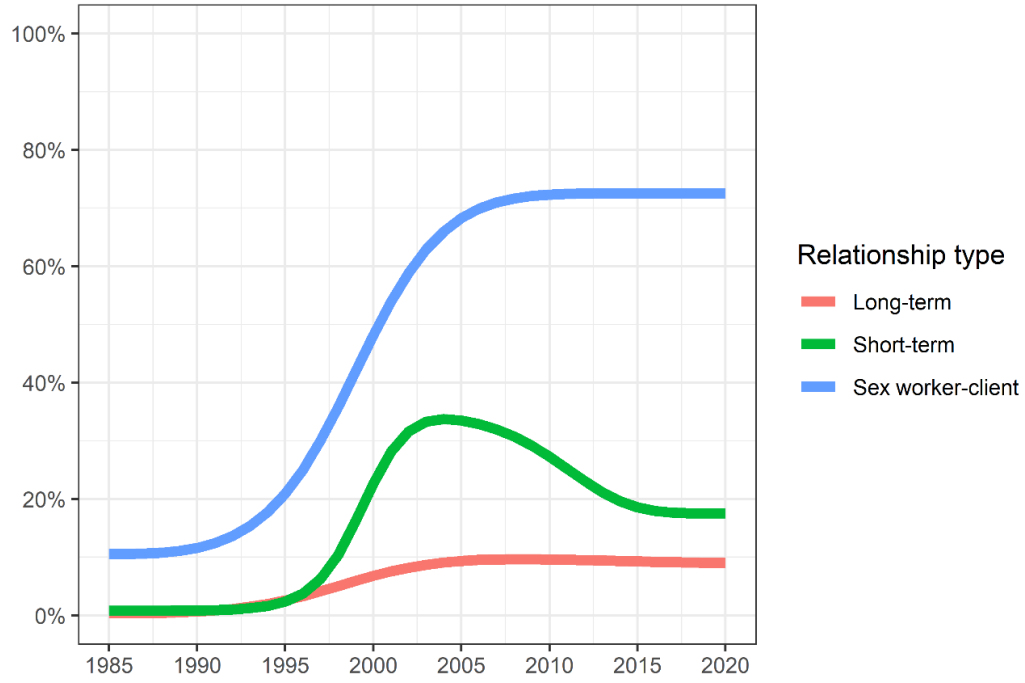


Figure A 4 – The proportion of sex acts protected by condoms among females aged 20-24 over time.

A.3 Starting conditions

As mentioned in the previous section, the simulation of individuals in MicroCOSM starts in 1985, and HIV is introduced in 1990. With the exception of 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.

A.4 Screening algorithm

A.4.1 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) (181). 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.

In 2010, NDoH published HIV management guidelines that included a section on cervical cancer screening (182). 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 (Chapter 5):

1) The same HIV-positive women can be counted 3 times in a ten-year interval, and they may be younger than 30.
2) 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.
3) 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 (188) 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 A 5 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.

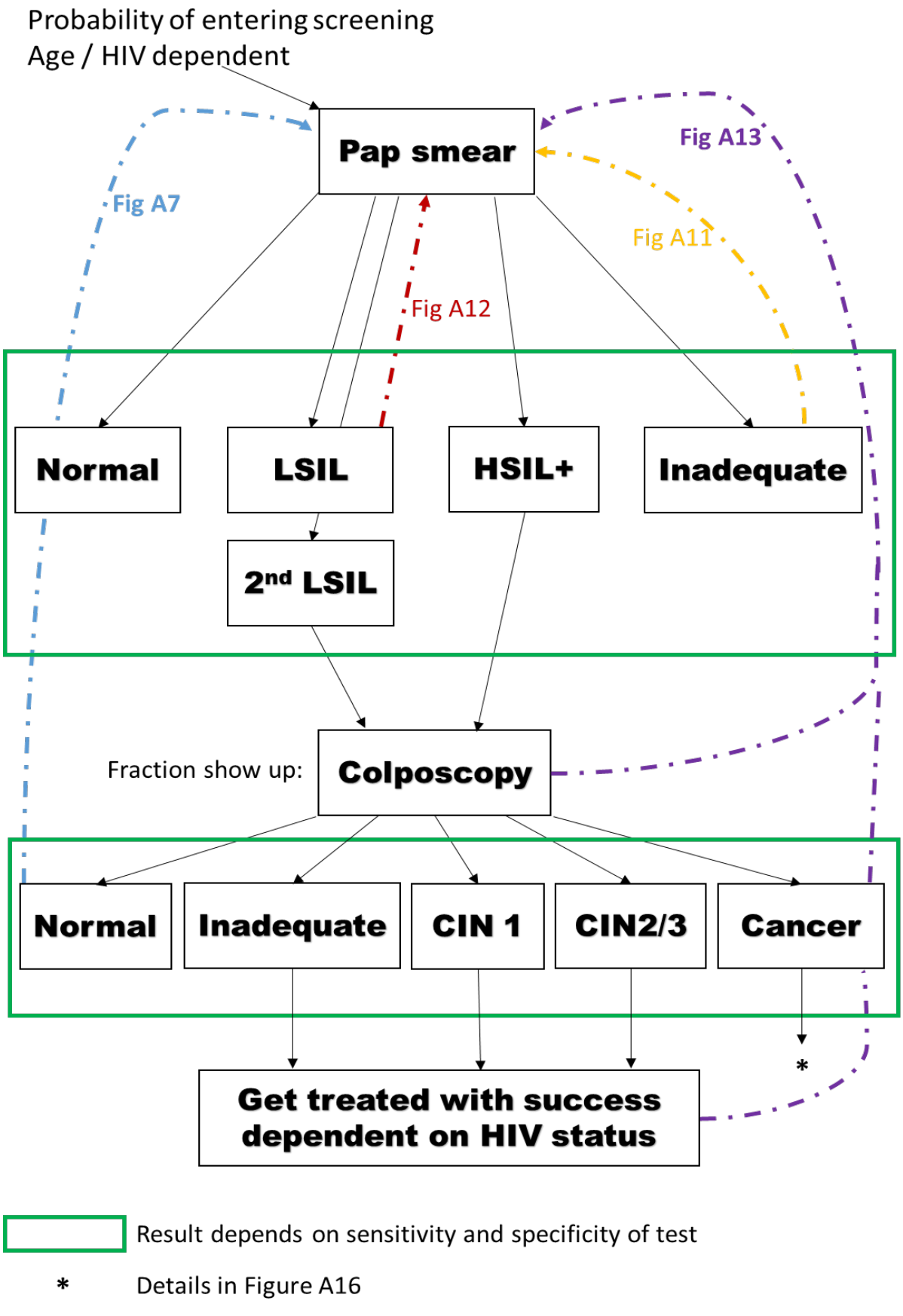


Figure A 5 - Screening algorithm in MicroCOSM

A.4.2 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.4.2.1 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 (215). 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 (Chapter 5). 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.

*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%

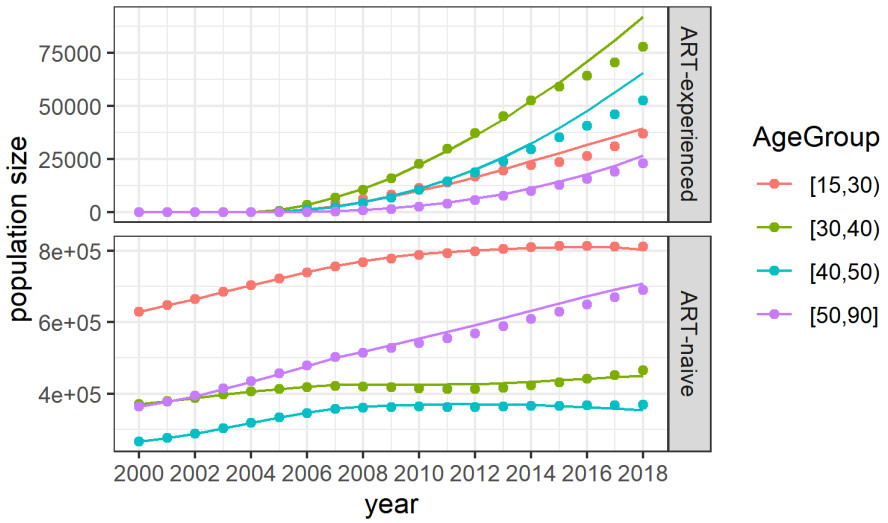


Figure A 6 - The number of ART-naïve women (HIV-positive or -negative) and ART-experienced in each age group in the Western Cape, as estimated by Thembisa (the dots) and the screening simulation model (the lines).

A.4.2.2 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.

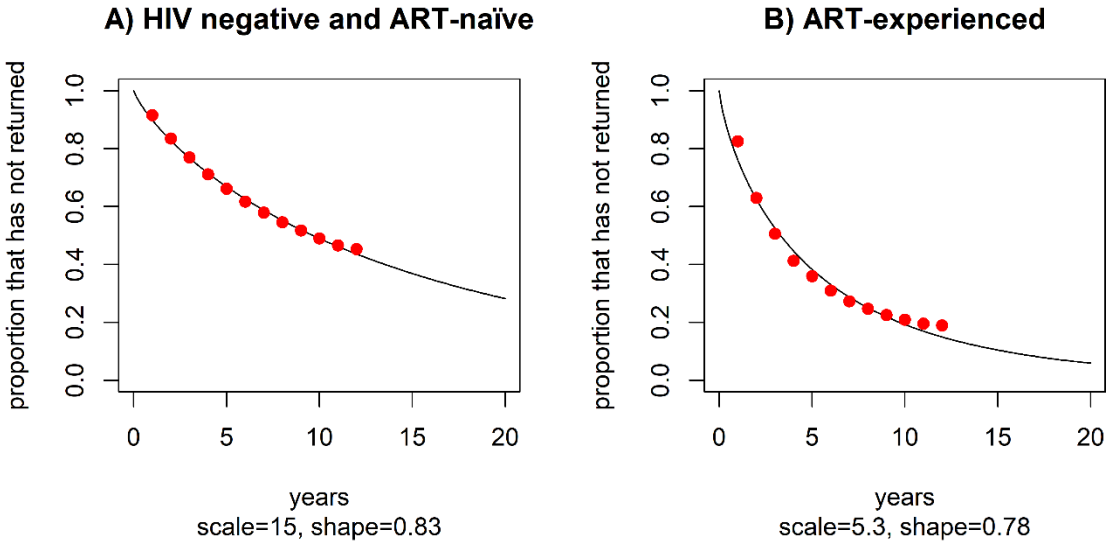


Figure A 7 - Distributions of time between a Pap smear with Normal cytology and next routine Pap smear. Red dots are data from the WC NHLS cytology database and black lines are best fitting Weibull distributions. These figures can be interpreted as follows: In A), around 23% of HIV-negative and ART-naïve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

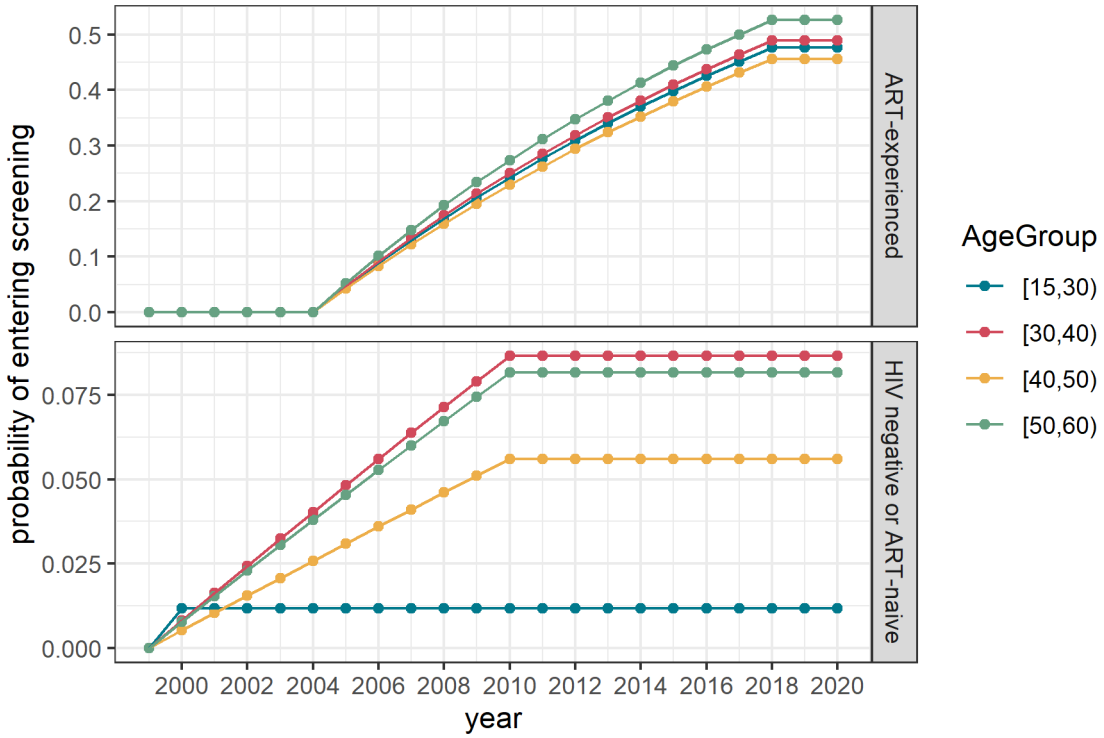


Figure A 8 – Yearly probabilities of entering the screening programme, by age and ART status.

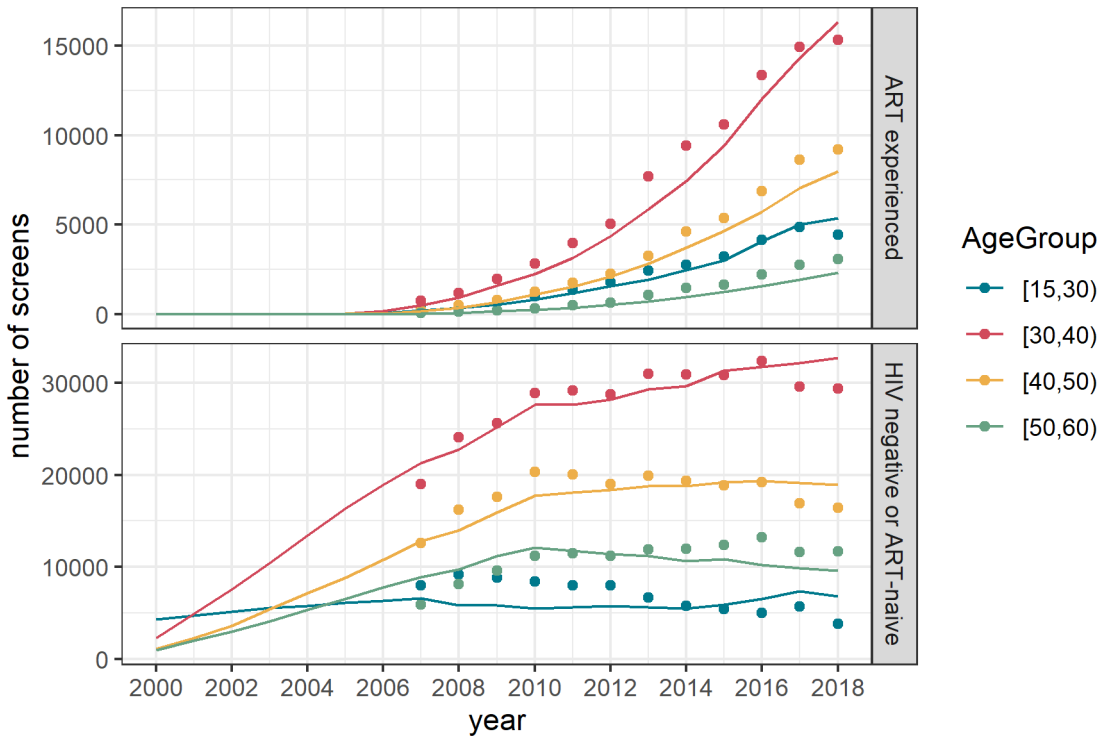


Figure A 9 - The number of ART-naïve women (HIV-positive or -negative) and on ART in each age group receiving routine Pap smears in the public sector of the Western Cape, as counted from the WC NHLS data (the dots) and the screening simulation model (the lines).

We calculate screening coverage from our simulated data using the same method as the estimate published in the District Health barometer (181) 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.

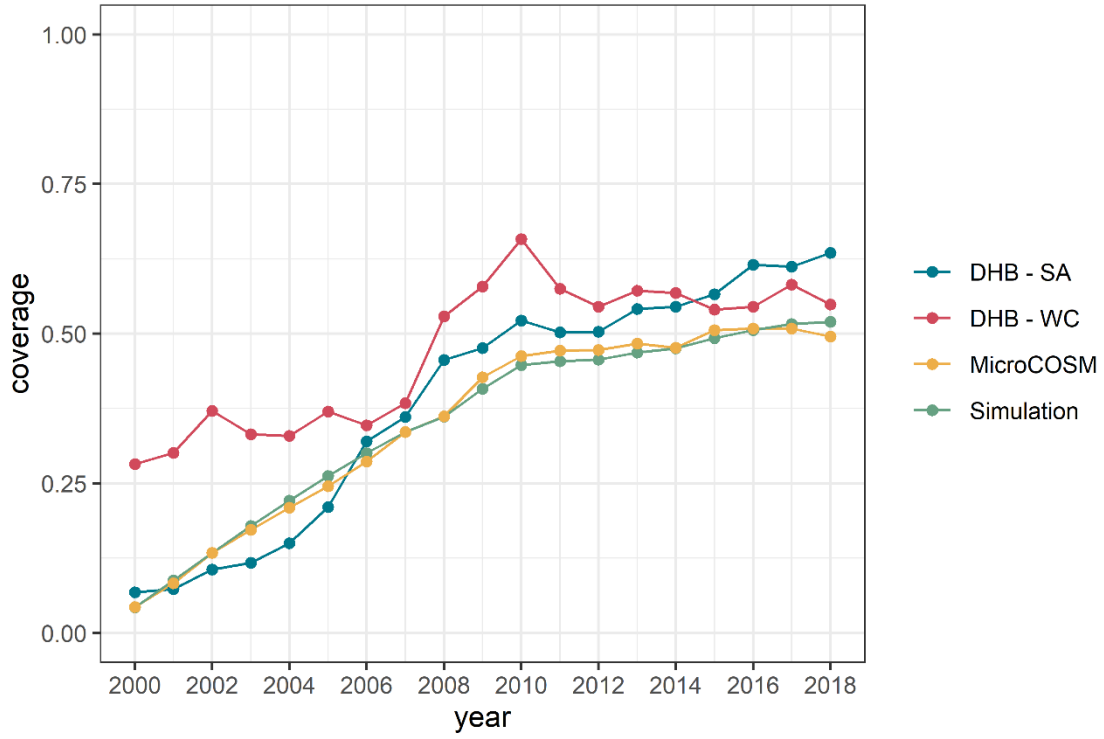


Figure A 10 - Screening coverage defined as the number of routine screens performed for women aged 30 and older, divided by a 10th of the population. In this definition, all screens are counted, with the exception of repeat smears and diagnostic smears.

A.4.3 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 (89), 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 (216). However, analyses of adequacy in South Africa as a whole, also using NHLS data, found very different results. Makura *et al.* (217) 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.* (218) 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% (39). A national level study performed in the early 2000s showed very high overall adequacy in all sites of 95% and above (219). 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 11). 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.

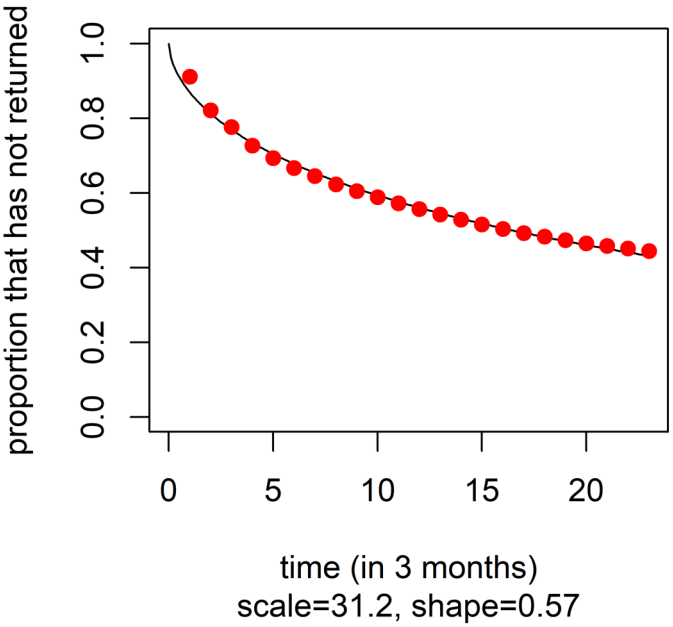


Figure A 11 - Distribution of time between an inadequate screen and the next screen, in three-monthly time steps. Red dots are data from the WC NHLS cytology database and the black line is the best fitting Weibull distribution. This figure can be interpreted as follows: 8.8% of women have returned for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

A.4.4 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 (33) and an HIV-positive woman should be rescreened after 3 years (182). 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 (Chapter 5). 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.4.5 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 (33). 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.

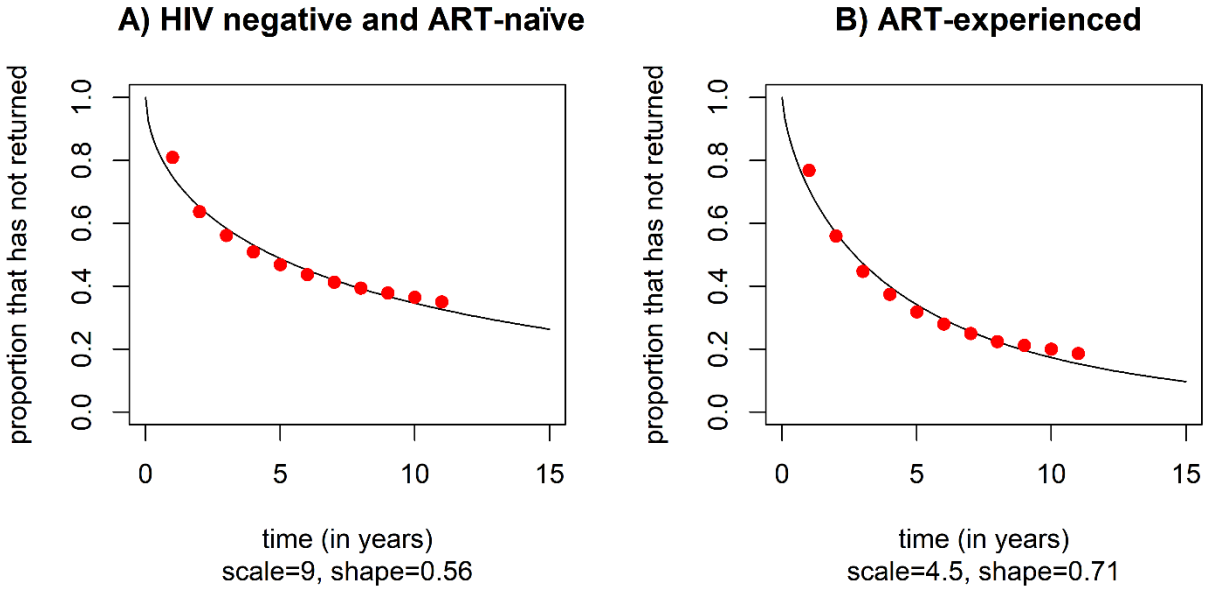


Figure A 12 - Distributions of time between a Pap smear with ASCUS/LSIL cytology and follow-up Pap smear. Red dots are data from the WC NHLS cytology database and black lines are best fitting Weibull distributions. These figures can be interpreted as follows: In A), around 19% of women unknown/negative 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.4.6 Sensitivity and specificity of screening tests

A.4.6.1 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% (90). 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% (6). When only considering studies in Nanda *et al.* (6) 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 (115,220,221). The overall sensitivity of a LSIL+ Pap smear to detect CIN1+ in these studies is higher than among the studies without verification bias in (6). These studies were performed among women in Cape Town, in communities with HIV prevalence of around 12% at the time (115,220,221). 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 (115,220,221).

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 (115,220,221).

	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 (220) (115). 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.4.6.2 HPV-DNA test

A recent meta-analysis estimated that the sensitivity/specificity of a HPV-DNA test to detect CIN2+ is 95%/92% (222). 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.

Study	HIV-negative		HIV-positive		Age	test
	Sensitivity	Specificity	Sensitivity	Specificity		
Kitchener 2007 (223)			100%	33%	median 29	HC2
Firnhaber 2013 (195)			92%	51%	18-65	HC2
McDonald 2014 (14)	85%	81%	99%	52%	17-65	HC2
Segondy 2016 (224)			92%	61%	25-50	careHPV
Segondy 2016 (224)			98%	31%	25-50	INNO-LiPA
Kuhn 2020 (28)	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.4.6.3 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 (225) 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.4.7 Fraction who visit colposcopy clinics

In the Western Cape, large fractions of women who should receive colposcopy services have no record of doing so (Chapter 5). 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.4.6. 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 (Chapter 5). 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.4.8 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 (33): "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" (226) 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 (226). 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.* (227) 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.* (228) 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.4.9 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 (91). 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 (226)	149	75.2%	266	30.8%	Normal at follow-up Pap smear, median 4 months later
Zeier 2012 (229)	335	73.1%	778	46.2%	Normal at follow-up Pap smear
Batra 2010 (230)	275	77.8%	219	47.9%	Normal Pap at 4 months post treatment
Noël 2015 (231)			259	41.3%	Normal at follow-up Pap smear
Smith 2017 (232)			83	19.7%	Normal/ASCUS Pap 6 months after treatment
Kabir 2012 (233)			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.* (226), 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 (92–94). 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.

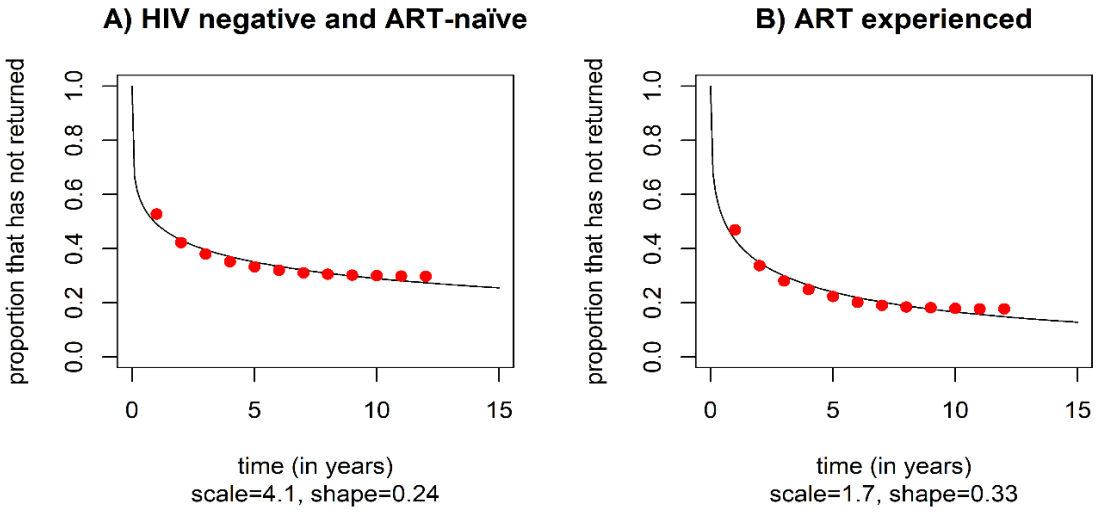


Figure A 13 - Distributions of time between colposcopy visit and follow-up Pap smear. Red dots are data from the WC NHLS cytology database and black lines are best fitting Weibull distributions.

A.5 Calibration method and data

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

A.5.1 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 parameter distributions derived in this calibration step were used for the analyses of Chapters 3 and 4.

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.* (234) 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:

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

Using the properties of the beta distribution, the mean and variance of θ_i are $E(\theta_i) = \alpha_i/(\alpha_i+\beta_i)$ 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 (146).

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.5.2 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). Parameter distributions derived in this step were used only in Chapter 6 of this thesis.

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 (6,235) 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 (14)	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 (196)	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 (115)	not tested	1996	Khayelitsha	CIN1+	35-65	2922	6.1%
Denny	not tested	1996	Khayelitsha	CIN2+|CIN1+	35-65	178	46.6%
Kuhn (28)	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 (14)	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 (14)	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 (236)	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 (237)	not tested	2011	Tshwane	253	5.7%	4.9%											
5	Snyman (238)	not tested	2012	Tshwane	160	4.4%	5.7%											
6	Adler (239)	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 (239)	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 (240)	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 (240)	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 (241)	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 (241)	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 (242)	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 (208)	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 (243)	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 (241)	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 (241)	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 (244)	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 (245)	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 (246)	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 (247)	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.5.3 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.5.3.1 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 (202). 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 (201). 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 (202) and Ekurhuleni (201)).

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.

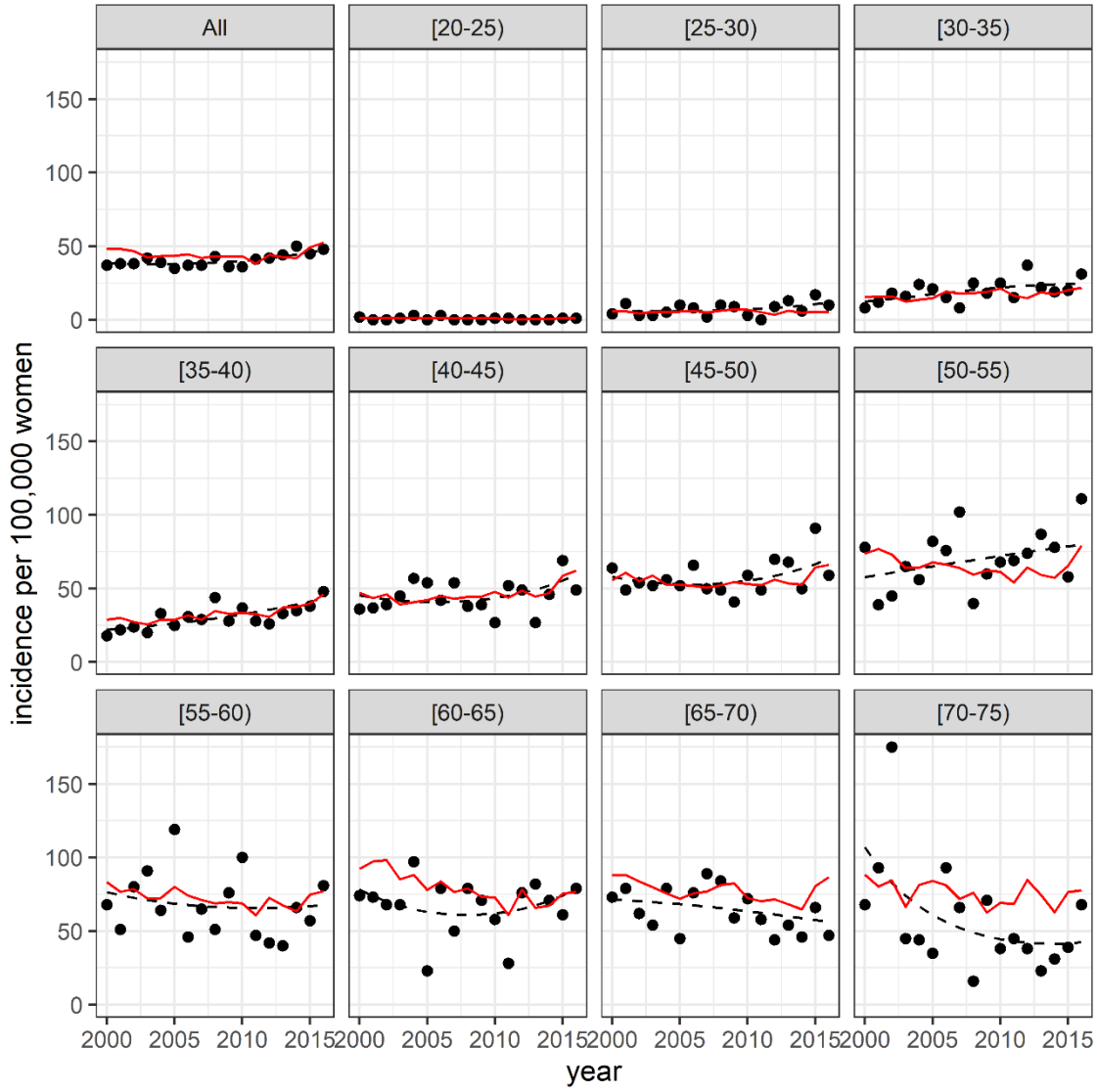


Figure A 14 - Cervical cancer incidence per 100,000 women as estimated by one parameter combination (black dots). The dashed line shows the fit of a Poisson regression through the black dots and the red line shows the crude estimate calculated using NCR data, adjusted by assuming that 10% of diagnosed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

A.5.3.2 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 (50). 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 (197)	2000	836	8.4%	35.6%	41.0%	15.0%	All women treated in Johannesburg hospital 1997-8
Mbodi (199)	2013	104	8.4%	22.1%	59.7%	9.6%	Sample of women treated in Chris Hani Baragwanath hospital 2013
Snyman (198)	2011	85	7.0%	17.6%	58.9%	16.5%	Sample of women treated in Kalafong hospital in 2011
Sabulei (200)	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 (248):

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

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.

A.6 Prior distributions of parameters

A.6.1 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

Section	Parameter	HPV 16	HPV 18	Other types
		Uniform prior distributions		
		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 (65,135)			
	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 (64,65)			
	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.6.1.1 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) (61) to almost 100% for male to female transmission in Van de Velde *et al.* (HPV-ADVISE) (169). 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.6.1.2 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 (249)	5	4 - 6	2500	all ages	One negative	7.3 (6.3-10.7)	6.9 (6-12)
Richardson (250)	2	6	621	university	One negative	19.4 (11.4-27.5)	9.4 (4.8-14)
Ho (251)	3	6	608	university	One negative	11 (7-12)	12 (6-17)
Woodman (252)	4	6	1075	15-19	One negative	10.3 (6.8-17.3)	7.8 (6-12.6)
Goodman (253)	5	4	972	all ages	One negative	9.7 (4.5-24.2)	14.3 (4.9-)
Munoz (254)	2.5	6	1610	13-85	One negative	13.7 (8.4-18.8)	11.9 (9.1-16.6)
Insinga (255)	4	6	1203	16-23	Two negatives	13.2 (12-17.6)	13.2 (11.7-17.7)
Insinga (256)	4	6	1788	16-23	Two negatives	17.1 (15.1-20.2)	12.4 (11-17.7)
Jaisamrarn (257)	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 (258). 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) (65,135). 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 (249,251,253,254,259).

A.6.1.3 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 (64,65). 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.6.2 Cervical pre-cancer and cancer parameters

Prior distributions assigned to the parameters of the cervical disease components of the model are given in Table 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.6.2.1 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.6.2.2 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 (255,260). 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% (251,257,261,262).

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 (260)							
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 (255)							
16	142	0.37	0.17	0.54	32%	12%	55%
18	62	0.48	0.08	0.56	14%	1%	42%

A.6.2.3 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 (263,264). 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.* (264), 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 (263). 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 (44,52,265,266).

Table A 16 – Regression and progression of CIN2 by 24 months from Tainio et al. (263)

	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.6.2.4 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 (18). 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.* (139), 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 (22). However, no study to our knowledge directly compared rates for women on ART to HIV-negative women. As shown in Rohner *et al.* (23), 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 (22).

HPV to CIN1: The relative risk for progression for untreated HIV+ vs HIV- women estimated in Liu *et al.* (139) of 3.73 (95% CI 2.62-5.32) were derived from two studies in the late 1990s. In Kelly *et al.* (22) 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.6.2.5 Proportion that clears HPV infection during CIN1 regression

In a study by Nobbenhuis *et al.* (267), 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.* (268), 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 (269,270). In the model, 90% of women are assumed to clear the HPV infection at the same time that CIN1 regression takes place.

A.6.2.6 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.

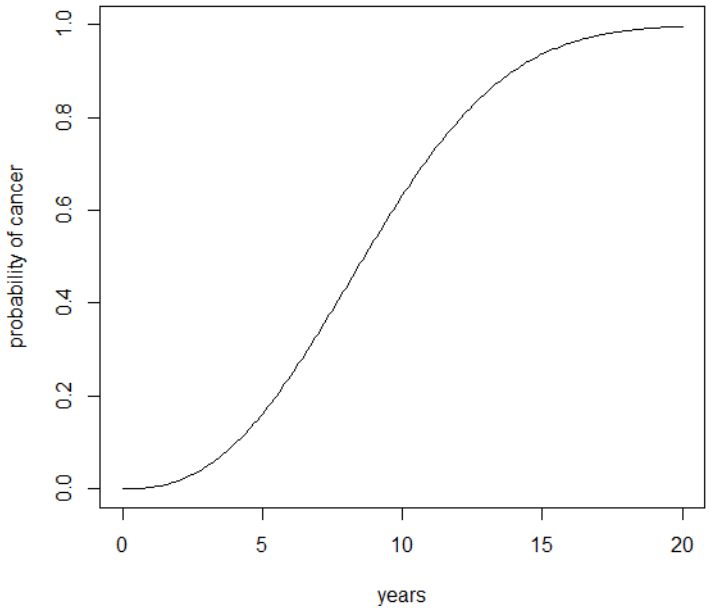


Figure A 15 – Cumulative probability function of a Weibull distribution with scale of 10 years and shape of 2.5. This curve represents the cumulative probability that an HIV-negative woman who has CIN3 will progress to cancer over time.

A.6.2.7 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.

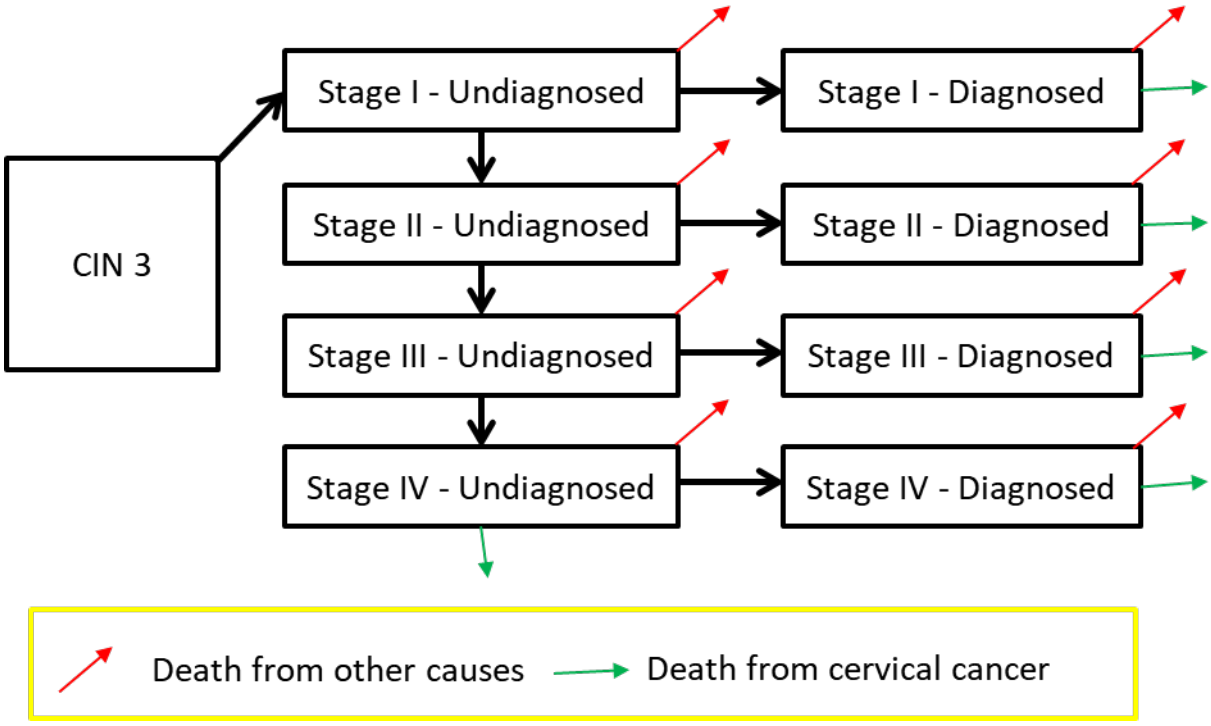


Figure A 16 – Progression and diagnosis of cervical cancer in our model. Undiagnosed women progress through all the stages before dying of cancer, while mortality in diagnosed women depends on the stage at diagnosis.

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 (51). A more recent analysis of the American National Cancer Institute's data found very similar estimates (271). 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 (30,272).

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.* (51), 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.* (51).

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

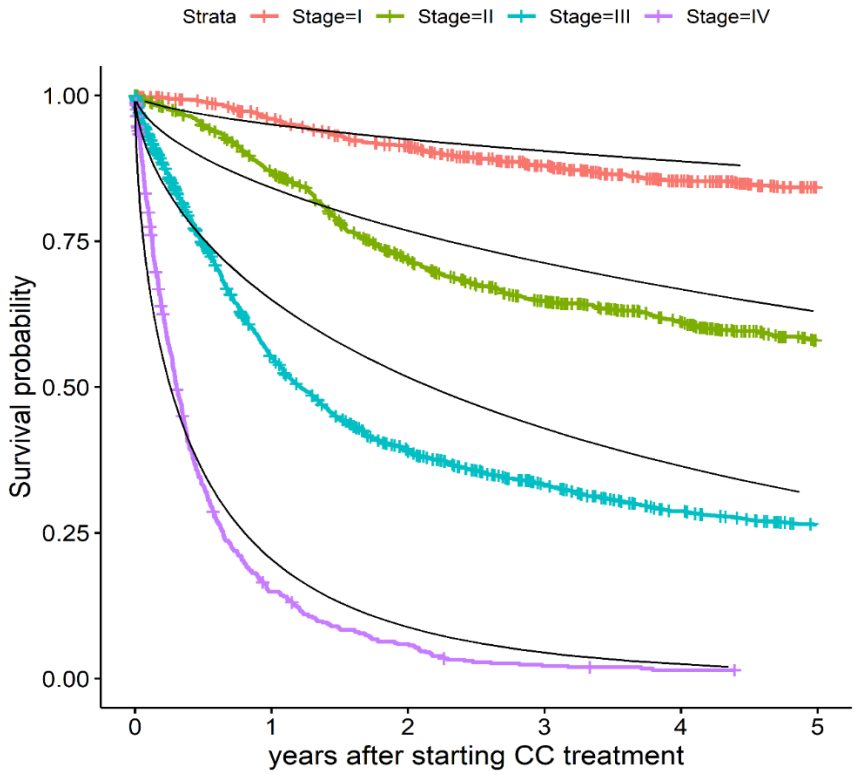


Figure A 17 – Cervical cancer survival probailities over time after treatment was started, by stage of diagnosis. The crosses represent the Groote Schuur Hospital data, while the black lines represent Weibull survival distributions.

A.7 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.

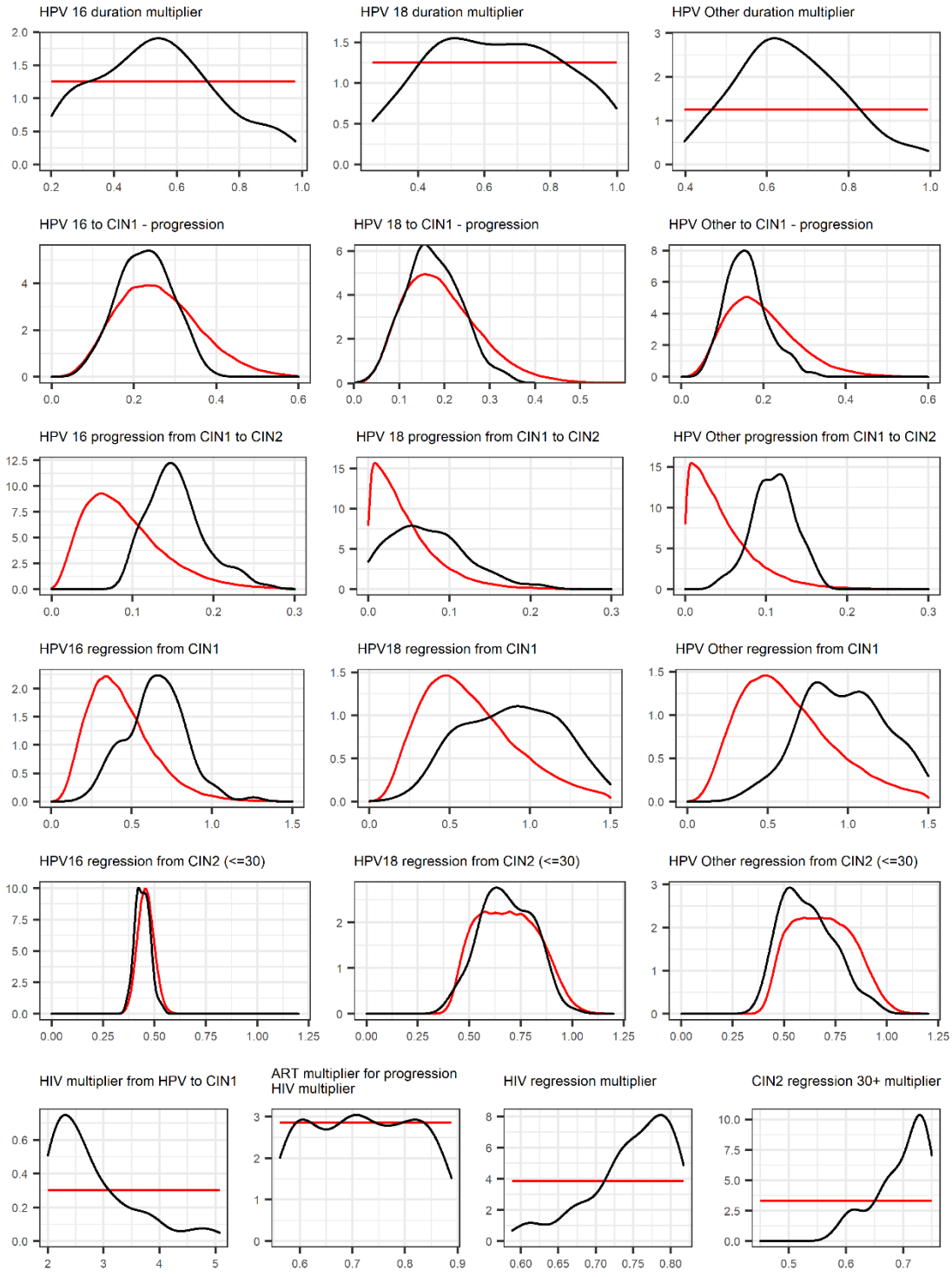


Figure A 18 – Prior and posterior distributions of cervical pre-cancer parameters. Red lines represent prior and black lines represent posterior distributions.

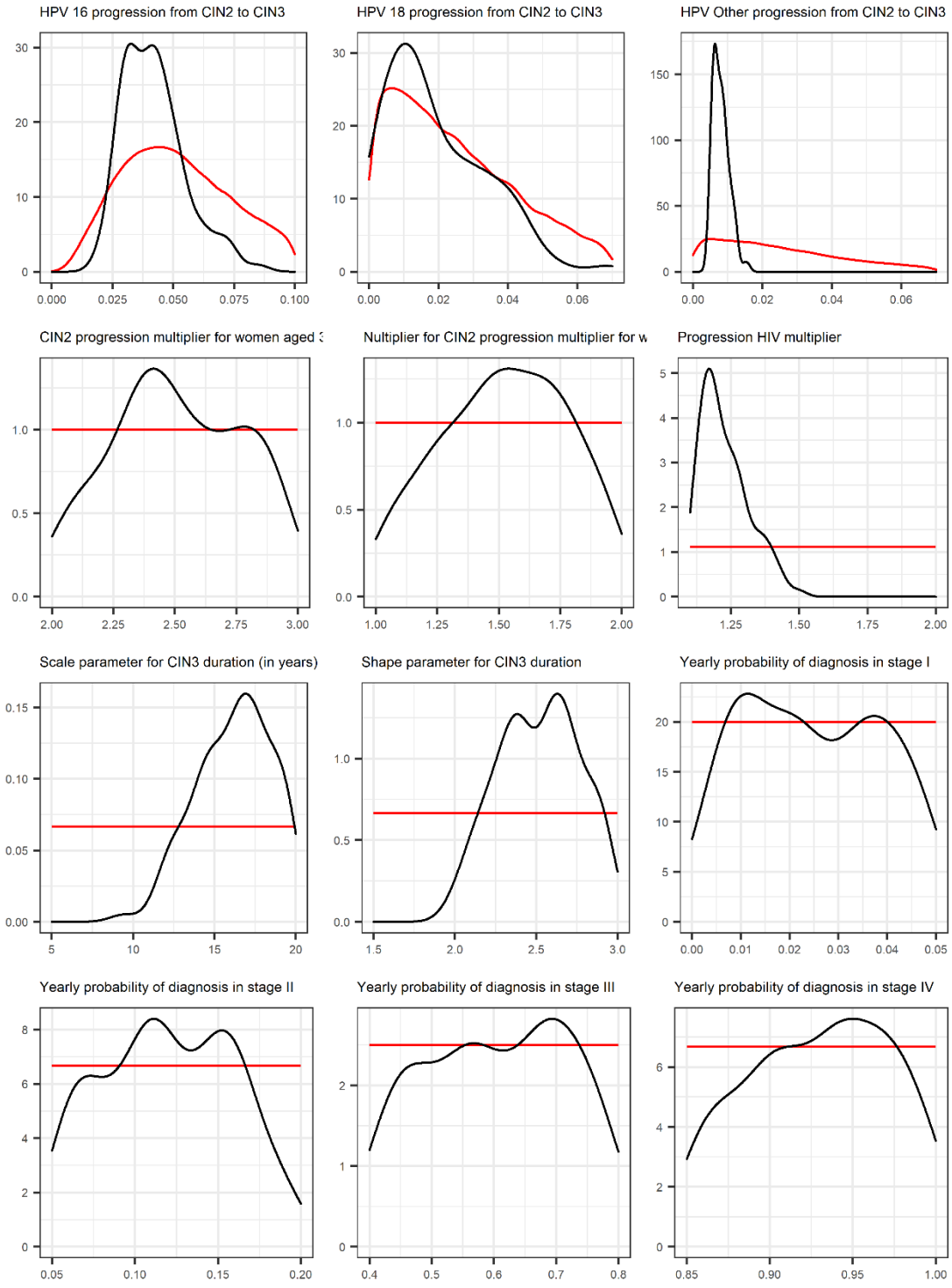


Figure A 19 - Prior and posterior distributions of cervical cancer parameters. Red lines represent prior and black lines represent posterior distributions.

A.8 Model fits to data

A.8.1 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)

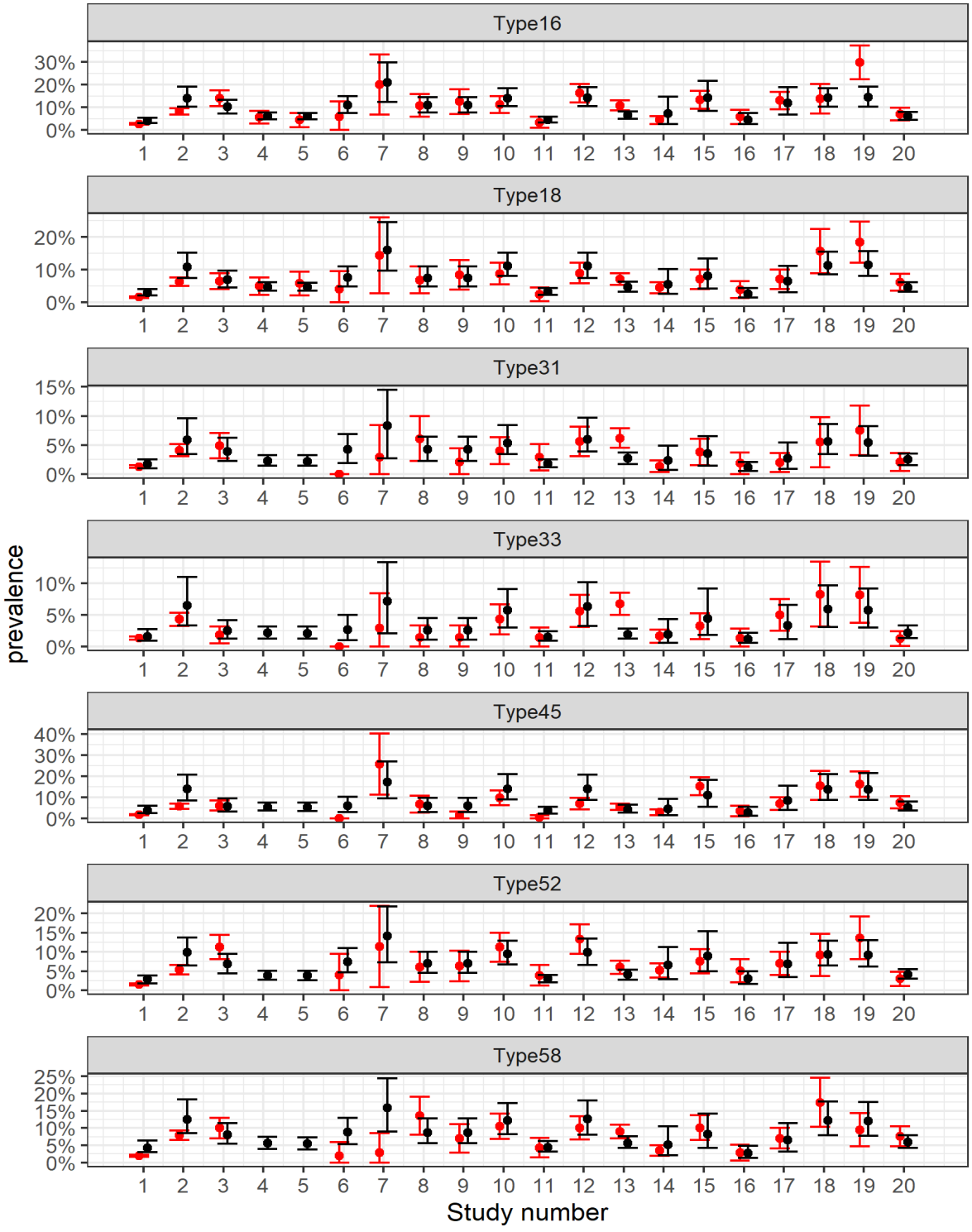
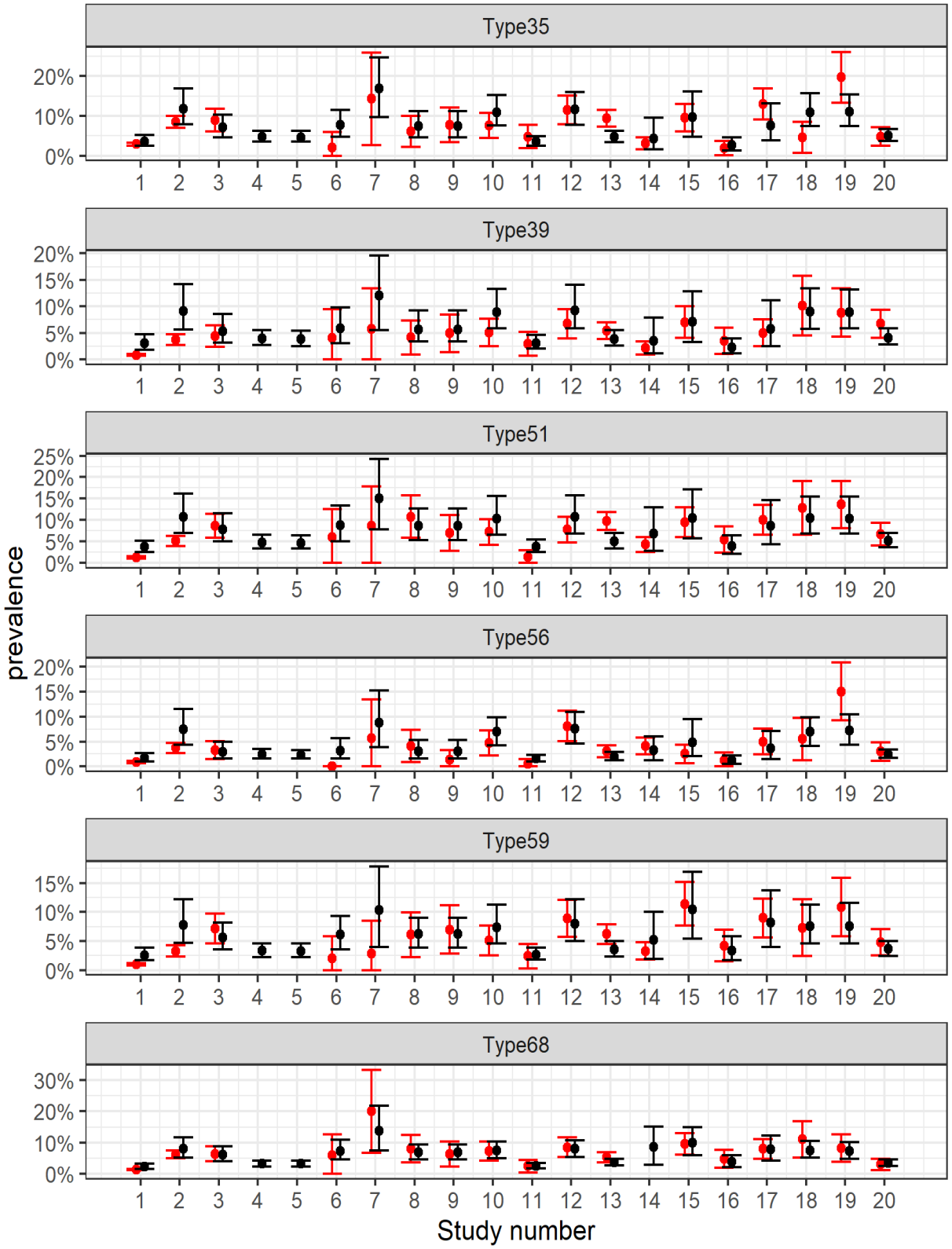


Figure A 21 - Model fits to HPV types 35, 39, 51, 56, 59, 68 (non-vaccine types)



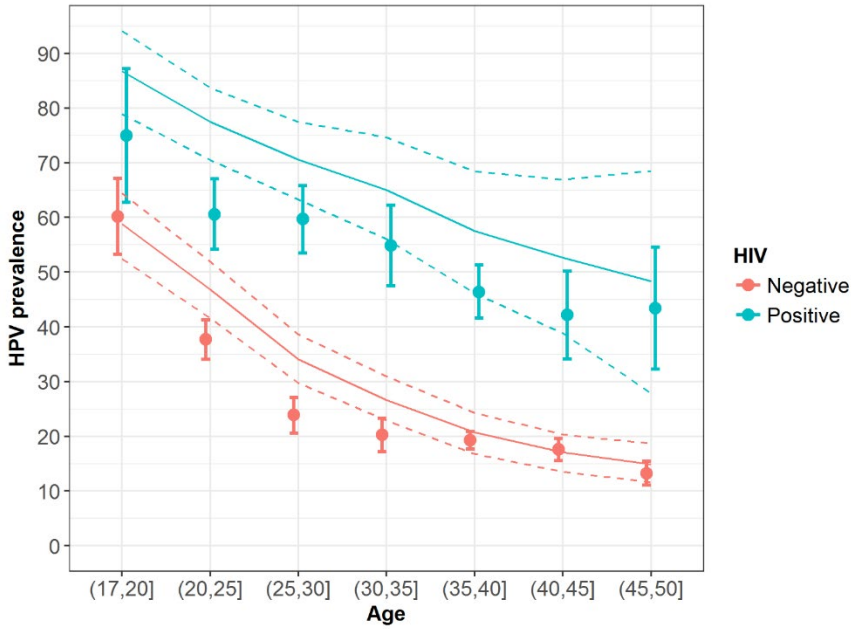


Figure A 22 - The mean overall high risk HPV prevalence in 2000 from 500 populations simulated using the sample of parameter combinations from the posterior distributions (solid lines). The confidence bands (2.5th and 97.5th percentiles) around these estimates are shown in dashed lines. Data points (in closed circles) represent HPV prevalence results from a population level study in Khayelitsha, Cape Town (273).

A.8.2 Cervical pre-cancer

Table A 20 and Figure A 23 show the model fits to the data in Table A 9. Although the data on HIV-positive women in the McDonald study (14) 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 (196)	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 (115)	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 (28)	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+)

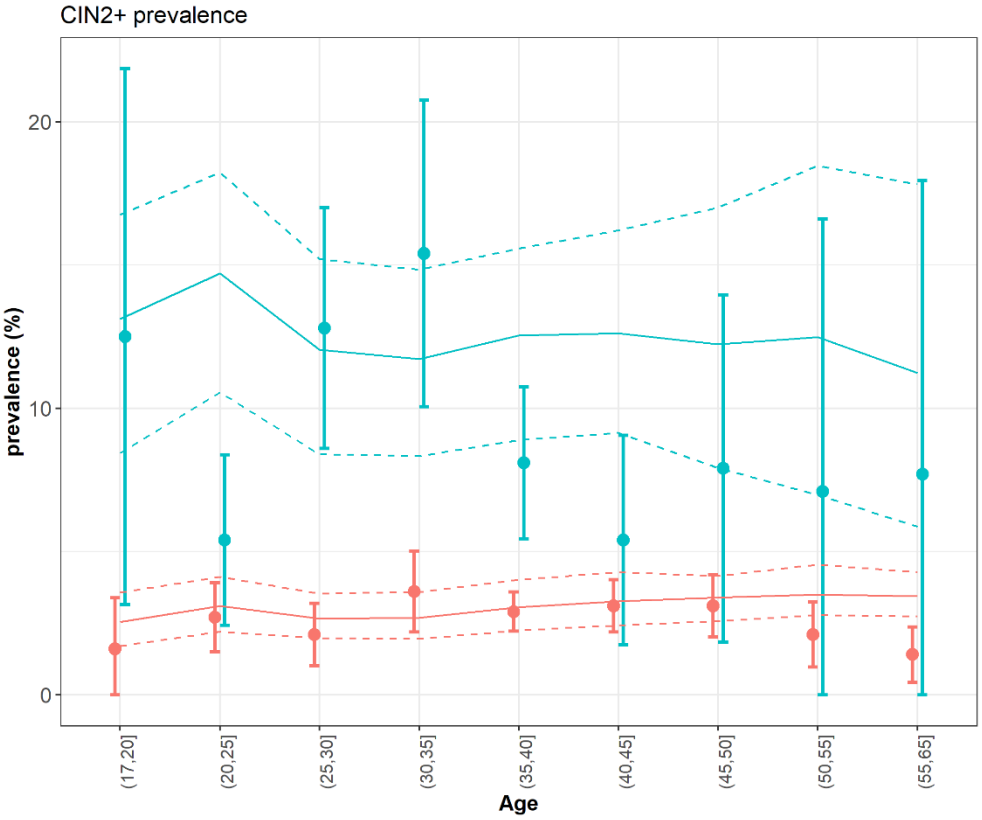


Figure A 23 – CIN2+ prevalence by age and HIV status (red=HIV-negative and blue=HIV-positive) from McDonald et al. (14)

A.8.3 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 24 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).

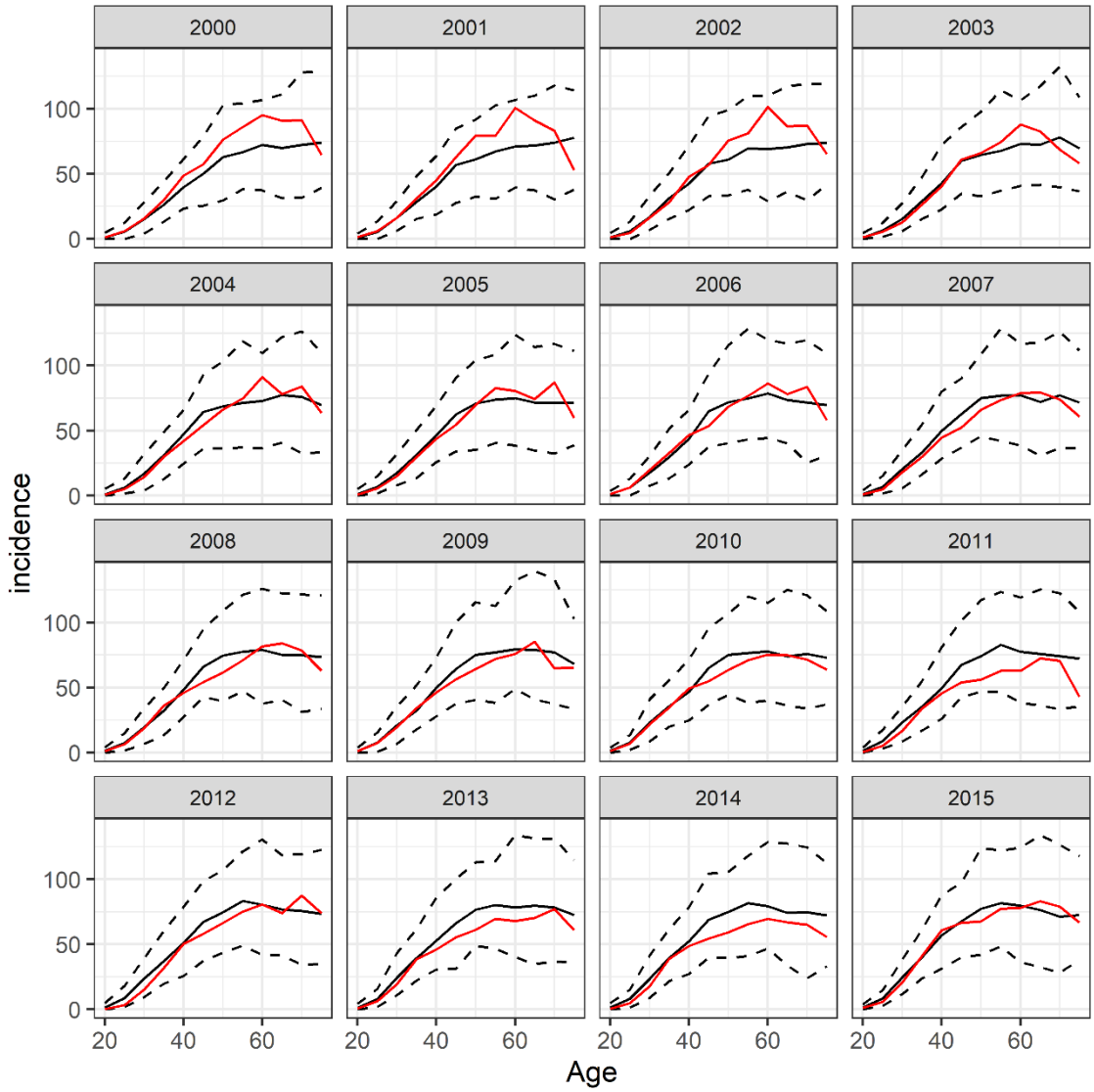


Figure A 24 – Age specific diagnosed cervical cancer incidence per 100,000 women as calculated from NCR data (red lines) and the 100 best fitting parameter combinations (black lines show mean of 100 estimates, dashed lines show 95% percentile intervals).

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.

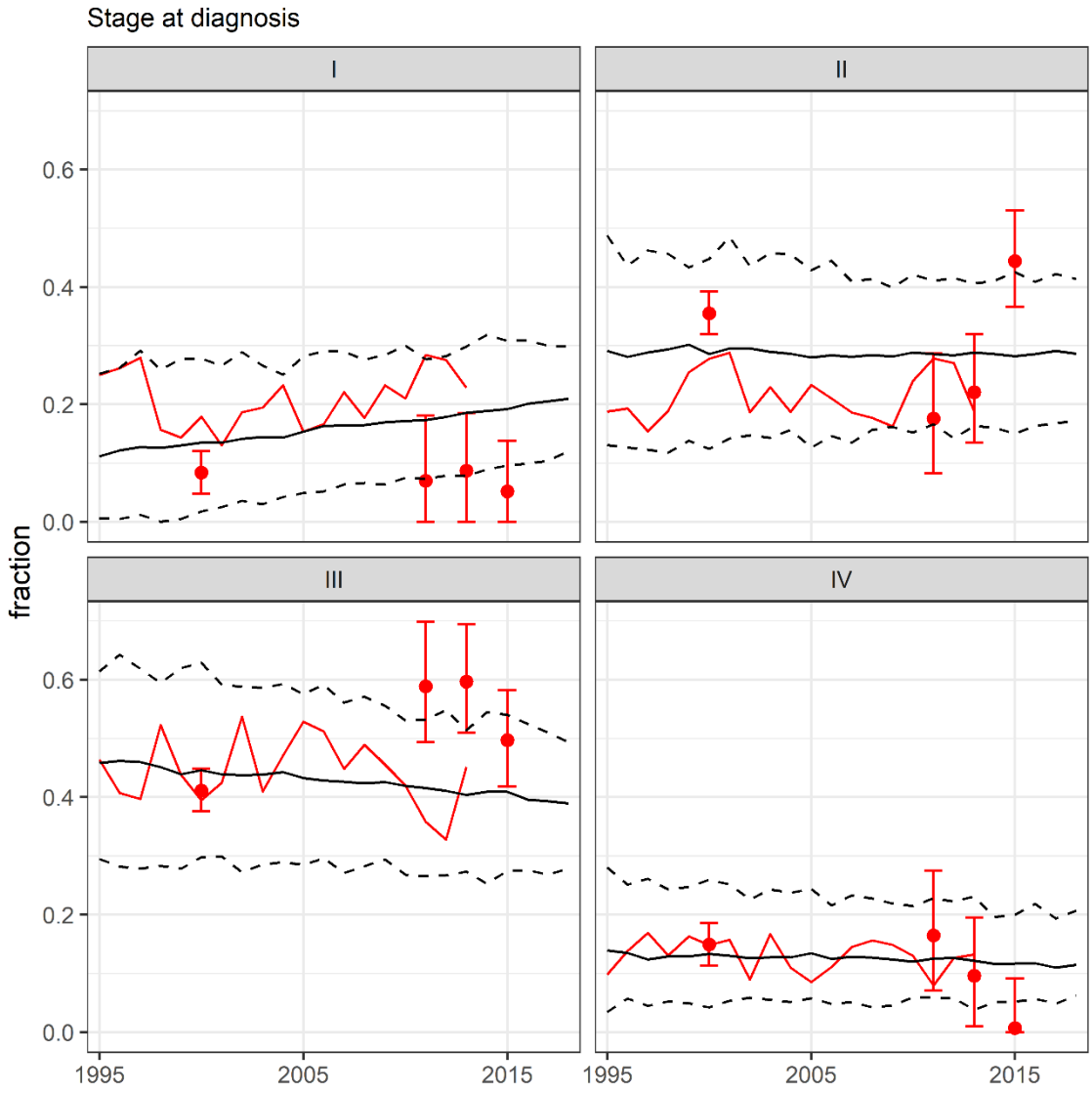


Figure A 25 – The fraction of cervical cancer cases diagnosed in each stage of the disease. The black lines show mean model estimates and the dashed lines show 95% percentile intervals. The red lines represent GSH data and red dots represent data as shown in Table A11 (Lomalisa 2000, Snyman 2011, Mbodi 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.* (206) 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 26 (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.*(41). 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.

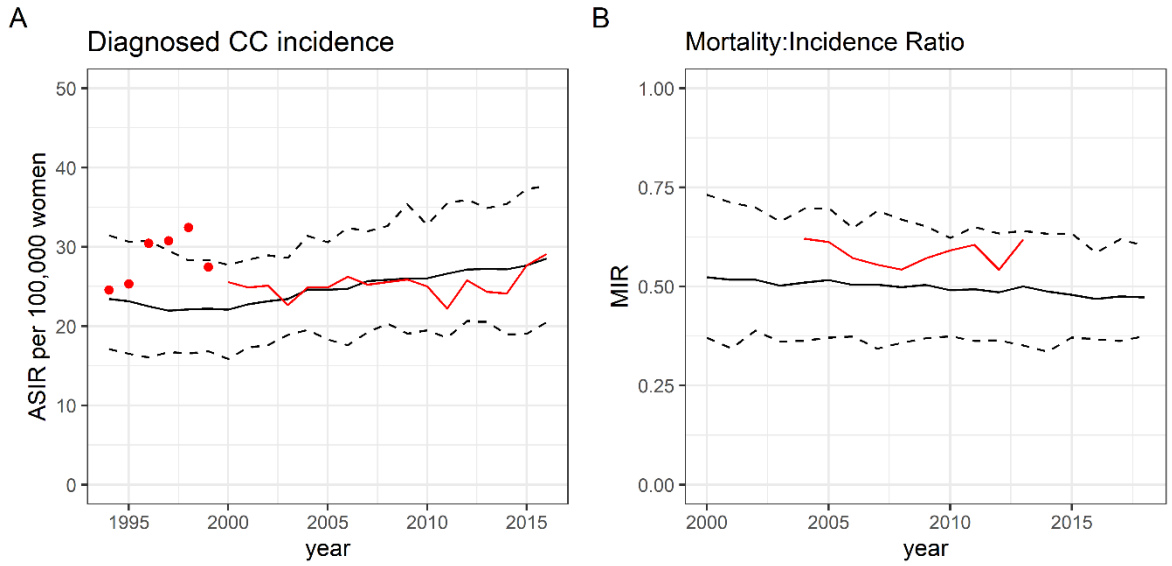


Figure A 22 – Diagnosed ASIR and mortality to incidence ratio using 100 best fitting parameter combinations. The black line represents the mean and the dashed lines represent the 95% percentile intervals. Red lines and points represent data from Olorumfemi et al.(41).

In Denny *et al.* (274), 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.* (275), prevalence of HPV16/18 among cancer patients was 63.2% (red dot in Figure A 27). 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.* (276) 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).

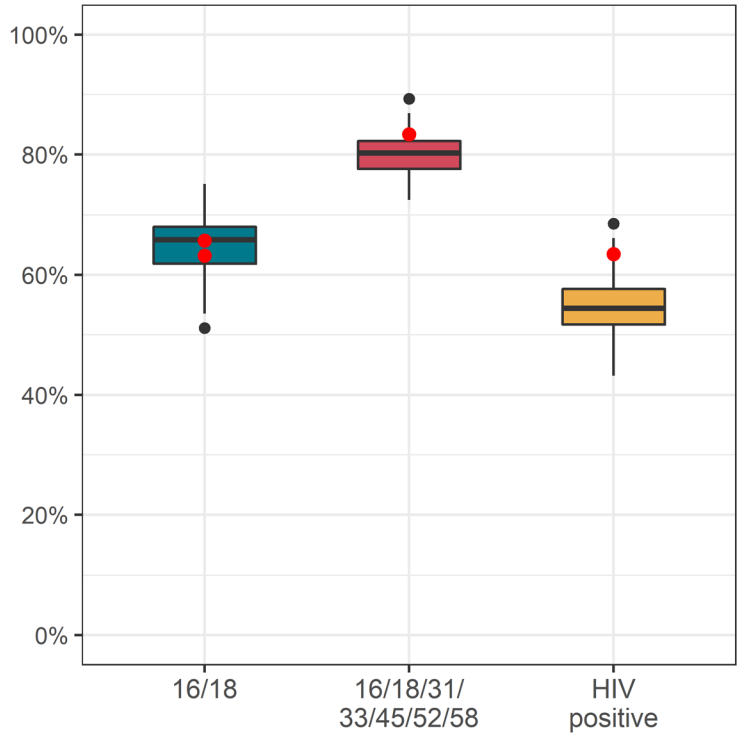


Figure A 27 – Fractions of CC cases that was caused by HPV16/18, HPV16/18/31/33/45/52/58, and fraction of CC cases who were co-infected with HIV. Boxplots represent model estimates and red points represent data. Data are from Denny et al. (274) and Van Aardt et al. (275) for cancer caused by HPV16/18, from Denny et al. (274) for cancer caused by other vaccine types, and from Stelzle et al. (276) 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 (197)	1997	776	60	53	44	56 (54-59)	38 (30-48)
Moodley (272)	1999	522	138	55.2	39.8	56 (54-60)	39 (33-43)
Moodley (277)	2000	457	29	46	40	56 (53-58)	40 (35-45)
Van Aardt (275)	2008	154	77	55.8	41.3	57 (55-59)	43 (40-46)
Van Bogaert (31)	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 (2) (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.

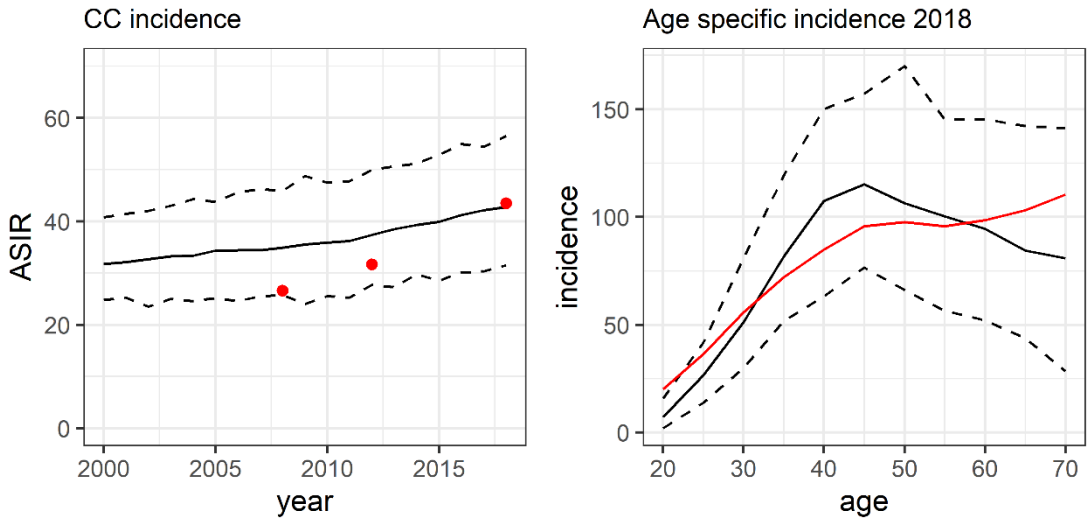


Figure A 28 – Age-standardised and age-specific cervical cancer incidence as calculated by GLOBOCAN (2) *(red) and the 100 best fitting parameter combinations (black line show mean of 100 estimates, dashed lines show 95% percentile intervals).*

Appendix B – Supplementary material to Chapter 3

B.1 HIV and HPV transmission associations

Associations, unadjusted for any behavioural factors, are calculated using the simulated data and are compared to unadjusted estimates from independently published longitudinal studies. The approach of comparing crude estimates is followed since studies adjusted for different demographic and STI information and sexual behaviour indicators were defined differently. For this reason, only studies that showed unadjusted estimates are included in the comparison. The model only simulates high risk HPV, and where possible, study results relating to high-risk HPV are used. For the association between HPV infection and HIV acquisition, studies were identified from two meta-analyses (20,21) and later studies that cited these meta-analyses. Appropriate studies that estimated the association between HIV status and new HPV detection were identified using a keyword search in Google Scholar and the citations of and references in identified studies. Tables B1 and B2 show the same information as Figures 5 and 6 in Chapter 3, with some extra details.

Table B 1 - Associations between HPV status and HIV acquisition. The reference category is HPV uninfected in every case, unless otherwise noted.

Study	**HPV types**	**Measure**	**Unadjusted estimate (95% CI)**	**Model estimate (95% CI)**
Effect of HPV prevalence at baseline				
Auvert (147)	High risk	Hazard ratio	1.5 (1.1-2.1)*	1.3 (1.1-1.4)*
Smith-McCune (151)	High risk	Hazard ratio	1.8 (1.1-2.9)	2.5 (2.0-3.1)
Tobian (152)	All	Odds ratio	3.9 (1.7-8.8)	2.5 (1.8-3.3)
Veldhuizen (133)	High risk	Odds ratio	4.9 (1.2-19.7)	1.7 (1.2-2.4)
Effect of HPV prevalent at visit *t-1*				
Gallagher (132)	Only high risk	Odds ratio	1.2 (0.7-2.0)	1.8 (1.2-2.5)
Rositch (153)	High risk	Hazard ratio	2.7 (1.7-4.8)	2.4 (1.8-3.0)
Effect of Clearance of an HPV type between visits *t-1* and *t*				
Gallagher (132)	High risk	Odds ratio	1.3 (0.8-2.1)	1.3 (0.6-2.6)
Smith-McCune (151)	High risk	Hazard ratio	2.0 (1.5-4.0)**	1.9 (1.3-2.5)**
Rositch (153)	High risk	Hazard ratio	2.8 (1.4-5.4)	1.8 (1.2-2.6)
Tobian (152)	High risk	Odds ratio	4.1 (1.6-10.4)	1.9 (1.0-3.2)

* For every unit increase in number of HR-HPV types present.
** Reference category includes those in whom HPV persisted.

Table B 2 - Associations between HIV status and new HPV detection

Study	HPV types	Measure	Unadjusted estimate (95% CI)	Model estimate (95% CI)
HIV-positive at baseline				
Adhieh (135)	High risk	Incidence rate ratio	1.8 (1.3-2.7)	2.1 (1.4-2.8)
Mbulawa (females) (17)	All	Incidence rate ratio	3.0 (2.1-4.3)	2.1 (1.4-2.8)
Minkoff (137)	High risk	Incidence rate ratio	2.7 (1.6-4.4)	2.1 (1.4-2.8)
Blitz (136)	High risk	Hazard ratio	2.3 (1.1-4.8)	2.4 (2.1-2.7)
Strickler (65)	All	Hazard ratio	(1.7-5.0)*	2.4 (2.1-2.7)
Whitham (148)	All	Hazard ratio	1.6 (1.3-1.9)	1.7 (1.4-2.0)
Mbulawa (males) (17)	All	Incidence rate ratio	2 (1.5-2.7)	2.2 (1.4-3.2)
Recent HIV				
Nowak (138)	All	Odds ratio	2.8 (1.4-5.5)**	2.8 (1.9-4.0)
Wang (16)	All	Hazard ratio	4.4 (2.5-7.6)	2.9 (2.5-3.3)

*Point estimates for different stages of HIV infection
**For one HPV type, at three months after first HIV detection

As described in the main manuscript, model results were adjusted for sex, age, number of new sexual partners in the preceding 6 months and marital status. These are variables that epidemiological studies typically control for. Our adjusted results remained significantly greater than 1 (Chapter 3 and Table B 3). In further analysis, we adjusted our results for the index of concurrency in the sexual network, a measure of the size of an individual's current sexual network (149). Epidemiological studies will not be able to accurately measure the size of each individual's sexual network, but in simulated cohorts this can be calculated exactly.

Amongst the 500 simulated cohorts, the mean hazard ratio for the association between an oncogenic HPV type at $v_d - 1$ and HIV acquisition at v_d, adjusted with the index of concurrency in addition to other variables, is 1.3 (95% CI 1.1–1.6). The corresponding mean hazard ratio for the association between HPV status at baseline and subsequent HIV acquisition is 1.2 (95% CI 1.0-1.5). The association between HPV clearance and HIV acquisition between $v_d - 1$ and v_d is 1.1 (95% CI 0.8-1.3).

The mean unadjusted hazard ratio for the association between HIV status at baseline and newly detected HPV, adjusted with the index of concurrency in addition to other variables, is 1.7 (95% CI 1.4–1.9). The corresponding mean hazard ratio for the association between new HIV detection at $v_d - 1$ and new HPV detection at v_d is 1.7 (95% CI 1.5-1.9).

As expected, this further adjustment brought the model results closer to the null, but only the association between HPV clearance and HIV acquisition did not remain significantly greater than one. The index of concurrency thus accounts for much – but not all – of the observed association between HIV and HPV when individual-level risk factors are controlled for.

Table B 3 - Associations between HIV and HPV in the general population aged 15-49

Association	Unadjusted HR (95% CI)	Adjusted HR* (95% CI)	Adjusted HR** (95% CI)
HPV at v_d - *1* and HIV acquisition at v_d	2.6 (2.2-3.1)	1.7 (1.4-2.1)	1.3 (1.1-1.6)
HPV at v_0 and HIV acquisition at v_d	2.4 (2.0-3.0)	1.6 (1.3-2.0)	1.2 (1.0-1.5)
HPV clearance between v_d - *1* and v_d and HIV acquisition at v_d	1.9 (1.4-2.4)	1.3 (1.0-1.7)	1.1 (0.8-1.3)
HIV at v_0 and new HPV detection at v_d	2.5 (2.2-2.8)	2.1 (1.8-2.3)	1.7 (1.4-1.9)
HIV acquisition at v_d-*1* and new HPV detection at v_d	2.9 (2.5-3.3)	2.2 (1.8-2.5)	1.7 (1.5-1.9)

* Adjusted for sex, age, number of new sexual partners in the preceding 6 months and marital status.
** Adjusted for sex, age, number of new sexual partners in the preceding 6 months, marital status, and index of concurrency.

B.2 Sensitivity analysis

B.2.1 HPV natural history

Different HPV natural history model structures were considered. In the base model, a proportion of individuals will acquire natural immunity (with an exponentially distributed duration) upon clearance and the remaining proportion becomes latently infected and can reactivate infection (this model structure corresponds to the methods and results of Chapter 3). As a sensitivity analysis, the latency assumption of this model was changed in three ways: 1) no-one enters a latent state, 2) everyone can enter a latent state, but only HIV-positive individuals can reactivate and 3) only women can enter and reactivate from the latent state.

Then, keeping the base assumption of latency, two different natural immunity processes were considered: 1) a proportion of individuals have lifelong natural immunity, and a proportion are immediately susceptible to new infection upon clearance and 2) all individuals have a lifelong reduced risk of re-infection. Due to time constraints on computing time, the model for each structure was calibrated to the HPV prevalence data using 120,000 simulations and the 100 best fitting (highest likelihoods) parameter combinations were used to simulate 100 cohorts. Table B 4 shows the unadjusted mean hazard ratios (and standard deviations) for both sexes in the general population for 1) HIV acquisition following detection of an oncogenic HPV type and 2) new detection of HPV by baseline HIV status.

Table B 4 - Mean hazard ratio (standard deviation) of 100 cohorts simulated using the best fitting parameter estimates for 5 different HPV natural history structures.

Model Structure	Effect of HPV infection at visit prior to HIV acquisition	Effect of HIV infection at study enrolment on new HPV detection
Waning immunity, reactivation for all (base model)	2.6 (0.3)	2.5 (0.2)
Waning immunity, no latency	3.6 (0.3)	2.2 (0.1)
Waning immunity, reactivation for HIV+ people	3.1 (0.3)	3.3 (0.3)
Waning immunity, latency and reactivation only for women	2.5 (0.2)	2.5 (0.2)
Proportion lifelong immunity, reactivation for all	2.2 (0.2)	2.5 (0.2)
Partial lifelong immunity, reactivation for all	2.3 (0.2)	2.7 (0.2)

Although the results vary for the different model structures, all the models show substantial relative risks of the one infection in the presence of the other, without biological transmission effects. This confirms that network level sexual behaviour effects can explain associations between the infections, regardless of the HPV natural history assumptions. Reducing or removing the extent of latency leads to stronger associations between HPV infection and HIV acquisition because the HPV infections are more likely to have been recently acquired in these scenarios and thus HPV infection is a stronger marker of recent risky behaviour. Removing latency from the base model leads to a weaker association between HIV infection and new HPV detection since HIV-positive individuals had higher reactivation rates than HIV-negative individuals. This association is however still greater than one.

B.2.2 Parameter uncertainty

To check sensitivity of the results due to parameter uncertainty, cohorts were simulated using 1) the medians of the best fitting HIV parameters from Johnson & Geffen (144) and the means of the sample from the posterior distributions of the HPV parameters. These parameter estimates were then increased or decreased by 50%, one at a time. Similarly, the degree to which individuals randomly select between high and low risk partners and the parameter that corrects for bias in reporting of condom use is increased and decreased by 50%. In each case, the mean estimate of association, with its standard deviation, is shown in Table B5. Note that only 10 cohorts were simulated for each change, since stochastic variation is the only source of variability in this case.

Table B 5 - Mean hazard ratio (standard deviation) of 10 cohorts simulated using the means of all the parameters and then halving or doubling the mean estimates for each parameter one at a time, but for all 13 HPV types simultaneously.

	Effect of HPV infection at visit prior to HIV acquisition		**Effect of HIV infection at study enrolment on new HPV detection**	
All means	2.4 (0.16)		2.5 (0.04)	
	50%	**150%**	**50%**	**150%**
HIV parameters				
Increased infectiousness during acute phase	2.2 (0.16)	2.4 (0.2)	2.4 (0.05)	2.5 (0.09)
Increased infectiousness during late phase	2.4 (0.17)	2.2 (0.13)	2.5 (0.03)	2.4 (0.07)
F-to-M transmission probability (non-spousal)	2.1 (0.33)	2.4 (0.14)	2.3 (0.06)	2.5 (0.07)
F-to-M transmission probability (spousal)	2.3 (0.13)	2.4 (0.14)	2.4 (0.06)	2.5 (0.08)
M-to-F transmission probability (non-spousal)	2.3 (0.16)	2.4 (0.15)	2.5 (0.05)	2.5 (0.07)
M-to-F transmission probability (spousal)	2.4 (0.07)	2.2 (0.13)	2.5 (0.04)	2.5 (0.06)
HPV parameters				
Transmission probabilities (per sex act)	2.7 (0.17)	2.3 (0.15)	2.6 (0.07)	2.4 (0.04)
Increase in HPV duration during acute/late HIV	2.3 (0.19)	2.5 (0.2)	2.4 (0.05)	2.4 (0.04)
Increase in HPV duration during latent HIV	2.4 (0.14)	2.4 (0.15)	2.4 (0.05)	2.4 (0.03)
Duration of viral latency	2.3 (0.21)	2.4 (0.23)	2.3 (0.04)	2.4 (0.06)
Proportion that progress to latency	2.6 (0.26)	2.1 (0.1)	2.1 (0.08)	2.8 (0.05)
Increase in reactivation of HPV during latent HIV	2.3 (0.16)	2.5 (0.14)	2.1 (0.04)	2.7 (0.06)
Increase in reactivation of HPV during acute/late HIV	2.4 (0.16)	2.4 (0.18)	2.2 (0.05)	2.6 (0.05)
Duration of natural immunity	2.6 (0.13)	2.3 (0.18)	2.5 (0.04)	2.4 (0.04)
Duration of HPV infection	2.7 (0.25)	2.3 (0.27)	3.2 (0.11)	2.2 (0.06)
Sexual Behaviour parameters				
Sexual mixing between risks group	2.6 (0.23)	2.1 (0.15)	2.7 (0.07)	2.3 (0.08)
Bias in reporting of condom use	2.3 (0.1)	2.2 (0.23)	2.5 (0.08)	2.5 (0.06)

Similar to the changes in model structure, changing the parameters that control the extent of and reactivation rate of viral latency leads to changes in the association between HIV status and new HPV detection. Reducing the proportion of individuals who acquire viral latency upon clearance of detectable infection leads to a weaker association and increasing this proportion leads to a stronger association, since confounding through reactivated infections influences the association less or more, respectively. Reduction in the duration of HPV infections leads to a much stronger association between HIV status and new HPV detection, presumably since individuals clear HPV infections at a higher rate (more so for HIV-positive individuals), i.e. more individuals enter the latent phase in the cohort time frame and reactivate infections (also at higher rate for HIV-positive individuals).

Increasing the degree of sexual mixing, which has the effect of reducing the heterogeneity in HIV acquisition risks and network effects, leads to weaker, but still significant associations.

B.2.3 Increased susceptibility and infectivity

Since our results without any co-factors match empirical studies well and we would like to avoid recalibration of the HIV parameters, we use a constant co-factor value along with parameter combinations from the calibrated model. Cohorts were simulated with 1) infectivity of HIV doubled in the presence of HPV infection; 2) susceptibility to HIV doubled in the presence of HPV; 3) infectivity of HPV doubled in the presence of HIV and 4) susceptibility to HPV doubled in the presence of HIV.

The choice of the co-factor value of two is based on meta-analyses (20,21) that estimated the effect of HPV infection on HIV acquisition. No meta-analysis[b] has been performed for the effect of HIV on new HPV detection, but we chose the value two since the estimated effect in the studies we included are close to two.

The mean of the unadjusted hazard ratios of HIV acquisition following detection of an HPV type calculated for each simulated cohort is 2.6 (95% CI 2.2–3.1) in the model without co-factors. When HIV infectivity is doubled in the presence of an HPV infection, the mean hazard ratio increases to 3.3 (95% CI 2.7-3.9) and to 4.6 (95% CI 3.8-5.4) when the susceptibility to HIV is doubled in the presence of an HPV infection. Figure B 1 shows that although the estimates of the models that include co-factors are higher than the model that does not, the estimates are in most cases still comparable to study estimates.

The mean of the unadjusted hazard ratios for the effect of HIV on newly detected HPV is 2.5 (95% CI 2.2–2.8) in the model without co-factors. When HPV infectivity is doubled in the presence of HIV, the mean hazard ratio is 2.5 (95% CI 2.2-2.9) and 2.6 (95% CI 2.3-3.0) when the susceptibility to HPV is doubled in the presence of HIV. Figure B 2 shows that the estimates of the models that include co-factors are very similar to those of the model that does not.

[b] A meta-analysis by Looker *et al.* was published after our analysis was completed (101).

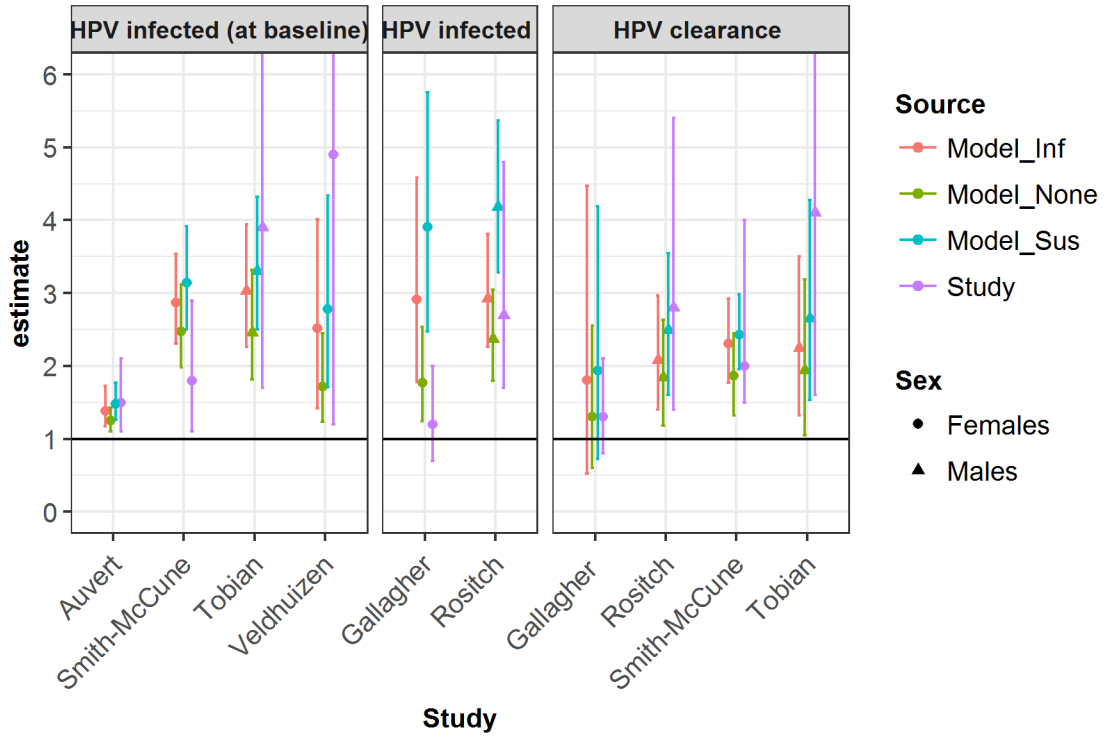


Figure B 1 - Associations between HPV status and HIV acquisition. Results shown are study estimates (Study), estimates using the model without co-factors (Model_None), estimates using the model where infectivity of HIV in presence of HPV was doubled (Model_Inf) and estimates using the model where susceptibility to HIV in presence of HPV was doubled (Model_Sus).

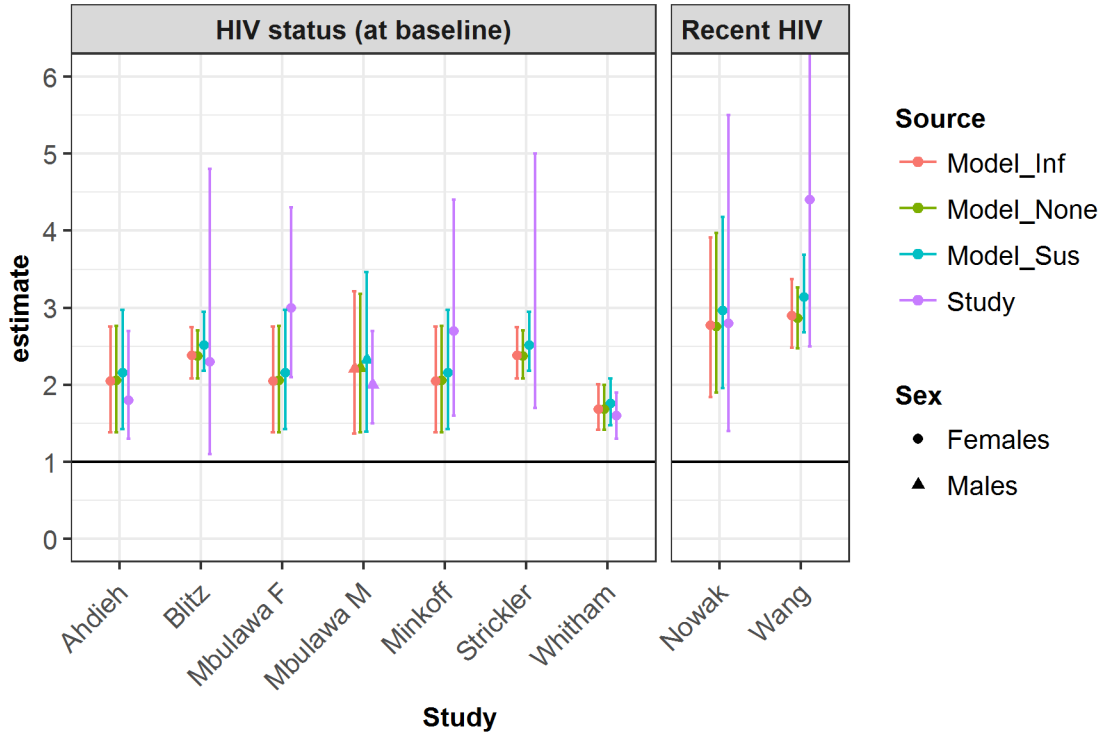


Figure B 2 - Associations between HIV status and new HPV detection. Results shown are study estimates (Study), estimates using the model without co-factors (Model_None), estimates using the model where infectivity of HPV in presence of HIV was doubled (Model_Inf) and estimates using the model where susceptibility to HPV in presence of HIV was doubled (Model_Sus).

Appendix C – Supplementary material to Chapter 4

C.1 Posterior distributions of HPV16/18 parameters

Table C 1 - HPV 16: the mean and interquartile range of the 500 samples from the posterior distributions for each parameter, for each model structure.

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transmission probability (per sex act)						
Male to Female	0.6 (0.31-0.77)	0.4 (0.16-0.67)	0.47 (0.18-0.74)	0.46 (0.23-0.72)	0.48 (0.25-0.7)	0.47 (0.25-0.67)
Female to Male	0.09 (0.04-0.25)	0.06 (0.03-0.19)	0.06 (0.03-0.1)	0.07 (0.04-0.21)	0.09 (0.04-0.22)	0.09 (0.04-0.28)
Relative HPV duration in HIV infection						
Latent HIV vs HIV-negative	2.2 (1.8-2.6)	2.5 (2.2-2.7)	2.2 (1.8-2.6)	2.2 (1.7-2.6)	2.1 (1.6-2.5)	2 (1.6-2.6)
Acute HIV/late HIV/recent ART* vs HIV-negative	1.3 (1.1-1.6)	1.5 (1.2-1.8)	1.3 (1.1-1.6)	1.4 (1.2-1.7)	1.3 (1.1-1.5)	1.3 (1.1-1.6)
Time to reactivation (in years) if HIV-negative						
Males	17.5 (9.4-23.2)		18.8 (11.1-23.2)		17.3 (11.2-23.1)	18.6 (13-24)
Females	19.4 (13.9-24.1)		20.1 (16.1-25)	19.4 (13.6-25)	21.5 (17.1-26.6)	22.5 (17.7-26.9)
Proportion who become latently infected after clearance	0.56 (0.42-0.74)		0.37 (0.23-0.51)	0.67 (0.47-0.82)	0.67 (0.52-0.81)	0.73 (0.56-0.87)
Relative HPV reactivation rate						
Latent HIV vs HIV-negative	2.4 (1.8-3.4)		2.1 (1.7-3.1)	2.3 (1.6-3.5)	2.3 (1.6-3.1)	1.8 (1.4-2.2)
Acute HIV/late HIV/recent ART* vs latent HIV	2.1 (1.6-2.5)		1.9 (1.4-2.6)	2 (1.6-2.5)	2.1 (1.6-2.6)	2 (1.6-2.5)
Duration of immunity (in years)						
Males	16.6 (9.3-22.9)	11.2 (4.9-17.1)	15.2 (9.8-23.1)	12.5 (6.8-20.2)	16.3 (7.9-22.8)	16.1 (6.6-21.6)
Females	15.7 (9.4-22.3)	8.3 (5-13.6)	14.6 (8.6-20.9)	17 (10.7-23.2)	13.7 (6.3-22.1)	14.5 (6.2-23)
Duration of HPV infection (in months) if HIV-negative						
Males	12.1 (8.9-16.1)	17 (11.9-23.1)	15.8 (12.6-19.1)	17 (11.9-22)	10.8 (8.3-14.7)	10.9 (7.5-15.3)
Females	11.4 (9.9-13.4)	14 (10.7-17.4)	15.8 (13.6-18.3)	11.6 (9.9-13.5)	11.1 (9.3-13)	9.8 (8.3-11.8)
Standard deviation of study effect	0.16 (0.11-0.2)	0.19 (0.15-0.24)	0.16 (0.13-0.2)	0.16 (0.11-0.21)	0.18 (0.12-0.22)	0.15 (0.11-0.2)

*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 have the same HPV duration as HIV-negative people.

Figure C 1 - HPV 16: Prior and posterior distributions of parameters

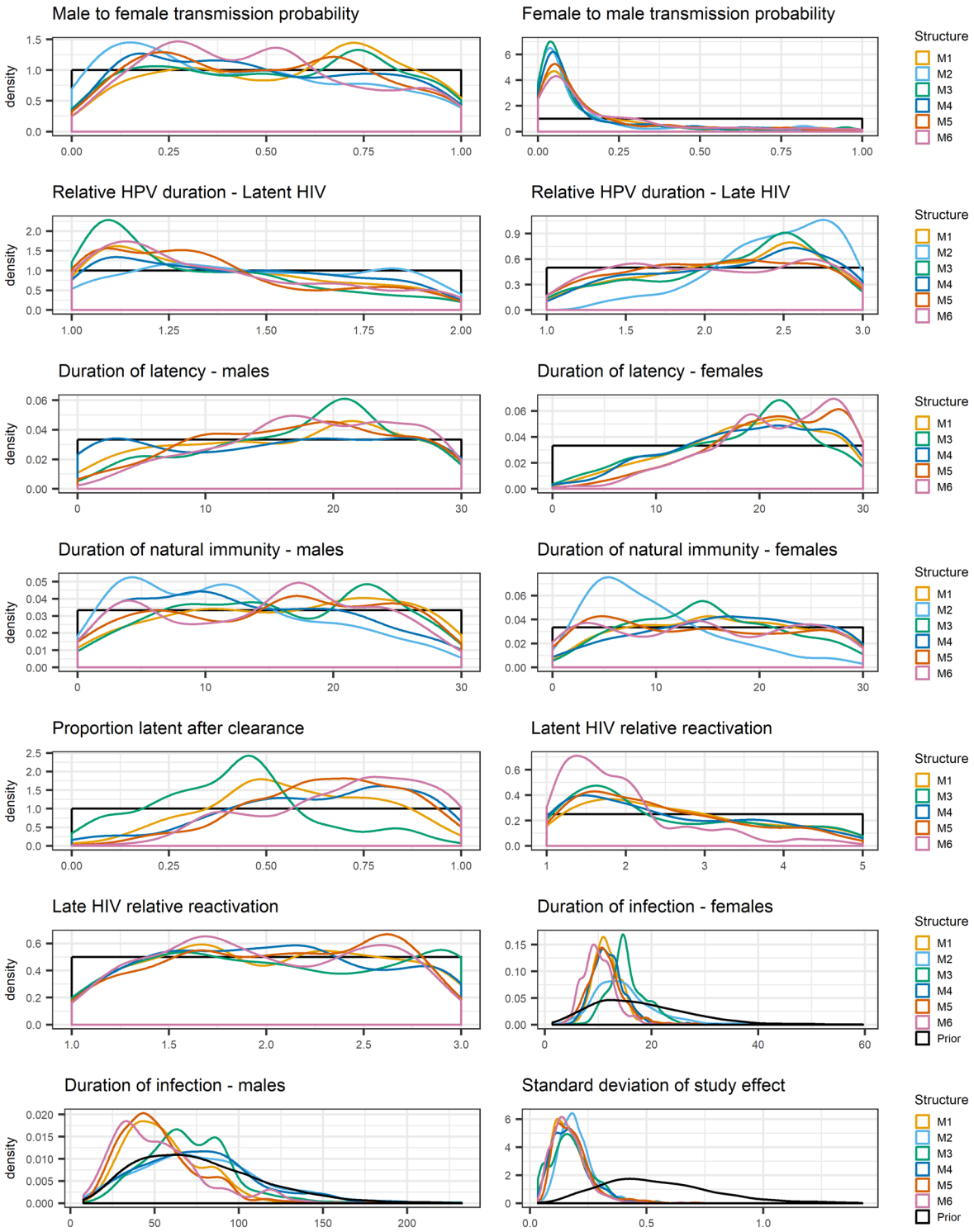
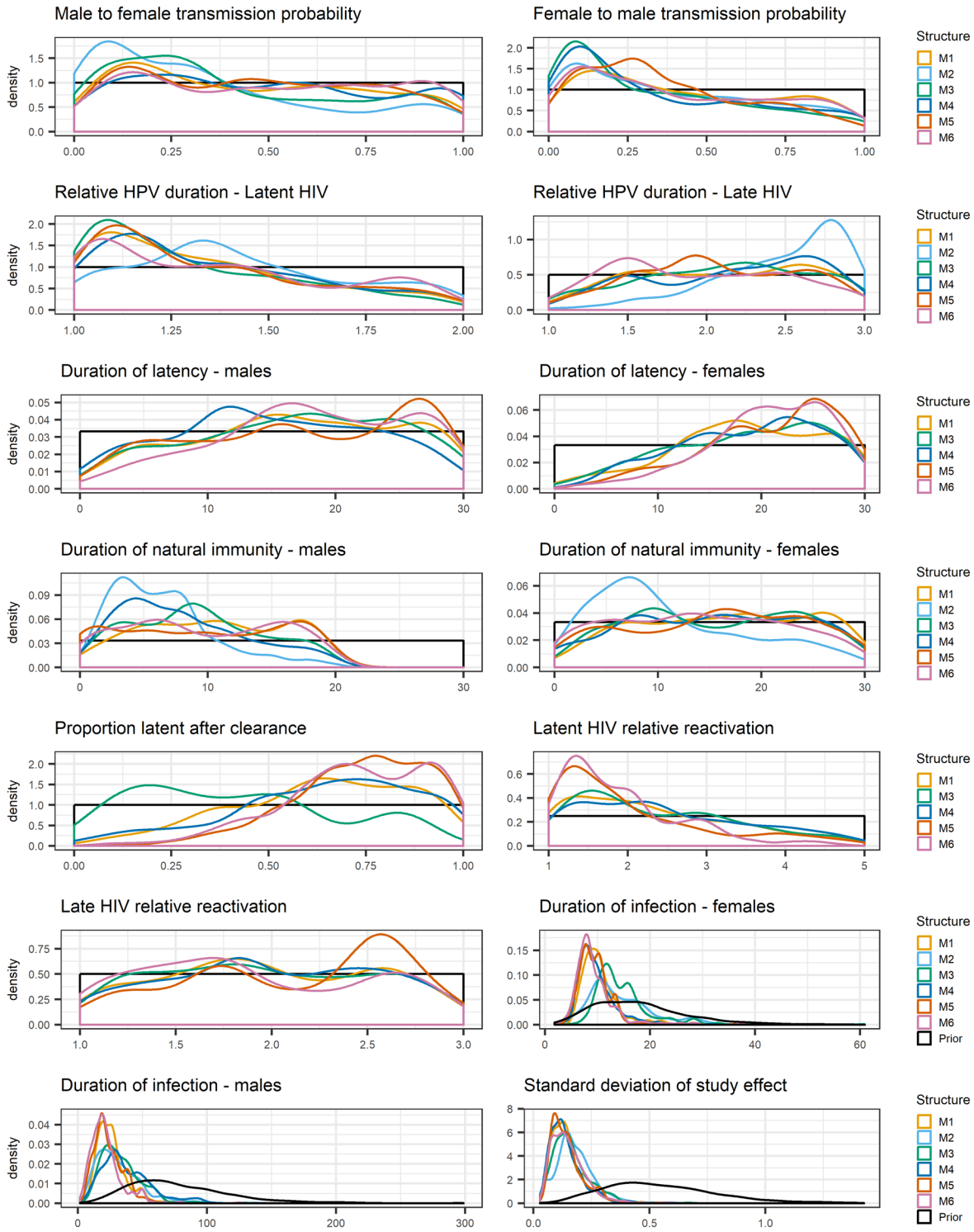


Table C 2 - HPV 18: the mean and interquartile range of the 500 samples from the posterior distributions for each parameter, for each model structure.

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transmission probability (per sex act)						
Male to Female	0.4 (0.18-0.7)	0.29 (0.09-0.51)	0.38 (0.14-0.75)	0.46 (0.23-0.74)	0.44 (0.19-0.71)	0.5 (0.2-0.78)
Female to Male	0.38 (0.18-0.68)	0.32 (0.1-0.62)	0.3 (0.1-0.48)	0.26 (0.09-0.57)	0.3 (0.16-0.54)	0.33 (0.13-0.67)
Relative HPV duration in HIV infection						
Latent HIV vs HIV-negative	2.1 (1.6-2.6)	2.6 (2.2-2.8)	2.3 (1.9-2.7)	2.2 (1.7-2.6)	2 (1.7-2.5)	1.9 (1.5-2.4)
Acute HIV/late HIV/recent ART* vs HIV-negative	1.3 (1.1-1.5)	1.4 (1.2-1.6)	1.2 (1.1-1.4)	1.3 (1.1-1.6)	1.3 (1.1-1.5)	1.3 (1.1-1.6)
Time to reactivation (in years) if HIV-negative						
Males	16.5 (10.8-23.7)		16.2 (10.8-22)		16.7 (10.1-26.1)	17.8 (13.2-24.9)
Females	18.5 (13.6-24.5)		19.4 (13.2-25.8)	19.6 (14.2-24.2)	22.1 (16.8-25.7)	21.3 (17.4-25.2)
Proportion who become latently infected after clearance	0.65 (0.48-0.81)		0.33 (0.18-0.51)	0.67 (0.49-0.82)	0.76 (0.63-0.87)	0.74 (0.59-0.88)
Relative HPV reactivation rate						
Latent HIV vs HIV-negative	2.1 (1.5-3)		2.2 (1.7-3.2)	2.3 (1.6-3.2)	1.7 (1.3-2.5)	1.7 (1.4-2.2)
Acute HIV/late HIV/recent ART* vs latent HIV	1.9 (1.6-2.5)		2.1 (1.5-2.6)	2 (1.6-2.5)	2.2 (1.7-2.6)	1.8 (1.4-2.5)
Duration of immunity (in years)						
Males	10.5 (5.8-15.6)	5.8 (3.4-8.2)	11.5 (5.5-18.5)	7.5 (4.1-11.2)	9.6 (4.3-15.7)	9.1 (5-15.4)
Females	17.5 (9.5-24.2)	9.1 (5.7-16.7)	16 (8.7-22.2)	15.8 (8.2-22.7)	15.5 (8.9-23)	14.1 (7-21)
Duration of HPV infection (in months) if HIV-negative						
Males	5.6 (4.1-7)	6.2 (4.1-9.2)	6.5 (4.4-8.9)	7.4 (5.6-10.7)	4.7 (3.7-6.8)	4.5 (3.2-5.9)
Females	9.7 (8.2-11.7)	12.9 (10.4-17.6)	12.9 (11.1-15.3)	8.9 (7.5-11.2)	9.2 (7.5-10.7)	8.4 (7.1-10.7)
Standard deviation of study effect	0.13 (0.1-0.18)	0.17 (0.13-0.22)	0.15 (0.12-0.19)	0.12 (0.09-0.17)	0.15 (0.09-0.17)	0.13 (0.09-0.19)

*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.

Figure C 2 - HPV 18: Prior and posterior distributions of parameters



C.2 Model fits to prevalence data

Table C 3 shows the means of the 500 log-likelihoods generated by the posterior sample. Models 1 and Model 4 have the highest mean log-likelihood and Model 2 has the lowest, indicating that models with latency fit better to the data.

Table C 3 - Mean log-likelihood of 500 samples from posterior distributions

	Type 16	**Type 18**
Model 1	-2794.4	-2020.42
Model 2	-2796	-2022.54
Model 3	-2795.14	-2022.48
Model 4	-2794.57	-2020.34
Model 5	-2795.02	-2020.58
Model 6	-2794.45	-2020.65

In Figures C 3 and C 4, the fits to HPV-16 and -18 data are shown for each of the six model structures. The mean prevalence estimate produced by the sample of 500 parameter combinations from the posterior distributions is shown, along with the 2.5th and 97.5th percentiles of the prevalence estimates. The number of each box represents the study number as shown in the first column of Table A 10. Although the models do not always match the data closely, this is to some extent because of inter-regional and inter-study variation in HPV prevalence, which is allowed for through the inclusion of random effect terms in the specification of the likelihood function. It is also worth noting that no model structure stands out as being clearly preferable to the others in terms of overall goodness of fit criteria.

*Figure C 3 - **Model fits to Type 16 data**. The solid horizontal lines represent the point estimate of prevalence from each study, while the dashed horizontal lines represent the confidence intervals.*

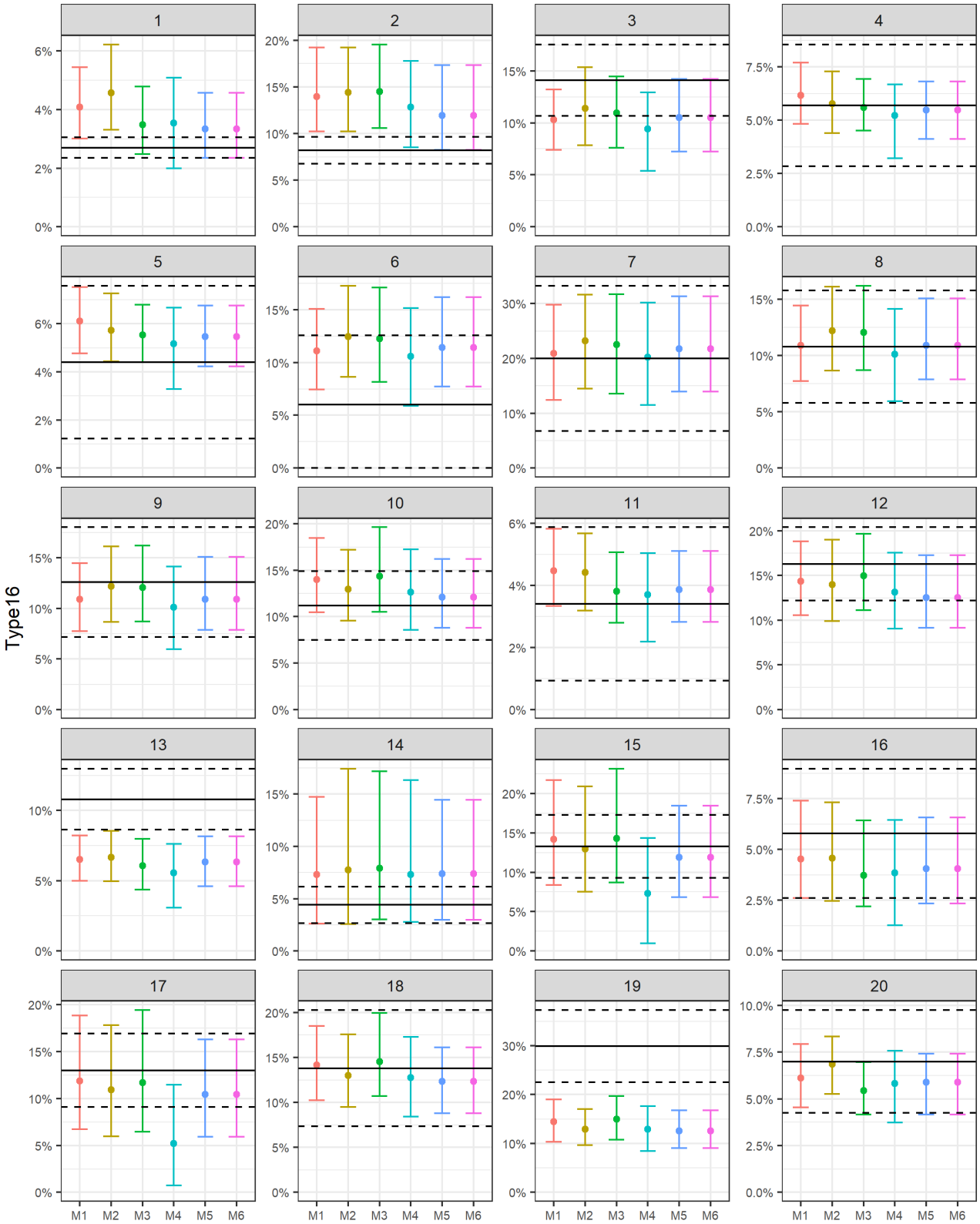


*Figure C 4 - **Model fits to Type 18 data.** The solid horizontal lines represent the point estimate of prevalence from each study, while the dashed horizontal lines represent the confidence intervals.*

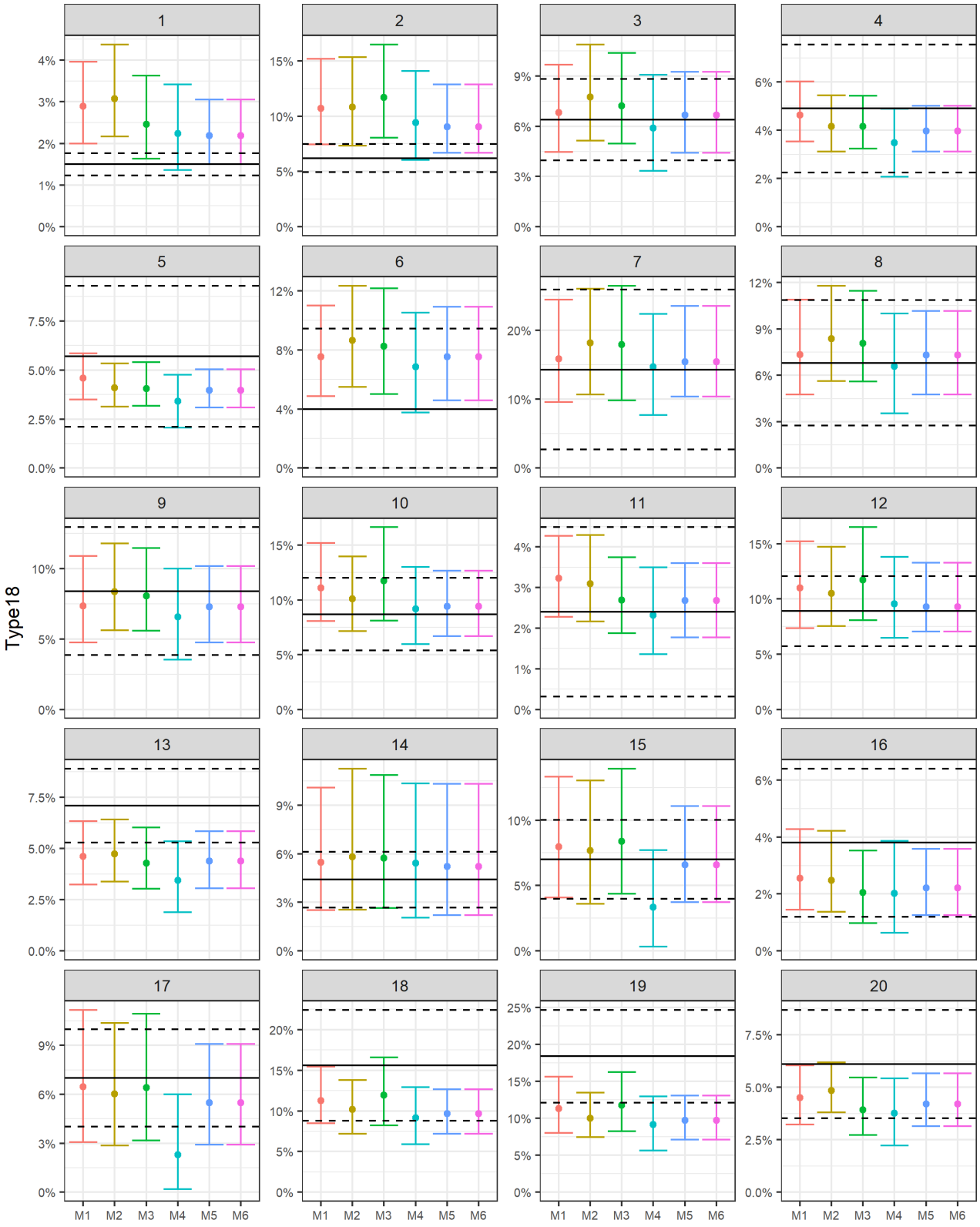
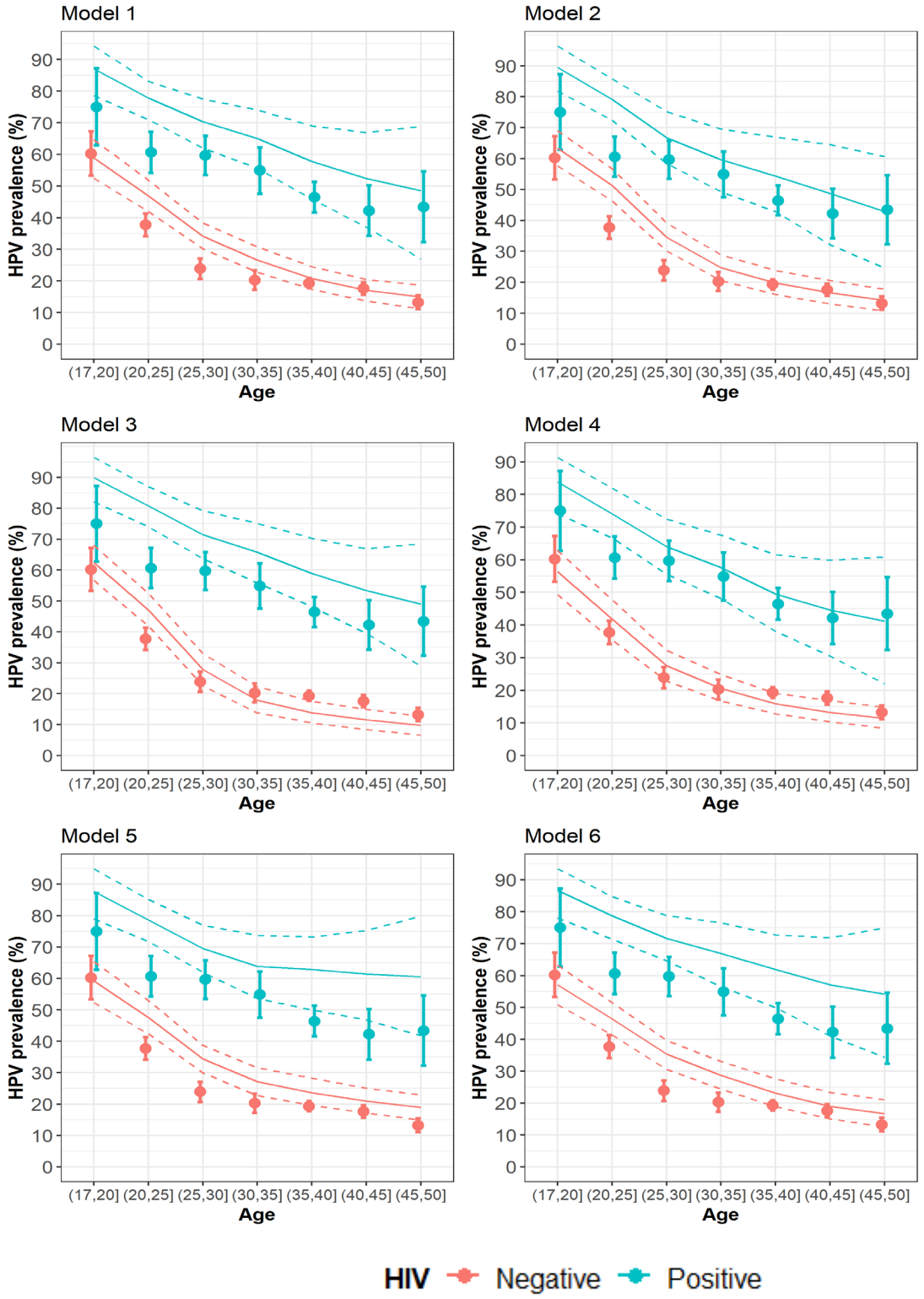


Figure C 5 - The mean overall high risk HPV prevalence in 2000 using the sample of 500 parameter combinations from the posterior distributions of each model structure (solid lines). The confidence bands (2.5th and 97.5th percentiles) around these estimates are shown in dashed lines. Data points (in closed circles) represent HPV prevalence results from a population level study in Khayelitsha, Cape Town (273)*. Type specific HPV prevalence in this study (for all ages combined) was used to calibrate the model (Table A 10). In this figure HPV types are aggregated, but age groups are disaggregated. The figure therefore serves as a validation of the modelled age pattern of overall oncogenic HPV prevalence.*



C.3 Simulation of RCTs

In this individual-based model, we can "survey" characteristics of each individual in any week between 1985 and the end of the model run. To simulate the RCTs, we vaccinate in the first week of 2014, at random, half of the 15-25 year old women in the model or a quarter of the women older than 25. We enrol the women in our model in the first week of 2014. The women who are included in the analysis match the following characteristics of the women enrolled in the empirical RCTs.

In the 15-25 analysis:

- HPV status: HPV-16 and -18 DNA negative at enrolment
- HIV status: no one HIV-positive at enrolment
- Lifetime partners: six or less (in additional analyses, this exclusion was not applied)

In the 25+ analysis:

- HPV status: 15% of women could have evidence of previous HPV 16/18 (DNA- or seropositive)
- HIV status: no one HIV-positive at enrolment
- Age: Only 10% of women in the analysed sample are older than 45
- Lifetime partners: 75% of sample had less than 6 (in additional analyses, this exclusion was not applied)

Women's characteristics are saved every six months from January 2014 for four years, resulting in longitudinal datasets similar to those produced by empirical studies. The saved longitudinal variables are: the model identification number, age, type-specific HPV status (13 types), HIV status, total lifetime number of sexual partners (LTP) and vaccination status. We simulate 500 datasets for women aged 15-25 using the 500 parameter combinations from the posterior distributions and then 500 datasets for women older than 25, vaccinating only the group of interest for each set of simulations.

C.4 Estimated HPV-16 and -18 prevalence in 2014

Table C 4 - Mean, 2.5th and 97.5th percentiles of HPV-16 and -18 prevalence (ages 15+) in 2014

	Type 16		Type 18	
	Female	Male	Female	Male
Model1	5.7 (4.5-7.3)	5.4 (3.2-8.6)	4.3 (3.3-5.6)	3.1 (1.7-4.9)
Model2	5.4 (4.0-6.8)	5.2 (2.9-8.0)	3.8 (2.8-5.1)	2.8 (1.5-4.7)
Model3	5.1 (4.1-6.4)	4.6 (2.8-7.6)	3.7 (2.8-4.8)	2.5 (1.3-4.3)
Model4	5.6 (4.3-6.8)	5.5 (3.3-8.4)	4.1 (3.1-5.3)	3.2 (1.8-5.2)
Model5	5.2 (4.0-6.5)	5.0 (3.0-7.6)	3.8 (2.8-4.8)	2.8 (1.5-4.3)
Model6	5.9 (4.6-7.3)	5.7 (3.3-8.8)	4.3 (3.2-5.7)	3.0 (1.8-5.0)

C.5 Long-term vaccine impact sensitivity analyses

We perform three additional analyses to confirm the robustness of our conclusion. In our base model, we assume 100% lifelong prophylactic efficacy, for those who stay HIV-negative and those who become HIV-positive. We assess long-term impact of the vaccine for three independent changes to the base model:

- 95% of women vaccinated have lifelong protection and the other 5% have no protection (Figure C 6).
- 100% of women are protected for a period that is randomly drawn from a Weibull distribution with scale of 30 years and shape such that 95% of women are still protected after 10 years (Figure C 7).
- 20% of women lose protection following HIV seroconversion (Figure C 8).

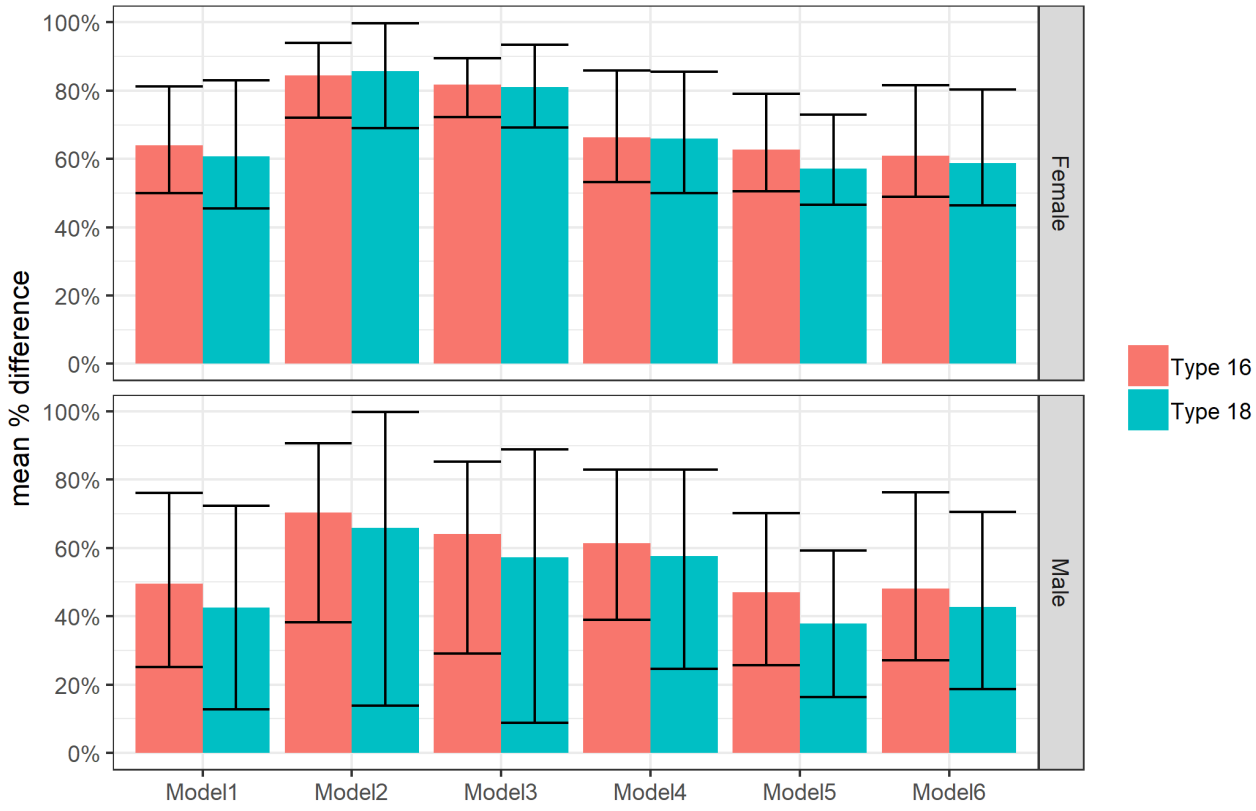


Figure C 6 - Mean percentage reduction in HPV16/18 prevalence in 2045 for individuals aged 15+ with lifelong protection for 95% of vaccinated women. Vaccination coverage of 9-year-old girls is assumed constant between 2014 and 2045 at 90%.

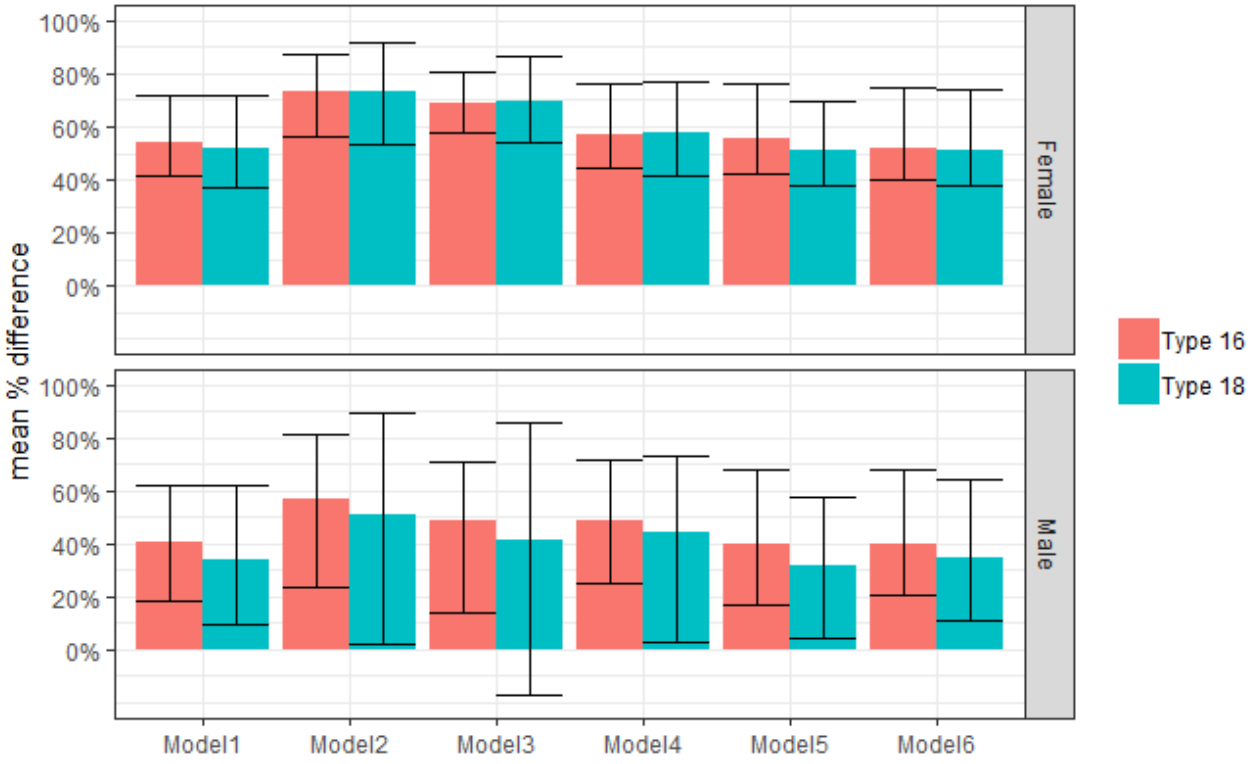


Figure C 7 - Mean percentage reduction in HPV16/18 prevalence in 2045 for individuals aged 15+ with prophylactic vaccine efficacy of 100% for a period that is randomly drawn from a Weibull distribution. Vaccination coverage of 9-year-old girls is assumed constant between 2014 and 2045 at 90%.

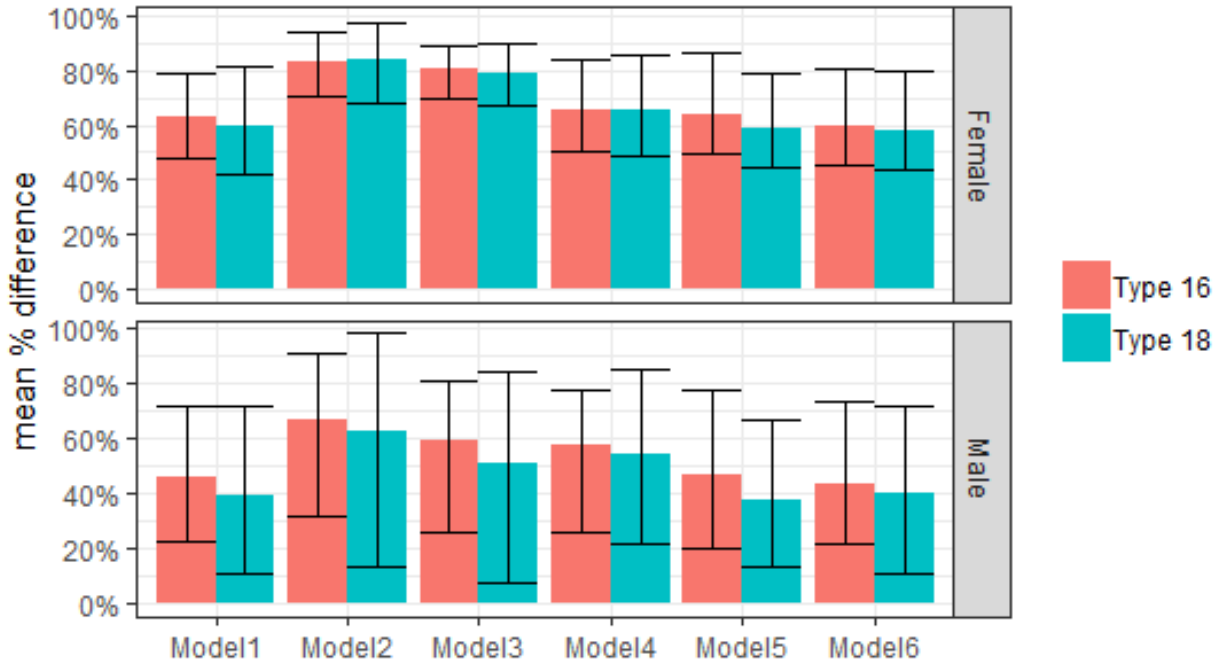


Figure C 8 - Mean percentage reduction in HPV16/18 prevalence in 2045 for individuals aged 15+ with lifelong prophylactic vaccine efficacy of 100%. Twenty percent of individuals lose protection after becoming HIV infected. Vaccination coverage of 9-year-old girls is assumed constant between 2014 and 2045 at 90%.

C.6 Natural history sensitivity analysis

Tables C5 and 6 show that this chapter's findings relating to vaccine efficacy are robust to changes in the model's structure, from simulating on HPV infection stages, to simulating these stages and cervical disease stages.

Table C5 - Vaccine effectiveness against persistent HPV 16 or 18 infection among women aged 15-25. Shown here are results for Models 1 and 2, comparing natural history structures with and without cervical disease. Mean effectiveness among the 500 simulated trials is shown, along with the 2.5th and 97.5th percentiles.

Model Structure		<=6 Lifetime partners			No limit on number of lifetime partners		
		m-TVC	n-TVC*	n-TVC**	m-TVC	n-TVC*	n-TVC**
	Kreimer (66)	89.1 (86.8;91.0)	93.6 (91.2;95.5)	93.6 (91.2;95.5)	89.1 (86.8;91.0)	93.6 (91.2;95.5)	93.6 (91.2;95.5)
	Harper (173)		96.0 (75.2;99.9)	96.0 (75.2;99.9)		96.0 (75.2;99.9)	96.0 (75.2;99.9)
100% prophylactic efficacy against HPV16/18 infection							
Only HPV	Model 1	95.9 (90.1;99.3)	97.2 (91.7;100)	98.5 (94.6;100)	86 (72.5;95)	90 (78.3;97.7)	94.4 (86.2;99.2)
HPV and CC		96 (90.8,99.4)	97.5 (93.4,100)	98.6 (95.1,100)	87.5 (76.7,95.4)	91.8 (82,97.7)	95.5 (89.4,99.2)
Only HPV	Model 2	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)
HPV and CC		100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)	100 (100;100)
95% prophylactic efficacy against HPV16/18 infection							
Only HPV	Model 1	91.3 (85.4;96.5)	92.6 (86.5;98)	93.9 (88.5;98.3)	81.9 (68.6;91.4)	85.7 (74.8;94.2)	89.9 (82.1;96.3)
HPV and CC		91.5 (84.5,96.3)	92.9 (86.2,97.3)	94 (88.5,98.1)	83.4 (73.2,91.8)	87.5 (76.9,94.8)	91.1 (83.5,96.7)
Only HPV	Model 2	95.5 (91.1;98.8)	95.5 (90.8;99.2)	95.5 (90.8;99.2)	95.4 (92.5;97.9)	95.4 (90.7;98.9)	95.4 (90.7;98.9)
HPV and CC		95.4 (91.0,98.4)	95.3 (90.9,99.3)	95.3 (90.9,99.3)	95.3 (92.4,97.9)	95.3 (91.6,98.9)	95.3 (91.6,98.9)

Table C6 - Vaccine effectiveness against persistent HPV 16 or 18 infection among women aged 25 and older. Shown here are results for Models 1 and 2, comparing natural history structures with and without cervical disease. Mean effectiveness among the 500 simulated trials is shown, along with the 2.5th and 97.5th percentiles.

Model Structure		Matching LTP distribution			No limit on number of LTP		
		TVC	n-TVC*	n-TVC**	TVC	n-TVC*	n-TVC**
	Skinner (74)	47.0 (25.4;62.7)	82.9 (53.8;95.1)	82.9 (53.8;95.1)	47.0 (25.4;62.7)	82.9 (53.8;95.1)	82.9 (53.8;95.1)
100% prophylactic efficacy against HPV16/18							
Only HPV	Model 1	48.4 (-16.2;88.2)	56.3 (-21.4;100)	72.4 (-5.3;100)	40.2 (13.4;66.3)	46.9 (14.0;78.1)	62.6 (22.6;88.1)
HPV and CC		45.8 (-47.3,92)	59.1 (-35.9,100)	74.1 (-4,100)	35.7 (6.7,62.1)	46.4 (12.6,77.1)	62.7 (23.1,89)
Only HPV	Model 2	80.2 (21.4;100)	100 (100;100)	100 (100;100)	76.3 (57.2;90.3)	100 (100;100)	100 (100;100)
HPV and CC		72.5 (1.5,100.0)	100 (100,100)	100 (100,100)	68.8 (46.7, 85.4)	100 (100,100)	100 (100,100)
95% prophylactic efficacy against HPV16/18							
Only HPV	Model 1	46.4 (-19.1;87.6)	53.5 (-24.5;100)	68.6 (-16.3;100)	38.4 (12.2;64.7)	44.8 (11.5;74.1)	59.7 (18.5;85.3)
HPV and CC		43 (-56.6,91.1)	56.2 (-44.6,100)	69.3 (-30.8,100)	33.9 (4.9,59.9)	44.3 (10.7,73.6)	59.9 (20.3,87.1)
Only HPV	Model 2	76.8 (18.9;100)	95.9 (63.4;100)	95.9 (63.4;100)	73.0 (54.1;87.6)	95.6 (86.2;100)	95.6 (86.2;100)
HPV and CC		68.7 (-8.2,100)	94.2 (47.1,100)	94.2 (47.1,100)	65.3 (41.4,83.1)	94.8 (79.3,100)	94.8 (79.3,100)

Appendix D – Supplementary material to Chapter 5

D.1 Additional screening information

Table D 1 - Total number of Pap smears performed and cytological results by year. The fraction of smears with no result changed substantially after implementation of the Bethesda 2014 classification system, in which smears without an endocervical component is defined as adequate. Also shown in this table are the fractions of women who attended screening services because of symptoms (i.e., not routine) and the fractions of smears that could not be linked to a PMI.

Year	Total	Symptomatic screening	Not linked to PMI	Normal	ASCUS/LSIL*	ASCH/HSIL*	Cancer	Inadequate
2007	74476	32.1%	12.4%	51424 (69%)	12100 (16.2%)	3153 (4.2%)	149 (0.2%)	7650 (10.3%)
2008	95482	32%	10.5%	67022 (70.2%)	13092 (13.7%)	3842 (4%)	171 (0.2%)	11355 (11.9%)
2009	102984	31%	9.4%	72128 (70%)	14289 (13.9%)	4533 (4.4%)	163 (0.2%)	11871 (11.5%)
2010	117710	31.1%	6.3%	83171 (70.7%)	16645 (14.1%)	4851 (4.1%)	160 (0.1%)	12883 (10.9%)
2011	120570	30.9%	5.6%	86757 (72%)	15563 (12.9%)	4513 (3.7%)	138 (0.1%)	13599 (11.3%)
2012	120400	30.5%	5%	86036 (71.5%)	14453 (12%)	5047 (4.2%)	156 (0.1%)	14708 (12.2%)
2013	130043	29.5%	3.2%	92733 (71.3%)	15287 (11.8%)	5540 (4.3%)	176 (0.1%)	16307 (12.5%)
2014	132902	29%	2.6%	96324 (72.5%)	15176 (11.4%)	5570 (4.2%)	190 (0.1%)	15642 (11.8%)
2015	128246	24.7%	2.8%	99928 (77.9%)	13698 (10.7%)	5291 (4.1%)	200 (0.2%)	9129 (7.1%)
2016	137165	23.7%	2.3%	115009 (83.8%)	12605 (9.2%)	5872 (4.3%)	167 (0.1%)	3512 (2.6%)
2017	131233	21.7%	1.6%	109000 (83.1%)	11057 (8.4%)	5654 (4.3%)	130 (0.1%)	5392 (4.1%)
2018	128988	21.4%	1.3%	107409 (83.3%)	11190 (8.7%)	7004 (5.4%)	212 (0.2%)	3173 (2.5%)

*ASCUS - Atypical squamous cells of undetermined significance; LSIL – low grade squamous intraepithelial lesions; HSIL – high grade squamous intraepithelial lesions; ASCH - Atypical squamous cells, HSIL cannot be ruled out.

D.2 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 (278). South African HIV estimates for the official UNAIDS reports have been based on Thembisa output since 2017 (279). For the analyses in Chapter 5 we use population and HIV estimates from the Western Cape and for the analyses in Chapter 6 we use national level estimates. At the time we performed these analyses, Thembisa version 4.2 was the latest available version.

The model population is stratified by sex and 5-year age groups and stratification for sexual behaviour is similar to that of MicroCOSM (Appendix A.2), 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.

Appendix E – Supplementary material to Chapter 6

E.1 Age standardisation

We use two world- standard populations in this paper. 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 (2,41). To be consistent with the Brisson *et al.* paper, we age-standardise according to the United Nations Development Programme's 2015 (UNDP 2015) world population (11). 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).

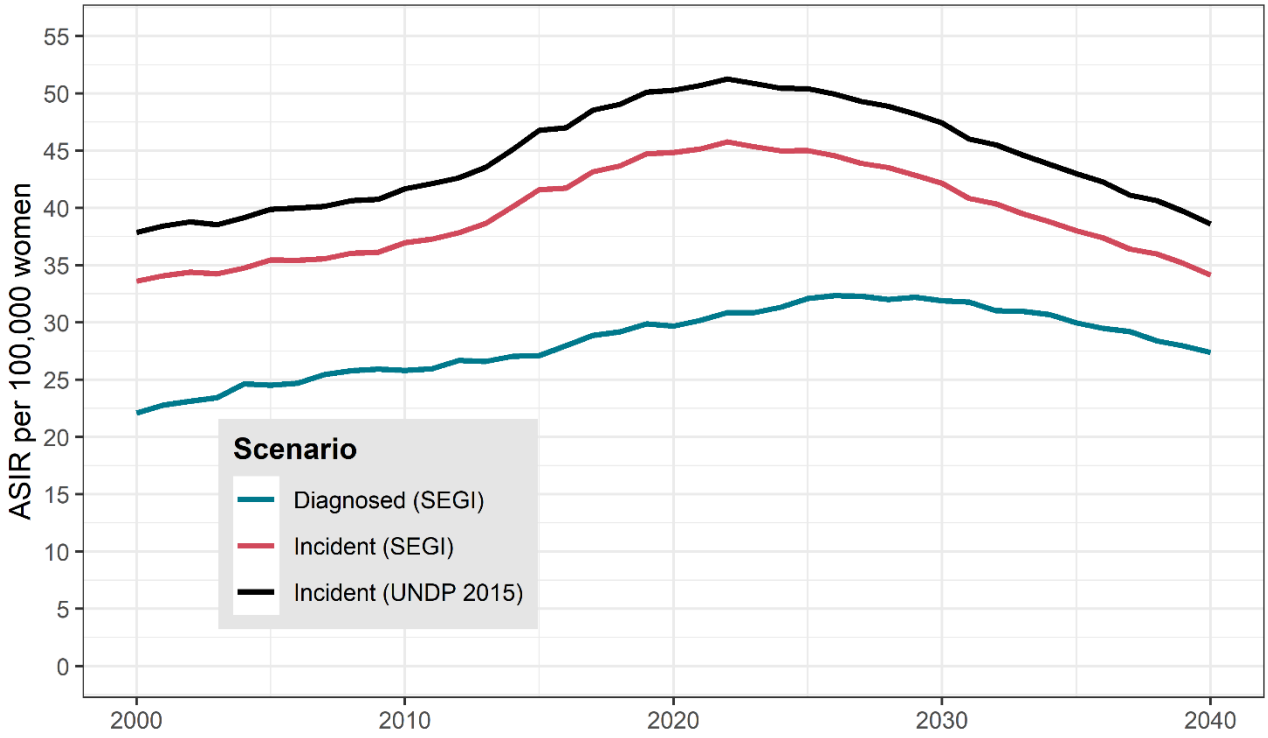


Figure E 1 – Diagnosed and incident cervical cancer calculated using different world standard populations.

E.2 Additional screening assumptions

When simulating the screening strategies suggested by the WHO's CCEMC (Scenarios G) to I) in Table 7), we need to make extra assumptions regarding screening. If our baseline screening is abruptly stopped in 2020 and replaced by an HPV-DNA screen at ages 35/45 or 3-yearly for HIV-positive women between the ages of 25 and 50, cervical cancer incidence may initially increase in those age groups not receiving Pap smear screening any more, and not receiving HPV-DNA based screening. We therefore follow the following steps:

E.2.1 Twice in a lifetime screening with HPV-DNA test

This scenario suggests one screen for all women aged 35 and one screen for all women aged 45. We assume that one of these two HPV-DNA screens can happen at any time between ages 30 and 40 and the other one at any time between ages 40 and 50. Women younger than 30 can still enrol in Pap smear screening at the same rates as the status quo, but won't receive any Pap smears after 30. Women aged 50-60 can still receive Pap smear screening at the same rate as the status quo, but this rate linearly decreases to zero over ten years, by which time all women who had the opportunity to receive an HPV-DNA test will be aged 50-60. Rates of Pap smear screening are dependent on ART status and age, as described in Appendix A.4.2. Initial rates of HPV-DNA screening are chosen so that coverage of screening among women aged 30-50 (by ART status) does not reduce after 2020, and these rates linearly increase to 70% coverage in 2030 and 90% in 2045.

E.2.2 Extra screening for HIV-positive women

This scenario suggests one screen for HIV-negative women aged 35; one screen for HIV-negative women aged 45 and three-yearly screens for HIV-positive women aged 25 to 50. Our assumptions for HIV-negative women are the same as in the previous scenario. HIV-positive women aged between 15 and 25 can receive Pap smear screening at the same rates as the status quo, but won't receive any Pap smears after 25. HIV-positive women aged 50-60 can still receive Pap smear screening at the same rate as the status quo, but this rate linearly decreases to zero over ten years, by which time all women who had the opportunity to receive an HPV-DNA test will be aged 50-60.

E.3 Sensitivity analyses

E.3.1 Fraction who receives only clinical diagnosis

As described in Appendix A.5.3.1 and A.8.3, data on the fraction of cervical cancer cases who only receive a clinical diagnosis (no pathology) are inconsistent and scarce, and our model fits only marginally better to the assumption that 10% do not receive a pathological diagnosis, compared to 7% or 14%. In Figure E2 we show that although the 3 scenarios lead to different estimates of CC ASIR, the impact of scaling up our current prevention programme will be similar across the three scenarios. Under all three assumptions, CC ASIR will reduce by 76% between 2019 and 2120 if vaccination, screening, and linkage to treatment are maintained at current levels, and with all three assumptions CC ASIR will reduce by 85% if the 90-70-90 targets are met by 2030.

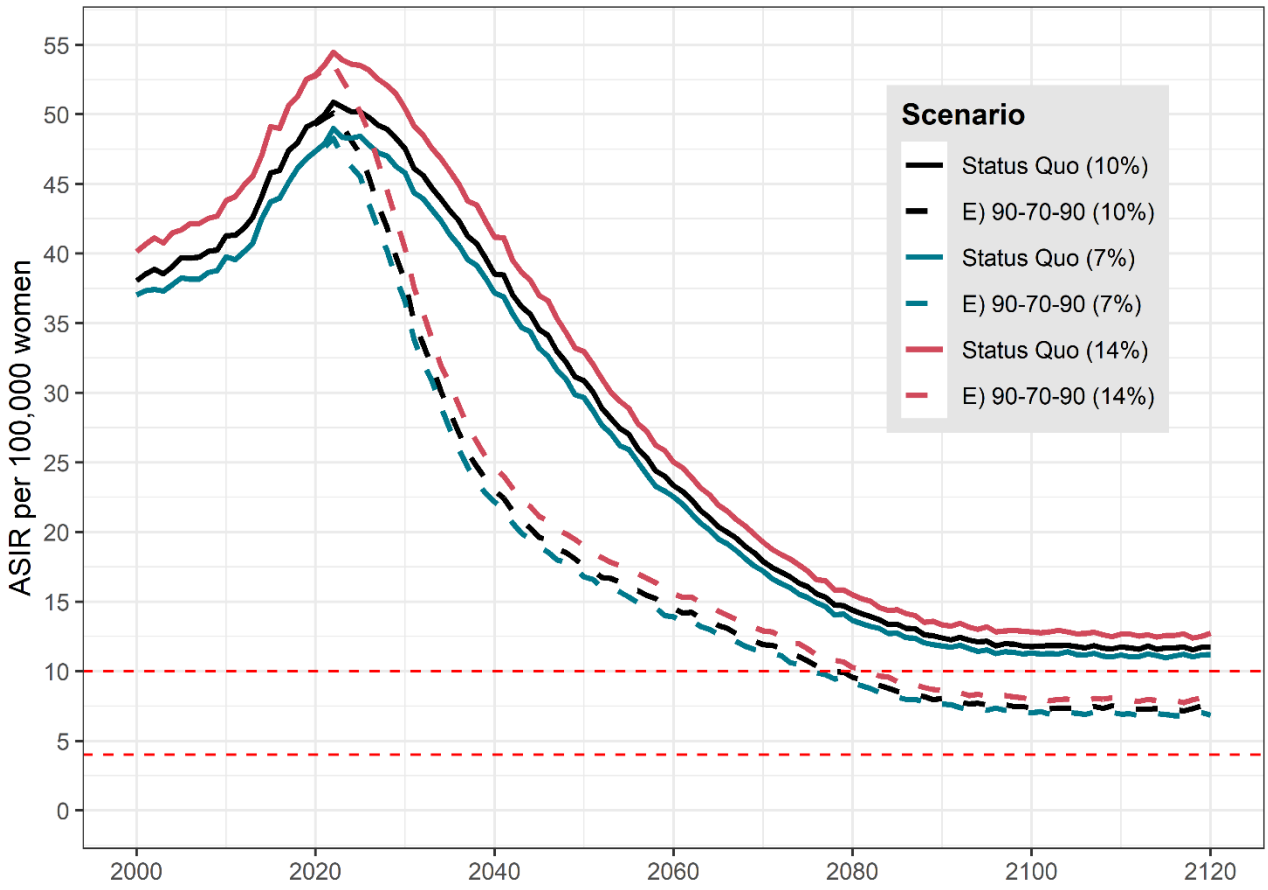


Figure E 2 – Mean model estimates of age-standardised cervical cancer incidence assuming three different fractions of clinically diagnosed cases. Estimates are shown for the status quo (current levels of vaccination, screening and linkage to treatment) and the scenario where our current prevention programme is scaled-up to meet the 90-70-90 targets by 2030.

E.3.2 Vaccine efficacy

In the main analysis of Chapter 6, we assume that all vaccinated women (regardless of the number of doses) receive 100% life-long protection against infection with HPV types 16 and 18, as well as against types 31/33/45 in 50% of women. Although there is currently no evidence that suggests lower efficacy or waning effectiveness of a one-dose schedule (66,203,204), we show here projections for the worst case scenario: that women who receive only one dose will have no protection against infection. We assume that one-dose coverage of 9-year-old girls will remain at 80%, and two-dose coverage at 60%. In Figure E4, "80% vaccinated" implies 80% coverage of one dose that gives 100% lifelong protection and "60% vaccinated" implies 60% coverage of two doses that give 100% lifelong protection, and that women with only one dose receive no protection.

In the best-case scenario (80% protected lifelong and some cross-protection), mean CC ASIR is estimated to be 11.7 per 100,000 women in 2120 (95% CI 7.8-16.8). In the worst-case scenario (assuming those with only one dose have zero protection, and no cross-protection), mean CC ASIR is estimated to be 16.7 per 100,000 women (95% CI 11.3-24.7) in 2120.

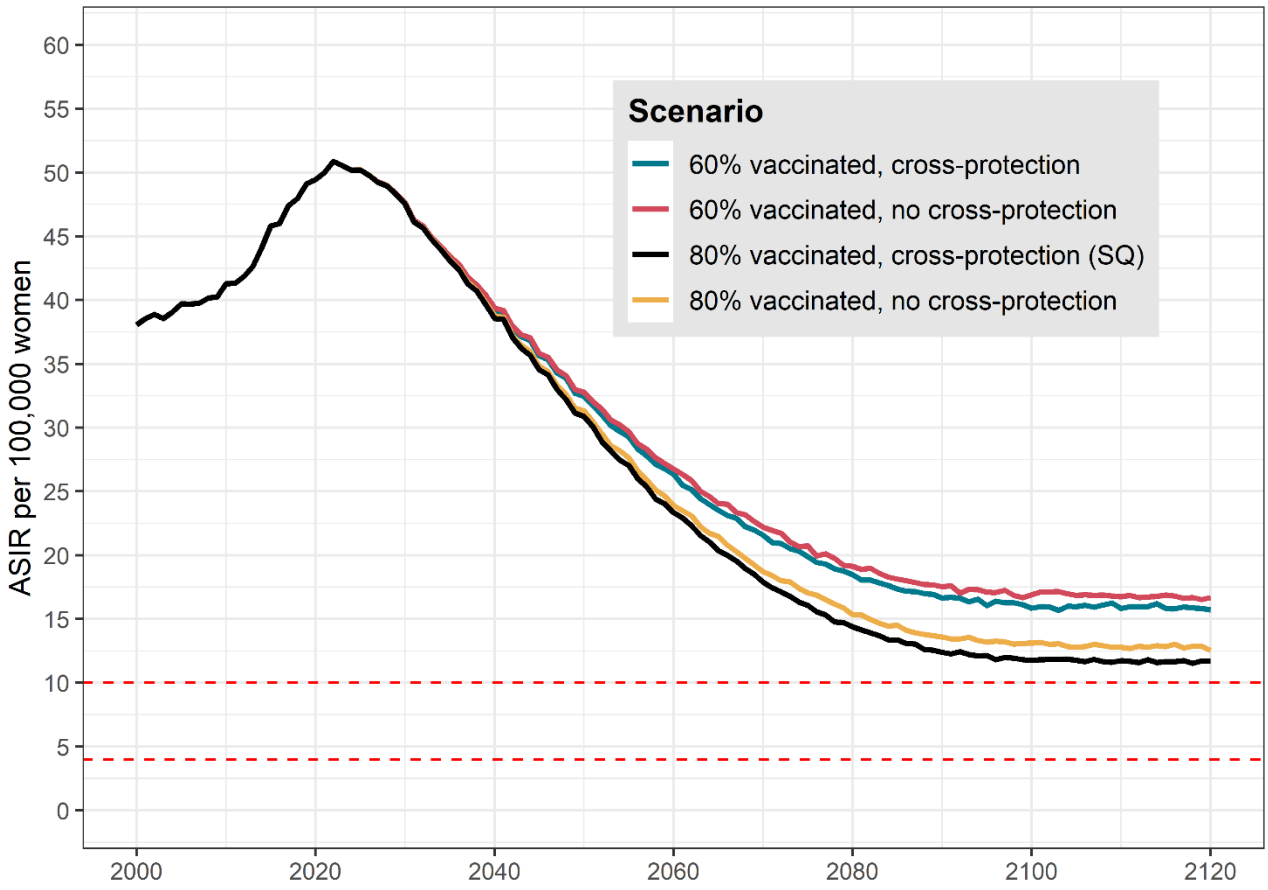


Figure E 4 – Mean model estimates of age-standardised cervical cancer incidence under different assumptions about vaccine efficacy.

E.3.3 HIV prevention

We reweight the population totals in our model using the projected population demographics of the Thembisa model, on the assumption that the Thembisa model estimates future HIV and demographic trends more realistically. The status quo HIV prevention efforts will lead to 96% of HIV-positive individuals diagnosed in 2030, 77.6% of diagnosed individuals on ART, 91.4% of those on ART virologically suppressed, and 76% of men aged 15-49 circumcised. As sensitivity analyses, we 1) increased ART coverage to reach 90% by 2030 and stay constant after ('90-90-90 targets' Scenario in Figure E5), and 2) kept all levels of coverage at 2020 levels (93-73-91, and 59% circumcised).

The curves in Figure E5 show that the impact of scaling up our current cervical cancer prevention programme to meet the 90-70-90 targets is not dependent on assumptions about HIV prevention.

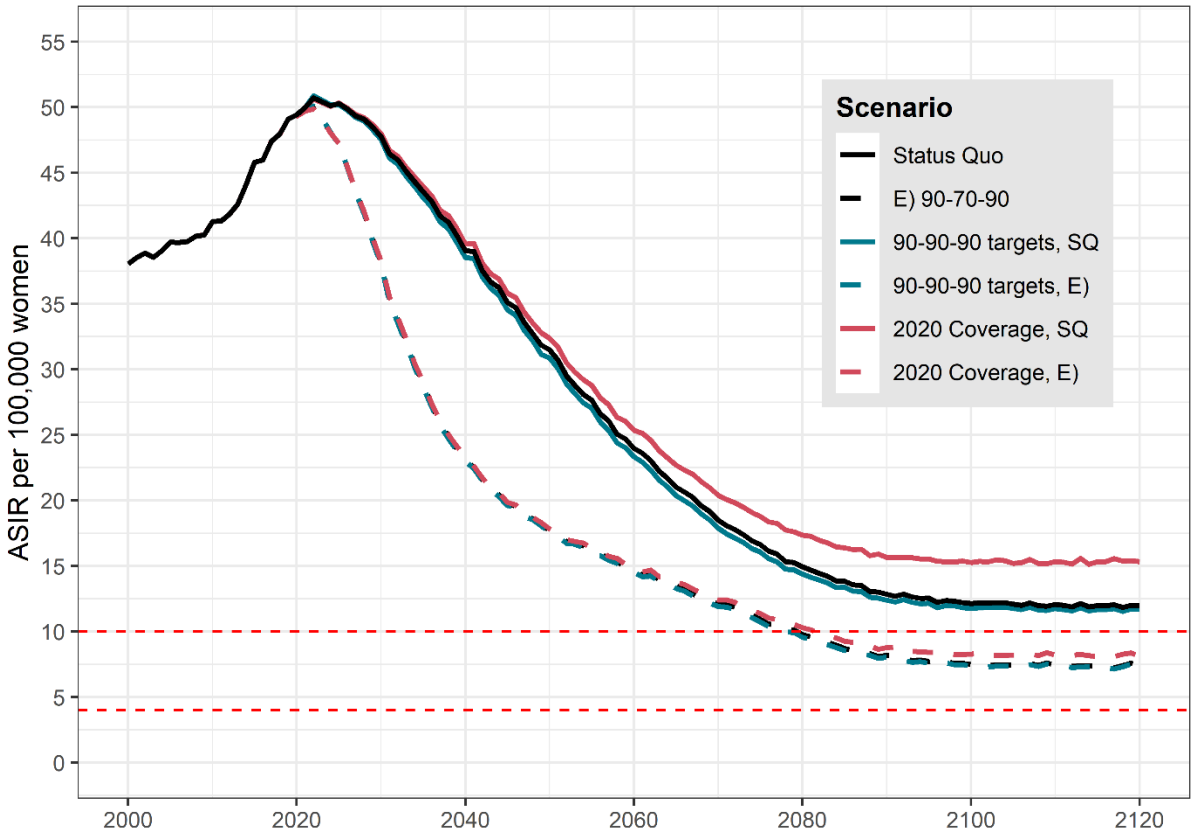


Figure E 5 – Mean model estimates of age-standardised cervical cancer incidence under different assumptions about future HIV prevention efforts.

E.3.4 Viral latency and reactivation of latent infections

In Chapter 4 we showed that natural history structures with and without viral latency and reactivation of latent infection can fit equally well to data, but that models without reactivation of latency do not match the difference in vaccine effectiveness between different risk groups. For this sensitivity analysis, we fit a model without viral latency and reactivation of latent infection to the cervical pre-cancer and cancer data. We use the same prior distributions of parameters as in Table A14, and calibrate to the same data and using the same method as described in Section A.5. This model does not fit well to cancer incidence at the older age groups (Figure E6) and predicts lower levels of cervical cancer incidence over time when assuming the status quo vaccination, screening, and treatment scenario (Figure E7).

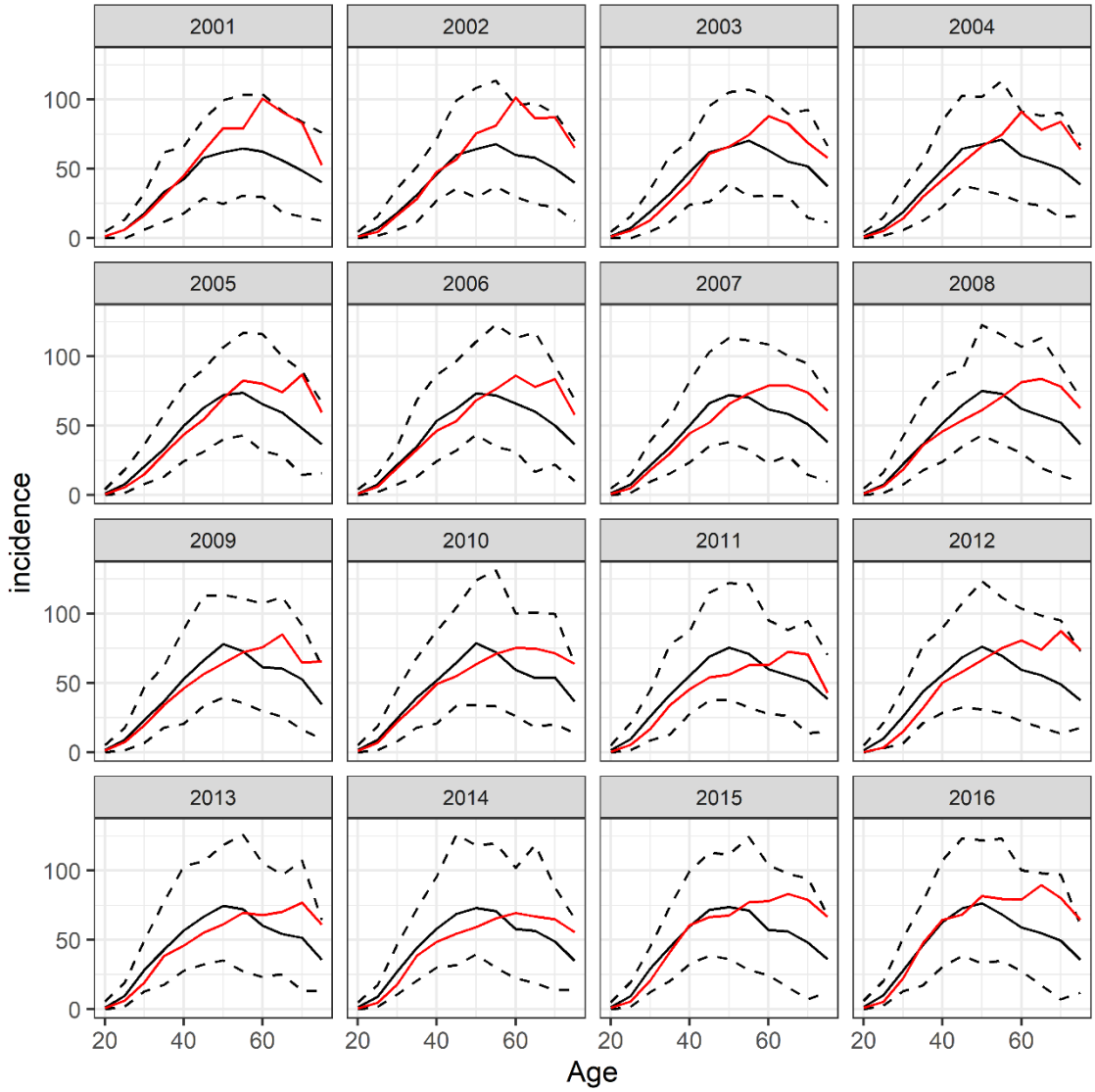


Figure E6 – Age specific diagnosed cervical cancer incidence per 100,000 women as calculated from NCR data (red lines) and the 100 best fitting parameter combinations for the model that assumes no viral latency and reactivation of latent infection (black lines show mean of 100 estimates, dashed lines show 95% percentile intervals).

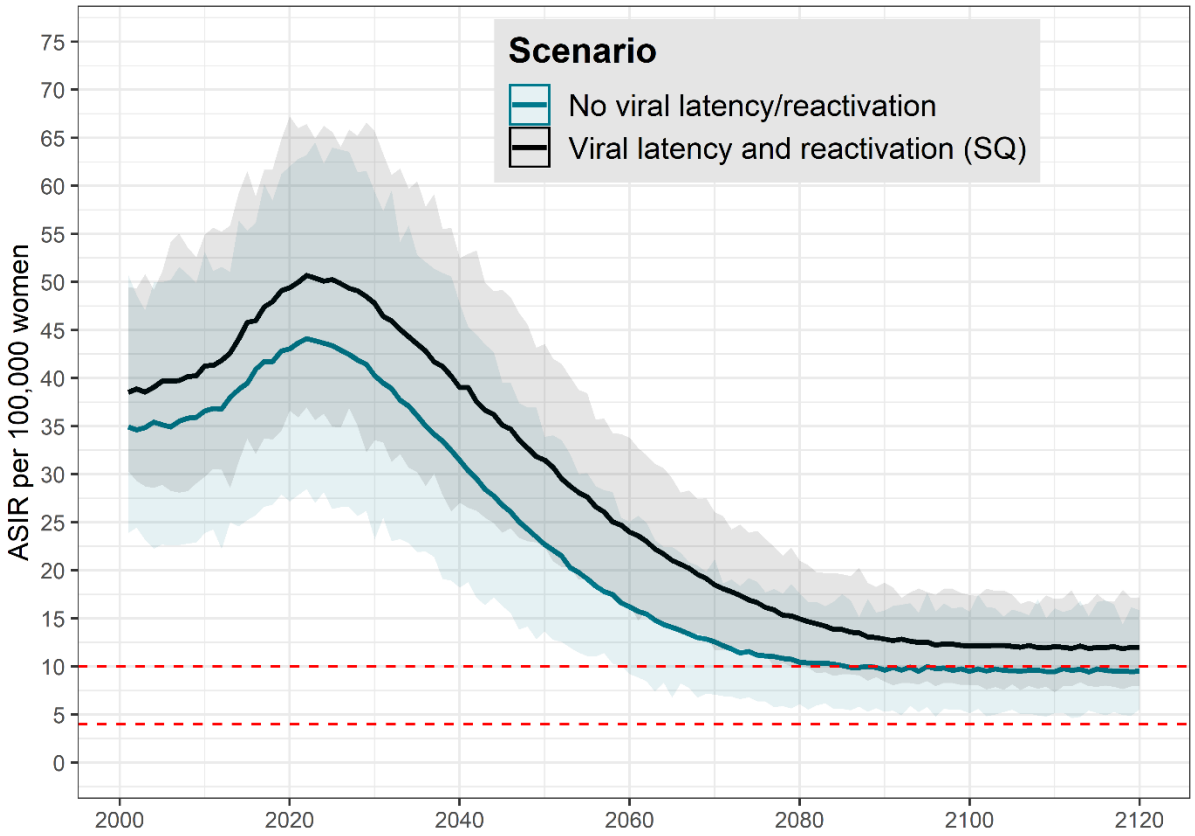


Figure E7 – Mean model estimates (and 95% percentile intervals) of age-standardised cervical cancer incidence including/excluding a state for viral latency and reactivation of latent infections in the natural history of cervical cancer.