Age-period-cohort analysis and prediction of Human Papillomavirus-related cancer incidence in South Africa

FACULTY OF HEALTH SCIENCE

SCHOOL OF PUBLIC HEALTH

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Sandile Ndabezitha (1232337)

Supervisor: Dr. Innocent Maposa, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Co-Supervisor: Dr. Mazvita Muchengeti, National Cancer Registry, National Health Laboratory Services, Johannesburg, South Africa

A research report submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in partial fulfillment of the requirements for the degree of Master of Science in Biostatistics.

OCTOBER 2021

DECLARATION

I, Sandile Ndabezitha, declare that is work is my original work. This is a research report submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in partial fulfillment of the requirements for the degree of Master of Science in Biostatistics. It has never been submitted to any University for examination or award of a degree. Thoughts and ideas from other authors have been correctly referenced.

Signed:

Date: 08/10/2021

DEDICATIONS

This piece of work is dedicated to my wife, my mom, and my siblings.

ACKNOWLEDGEMENTS

First of all, I would like to pass my sincere gratitude to my research supervisors, Dr. Innocent Maposa and Dr. Mazvita Muchengeti. Thank you for your training, mentoring, supervision, and guidance throughout this project.

Thanks and appreciation also goes to all the School of Public Health lecturers and staff for their time and effort in helping us with all the concepts in class. I would also like to thank the National Cancer Registry (NCR) leadership for allowing me to use their data and facility, the staff, and fellow students for sharing their knowledge and experience on cancer research with me.

I would also like to extend my appreciation and thanks to the Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) training programme, for the financial support for undertaking my studies and research.

To my family, thank you for the love and support while away from home. To, my mom, your prayers and well wishes go a long way in my life, more than you can imagine, thank you, Masinga. My love, Andile, thank you for everything. Your love, support, sacrifices, and prayers kept me going even when there was no strength left to continue. Gracious thanks to my friends for their helping hand whenever I reached out. Thanks to the three and a half men for always being there for each other in all aspects of life while undertaking our studies.

TO GOD, my source- Glory, and Honour to You, now and forever, Amen.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

HPV-Human Papillomavirus

APC- Age-Period-Cohort

NCR- National Cancer Registry

STATS SA- Statistics South Africa

ASIR- Age-Standardized Incidence Rate

CI- Confidence Interval

HIV-Human Immune Virus

ART- Anti Retroviral Therapy

PLHIV- People living with HIV

AAPC- Average Annual Percentage Change

APC- Annual Percentage Change

AGC- Anogenital Cancers

OPC- Oropharyngeal Cancers

CHAPTER 1: INTRODUCTION

This chapter gives a global and local overview of cancers in general and for HPV-related cancers. Study justification, problem statement, research question, and objectives are also detailed. A review of related literature was also detailed in this chapter.

1.1 BACKGROUND:

1.1.1 Cancer

Cancer is one of the major public health concerns in the world today and it is among the most common causes of mortality and morbidity. Non-communicable diseases are now considered the main contributors to global deaths. In the future, cancer is projected to be the leading cause of death and decreased life expectancy for almost every country across the globe in the 21st century (1). In 2018, there were 18 million cancer cases reported, and 9.6 million deaths in the world (1,2). The perception of cancer has also shifted from being viewed as a disease for the rich to a disease everyone can get. A huge burden of cancer cases and deaths fall on low and middle-income countries. About 70% of cancers diagnosed worldwide in 2018 were from low and middle-income countries (3–5). In South Africa, there were 107 467 cancer cases reported, and 57373 lives lost due to cancer (6) in 2018. In 2020, the world recorded about 19 million cancer cases and about 9 million deaths. The Southern Africa region had 116 391 cases and 61 659 deaths, with more than 90% of these cases coming from South Africa (108 168) and 56 802 deaths (7).

1.1.2 HPV-related cancers

HPV-related cancers are cancers mainly caused by the Human Papillomavirus (HPV) and these include cervical, penile, anal, vaginal, vulva, and oropharyngeal cancers. The incidence of Human Papillomavirus (HPV) associated anogenital and oropharyngeal cancers is on the rise in developed countries and it is an important cause of mortality and morbidity in all low and middle-income countries (8). Annually, there are about 600 000 cases of cancers of the cervix, oropharyngeal cancers, vulvovaginal and penile cancers, as well as genital warts and recurrent papillomatosis of the lungs worldwide, all accounted to HPV (9). Duncan et al found that HPV is attributed to 4% of all cancers, 2% in high-income countries, and about 8% in the low and middle-income countries (8). In 2018, the incidence per 100 000 per year of HPV-related cancers in South Africa was 44.4 in cervical, 0.2 in anal cancer, 0.8 in vulva, 0.2 in vaginal, 0.6 of penile and oropharyngeal (1.1 males, 0.3 females) cancers (10). South Africa has a high prevalence of HPV. In a study conducted in Soweto and Cape Town, HPV was found in 66.7% of the participants, with HPV 16 and 18 more dominant (11). Young adults are most at risk of developing HPV-related cancers compared to other age groups. Mbulawa et al found that HPV prevalence was higher among the 18-25 years group for both males and females (12).

1.2 PROBLEM STATEMENT:

Globally, the burden of cancer is expected to increase in the coming years. Cancer Research UK estimated that there will be 27.5 million cancer cases by 2040 (13) in the whole world. Also, by 2030 cervical cancer alone is estimated to result in 443000 deaths worldwide and 98% of these deaths would happen in low and middle-income countries and sub-Saharan Africa having the highest number of deaths (14).

In South Africa, the cancer-related mortality rate for the year 2014 was 70/100 000 population. Differences are noticeable between the nine provinces. The Eastern Cape had 75/100 000, Free State had 82/100 000, Gauteng with 65/100 000, KwaZulu Natal had 52/100 000, 46/100 000 for Limpopo, 50/100 000 for Mpumalanga, Northern Cape had 124/100 000, Northwest had 66/100 000 and 131/100 000 for Western Cape. These statistics suggest the existence of different cancer profiles for different provinces (15).

HPV is the most common sexually transmitted infection (STI) and every person acquires it at some point, as long as they are not vaccinated against it (16). HPV is also the main cause of most cancers, especially in countries where HIV prevalence is high (17). HPV is also preventable, and preventing it can also help prevent a lot of HPV-related cancers.

Cancer projections and estimates are available for cancers in general but currently, in South Africa, there is a lack of vital cancer-specific projections of HPV-related cancers for the future. This then means that there is a lack of information that can help influence policy making on prevention and control.

1.3 JUSTIFICATION:

It is of public health importance to detect the risk of any disease or illness early and then engage relevant interventions so that the community is protected from severe effects of that disease. To fully understand the problem of HPV-related cancers, there is a need to study current incidence rates and changes that may have taken place in the past, and use this information to estimate what might happen in the future. Extrapolating trends might inform programmes on the impact of risk factors, and planned or unplanned interventions, on the incidence of cancer in the years that follow (18). The following two areas are of great importance;

Future planning. Accurate estimates of expected cancer cases would be useful for healthcare providers in planning for the distribution and allocation of resources for cancer control from primary prevention, screening, and early diagnosis, treatment, rehabilitation to palliative care. It is also important to be aware of and understand future changes or interventions that might affect future estimates when basing decisions on cancer predictions (18).

Evaluation of cancer prevention and control. Predictions can also be used to create awareness for public health workers on the need for prevention measures that would help avoid the predicted situation. Predictions also help in evaluating planned interventions if they will be effective by comparing future scenarios with and without them in place (18). There is also a need to calculate the anticipated benefits of the vaccination programme started in 2014. Future projections can also help reveal if such programs are indeed going to benefit the population in the future or not.

This study, therefore, will help profile HPV-related cancers and produce evidence-based information that will be used for future planning, evaluation of cancer prevention policies, and control of HPV-related cancers in the country.

1.4 LITERATURE REVIEW:

1.4.1 HPV and Cancer

Human Papillomavirus (HPV) is a small non-enclosed deoxyribonucleic acid virus that mostly infects epithelial cells (14). HPV infection commonly causes skin or mucous membrane growth (warts) (19). HPV is one of the most common sexually transmitted infections in the world. There are more than 100 strains of this virus and the common and high-risk types are type 16 and 18 which are deemed necessary causes for cervical and other anogenital cancers (20).

Just like most common sexually transmitted infections, HPV acquisition generally occurs soon after first sexual activity and a majority of new HPV infections occur in adolescents and young adults. Meites et al found that a majority of sexually active adults have already been exposed to the virus and new infections easily occur with a new sexual partner (21). In 2012, there were about 35.3 million HPV infections worldwide, with 25 million in sub-Saharan Africa and 6.3 million in South Africa (12).

If left untreated, HPV infection may sometimes cause cancer (HPV-related cancers). HPV is believed to cause all cervical cancers, 90% of anal cancers, 40% of penile, vaginal, and vulva cancers, 25% of oral cavity cancers, and 35% of oropharyngeal cancers (22).

1.4.2 Factors associated with HPV and HPV-related cancers

1.4.2.1 HPV Vaccination

Three different vaccines for HPV are available and in use across the globe. These are Bivalent (Cervarix) which protects against HPV types 16 and 18, the Quadrivalent (Gardasil) which offers protection against HPV types (6, 11,16 and 18), and also the 9-valent which protects against HPV types (6, 11,16,18,31,33,45,52,58) (23). These vaccines are administered and recommended for both males and females aged 9-26 years, as long as individuals are not sexually active (HPV-naïve). Two doses are recommended for boys/girls who are less than 15 years, and three doses are administered for individuals 15-26 years. Doses are given in 6 months intervals (24).

HPV vaccines are effective and offer more than 10 years of protection and with no signs of waning immunity (24,25). Studies also show that HPV vaccination is positively essential in preventing cancer-causing infections and precancers (25,26), reduce the detection of HPV and provide herd immunity (27). The majority of these studies showed vaccine efficacy of more than 90% (28–32).

In South Africa, the Medicines Control Council of South Africa approved two doses of the Bivalent HPV vaccine in 2008. In 2014 the South African National Department of Health implemented the first school-based HPV vaccine, where girls who were 9 years and older, doing grade 4 from public schools were targeted (33,34). About 500 000 girls were targeted and the campaign reached about 85% of the target (33). Because cancers take time to develop, the early effects of this campaign will only be seen after 10-20 years.

1.4.2.2 Cervical Cancer Screening

Cervical cancer screening is a test done to detect the presence of abnormal cervical epithelium or HPV, intending to identify precancerous lesions and remove them before developing into cancer, and for early detection and treatment of cancer (35,36). Methods for screening used around the globe include;

Table 1.4.2.2 Cervical Cancer Screening Methods Used

Method	Description	Advantages	Disadvantages
Visual Inspection with acetic acid (VIA)	The healthcare worker inspects the cervix of a patient by applying acetic acid and observe the reactions to see if there are any abnormalities	1. Cheap 2. Easy to administer (needs less equipment)	1. Low sensitivity and non-specificity 2. Bound to human error
PAP Test (Smear)	Cells are collected from the cervix and sent to the lab for testing using a glass microscopy slide	1. Good specificity	1. Poor sensitivity 2. Labour intensive 3. Needs special equipment to administer

Liquid-based cytology	This is a variation of the pap test. It uses a liquid transport medium to store the specimen from collection to the laboratory.	1. High sensitivity compared to a pap smear. 2. Simultaneous testing is possible on the same specimen collected (HPV, chlamydia) 3. Improved specimen quality	1. Labour intensive 2. Low specificity
HPV Testing	Samples removed from the cervix are sent to the lab for HPV testing. HPV testing is also done on samples from the vagina, which women collect on their own	1. More effective for early detection of HPV/Cancer. 2. More accurate	1. Expensive 2. Need of a more improved infrastructure

In South Africa, both cytology and HPV testing are recommended and due to the available resources at present, cytology testing remains the primary method used (37). Three (3) pap smears per woman are recommended, administered from 25 years of age the earliest or at the time of HIV diagnosis, and screening is stopped at age 65 if the individual is HIV-negative.

1.4.2.3 Male Circumcision

The health benefits of male circumcision include improved hygiene, lower risk of urinary tract infections, decreased chances of sexually transmitted infections, and prevention of foreskin retraction problems (38). Male circumcision is also associated with a decreased rate of HPV transmission, which also has an impact on HPV-related cancer incidence. Studies conducted show that there is enough statistical evidence that male circumcision is inversely associated with HPV infection (39–41), penile HPV (42), cervical cancer (for women with circumcised partners) (41,42).

A study conducted in South Africa in 2014 found that the prevalence of male circumcision was 42.8%(43). The prevalence was 48.2% among Black African men, 24% among Coloured men, 25.5% among White men, and 32.6% among Indian/Asian men. Different patterns were also seen among provinces (43). Limpopo was the highest with 74.9%, Eastern Cape had 68.4%, Gauteng 43.6%, Western Cape 38.9 %, Mpumalanga 37.6%, Free States 34.5%, North West 30.5%, KwaZulu Natal 20.7%, and Northern Cape 19.3% (43).

1.4.2.4 Cancer and HIV

South Africa has the highest number of people living with HIV. About 20% of all people living with HIV in the world reside in South Africa (44,45). In 2018, South Africa recorded a high 19% prevalence rate among adults 15-49 years (44,45).

Most cancers, including AIDS-defining, conjunctival and HPV-related cancers occur most in HIV-positive individuals (46–49). Silverberg et al found that the median age at diagnosis for some cancers was younger for HIV-positive people with CD4 counts less than 200 ml/μl. This suggests that there is a vital role of immunosuppression on cancer development. In a study in South Africa, women who were HIV-positive were found to have a higher risk of HPV infections, precancerous lesions, and cervical cancer (50). These women also had a high viral load, evidence of compromised immune systems, promoting opportunities of comorbidities including cancer and other non-communicable diseases. The introduction of ART is also another factor, as it reduced AIDS-defining cancers and PLHIV had prolonged life, but this increased non-AIDS defining cancers (51).

HIV-positive individuals have a higher risk of HPV cancers compared to HIV negatives because of the modifying effect of HIV on HPV pathogenesis (47). The surveillance of cancer cells is impaired by HIV, lowering the clearance of HPV in the body (52) and perpetuating the progression of HPV tumors to cancer (17)

Being infected with more than one HPV type is more common among HIV-positive people compared to those who are HIV negative (12). Other cancers like anal are non-AIDS defining cancers but they have a significant association with HIV. In 2014, Chia-Ching et al found a 1% increase in HIV-positive women and a 28% increase in HIV-positive men for anal cancer compared to HIV-negative individuals (53). Therefore, HIV prevalence plays a vital role in HPV-related cancer incidence, especially in a high HIV burden setting like South Africa.

1.4.3 Spatial variation in cancer cases

Several studies have been done to investigate geographical variations in cancer incidence. In a study on HPV-related oropharyngeal cancer, there was significant geographical variation in the prevalence of biologically relevant oncogenic HPV in oropharyngeal cancer (54). In 2018, differences in cancer cases and deaths were noticed between continents. Europe contributed 23.4% of the overall cases and 20.3% of deaths, Americas had 21% cases and 14.4% of deaths, Africa contributed 5.8% of cases and 7.3% of deaths and Asia had 48% cases and 57.3% deaths (55). Bray et al added that the different cancer profiles within a country or between regions suggest that there are still differences caused by geographical separations of the population and risk factors differ from place to place and are experienced at different times (55).

In South Africa, we know that provinces differ by population size, healthcare practices (traditional healthcare vs modernized), way of living (smoking habits, diet, exercise, sexual activities). In this way, cancer risk factors are also bound to be different across provinces and this suggests a possibility of spatial variation of cancer cases between provinces in the country(15,56,57)

1.4.4 Models used for predicting cancer incidence

Besides investigating factors associated with cancer incidence, it is important to understand the effects of different time scales in cancer incidence. These time scales are surrogate measures that can give an understanding of the causes of cancers. Different models concentrated on;

Age Effects-Even though anyone can acquire cancer, the probability of acquiring cancer increases with increasing age. Age effects are, therefore, the variations in incidence rates for the different age groups as they move through time. Increasing age is the major risk factor in cancer and for many individual cancer types. The median age for cancer diagnosis is 66 years and one-quarter of new cancer cases are diagnosed in individuals aged 65 to 74 years (58). HPV-related cancers show a similar trend in terms of mean age at diagnosis, except for cervical cancer which has the lowest (59,60). The median age at diagnosis for HPV-related cancers is between 49 and 53 for cervical, 68 for vaginal, 66 for vulva, 69 for penile, 62 among women and 59 among men for anal, and 63 among women and 61 among men for oropharyngeal cancers (59,61).

Period Effects-These effects account for all factors which affect every person at a time in the past such as pollution or medical advances (62). Period effects are the same for all age groups at a particular time as they result from external factors (social, economic, and environmental) such as economic meltdown, famine, war (21).

Cohort Effects-Individuals are grouped according to their year of birth to form a birth cohort (commonly defined in epidemiology). Cohort effects arise from a unique exposure that affects a particular cohort as it advances in time (years) (58). Theodore found that a particular risk factor affects different cohorts differently, and this causes a difference in incidence between cohorts (63). If we take a group of people vaccinated, for example, against say HPV infection, and if there is any exposure of HPV infection as different cohorts move across time, the vaccinated group would experience fewer effects of the exposure compared to the unvaccinated cohort.

1.4.4.1 APC

The Age-Period-Cohort (APC) model has been used in most cancer prediction studies conducted in the past. This is a descriptive tool that seeks to profile the effects of age, period, and cohort for a disease over time. APC models are based on fundamental generalized linear model theory. They serve as a general methodology for cohort analysis when all three factors are potentially of interest (64). These models are also based on the Poisson log-linear model for the expected rates with additive effects of age, period, and cohort (65).

Even though this is a model widely used in the past, it has also received criticism over its identification problem. This is because the three important covariates under study are linearly dependent. Age = Period – Cohort, so if you know someone's age, you also know their birth cohort. Smith and Wakefield pointed out that direct interpretation of these effects is difficult because the model is over parameterized and two sources of identifiability to consider were;

- For every model with factors and with an intercept in the model, there was one more level than is estimable and so a constraint will be required. A typical solution was to impose a sum-to-zero constraint (66).
- Linear dependence between age, period, and cohort also causes identifiability in the model. The use of assumptions can help in direct interpretation of the equation parameters, but these assumptions can not be verified from the original data alone (66).

With these issues in the classical APC model, other models were formulated to retain the forecasting ability of the model while providing a solution for the identification problem in the model.

1.4.4.2 Other models (APC extension)

The other models mentioned below present an alternative to the APC model by addressing the identification problem through their setup.

Stratified APC This model groups observations by factors not directly affected by time, e.g., geographical location, sex, disease type. In this model, relative risks between strata are identifiable if a set of effects is common between strata. For example, if age effects are shared between strata, the period and cohort effects' trends in relative risk are identifiable even if they are within the same time, as long as they are between strata (67–69).

Bayesian APC- Some studies used this model in the past in cancer prediction studies (70–72). This model allows for incidence count data analysis by using Bayesian priors (Random walk priors). Unlike the classical APC model, with the Bayesian approach, it is possible to use input data with different time scales. Also, retrospective predictions are possible for model testing purposes.

Constrained Coefficients (CGLM)-This model enforces additional constraints on one of the categories of at least one predictor to estimate the age, period, and cohort effects at the same time. It is possible to separately estimate the effects of age, period, and cohort by assuming that some categories of the independent variables have equally the same effects on the outcome variable (73). The results of the analysis depend on the chosen constraints and the validity of the constraints depends on the theory about the categories of parameters that are identical. This is often subjective and there is no empirical way to confirm the validity of the chosen constraints (74).

Intrinsic estimator method (IE)- This method draws its strength from principal component analysis and is suitable for sorting out the identification problem when explanatory variables are highly correlated (75). It erects rules on parameters similar to CGLM but the difference is that these rules are less subjective and do not have any effect on the estimation of regression parameters for age, period, or cohort (74,76). Studies on model authentication confirm the strength of the statistical characteristics of the IE model. When the same set of data is analyzed, the IE findings outweigh the results of any other model that does not use the same properties (76).

1.5 RESEARCH QUESTION(S):

We postulated that the impact of HPV-related cancers differed between age groups, different periods, and birth cohorts for individuals. Also, cancer incidence differed between geographical locations (province of diagnosis) of individuals. The research question, therefore, was;

1. What would be the burden of HPV-related cancers from 2016 to 2040?

1.6 AIM AND OBJECTIVES:

The aim of the study was to estimate the age, period, and cohort-specific incidence of HPV-related cancers in South Africa over the next years from 2016 to 2040. This was achieved through the following objectives:

1. Describe the demographic characteristics of HPV-related cancers for the years 1996-2015.
2. Estimate the incidence of HPV-related cancers for 1996-2015.
3. Forecasting of the age, period, and cohort-specific incidences of HPV-related cancers for the years 2016 to 2040.

CHAPTER 2: METHODS

This chapter gives details of the methods used, the study design, site of study, study population, data sources, study limitations, data management and statistical analysis, and details of ethics clearance for conducting the study.

2.1 Description of the previous study/data sources

Two data sources were used in this study.

The NCR data: The South African National Cancer Registry (NCR) is a national pathology-based cancer registry. Private and public laboratories across the country submit laboratory reports of cancer cases to NCR. The data are cleaned, analyzed, reported annually, disaggregated by sex, age, and population groups (race). Information is used for the development of cancer prevention and control policy (77).

Stats SA data: The Statistics South Africa (Stats SA) is a national South African statistical service that produces, stores, and shares relevant and accurate statistical information whenever needed. Some information is available upon request, and some are publicly available and can be downloaded on the online reserves. South African mid-year population estimates are generated annually and are made available online for everyone who may want to use them. Data are disaggregated by sex, age group, population groups (race), province, and district, among other variables (78).

2.2 Study design

This study employed a cross-sectional study design. HPV-related cancer incidence rates (Cases reported per year over the mid-year population projections) for the years 1996-2015 were calculated. Trend analysis was conduction for 1996-2015 period. Predictions of HPV-related cancer incidence rates were conducted for the years 2016-2040.

2.3 Study site

Data for all reported HPV-related cancer cases in South Africa as recorded in the National Cancer Registry were used in this study.

2.4 Study population

The study used pathology-confirmed HPV-related cancer cases reported in the NCR database between the years 1996-2015 in South Africa

2.5 Data collection

Secondary data were used for this study. Data on de-identified pathology-confirmed HPV-related cancer cases reported to the NCR were used after obtaining ethics approval. These data formed the numerator of the incidence rates to be calculated. Publicly available mid-year population estimates data from STATS SA were then used as denominators for incidence rates. Five-yearly age groups' data (0-4 to 85+) for the different years (1996-2015) and the 9 provinces of South Africa were used.

The following cancer types with their International Classification of Diseases for Oncology Version 3 (ICD-03) codes were classified as HPV-related cancers in the study (79);

Table 2.5 Cancer types classified as HPV-related cancers in the study and their ICD-03 codes

Cancer Type	ICD-03 code
Cervical	C53.9
Penile	C60.9
Vaginal	C52.9
Oropharyngeal	C09.0, C09.1, C09.8, C09.9, C10.0, C10.1, C10.2, C10.3, C10.9, C11.1, C11.2, C11.3, C11.8, C11.9, C12.9, C13.0, C13.1, C13.2, C13.8, C13.9, C30.0, C30.1, C31.0, C31.1, C31.2, C31.3, C31.8, C31.9
Vulva	C51.9
Anal	C21.0, C21.1, C21.8

2.6 Data management and statistical analysis plan

2.6.1 Data Management

All pathology-confirmed HPV-related cancer cases within South Africa, reported to the NCR, diagnosed between 1996-2015 were considered for this study. Only cancers listed in section 2.5 above were included in the study. Duplicate checks were also conducted in Stata and there were no duplicates from the data we received. Cases with unknown sex (62), unknown race (3714), missing age (3879) were dropped from the analysis. We also checked for consistency in the data, e.g., we checked if there were no males who were classified as having cervical cancer, or females diagnosed with penile cancer, and all the observations were consistent. The Stats SA data was obtained from the Stats SA website online (78). The data were merged to the NCR data to form one dataset that was used for the entire analysis in the study. Computer programmes used for data cleaning and analysis were R, Stata, Joinpoint, and QGIS.

2.6.2 Statistical Analysis

2.6.2.1 Statistical Inference

Different models of the APC were fitted and summarized. Model comparison was also carried out between these models discussed below.

APC model

The model is based on a Poisson log-linear model for the expected rates with additive effects of age, period, and cohorts. The model has the form (66);

$$\log(M_{ij}) = \log\left(\frac{\theta_{ij}}{N_{ij}}\right) = \mu + \alpha_i + \beta_j + \lambda_k$$

where M_{ij} is the HPV-related cancer incidence rate for individuals in age group i at calendar time j

θ_{ij} is the number of incident cases among individuals in age group i at calendar time j

N_{ij} is the number of individuals in age group i at calendar time j

μ is the overall expected mean log incidence rate

α_i, β_j, and λ_k are the differential rates from the overall mean log incidence rate for individuals in age group i, calendar time j, and birth cohort k, respectively.

Bayesian APC Model

The Bayesian model offers a solution to the identification problem in the classical APC model. The basic model used in Bayesian APC can be expressed as (66);

$$M_{ij} = \mu + \theta_i + \alpha_j + \beta_k + z_{ij}$$

where M_{ij}= HPV-related cancer incidence rate in age-group i and year j

μ = overall level

θ_i= age effects in age-group i

α_j = period effects in year j

β_k = cohort effects in cohort k

z_{ij} = unstructured heterogeneity in age-group i and year j not captured by θ,

α and/or β are considered to be due to unobserved observation-specific covariates

The parameter for unstructured heterogeneity (assumed to relate to unobserved covariates) can be thought of as analogous to adjustments made for overdispersion in a frequentist setting. However, within a Bayesian framework, it ensures a good fit for the data. This model improves the classical APC model by including time-varying covariates effects, with the use of random walk priors.

Stratified APC model

This is another extension to the classical APC model which includes the stratification effects (e.g. gender, place of residence, province). The model is presented as (66);

$$\log(M_{ijr}) = \log\left(\frac{\theta_{ijr}}{N_{ijr}}\right) = \mu_r + \alpha_{ir} + \beta_{jr} + \lambda_{kr}$$

where M_{ijr} is the HPV-related cancer incidence rate for individuals in age-group i, at calendar time j in stratum r

θ_{ijr} is the number of HPV-related cancer incident cases among individuals in age-group i, at calendar time j in stratum r.

N_{ijr} is the number of individuals in age-group i at calendar time j in stratum r.

μ_r is the expected mean log incidence rate for stratum r.

α_{ir} , β_{jr} *and* λ_{kr} are the differential rates from the overall mean log incidence rate for individuals from stratum r, in age group i, calendar time j, and birth cohort k, respectively

2.6.2.2 Analysis plan for specific objectives

Objective 1 *(Describe HPV-related cancers according to demographic and clinical characteristics for years 1996-2015)*

To get a better understanding of the data and the study, descriptive data analysis was performed for the age of the patient, birth cohort, period of diagnosis, and sex. Frequency tables (Numbers and percentages) for categorical variables were provided.

A Choropleth map, with different colour coding, was used to display the distribution of HPV-related cancers according to the province of diagnosis. The distribution of the HPV-related cancer cases was calculated for all provinces and categorized into 5 classes, < 4000, 4000-6000, 6000-8000, 8000-10 000, and >10 000, and then were plotted on the choropleth map.

Objective 2 *(Estimate the incidence of HPV-related cancers for 1996-2015)*

The outcome variable was HPV-related cancer incidence rate and was compared over the years between 1996-2015. The total number of cases per year was used as the numerator and the mid-year population estimates for each year were used as a denominator for this objective (total cases/population = incidence). **Age-standardized** HPV-related cancer incidence rates were calculated. Age-standardized rates are summary rates calculated to incorporate the age composition of a reference population (World standard population) (80). Recommended by WHO, the use of standardized rates allows estimates and findings generated to be comparable to other countries' findings across the world. The formula was used for calculating the ASIR in the study (80);

ASIR = $\sum_{i=1}^{n} d_i w_i / y_i$

Where: $_i$= age groups

d_i= Number of HPV-related cancer cases in age-group i

w_i= World standard population for age-group i

y_i= population at risk, for age-group i

The **average annual percentage change (AAPC)** was then used to summarize the trend performance of the incidence over time. The AAPC uses a single measure to give an aggregated summary of the changes in incidence rates over the period under study(81,82). The statistical significance of this trend analysis was determined using confidence intervals and p-values. Linear regression analysis and the formulae below were used for calculating both the percentage change and the 95% confidence interval (81,82);

AAPC = $[e^{\beta} - 1]$ x 100

95 % CI = $\beta \pm \left[Z_{1-\alpha/2}\right] x \; SE_{\beta}$

Where: β = diagnosis year coefficient from the regression analysis

$Z_{1-\alpha/2}$= Standard normal table corresponding to 95 % confidence level

SE_{β} = Standard error of the diagnosis year coefficient from the regression analysis

Objective 3 *(Projection of the age, period, and cohort-specific incidences of HPV-related cancers for the years from 2016-2040)*

The three models were fit and model comparison was performed. Models were compared using three measures; the Model's ability to estimate parameters/coefficients, the ability to make predictions for the future (Based on the data used in the study), and considerations of the AIC/

DIC. The best model was selected and used for prediction. Data from 1996-2015 were used to produce forecasts for 2016-2040.

2.6.3 Ethics Consideration

The researcher applied for and obtained ethics approval (M200522) to undertake this study from the University of the Witwatersrand Human Research Ethics (Medical) Committee (HREC). Permission from the National Cancer Registry to use their data was also obtained. To obtain confidentiality, de-identified data were given to the student in a password-protected Stata format, and the student stored the data in a password-protected laptop.

CHAPTER 3: RESULTS

This chapter gives a summary of the main results of the study.

3.1 Description of the demographic and clinical characteristics of HPV-related cancer cases

Table 3.1 Demographic Characteristics of data used

Variable		HPV-related cancer cases	
		Number	Percentage
Age-group			
	0-4	62	0.05
	5-9	45	0.04
	10-14	112	0.10
	15-19	275	0.24
	20-24	742	0.65
	25-29	2830	2.49
	30-34	6944	6.11
	35-39	11195	9.85
	40-44	13996	12.31
	45-49	14580	12.82
	50-54	14410	12.68
	55-59	12996	11.43
	60-64	11766	10.35
	65-69	9049	7.96
	70-74	6722	5.91
	75+	7963	7.00
Period of diagnosis			
	1996-2000	26100	22.96
	2001-2005	24873	21.88
	2006-2010	28388	24.97
	2011-2015	34326	30.19
Sex			
	Males	9119	8.02
	Females	104568	91.98
Race			
	Black African	90838	79.90
	Coloured	9206	8.10

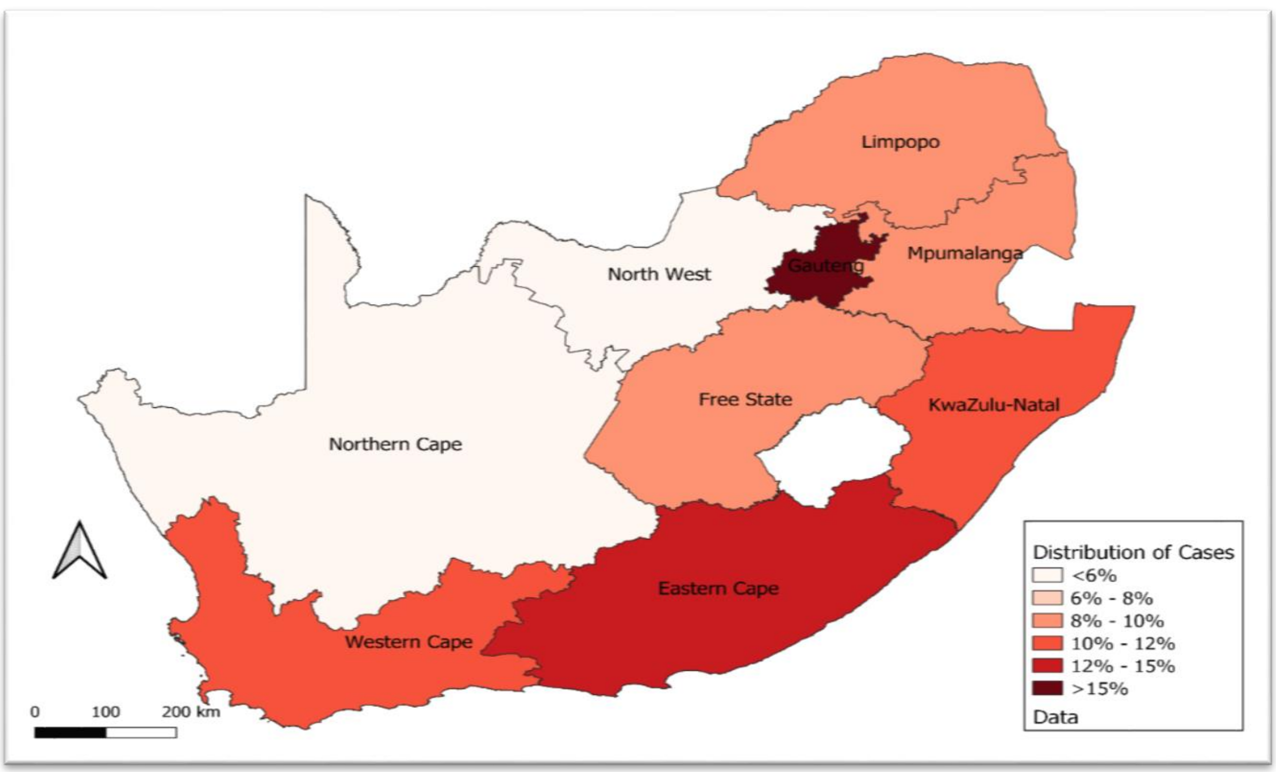
Indian/Asian	1883	1.66
White	11760	10.34

Table 3.1.1 above displays the demographic characteristics of the data used in the study. A total of 113 687 HPV-related cancer cases were used for the analysis. About 12.82 % (14580) of these were aged 45-49 years old, followed by the 50-54 and 40-44 year-olds having (12.68%; 14410) and (12.31%; 13996) cases respectively. There were 229 (0.9%) children who were less than 14 years old, while there were 7963 (7%) adults who were 75 years old and above.

A majority of these cases were female 104568/113 687, and males taking up only 8%. The Black-African population group was commonest, making up about 80% of the cases, with a combination of Coloured and Indian/Asian making only 10% of the cases, and the remaining 10% being of the White population group.

More cases were reported in the period 2011-2015 (34 326), compared to 26 100 in 1996-2000, 21 873 in 2001-2005, and 28 388 in 2006-2010.

Figure 3.1.1 Choropleth map of the province of diagnosis of HPV-related cancer



The map above displays the distribution of HPV-related cancer cases according to the province of diagnosis, for the years 1996-2015. About 65 917 records of the data had a documented province of diagnosis. A majority of cases were diagnosed in Gauteng province (with close to 30% cases diagnosed), followed by the Eastern Cape province. Fewer cases were diagnosed in the Northern Cape and North West provinces.

3.2 Estimating the HPV-related cancer incidence from 1996-2015

3.21 Age-standardized HPV-related cancer incidence rates (1996-2015)

Table 3.21 Age-Standardized Incidence rates for Anogenital (AGC) and Oropharyngeal OPC) Cancers

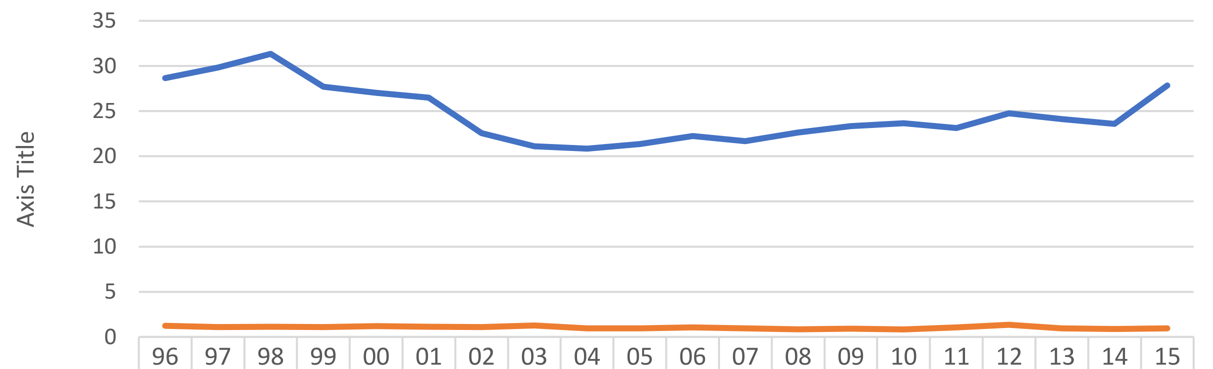
Age group	World Standard Population	OPC cases	OPC ASR	AGC cases	AGC ASR
0-4	12000	37	0.15	25	0.10
5-9	10000	38	0.14	7	0.03
10-14	9000	84	0.31	28	0.10
15-19	9000	154	0.58	121	0.45
20-24	8000	159	0.49	583	1.78
25-29	8000	144	0.42	2686	7.91
30-34	6000	200	0.49	6744	16.82
35-39	6000	288	0.90	10907	34.76
40-44	6000	456	1.69	13540	50.12
45-49	6000	701	3.30	13879	58.11
50-54	5000	1026	4.70	13384	53.03
55-59	4000	1058	4.70	11938	45.34
60-64	4000	1051	5.96	10715	50.37
65-69	3000	763	4.55	8286	39.05
70-74	2000	500	3.01	6222	28.09
75+	2000	564	2.84	7399	22.43
Total	100 000	7223	**34.23**	106464	**408.51**

Table 3.21 above displays the age-standardized incidence rates for Oropharyngeal and Anogenital cancers for the different age groups. The general idea from the above table, as postulated by literature, is that cancer incidence increases with increasing age, with the age group 45-49 years

having the highest Age Standardized incidence rate for AGC and the 60-64 having the highest for OPC.

3.22 Average Annual Percentage Change (1996-2015)

Figure 3.221 HPV-related cancer age-standardized incidence rates for period 1996-2015



	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
AGC	28.6	29.8	31.3	27.7	27.0	26.4	22.5	21.0	20.8	21.3	22.2	21.6	22.6	23.3	23.6	23.1	24.7	24.1	23.5	27.8
OPC	1.22	1.09	1.13	1.10	1.21	1.12	1.10	1.26	0.96	0.93	1.04	0.96	0.85	0.91	0.83	1.04	1.34	0.96	0.87	0.96

Figure 3.221 above shows the trend of Age-standardized incidence rates for AGC and OPC for 1996-2015. The graph shows an increase in AGC incidence rates from 28.6 to 31.3 cases per 100 000 world population, for the years 1996-1998. This was followed by a decline to 20.8 cases per 100 000 in 2004. An up-and-down trend can be noticed from 2005 until 2015.

The incidence of OPC moved between 0.83 and 1.34 over the years. It was lowest in 2010 (0.83) and it peaked in 2012 (1.34).

AAPC (appendix 1)

The Joinpoint software (version 4.9) was used for the analysis. The software runs a model selection where several join points (segments) that best fit the data are determined. For AGC, a model with 3 segments was selected. There was an increasing trend for 1996-1998 even though it was not statistically significant (APC 4.8%, p=0.434). There was a significant downward trend for the 1998 to 2003 period (APC -7.8%, p=0.001), and an upward trend for 2003 to 2015 (APC 1.7%, p<0.001). The full range period showed a downward trend in the AGC incidence (APC -0.9%, p=0.056) although it was marginally significant.

For OPC, a full range model was selected and it showed a significant decreasing trend in OPC incidence rate (APC -1.2%, p=0.019).

3.3 Projection the HPV-related cancer incidence from 2016-2040

3.31 Model Comparison and selection of preferred model

Four models were considered for the analysis and prediction of HPV-related cancer cases/incidence for the years 2016-2040. These models were the classical APC, Stratified APC (Sex as strata), Stratified APC (Race as strata), and the Bayesian APC. Models were compared using three measures; the Model's ability to estimate parameters/coefficients, the ability to make predictions for the future (Based on the data used in the study), and considerations of the AIC/ DIC.

Table 3.31 Models considered in the study

Model	Parameter estimation	Prediction/Forecasting	AIC
Classical APC	Yes	Yes	706.18
Stratified APC-Sex			
Male	Yes	Yes	509.98
Female	Yes	Yes	677.57
Stratified APC-Race			
Black	Yes	Yes	687.6
White	Yes	Yes	491.18
Indian/Asian	No	No	377.68
Coloured	Yes	Yes	461.69
Bayesian APC	Yes	No	273.98

Table 3.31 above shows the scoring or rating of the models under these different criteria. The Bayesian model had the lowest AIC compared to all the other models. It was also possible to estimate the coefficients of the three explanatory variables of the model. However, based on the data used in the study (missing data in some variables in the predicted values (prediction model produced NaN values which are undefined or not representable)), it was not possible to perform predictions using this model.

The Stratified APC using race as strata, had the second-lowest AIC, with the ability to estimate coefficients and also do predictions into the future, except for one stratum. The Indian/Asian stratum model could not estimate coefficients and also do predictions. Just like the Bayesian APC, this was due to missing data in some variables.

Both the classical APC and the Stratified APC using sex as strata passed all the measures set. Model comparison and selection were based on the AIC scores. The Stratified APC had the lowest AIC and thus was selected among the four models.

3.32 The Selected Model (Stratified APC (Sex as Strata))

Table 3.32 Input data for the model

	Females				Males			
	Period				Period			
Age-group	**1996-2000**	**2001-2005**	**2006-2010**	**2011-2015**	**1996-2000**	**2001-2005**	**2006-2010**	**2011-2015**
0-4	14	11	13	7	7	1	6	3
5-9	12	2	3	2	12	6	6	2
10-14	13	16	20	10	10	14	9	20
15-19	59	34	33	35	36	31	25	22
20-24	167	135	182	128	33	24	38	35
25-29	594	532	737	826	32	29	48	32
30-34	1358	1236	1761	2275	63	55	78	118
35-39	2352	2155	2605	3599	109	101	109	165
40-44	3091	2828	3259	4083	184	137	161	253
45-49	3083	3070	3361	4106	238	215	225	282
50-54	2870	3041	3293	3899	308	317	316	366
55-59	2733	2502	2975	3556	239	293	312	386
60-64	2643	2473	2501	2890	262	306	298	393
65-69	2046	1859	2014	2197	201	196	244	292
70-74	1437	1436	1498	1696	179	141	139	196
75+	1519	1473	1950	2260	196	204	169	192

The input data were aggregated according to age and period of diagnosis only, for both sexes. The APC model (in R) uses this data to calculate the cohort variable (period-age = birth cohort).

3.33 Age-Period-Cohort Analysis Of HPV-related Cancer 1996-2015

Figure 1.33a: Age-Period-Cohort analysis of HPV-related cancer cases in Females

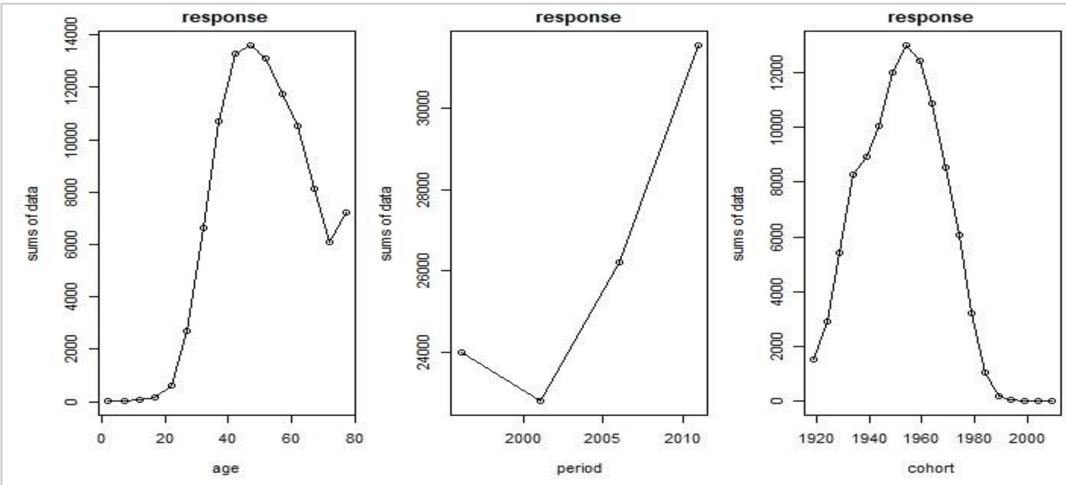
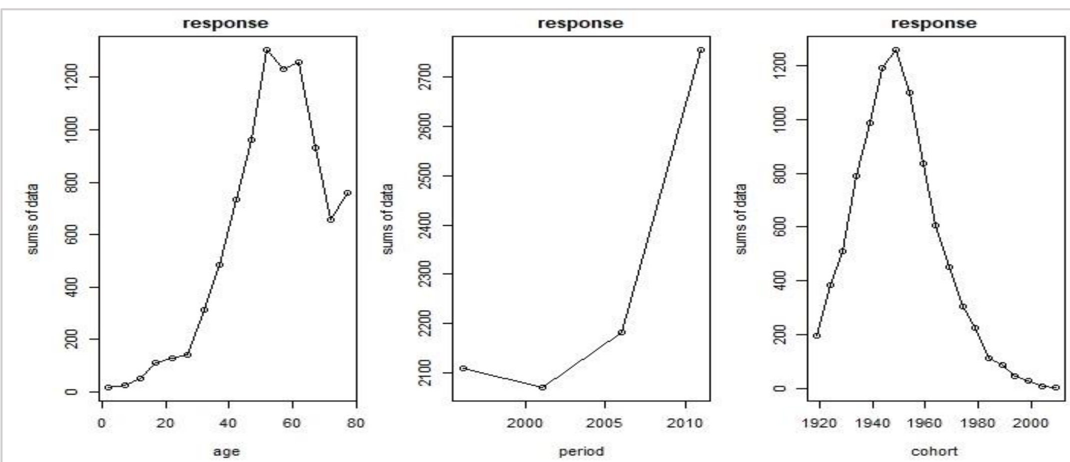


Figure 3.33b: Age-Period-Cohort analysis of HPV-related cancer cases in Males



Figures 3.33a and 3.33b display the age, period, and cohort effects on HPV-related cancer cases for females and males respectively. The general idea is the same for both males and females. For age effects, HPV-cancer cases increase with increasing age. The cases increase for both females and males, until they reach a peak at around 50 years, then starts to drop. In females, there was a slow increase in cases, until age 25, where cases start to increase drastically. In males, it is a slow increase until age 30.

HPV-related cancer cases were also increasing with the period, for both females and males. There is a drop from 1996 to 2001, then and an increase is seen until 2015.

For cohort effects, there is an increase in HPV-related cancer cases from the earlier cohorts until a peak in cases on the 1955 and 1950 cohorts, for both females and males respectively. The numbers drop for the younger cohorts.

3.34 Projection of HPV-related cancer cases using the Stratified APC model:

Figure 3.341 Forecasted numbers of HPV-related cancer cases 2016-2040

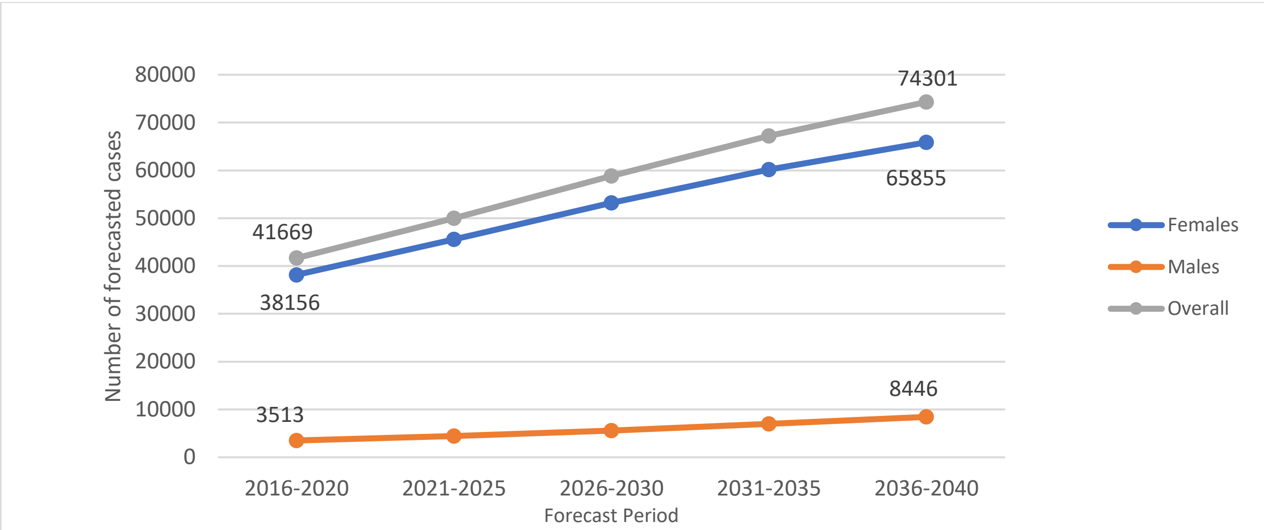


Figure 3.341 above displays projected numbers of HPV-related cases for the forecasting period 2016-2040. In total, HPV cases are expected to increase over the years, rising from 41 669 cases in 2016 up to 74 301 cases in 2040. Both sexes follow the same pattern, with females reaching 65 855 cases in 2040, from 38 156 cases in 2016, and male cases rising from 3 513 cases in 2016 to 8 446 cases in 2040.

Table 3.342a Forecasting of HPV cases by birth cohort and age-group-Females

Age	Birth Cohorts																		
	1919-1923	1924-1928	1929-1933	1934-1938	1939-1943	1944-1948	1949-1953	1954-1958	1959-1963	1964-1968	1969-1973	1974-1978	1979-1983	1984-1988	1989-1993	1994-1998	1999-2003	2004-2008	2009-2014
0-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	17	10	11	7
5-9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	5	4	4	3
10-14	-	-	-	-	-	-	-	-	-	-	-	-	-	21	14	14	10	12	7
15-19	-	-	-	-	-	-	-	-	-	-	-	-	49	43	34	35	26	29	19
20-24	-	-	-	-	-	-	-	-	-	-	-	151	158	168	136	142	106	119	76
25-29	-	-	-	-	-	-	-	-	-	-	537	590	739	822	664	694	517	582	374
30-34	-	-	-	-	-	-	-	-	-	1329	1313	1728	2260	2514	2029	2123	1581	1778	1143
35-39	-	-	-	-	-	-	-	-	2300	2207	2613	3590	4697	5225	4217	4412	3285	3695	2375
40-44	-	-	-	-	-	-	-	3073	2853	3281	4055	5572	7289	8108	6545	6848	5098	5735	3685
45-49	-	-	-	-	-	-	3138	3039	3382	4061	5019	6897	9023	10036	8101	8476	6310	7098	4561
50-54	-	-	-	-	-	2943	2895	3361	3905	4688	5795	7963	10417	11587	9353	9786	7285	8195	5266
55-59	-	-	-	-	2834	2476	2919	3538	4110	4935	6100	8382	10966	12198	9846	10301	7669	8627	5544
60-64	-	-	-	2644	2357	2468	3038	3682	4278	5137	6349	8724	11413	12695	10248	10722	7982	8979	5770
65-69	-	-	2065	1870	1997	2184	2688	3258	3785	4545	5618	7719	10098	11233	9067	9487	7062	7945	5105
70-74	-	1364	1427	1549	1727	1888	2324	2817	3274	3931	4858	6676	8733	9714	7841	8204	6107	6871	4415
75+	1519	1546	1939	2197	2451	2679	3298	3998	4645	5577	6893	9472	12391	13783	11126	11641	8666	9749	6264

Figure 3.342a shows forecasted HPV-related cancer cases for the different birth cohorts for females as they move up the different age groups. HPV-related cancer cases are increasing with increasing age, with the 75 and above age group having the highest number of cases, for the cohorts after 1944-1948. Cases increase with cohorts, until they reach a peak in the 1984-1988 cohort, then drops again.

Table 3.342b Forecasting of HPV cases by birth cohort and age-group-Males

Age	Birth Cohorts 1919-1923	1924-1928	1929-1933	1934-1938	1939-1943	1944-1948	1949-1953	1954-1958	1959-1963	1964-1968	1969-1973	1974-1978	1979-1983	1984-1988	1989-1993	1994-1998	1999-2003	2004-2008	2009-2014
0-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	6	3	3
5-9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9	6	7	5	4
10-14	-	-	-	-	-	-	-	-	-	-	-	-	-	15	14	10	14	10	9
15-19	-	-	-	-	-	-	-	-	-	-	-	-	35	27	28	23	34	23	21
20-24	-	-	-	-	-	-	-	-	-	-	-	30	35	29	35	30	43	29	26
25-29	-	-	-	-	-	-	-	-	-	-	29	32	41	40	48	40	57	39	35
30-34	-	-	-	-	-	-	-	-	-	58	64	77	115	112	135	112	162	110	100
35-39	-	-	-	-	-	-	-	-	99	100	119	167	249	243	293	244	352	238	217
40-44	-	-	-	-	-	-	-	174	155	169	237	331	495	483	582	486	699	473	432
45-49	-	-	-	-	-	-	236	227	218	280	391	547	818	798	962	802	1155	781	713
50-54	-	-	-	-	-	304	312	325	366	470	657	919	1374	1341	1615	1347	1940	1312	1198
55-59	-	-	-	-	273	276	307	375	422	541	757	1059	1583	1545	1861	1553	2236	1512	1380
60-64	-	-	-	257	284	312	406	496	558	716	1002	1401	2095	2045	2463	2055	2959	2000	1827
65-69	-	-	207	194	232	299	389	475	535	687	961	1343	2008	1961	2362	1970	2837	1918	1751
70-74	-	181	138	140	196	252	328	401	451	579	810	1133	1694	1654	1992	1662	2393	1618	1477
75+	196	202	166	197	277	356	463	566	637	818	1144	1600	2392	2335	2812	2346	3378	2284	2085

Displayed in figure 3.342b above are forecasted HPV-related cancer cases for the different birth cohorts for males as they move up the different age groups. Just like all other cancers, the cases are increasing while moving to the next age group. Cases increase until the 60-64 age group, then drops before reaching a peak in the 75 and above age group. The 1999-2003 cohort is expected to have the highest cases among all the cohorts

3.4 Forecasting of HPV-related cancer incidence-Under 3 scenarios

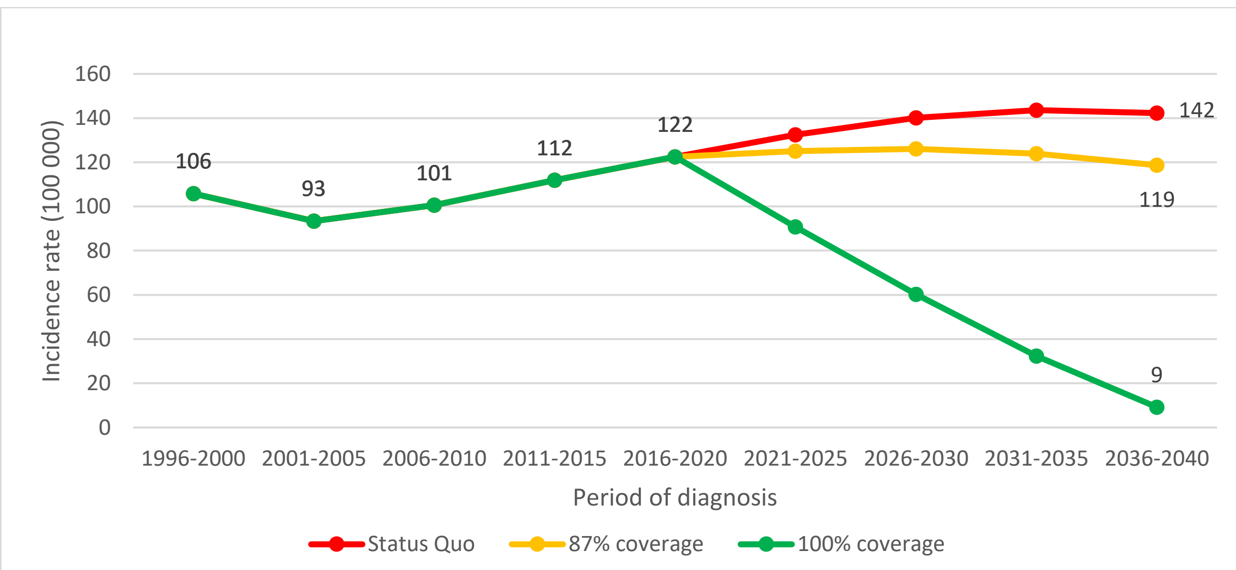
The sex-stratified APC was used for forecasting the HPV-related cancer incidence for the years 2016 to 2040. Three scenarios were examined; No intervention, HPV-vaccination of 87% (425 000-using 2014 coverage) girls who are 10-14 years old, and lastly, HPV vaccination of all girls who are 10-14 years (100% coverage), beginning from 2014.

From the review of literature, the following assumptions were established for this study;

1. Vaccinated girls were only aged 10-14 years
2. Only girls were vaccinated
3. Vaccine efficacy of 90% assumed

The forecasted cases from the sex-stratified APC model (Figure 3.341), factoring in the above assumptions, were used in the calculation of incidence rates (Figure 3.41). Because only girls are being vaccinated in South Africa HPV vaccine campaign, forecasting was done for females, then factored in the effects for the whole population.

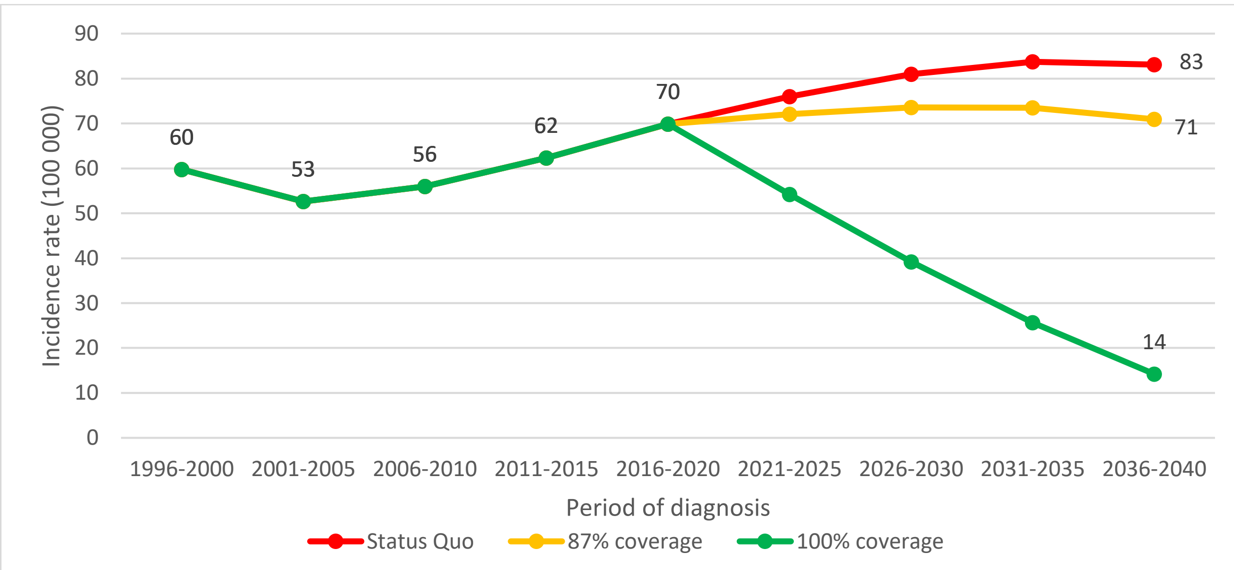
Figure 3.41 Forecasted incidence rates for females under the scenarios reviewed



Generally, an upward trend in HPV-related cancer incidence is observed for the years of the forecast. The incidence was 109 cases per 100 000 population, in 2016 and rose to 142 cases per

100 000 population in 2040. The HPV vaccination was considered from 2014, and the first effects are first seen in the 2021-2025 period. In 2021-2025, without any vaccination, the incidence stood at 132 cases per 100 000 population, with an 87% (425 000) vaccine coverage on girls aged 10-14 years old, the incidence decreases to 125 cases per 100 000 population, and further decreases to 91 cases per 100 000 population with a 100 % HPV vaccination of girls aged 10-14 years old. At the end of the forecasting period, in 2040, if no vaccine was administered, the incidence would reach 142 cases per 100 000 population. Vaccine coverage of 87% decreases the incidence to 119 cases per 100 000 population, and further reduces to 9 cases per 100 000 population, with a 100% vaccine coverage.

Figure 3.42 Forecasted incidence rates for both sexes under the 3 scenarios



In figure 3.42 above, we looked at the effects of HPV vaccination of girls age 10-14 years on the incidence rate of HPV-related cancer for both sexes combined. Again, if nothing is done, an upward trend in the incidence is projected, with the incidence rising from 70 cases per 100 000 population in 2016 to 83 cases per 100 000 population in 2040. There is a slight decrease in incidence rate when 87% coverage is achieved, as it rises from 70 cases per 100 000 population in 2016 to 71 cases per 100 000 population in 2040. A 100% yields a further reduction on the incidence at it decreases to 14 cases per 100 000 population in 2040, from the 70 cases per 100 000 population in 2016.

CHAPTER 4: DISCUSSION

This section details the results and findings of this study, comparing them to previous studies and what has already been published. The study's strengths and limitations are also discussed.

4.1 Summary of the main findings

The aim of the study was to estimate the age, period, and cohort-specific incidence of HPV-related cancers in South Africa over the next years from 2016 to 2040. A trend analysis was done for both AGC and OPC for the period of 1996 to 2015. There was a significant downward trend in AGC incidence for 1998 to 2003 and a significant upward trend from 2003 to 2015. The OPC showed a decreasing trend for 1996 to 2015 period.

The sex-stratified APC was used for predicting HPV-related cancer incidence from 2016 up to 2040. Our analysis predicted an increase in HPV cancer incidence for both males and females. We also predicted HPV-related cancer incidence under three scenarios; No intervention, HPV vaccination of 87% coverage, and a 100% HPV vaccine coverage. The incidence decreased when an 87% vaccine coverage is administered, a 100% coverage reduced the incidence further.

4.2 Age-Period-Cohort analysis of HPV-related cancer

Age:

Our model results showed an increase in HPV-related cancer cases with advancing age. A majority of studies pointed out that advancing age was the main and most important risk factor for any cancer type (50,83). This was the case for both males and females, where cases peaked around age 50, and then starts to drop. The mean age at diagnosis for the study was 52 years, a bit younger compared to all types of cancers mean age at diagnosis (61). This may be because a majority of HPV-related cancer cases are cervical cancer, which has 49 years as the mean age at diagnosis (61).

Period:

A majority of studies on cancer predictions projected an increase in cancer incidence/cases in the future (2,84). It only makes sense that our study found an increase in HPV-related cancer cases as we progress in the period of diagnosis.

Cohort:

Ideally, earlier cohorts are supposed to have the highest incidence rates/largest number of cancer cases because of the association between cancer and advancing age (61). This, however, holds if all cohorts are experiencing the same set of conditions e.g. HIV incidence, HPV vaccination. Our study shows an increase in HPV-related cancer cases as birth cohort advances, until around the 1955 and 1950 cohorts for females and males respectively. Cases then drop for the "younger" cohorts.

The results about the three main variables for our study show a similar trend to other APC studies conducted in other parts of the world.

4.3 Analysis of HPV-related cancer trend (AGC and OPC)

AGC had higher incidence rates than OPC for period under study. For AGC, there was a downward trend from 1998-2003 (APC -7.8%, p=0.001), following an increase in incidence rates from 1996-1998 although it was not statistically significant (APC 4.8%, p=0.434). An upward trend is seen from 2003 until the end of the period (APC 1.7%, p0.001). OPC also showed a downward trend for the years 1996-2015 (APC -1.2%, $p<0.001$). Chikandiwa et al also found a downward trend for OPC for 1996-2005 (APC -4.4%, p=0.01) (85). They also found different patterns for the individual AGC's as there were increasing trends for anal, vulva, and vaginal (although it was not significant), and declines in penile cancer (85).

4.4 Province of diagnosis of HPV-related cancers

Our analysis showed that a majority of HPV cancers are diagnosed in the Gauteng province, followed by the Eastern Cape. This is confirmed by a recent study done which showed that the Gauteng province had the highest screening rate for cancers and other conditions (86). This is true for the South African setup as cancer patients are referred from their Province of residence to these two Provinces for diagnosis and/or treatment. This then can only point out to lack of specialists, facilities, and equipment in the other Provinces.

Because cancer patients are being referred from their province of origin to another for diagnosis, it is worth noting that the province of diagnosis is not always the same as the province of residence. This then makes it difficult to estimate the distribution of cancer cases by province.

4.5 Prediction of Cancer Cases-2016-2040

Five-year projections of HPV-related cancer cases were produced for 2016-2040. Overall, HPV-related cancer cases are expected to increase from 2016 up to 2040. Like the previous analysis of the trend (1996-2015), females are expected to have the highest number of HPV-related cancer cases compared to men. This is true because HPV-related cancer consists of around 90% of cervical cancer, affecting females only. Although the study focused on HPV-related cancer alone, the projections are in line with other cancer estimates in general, which also projected an increase in cancer cases by the year 2040 (2,84,87).

4.6 Prediction of HPV-related Cancer Incidence-2016-2040 under 3 scenarios

The first scenario of projection was when no intervention is being implemented. When nothing is done, HPV-related cancer incidence is expected to increase from 70 to 83 cases per 100 000 population, from 2016 to 2040. This is in line with the majority of estimates for determining future cancer incidence, where these studies predicted an upward trend in cancer incidence up to 2030-2040 (2,88)

The second scenario studied concentrated on the estimation of HPV-related cancer incidence by taking into consideration the HPV vaccination campaign efforts. In 2014, HPV vaccination was administered to 87% (425 000) of school-going girls, who were less than 14 years old. Then assuming that on average, 425 000 girls are reached each year from 2014, with a vaccine efficacy of 90%, the HPV-related cancer incidence only drops to 71 cases per 100 000 population, by 2040.

The third and last scenario considered was an HPV vaccine campaign of all (100%) girls who are 10-14 years old. This means that on average, 2, 404, 691 girls would be vaccinated per year. This initiative would bring the incidence rate further down from 83 to 14 cases per 100 000 population, by 2040.

Other studies conducted also found out that HPV vaccination decreases future cancer incidence (28,29). The vaccine campaign coverage and the period the vaccine was first administered also contributed to the overall incidence estimation. The higher the coverage, the lower the expected incidence rate, and the earlier a strategy is implemented, the lower the incidence in the future (89).

4.7 Limitations

This study used secondary data for all the analyses conducted. This then means that only the data collected for the primary study could be used in our analysis. There was no simple and straight way of dealing with incomplete, missing information. Also, there was no way of correcting misclassified information from the primary study. This means that there might have been misclassification bias tolerated.

Because the primary studies are used for other purposes, other than research, key socio-demographic information was not collected. This limited the descriptive analysis part of the study, e.g, it would be incorrect to use province of diagnosis for Spatio-temporal analysis as this is different from the province of residence.

Data from STATS-SA are only aggregated and not available for the individual level. The analysis was limited to aggregated methods and techniques only as both STATS-SA and NCR data were key for the study. The NCR data were only available until 2015. There might have been changes in the distributions and patterns in HPV-related cancers at the time of analysis, which the study omitted.

It is of good practice to provide uncertainty bounds for predictions as this ascertains reliability of the predicted values. Confidence intervals were not provided for the predictions in this study. The model used only produced the estimated figures and there were not enough statistics for calculating the confidence intervals.

4.8 Strengths

The data used in the study were obtained from the NCR, where it undergoes strong quality checks and verifications before it can be used in any analysis. These processes include, but are not limited to; verifications before capturing, capturing, verifications, cleaning and removal of noise data, storage. This provided the study with rich quality data, from a good data management environment from the primary study.

This study also used a larger sample-sized data collected across the whole of South Africa. This, therefore, gives strong generalizability of the results obtained.

To our best knowledge, this is the first study to use the age-period-cohort model in South Africa.

CHAPTER 5: RECOMMENDATIONS AND CONCLUSION

This section details lessons learned from carrying out the study, study implications, and then provide a take-home message.

5.1 Recommendations

There is a need for revision of data collection tools from the NCR to cater for other important socio-economic variables that may be of research interest, e.g., province of residence to enable spatial analysis, lifestyle, educational and clinical factors to cater for risk factor analyses.

The province of diagnosis showed that the majority of cancers are largely diagnosed and treated in two or three provinces. Cancer patients would have to travel long distances to access up-to-standard healthcare, incurring travel and accommodation costs. In some cases, it may lead to premature cancer deaths as other patients may not afford to pay these bills. There is a need for decentralization of skills and efforts to cater to the whole country.

HPV vaccination provides one significant way of combating HPV-related cancers in the country. Increasing the coverage of girls being vaccinated would help reduce the incidence to lower rates. The inclusion of boys into the campaign could also present positive results. Other studies showed that synchronizing HPV vaccination and screening produced optimum results (89), this is another strategy that can be explored.

This study, backed by other studies around the globe, highlights that HPV-related cancer incidence is expected to increase until 2040. There is a need for planning for the future, in terms of cancer burden and ways of trying to bring this incidence down.

There are many APC models used in previous researches. Five (5) models were discussed in this study under literature review and three (3) were attempted. Future research can explore these other models.

In this study, we only explored HPV vaccination as a strategy to reduce HPV-related cancer incidence. For future research, many other strategies and risk factors (e.g., HIV incidence, HPV screening, etc.) need to be explored to give a precise estimation.

5.2 Conclusion

Results from our study show that age, period, and cohort are important covariates in studying patterns and distribution of cancer. Cancer incidence increased in advancing age, period and cohort, for both males and females. Our study estimations concluded that HPV-related cancer incidence is expected to increase in the future. This is in line with a lot of studies and predictions conducted by other bodies around the world.

We explored strategies that would help reduce this high predicted incidence. Our model shows that by 2040, for both sexes, HPV-related cancer incidence will increase from 70 cases per 100 000 cases in 2016 to 83 cases per 100 000 population. If a vaccine is administered to 87% percent of girls who are 10-14 years, the incidence will drop to 70 cases per 100 000 population; if 100% of girls who are 10-14 years are vaccinated the rate will drop to a low 14 cases per 100 000 population.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 Nov;68(6):394–424.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin [Internet]. 2021 Feb 4 [cited 2021 Feb 21];caac.21660. Available from: https://onlinelibrary.wiley.com/doi/10.3322/caac.21660

3. Cancer | WHO | Regional Office for Africa [Internet]. [cited 2021 Apr 10]. Available from: https://www.afro.who.int/health-topics/cancer

4. WHO report on cancer: setting priorities, investing wisely and providing care for all [Internet]. [cited 2021 Apr 10]. Available from: https://www.who.int/publications/i/item/who-report-on-cancer-setting-priorities-investing-wisely-and-providing-care-for-all

5. Kanavos P. The rising burden of cancer in the developing world. Ann Oncol. 2005;17:15–23.

6. (No Title) [Internet]. [cited 2020 Jan 15]. Available from: https://gco.iarc.fr/today/data/factsheets/populations/710-south-africa-fact-sheets.pdf

7. Cancer Today [Internet]. [cited 2021 Jan 5]. Available from: https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&group_cancer=1&include_nmsc=1&include_nmsc_other=1

8. Gilbert DC, Wakeham K, Langley RE, Vale CL. ARTICLE Increased risk of second cancers at sites associated with HPV after a prior HPV-associated malignancy, a systematic review and meta-analysis. Br J Cancer [Internet]. 2019 [cited 2019 Nov 18];120:256–68. Available from: https://doi.org/10.1038/s41416-018-0273-9

9. Arbyn M, de Sanjosé S, Saraiya M, Sideri M, Palefsky J, Lacey C, et al. EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease. Int J Cancer. 2012;131(9):1969–82.

10. Human Papillomavirus and Related Diseases Report SOUTH AFRICA [Internet]. [cited 2019 Nov 19]. Available from: www.hpvcentre.net

11. Mbulawa ZZA, Van Schalkwyk C, Hu NC, Meiring TL, Barnabas S, Dabee S, et al. High human papillomavirus (HPV) prevalence in South African adolescents and young women encourages expanded HPV vaccination campaigns. PLoS One [Internet]. 2018 Jan 1 [cited 2021 Jan 5];13(1). Available from: /pmc/articles/PMC5749739/?report=abstract

12. Mbulawa ZZA, Coetzee D, Williamson A-L. Human papillomavirus prevalence in South African women and men according to age and human immunodeficiency virus status. 2011;

13. UK CR. Worldwide cancer statistics | Cancer Research UK [Internet]. [cited 2019 Nov 19]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer#heading-One

14. Lekoane KMB, Kuupiel D, Mashamba-Thompson TP, Ginindza TG. Evidence on the prevalence, incidence, mortality and trends of human papilloma virus-associated cancers in sub-Saharan Africa: systematic scoping review. [cited 2019 Nov 19]; Available from: https://doi.org/10.1186/s12885-019-5781-3

15. Made F, Wilson K, Jina R, Tlotleng N, Jack S, Ntlebi V, et al. Distribution of cancer mortality rates by province in South Africa. Cancer Epidemiol. 2017 Dec 1;51:56–61.

16. STD Facts - Human papillomavirus (HPV) [Internet]. [cited 2021 Apr 21]. Available from: https://www.cdc.gov/std/hpv/stdfact-hpv.htm

17. Angeletti PC, Zhang L, Wood C. The Viral Etiology of AIDS-Associated Malignancies [Internet]. Vol. 56, Advances in Pharmacology. NIH Public Access; 2008 [cited 2021 Apr 1]. p. 509–57. Available from: /pmc/articles/PMC2149907/

18. Bray F, Møller B. Predicting the future burden of cancer. Nat Rev Cancer. 2006;6(1):63–74.

19. HPV infection - Symptoms and causes - Mayo Clinic [Internet]. [cited 2019 Nov 18]. Available from: https://www.mayoclinic.org/diseases-conditions/hpv-infection/symptoms-causes/syc-20351596

20. Steben M, Duarte-Franco E. Human papillomavirus infection: Epidemiology and pathophysiology. Gynecol Oncol. 2007;107(2 SUPPL.):S2.

21. Meites E, Szilagyi PG, Chesson HW, Unger ER, Romero JR, Markowitz LE. Human Papillomavirus Vaccination for Adults: Updated Recommendations of the Advisory Committee on Immunization Practices. MMWR Morb Mortal Wkly Rep. 2019 Aug 16;68(32):698–702.

22. Watson M, Saraiya M, Ahmed F, Cardinez CJ, Reichman ME, Weir HK, et al. Using population-based cancer registry data to assess the burden of human papillomavirus-associated cancers in the United States: Overview of methods. Vol. 113, Cancer. 2008. p. 2841–54.

23. Garland SM, Lee L yang. Human papillomavirus vaccination: The population impact [Internet]. Vol. 6, F1000Research. Faculty of 1000 Ltd; 2017 [cited 2021 Jan 5]. Available from: https://pubmed.ncbi.nlm.nih.gov/28663791/

24. HPV Vaccine Safety and Effectiveness | Human Papillomavirus | CDC [Internet]. [cited 2020 Feb 6]. Available from: https://www.cdc.gov/vaccines/vpd/hpv/hcp/safety-effectiveness.html

25. How effective is the HPV vaccine? [Internet]. [cited 2020 Feb 6]. Available from: http://www.hpvvaccine.org.au/the-hpv-vaccine/how-effective-is-the-vaccine.aspx

26. Markowitz LE, Hariri S, Lin C, Dunne EF, Steinau M, McQuillan G, et al. Reduction in human papillomavirus (HPV) prevalence among young women following HPV vaccine

introduction in the United States, National Health and Nutrition Examination Surveys, 2003-2010. J Infect Dis. 2013 Aug 1;208(3):385–93.

27. Spinner C, Ding L, Bernstein DI, Brown DR, Franco EL, Covert C, et al. Human papillomavirus vaccine effectiveness and herd protection in young women. Pediatrics. 2019 Feb 1;143(2).

28. 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 [Internet]. 2020 Feb 22 [cited 2021 Jan 5];395(10224):575–90. Available from: https://doi.org/10.1016/

29. Malagón T, Drolet M, Boily MC, 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.

30. Braaten KP, Laufer MR. Human Papillomavirus (HPV), HPV-Related Disease, and the HPV Vaccine. Rev Obstet Gynecol [Internet]. 2008 [cited 2021 Jan 5];1(1):2–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18701931

31. Heard I, Tondeur L, Arowas L, Demazoin M, Falguières M, Parent Du Chatelet I. Effectiveness of HPV vaccination on prevalence of vaccine genotypes in young sexually active women in France. J Infect Dis [Internet]. 2016 Dec 23 [cited 2021 Jan 5];215(5):jiw639. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiw639

32. Human papillomavirus vaccination - UpToDate [Internet]. [cited 2021 Jan 5]. Available from: https://www.uptodate.com/contents/human-papillomavirus-vaccination/print

33. Ngcobo NJ, Burnett RJ, Cooper S, Wiysonge CS. Human papillomavirus vaccination acceptance and hesitancy in South Africa: Research and policy agenda. S Afr Med J. 2018 Dec 13;109(1):13–5.

34. 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 [Internet]. [cited 2020 Feb 6]. Available from: www.ghspjournal.org

35. Fontham ETH, Wolf AMD, Church TR, Etzioni R, Flowers CR, Herzig A, et al. Cervical cancer screening for individuals at average risk: 2020 guideline update from the American Cancer Society. CA Cancer J Clin [Internet]. 2020 Sep 30 [cited 2021 Jan 5];70(5):321–46. Available from: https://onlinelibrary.wiley.com/doi/10.3322/caac.21628

36. Cervical Cancer Screening: Pap and HPV Tests – NCCC [Internet]. [cited 2021 Jan 6]. Available from: https://www.nccc-online.org/hpvcervical-cancer/cervical-cancer-screening/

37. Botha MH, Richter KL. Cervical cancer prevention in South Africa: HPV vaccination and screening both essential to achieve and maintain a reduction in incidence. Vol. 105, South African Medical Journal. South African Medical Association; 2015. p. 33–4.

38. Circumcision (male) - Mayo Clinic [Internet]. [cited 2020 Apr 18]. Available from: https://www.mayoclinic.org/tests-procedures/circumcision/about/pac-20393550

39. Nielson CM, Schiaffino MK, Dunne EF, Salemi JL, Giuliano AR. Associations between Male Anogenital Human Papillomavirus Infection and Circumcision by Anatomic Site Sampled and Lifetime Number of Female Sex Partners. J Infect Dis. 2009 Jan 1;199(1):7–13.

40. Larke N, Thomas SL, Dos I, Silva S, Weiss HA. Male Circumcision and Human Papillomavirus Infection in Men: A Systematic Review and Meta-Analysis. [cited 2020 Mar 10]; Available from: https://academic.oup.com/jid/article-abstract/204/9/1375/846860

41. 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.

42. Bosch FX, Albero G, Castellsagué X. Male circumcision, human papillomavirus and cervical cancer: From evidence to intervention. J Fam Plan Reprod Heal Care. 2009;35(1):5–7.

43. Peltzer K, Onoya D, Makonko E, Simbayi L. Prevalence and acceptability of male circumcision in South Africa. African J Tradit Complement Altern Med [Internet]. 2014 [cited 2021 Mar 17];11(4):126–30. Available from: /pmc/articles/PMC4202407/

44. Global HIV and AIDS statistics | Avert [Internet]. [cited 2021 Jan 5]. Available from: https://www.avert.org/global-hiv-and-aids-statistics

45. The World's Largest HIV Epidemic | Center for Strategic and International Studies [Internet]. [cited 2021 Jan 5]. Available from: https://www.csis.org/features/worlds-largest-hiv-epidemic-0

46. Goncalves PH, Montezuma-Rusca JM, Yarchoan R, Uldrick TS. Cancer prevention in HIV-infected populations [Internet]. Vol. 43, Seminars in Oncology. W.B. Saunders; 2016 [cited 2021 Apr 1]. p. 173–88. Available from: /pmc/articles/PMC4789150/

47. LIU G, SHARMA M, TAN N, BARNABAS R. HIV-positive women have higher risk of HPV infection, precancerous lesions, and cervical cancer: A systematic review and meta-analysis. AIDS. 2018;32(6):795.

48. 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 [Internet]. Vol. 14, Infectious Agents and Cancer. BioMed Central Ltd.; 2019 [cited 2021 Apr 1]. p. 12. Available from: https://infectagentscancer.biomedcentral.com/articles/10.1186/s13027-019-0228-7

49. Silverberg MJ, Chao C, Leyden WA, Xu L, Horberg MA, Klein D, et al. HIV infection, immunodeficiency, viral replication, and the risk of cancer. Cancer Epidemiol Biomarkers Prev. 2011 Dec;20(12):2551–9.

50. Taku O, Businge CB, Mdaka ML, Phohlo K, Basera W, Garcia-Jardon M, et al. Human papillomavirus prevalence and risk factors among HIV-negative and HIV-positive women

residing in rural Eastern Cape, South Africa. Int J Infect Dis [Internet]. 2020 [cited 2021 Jan 5];95:176–82. Available from: https://doi.org/10.1016/j.ijid.2020.02.051

51. Shiels MS, Engels EA. Evolving epidemiology of HIV-associated malignancies. Current Opinion in HIV and AIDS. 2017.

52. Rowhani-Rahbar A, Hawes SE, Sow PS, Toure P, Feng Q, Dem A, et al. The Impact of HIV Status and Type on the Clearance of Human Papillomavirus Infection among Senegalese Women. J Infect Dis [Internet]. 2007 Sep 15 [cited 2021 Apr 1];196(6):887–94. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1086/520883

53. Wang C ching J, Sparano J, Palefsky JM. Human Immunodeficiency Virus/AIDS, Human Papillomavirus, and Anal Cancer. Vol. 26, Surgical Oncology Clinics of North America. W.B. Saunders; 2017. p. 17–31.

54. Mehanna H, Franklin N, Compton N, Robinson M, Powell N, Biswas-Baldwin N, et al. Geographic variation in human papillomavirus-related oropharyngeal cancer: Data from 4 multinational randomized trials. Head Neck [Internet]. 2016 Apr [cited 2019 Nov 23];38(S1):E1863–9. Available from: http://doi.wiley.com/10.1002/hed.24336

55. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

56. Reddy P, Zuma K, Shisana O, Jonas K, Sewpaul R. Prevalence of tobacco use among adults in South Africa: Results from the first South African national health and Nutrition Examination Survey. South African Med J [Internet]. 2015 Aug 1 [cited 2021 Mar 21];105(8):648–55. Available from: http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S0256-95742015000800016&lng=en&nrm=iso&tlng=en

57. Mqoqi N, Kellett P, Sitas F, Phil D, Jula M. INCIDENCE OF HISTOLOGICALLY DIAGNOSED CANCER IN SOUTH AFRICA, NATIONAL CANCER REGISTRY. 1998.

58. Risk Factors: Age - National Cancer Institute [Internet]. [cited 2019 Nov 22]. Available from: https://www.cancer.gov/about-cancer/causes-prevention/risk/age

59. 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 [Internet]. 2018 Nov 1 [cited 2021 Mar 21];143(9):2238–49. Available from: http://doi.wiley.com/10.1002/ijc.31610

60. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Heal [Internet]. 2020 Feb 1 [cited 2021 Mar 21];8(2):e191–203. Available from: www.thelancet.com/lancetgh

61. HPV-Associated Cancer Diagnosis by Age | CDC [Internet]. [cited 2019 Nov 21]. Available from: https://www.cdc.gov/cancer/hpv/statistics/age.htm

62. Schmid VJ, Held L. Bayesian age-period-cohort modeling and prediction - BAMP. J Stat

Softw. 2007;21(8):1–15.

63. Holford TR. UNDERSTANDING THE EFFECTS OF AGE, PERIOD, AND COHORT ON INCIDENCE AND MORTALITY RATES [Internet]. Vol. 12, Annu. Rev. Publ. Health. 1991 [cited 2019 Nov 22]. Available from: www.annualreviews.org

64. Yang Y, Fu WJ, Land KC. A Methodological Comparison of Age-Period-Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models. Vol. 34, Methodology. 2004.

65. Bmed XZ. AGE-PERIOD-COHORT EFFECTS ON THE DEVELOPMENT OF COGNITIVE IMPAIRMENT AMONG THE ELDERLY. 2017.

66. Smith TR, Wakefield J. A Review and Comparison of Age-Period-Cohort Models for Cancer Incidence. Stat Sci [Internet]. 2016 [cited 2020 Jan 22];31(4):591–610. Available from: https://fortress.wa.gov/doh/wscr/

67. Smith T. A stratified age-period-cohort model for spatial heterogeneity in all-cause mortality [Internet]. [cited 2019 Nov 24]. Available from: http://www.mortality.

68. Li Y, Tiwari RC, Zou Z. An age-stratified Poisson model for comparing trends in cancer rates across overlapping regions. Biometrical J. 2008 Aug;50(4):608–19.

69. Riebler A, Held L, Rue H. Estimation and extrapolation of time trends in registry data-Borrowing strength from related populations. Ann Appl Stat. 2010 Mar;4(1):304–33.

70. Schmid L, Schmid VJ, Held L. Bayesian age-period-cohort modeling and prediction-BAMP [Internet]. Vol. 21, JSS Journal of Statistical Software. 2007 [cited 2019 Nov 24]. Available from: http://www.zora.uzh.chhttp//www.zora.uzh.chhttp://www.zora.uzh.chhttp://www.jstatsoft.org/

71. Chen W-Q, Zheng R-S, Zeng H-M. Bayesian age-period-cohort prediction of lung cancer incidence in China. Thorac Cancer [Internet]. 2011 Nov [cited 2019 Nov 24];2(4):149–55. Available from: http://doi.wiley.com/10.1111/j.1759-7714.2011.00062.x

72. Riebler A, Held L. Projecting the future burden of cancer: Bayesian age–period–cohort analysis with integrated nested Laplace approximations. Biometrical J. 2017 May 1;59(3):531–49.

73. Browning MJ, Crawford I, Knoef M, Browning M, Crawford I, Knoef M. The age-period cohort problem: set identification and point identification. 2012 Jan 23;

74. Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971-2006. Soc Sci Med. 2010 Apr;70(7):1100–8.

75. Age-Period-Cohort Analysis | Columbia University Mailman School of Public Health [Internet]. [cited 2019 Nov 21]. Available from: https://www.mailman.columbia.edu/research/population-health-methods/age-period-cohort-analysis

76. Land KC, Yang Y. Age-period-cohort analysis : new models, methods, and empirical applications. CRC Press; 2013.

77. National Cancer Registry | NICD [Internet]. [cited 2021 Mar 29]. Available from: https://www.nicd.ac.za/centres/national-cancer-registry/

78. Statistics South Africa | The South Africa I Know, The Home I Understand [Internet]. [cited 2021 Mar 29]. Available from: http://www.statssa.gov.za/

79. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin M, et al. ICD-O International Classification of Diseases for Oncology First Revision [Internet]. 2013 [cited 2021 Mar 29]. Available from: www.who.int

80. Bray, Freddie, Jacques F. Cancer Incidence in Five Continents: CHAPTER 7: AGE STANDARDIZATION Volume XI [Internet]. [cited 2021 Feb 24]. Available from: https://ci5.iarc.fr/CI5-XI/Pages/Chapter7.aspx

81. Average Annual Percent Change (AAPC) and Confidence Interval — Joinpoint Help System [Internet]. [cited 2021 Mar 29]. Available from: https://surveillance.cancer.gov/help/joinpoint/setting-parameters/method-and-parameters-tab/apc-aapc-tau-confidence-intervals/average-annual-percent-change-aapc

82. (No Title) [Internet]. [cited 2021 Jan 15]. Available from: http://rht.iconcologia.net/stats/sart/eapc/eapc_method.pdf

83. Risk Factors: Age - National Cancer Institute [Internet]. [cited 2019 Nov 21]. Available from: https://www.cancer.gov/about-cancer/causes-prevention/risk/age

84. Rahib L, Wehner M, Matrisian LM, Nead KT. Projection of cancer incidence and death to 2040 in the US: Impact of cancer screening and a changing demographic. J Clin Oncol. 2020 May 20;38(15_suppl):1566–1566.

85. Chikandiwa A, Pisa PT, Sengayi M, Singh E, Delany-Moretlwe S. Patterns and trends of HPV-related cancers other than cervix in South Africa from 1994–2013. Cancer Epidemiol. 2019 Feb 1;58:121–9.

86. Adonis L, An R, Luiz J, Mehrotra A, Patel D, Basu D, et al. Provincial screening rates for chronic diseases of lifestyle, cancers and HIV in a health-insured population. South African Med J [Internet]. 2013 [cited 2021 Feb 21];103(5):309–12. Available from: /pmc/articles/PMC3752603/

87. Sylla BS, Wild CP. A million africans a year dying from cancer by 2030: What can cancer research and control offer to the continent? [Internet]. Vol. 130, International Journal of Cancer. NIH Public Access; 2012 [cited 2021 Feb 21]. p. 245–50. Available from: /pmc/articles/PMC3244688/

88. Cancer Tomorrow [Internet]. [cited 2021 Feb 21]. Available from: https://gco.iarc.fr/tomorrow/en

89. Castanon A, Landy R, Pesola F, Windridge P, Sasieni P. Prediction of cervical cancer incidence in England, UK, up to 2040, under four scenarios: a modelling study. Lancet Public Heal [Internet]. 2018 Jan 1 [cited 2021 Feb 21];3(1):e34–43. Available from:

https://pubmed.ncbi.nlm.nih.gov/29307386/

APPENDICES

Appendix 1- Plagiarism Declaration Form

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Sandile Ndabezitha (Student number: 1232337) am a student registered for the degree of MSc Epidemiology in the academic year 2018.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: Date: 08/10/2021

Appendix 2- Additional results
<u>AAPC</u>

AGC

Segment	Lower Endpoint	Upper Endpoint	APC	Lower CI	Upper CI	Test Statistic (t)	Prob > |t|
1	1996	1998	4.8	-7.6	18.7	0.8	0.434
2	1998	2003	-7.8*	-11.4	-4.1	-4.5	0.001
3	2003	2015	1.7*	0.9	2.4	4.9	< 0.001

Range	Lower Endpoint	Upper Endpoint	AAPC	Lower CI	Upper CI	Test Statistic	P=
Full Range	1996	2015	-0.9	-1.8	0.0	-2.0	0.056

OPC

Range	Lower Endpoint	Upper Endpoint	AAPC	Lower CI	Upper CI	Test Statistic~	P= ~
Full Range	1996	2015	-1.2*	-2.1	-0.2	-2.6	0.019

Appendix 3-Human Research Ethics Committee (Medical) Clearance certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R14/49 Mr S Ndabezitha

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M200522

NAME: (Principal Investigator)	Mr S Ndabezitha
DEPARTMENT:	School of Public Health Medical School University
PROJECT TITLE:	Spatio-temporal age-period cohort analysis and prediction of Human Papillomavirus-related cancer incidence in South Africa
DATE CONSIDERED:	2020/05/29
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Drs I Maposa and M Muchengeti
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2020/07/13

This clearance certificate is valid for 5 years from the date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the 3rd Floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to submit details to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **May** and will therefore reports and re-certification will be due early in the month of **May** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

2020/07/22
Date