

UNIVERSITY OF KWAZULU-NATAL

**Age, Period and Cohort Analysis of Young Adult
Mortality due to HIV and TB in South Africa: 1997-2015**

By

Tshifhiwa Mildred Nkwenika

217079813

**A dissertation submitted in partial fulfilment of the
requirements for the degree
of
Masters in Statistics**

**Graduate School College of Agriculture, Engineering and
Science
Faculty of Science**

Supervisor: Prof. Samuel Manda

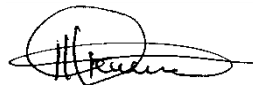
Co-Supervisor: Prof. Henry Mwambi

2019

Declaration

I TM Nkwenika, student 217079813 declare that

- i. The research reported in this dissertation/thesis, except where otherwise indicated, is my original research.
- ii. This dissertation/thesis has not been submitted for any degree or examination at any other university.
- iii. This dissertation/thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
- iv. This dissertation/thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a. their words have been re-written but the general information attributed to them has been referenced;
- v. This dissertation/thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation/thesis and in the References sections.



Tshifhiwa Mildred Nkwenika

01/03/2020

Date



Professor S. Manda

01/03/2020

Date

Prof Henry Mwambi

Date

Dedication

To my mother Mrs Nancy Nemaorani and my daughter Larhandzhekaka Nkwenika

Abstract

Young adult mortality is very important in South Africa with the impact of Human Immunodeficiency Virus /Acquired Immune deficiency Syndrome (HIV/AIDS), Tuberculosis (TB), injuries and emerging non-communicable diseases (NCDs). Investigation of temporal trends for adult mortality associated with TB and HIV has often based on age, gender, period and birth cohort separately. The overall aim of this study was to estimate age effect across period and birth cohort; period effect across age and birth cohort; and birth cohort effect across age and period on TB and HIV-related mortality. Mortality data and mid population estimates were obtained from Statistics South Africa for the period 1997 to 2015. Observed HIV/AIDS deaths were adjusted for under-reporting while adjustments for the misclassification of AIDS deaths and the proportion of ill-defined natural causes were made. Three-year age, period and birth cohort intervals for 15-64 years, 1997-2015 and 1934-2000 respectively were used. Age-Period-Cohort (APC) analysis using the Poisson distribution was used to compute effects of age, period and cohort on mortality due to TB and HIV. A total of 5, 825,502 adult deaths from the period 1997 to 2015, of which 910,731 (15.6%) were TB deaths while 252,101 (4.3%) were HIV deaths. A concave down association between TB mortality and period was observed while an upward trend was observed for HIV-related mortality. The estimated TB relative mortality showed a concave down association with age, a peak at 36-38 years was found. There was a concave down relationship between TB relative risk between 1997 and 2015. Findings showed a general downward trend between TB mortality and birth cohort, which 1934 cohort had higher rates of mortality. There was an inverse flatter U-shaped association between age and HIV-related mortality, where 30-32 years was more pronounced. An inverse U-shaped relationship between HIV-related mortality and period from 1997 to 2015 was estimated. An inverted V-shape relationship between birth cohort and HIV-related mortality was estimated. The study has found an inverse U-shaped association between TB-related mortality and age, period and general downward trend with birth cohort for deaths reported between 1997 and 2015. A concave down relationship between HIV-related mortality and age, period and inverted V-shaped with birth

cohort was found. The association between HIV-related mortality and period differs from the officially reported trend with adjustment, which shows an upward progression. Our findings are based on a slight advanced statistical model using Age-Period-Cohort. Using APC analysis, we found a secular trend in TB and HIV-related mortality rates which could contribute certain clues in long-term planning, monitoring and evaluation.

Acknowledgements

The fulfilment of this project has been through the guardianship and mentorship of my supervisors **Professors Samuel Manda and Henry Mwambi**. I am grateful for their support, entrain and advice.

To **Dr Steve Olorunju**, thank you for the support and guidance you accorded me.

I thank my family, friends and colleagues for their moral support.

I also acknowledge the support from my employer South African Medical Research Council

Lastly, I thank Almighty God for having given me strength from day to day so I could make it this far.

Table of Contents

Declaration	i
Dedication	ii
Acknowledgements	v
List of Figures	viii
List of Tables	ix
List of Abbreviations and Acronyms	x
CHAPTER 1: INTRODUCTION	1
1.1 Background	1
1.3 Related Studies (Age-Period-Cohort Analysis)	6
1.4 Overall Objectives.....	8
CHAPTER 2: METHODOLOGY	10
2.1 Introduction	10
2.2 Overview of Age-Period-Cohort Analysis.....	10
2.3 Modelling.....	12
2.4 Ethical considerations	17
3.1 Introduction	18
3.2 Data	18
3.3 Results.....	19
3.4 HIV-related mortality data	22
3.5 Analysis of the modelled data with reference to HIV/AIDS and TB related causes of death	26
3.6 Data exploration by three-year age, period and birth cohort intervals.....	26
3.7 Mortality relative risk (RR) for TB and HIV cause-specific.....	29
CHAPTER 4: DISCUSSION AND CONCLUSION	39
4.1 Introduction.....	39
4.2 Discussions.....	39
4.3 Limitations of the study	41
4.4 Recommendations for Further Research.....	42
4.5 Conclusions.....	42

REFERENCES	44
APPENDIX I	53
APPENDIX II	55
APPENDIX III	56
APPENDIX IV	58

List of Figures

Figure 1: Observed TB and HIV-related mortality trends by age groups	21
Figure 2: Observed and revised HIV/AIDS death rates for the 15 to 64-year-old South African Population	25
Figure 3: Percentage distribution of observed TB mortality by age groups over the 3-year period and 3-year birth cohorts	28
Figure 4: Percentage distribution of observed HIV-related mortality by age groups over a 3-year period and 3-year birth cohorts	29
Figure 5: Age, Period and cohort effects of TB mortality. Notes: dotted lines represent 95% confidence intervals.	34
Figure 6: Age, Period and cohort effects of TB mortality. Notes: dotted lines represent 95% confidence intervals.	35
Figure 7: Age, period and cohort effects of HIV-related mortality. Notes: dotted lines present 95% confidence intervals.	36
Figure 8: Age, period and cohort effects of HIV-related mortality. Notes: dotted lines present 95% confidence intervals.	37

List of Tables

Table 1: Distribution of deaths by cause-specific and age groups, 1997-2015.....	19
Table 2: Specific mortality rates calculated for each cause and age groups for each year	20
Table 3: Summary of AIDS mortality Adjustments in the 15 to 64-year-old South African Population: 1997-2015	25
Table 4: Age-period-cohort estimates for TB cause-specific mortality	32
Table 5: Age-period-cohort estimates for HIV cause-specific mortality	33
Table 6: Summary of the goodness of fit	38

List of Abbreviations and Acronyms

AIC	Akaike Information Criterion
AIDS	Acquired Immune Deficiency Syndrome
APC	Age-Period-Cohort Analysis
BIC	Bayesian information criterion
CGLM	Constrained Generalised Linear Models
DHA	Department of Home Affairs
DOH	Department of Health
HIV	Human Immunodeficiency Virus
ICD	International Classification of Diseases
IE	Intrinsic Estimator
PCA	Principal Component Analysis
SA	South Africa
SAMRC	South African Medical Research Council
STATSSA	Statistics South Africa
TB	Tuberculosis
UN	United Nations
WHO	World Health Organization

CHAPTER 1: INTRODUCTION

1.1 Background

Crucial indicators of mortality including age-specific, gender-specific mortality rates, expectation of life from birth together with the leading causes of death are useful for public health planning, implementation of monitoring and evaluation, distribution of resources as well as making major and minor decisions about the population (WHO 2018). This information allow government, at both national and subnational level, to develop strategic plans to administer necessary resources in order to reduce burden of diseases among limited resources settings (STATSSA 2015, STATSSA 2014 and Bradshaw et.al 2012). The knowledge of young adult mortality is very important in South Africa to provide reliable population estimates and projections that underlie planning in any sectors, especially with an influence of Human Immunodeficiency Virus /Acquired Immune deficiency Syndrome (HIV/AIDS), Tuberculosis (TB), injuries and emerging non-communicable diseases. In order to produce health outcome rates, population statistics are needed at both national and subnational level. However, the availability of vital registrations is usually the limiting factor (STATSSA 2014 and Bradshaw et.al 2012).

Mortality rates are often measured mathematically as the number of observed deaths over a studied population at a given period of time (Simpfendorfer, Bonfil, and Latour, 2005). These mortality rates are then determined using occurring death ratios during specified period of time to overall exposed population (persons-years) among which deaths occurred. In most cases, these mortality rates are multiplied by 1000, 10 000 or 100 000 depending on the frequency of deaths. Total observed deaths are often essential for comparisons of changes in time trends between age groups, gender and population groups at a given time. Therefore, to compare differences in mortality among age groups, gender and population groups over time,

total observed deaths must be associated to the exposed population at risk during the specific period.

Mortality can be modelled using survival models of analysis where a set of statistical approaches is used to investigate the lapsed time for death to occur (Collet 2015). The definite time of survival for an individual, t (*time to mortality*), could be considered as the value of the confounding variable, T (*observing Time*), which is an element of real numbers. T defines a random variable associated with an individual survival time and Probability distribution is formed as values of T are observed. Assuming that a random variable T has a probability distribution with probability density function $f(t)$, the cumulative distribution function of T is given by $F(t) = P(T < t) = \int_0^t f(u)du$, which implies that the probability of definite survival time is greater than the survival time. The function for survival time, $S(t)$, is defined to be the probability that the survival time is greater or equal to t , and so $S(t) = P(T \geq t) = 1 - F(t)$. Therefore, the probability that an individual survives from the time of existence to some observed period beyond the definite survival time t is represented by the survival function.

TB and HIV-related mortality have remained among the top 10 leading causes of deaths in South Africa for many years (STATSSA 2015, STATSSA 2017, Pillay-van Wyk 2016). According to Johnson and Dorrington (2017), deaths attributable to AIDS-related causes in 2017 were estimated at 230,000 of all average deaths while TB alone was responsible for over 33,063 deaths in 2016 (UNAIDS 2016). Mortality patterns associated with age strongly reflects a combination of a simple underlying ageing effect, which can be used as a good indicator of the epidemiological profile of a specific population (Ngom and Clark 2003). According to Herbest, Mafojane and Newell (2011), the majority of mortality attributable to TB and HIV are among age group 25-54 years, which is the most economically active population. Therefore, to mitigate the possible negative impact on human development, it is

fundamental to measure mortality levels and patterns of mortality due to TB and HIV-related causes during adulthood.

The influence of HIV/AIDS on the age pattern of adult mortality has been reported to be striking with a much-focused effect at reproductive age (Pillay-van Wyk et al. 2016). Statistics South Africa (2017) reported that 12.4% of mortality between age group 15-44 years was attributable to TB, while 10.5% were caused by HIV-related causes in the year 2017. Although there are implications taken to reduce and eliminate the burden of HIV and TB in South Africa, mortality due to TB and HIV/AIDS remain high (STATSSA 2017).

In South Africa, TB and HIV-related causes are ranked the leading infectious cause of death since 1997 and remained unchanged till date (STATSSA 2015, Pillay van Wyk et al. 2016, Adeiza, Abba and Okpapi 2014). The rates of mortality due to both HIV and TB combined were observed to have rapidly increased from 1997, reached a peak during the period 2006, and then declined with time periods (STATSSA 2015). The interaction between HIV and TB in the sub-Saharan countries has also been associated with higher mortality rate, where as a result, 25% of deaths were attributable to HIV in 2010 compared to 12% in the year 2000 (Adeiza, Abba and Okpapi. 2014). The downward trend of HIV and TB mortality rates evaluated by period in years has been observed in most African countries such as Ethiopia, Kenya, Mozambique and Nigeria during the year 2004-2008 (Bendavid et al. 2012). A study conducted in South Africa also found a rapid decline (from 56% to 39%) of mortality due to HIV during the year period 2000-2009 (Herbst, Mafojane and Newell 2011).

Estimations of mortality are often analysed as age-specific, gender-specific and cause-specific, as exemplified by the annual mortality report given by the Statistics South Africa, South African Medical Research Council (SAMRC), Department of Health (DOH) and World Health Organization (WHO). An investigation of mortality trends over a long period of time by various demographic characteristics provides some crucial information and portrays a better picture to understand mortality. Such crucial information on mortality trends by age-,

gender-, race-, and cause-specific classifications have been reported in South Africa (STATSSA 2015, Dorrington et al. 2018, Groenewald et al. 2017). Furthermore, an application of statistical analysis could also provide a better insight of the effects of demographic characteristics on mortality patterns at a specific time. 1.2 Statistical methods of analyses of HIV and TB mortality

Studies to investigate TB and HIV-related mortality have been undertaken globally including (Biset ,2017), Gesesew et al. (2016), Lima et al. (2016) Podlekareva et al. (2014)) since their epidemic were different statistical methods have been adopted. Some of the studies sought to describe epidemiological patterns, determine predictors of TB and HIV-related mortality rates and distinguished the relationship between TB and HIV-related mortality. A study conducted in Ethiopia assessed the predictors of mortality among HIV infected population who were known to be taking antiretroviral treatment using meta-analysis (Biset 2017). Factors like advance disease stage, low CD4 count, baseline weight and poor treatment adherence were significantly associated with HIV-related mortality. Other factors which were found statistically significant to HIV/TB co-infected patients in Southwest Ethiopia were age (35-44), being a female sex worker, bed ridden patients and disease stage, Gesesew et al. (2016). Gesesew et al. (2016) study adopted Cox regression model to assess the predictors of mortality. Furthermore, similar factors were reported by Aung et al. (2018) for the TB-HIV co-infected patients aged 15 years and above of Myanmar, using cox proportional hazard model.

In addition, cox proportional hazards regression has also been utilized to estimate the HIV-cause specific associated and the outcome of antiretroviral treatment to reduce mortality among TB patients in Kenya (Onyango et al. 2017). The study found that the probability of mortality out of patients who were not on treatment were four times greater compared to patients receiving treatment. Podlekareva et al. (2016) also adopted Cox models to study the TB-HIV co-infected mortality among the European and American population. The study showed that mortality differs in among this group of people differs by residential area.

Application of Cox proportional hazards regression models showed its significance in comparing hazards ratios and predicting associated factors with TB and HIV-related mortality in Rio de Janeiro, Brazil, da Silva Escada et al. (2017).

Logistic and Poisson regression models have also gained momentum to predict factors contributing to HIV-related and TB mortality. Podlekareva et al. (2014) studied short and long-term mortality and causes of death in HIV/TB-related mortality in Europe using Poisson regression while Takarinda et al. (2017) used logistic regression to determine the factors related with mortality among TB patients in Zimbabwe. Gaifer (2017) reported confounders for TB mortality in Oman for the period of 10 years (2006-2016) using logistic regression was applied. Risk factors reported in Zimbabwe were not different from risk factors reported in Oman. In Texas, a multiple logistic regression was utilised to develop and validate a prognostic score for predicting mortality during TB treatment in TB-HIV co-infected population (Nguyen, Jenkins and Graviss 2018). Time series analysis was found useful to describe the patterns and time trends of TB-HIV-related mortality in Brazil (Lima et al. 2016).

In South Africa, a structural equation modelling approach was applied to examine HIV/TB mortality determinants and the spatial allocation in the rural regions (Musenge et al. 2013). Spatial models in this study showed that areas without any health facilities experienced highest child HIV/TB mortality while Namosa et al. (2013) found that peri-urban communities of South Africa were associated with higher HIV-related mortality. A spatio-temporal analysis was used to examine HIV-associated mortality in rural western Kenya, (Sifuna et al. 2018) where hot spot analysis was used. Queiroz et al. (2018) described epidemiological profile, analysed the spatial patterns and investigated temporal trend of mortality due to TB in Northeast Brazil and geographical areas with high prevalence of TB mortality. Studies conducted across countries show available evidence that Spatio-temporal models are essential in studying TB and HIV-related mortality across geographical areas.

Models of analyzing count data have been found useful, especially in the field of epidemiology where rates of different health outcomes can be measured over time (Tetteh-Ahinakwa and Oduro 2017, Russo et.al 2012, Bourne et al. 2009, AA and Naing 2012, Osagie and Adebukola 2017). These models produce a formal technique to investigate potential interrelations on the average mortality rate and the specific risk factors with respect to the exposed population. Furthermore, they also allow comparisons of health outcome rates among different exposure groups under investigation controlling for confounding factors, where confounders are estimated independently. On the other hand, age, period, and cohort model (APC) have been found profitable to analyse the attributes of mortality trends over the period of time (Wang et.al 2015). This model estimates the effects age, period and birth cohort simultaneously on mortality rates, where the researcher investigate the manner in which mortality among the particular age groups changes with time and by cohort and how age-specific death rate pattern changes related to birth cohorts (Yang and land 2013). In addition, it may be used to better depict the demographical, environmental and social factors that jointly affect individuals and study population over a given period (Yang and Land, 2016). Furthermore, to yield a better overview of mortality over a range of studied populations; age, period to mortality and birth cohort should be taken into consideration. Nevertheless, there are no studies in South Africa that examined the trend of TB and HIV-related mortality using age-period-cohort (APC) model, which has been reported to be more effective than traditional cross-sectional analysis in analyzing trends (Chang et al. 2011, Gao et al. 2017). Therefore, this study used APC model to estimate age, period and birth cohort effects jointly on TB and HIV-related mortality using a Poisson distribution on the counts of cause-specific deaths, with log-linear link on the Poisson mean.

1.3 Related Studies (Age-Period-Cohort Analysis)

The utilisation of the APC models is a widely known tradition in medical sociology. It measures the variation of birth, corresponding with the shifts in the population that are exposed to risk factors over time. The APC models have been used for the analysis of the

trends in the long-term mortalities of various chronic diseases including cancer, diabetes, epilepsy etc. The significance of the APC analysis has been reported in measuring the incidences, prevalence and the mortality trends in different parts of the world. Suddenly, APC analysis is gaining momentum in analysing the cause of mortality worldwide (Ocaña-Riola et.al 2013, Chang et.al, 2011, Gao et.al, 2017).

An effect of overall mortality in Andalusia was reported by Ocaña-Riola, Mayoral-Cortés and Blanco-Reina (2013) where Lexis diagram was used to tabulate the death per individual's age and cohort, and the period effect was evaluated. The effects of age, period and cohort in all geographical areas of Andalusia were evaluated and a tangible evidence about the time trends of all age groups mortality was provided.

A study conducted in Taiwan employed the three phases of age-period-cohort analysis to compare the pattern of epilepsy cause-specific mortality in European Union and United Kingdom (Chang et al. 2011). Age and period effects were found to be statistically significant, though the cohort effect was not significant. Regardless of the age and period effects found, the overall effects of the epilepsy mortality in Taiwan were different from those of European Union countries compared to the United Kingdom countries.

Gao et al. (2017) employed the APC analysis to describe the time trends of mortality due to oesophageal cancer in China. Of the recorded deaths from 1989 to 2013, the effect of age in this cohort was reported from the age of 30 to 84 years while the period and cohort effect declined from 1989 to 2003 and 1910 to 1980s respectively.

The first study to use the APC analysis to compare the trends of breast cancer across East Asia to the United States trend was conducted by Wang et al. (2015). The APC models estimated when and how the three dependent covariates affected the breast cancer cause-specific mortality trends. At age 20 to 54, there was an increase in mortality due to breast cancer

associated with age in all areas. However, after age 55 the United States pattern differed from the East Asian.

Another study aimed to investigate the effects of age, period and cohorts on the temporal trend of type I diabetes in children under 15 in Italy (Bruno et al. 2010). Poisson regression models were used to model sex, age, period and cohort effects considering registry-level variance component. The incidence of type I diabetes among the children in Italy was reported to have increased within the time frame of the study. However, the increase could not be ascribed to either the calendar time or birth cohort. The use of APC model provided evidence on the ability to monitor disease incidence rate.

These models have not only gained popularity in studying effects of three factors on mortality and burden of diseases but have also been utilised to study the effects on incidence and prevalence based on social, cultural and environmental factors. Jee and Cho (2016) conducted a study to identify the contributions of age, period and birth cohort effects on smoking prevalence in young adult's men in Korea. Age-period-cohort analysis was applied with the aim of identifying independent effects, where the smoking prevalence was adjusted to 2008 population. Results of this study showed a significant increase with age, increasing with smoking prevalence and declining with the increasing of period and cohort.

1.4 Overall Objectives

Different statistical methods including logistic regression, poisson regression, cox proportional and spatio temporal analysis have been used to estimate mortality due to TB and HIV-related cause. Given the array of methods available, the main purpose of this study is to estimate age effect across period and birth cohort, period effect across age and birth cohort, and birth cohort effect across age and period on TB and HIV-related mortality. This was done using Age, period, cohort analysis which has not been applied to this type of data before. Mortality data and Mid population estimates acquired from statistics south Africa for the

period 1997-2015 will be used to describe time trends of TB and HIV-related mortality. This will be given by age specific mortality rates per 100 000 residents of South Africa. Estimations of the mortality are often analysed as age-specific, gender-specific or cause-specific separately, which may provide a better understanding of the mortality rates investigated. For this study, effects of age, period and birth cohorts will be estimated simultaneously. To estimate effect of age, period and birth cohort, the number of deaths over 3-year interval for age, period and birth cohorts for 15-64, 1997-2015, and 1934-2000 respectively will be drawn from the data. The effects of age, period and birth cohort will be estimated using a Poisson distribution on the counts of cause-specific deaths, with log-linear link on the Poisson mean.

CHAPTER 2: METHODOLOGY

2.1 Introduction

This chapter described an overview of the utilised methods, study site, and the population as well as the variables of interest of the data used for this study. Age-period-cohort (APC) models are described fully.

2.2 Overview of Age-Period-Cohort Analysis

APC models were firstly introduced by demographers and sociologists in the 1970s and have been proven to provide insights into the effects of age, period and cohort on observed changes over various health outcomes. Analysis of age, period and cohort models have widely been used in descriptive epidemiology to evaluate the mortality trends, prevalence, and incidences.

Age effects measures the variations related with various age groups created by aggregation of experience, and physiological diversity. Age effects measures the features of individual's ageing processes and developments observed during the lifetime. Furthermore, period effects evaluate changes of health outcomes over time which are influenced by all age groups simultaneously (Yang and Land 2008). Historical events and environmental factors including epidemics and pandemics of infectious diseases are represented by effect of period. Lastly, birth cohort effects determine the differences across a group of people who experiences the health outcome in the same year. Usually, birth cohort effect arises as a result of mortality distributions occurring from an exposed group separately impact age groups.

The APC models parameters combine time effects for age, period and cohort, where age effects are observed within the age groups while a birth cohort effect is considered an association due to period effect (Holford 1983). In most studies where the APC models have

been utilised, the available evidence suggested that the analysis can provide useful interpretation of the time trends of a disease and mortality rates (Ocala-Riola, Mayoral-Cortés and Blanco-Reina 2013, Rosenberg and Anderson 2011). Moreover, these models perform a crucial role in understanding the time-varying elements of burden of diseases and mortality.

Although APC models are vital to study time trends of incidence and mortality rates, the model suffers the identification problem (Holford 1983, Yang and Land 2008, Yang, Fu and Land 2004). Age, period and cohort are known to be highly correlated, which causes the identification problem between the three factors. That is, if the value of two of these variables are given, the third variable can be estimated ($\text{period} = \text{age} + \text{cohort}$). Several researchers have attempted to solve the identification problem where different methods have been proposed including, using prior information to impose parameter constraints, non-informative constraints causal models to approximate for age, period and cohort effects, and estimable functions of the APC model parameters, intrinsic estimator and application of Bayesian smoothing models (Yang and Land 2008, Luo 2013, Holford 1983). These attempts, were started decades ago, continue to date.

Conventional APC models such as Constrained Generalised Linear Models (CGLIM) have been used to resolve an identification problem by finding a relevant constraint of either age, period or birth cohort (Yang and Land 2008, Yang et al. 2008). However, the method has significant limitations; regarding finding the relevant constraint, which depends on the additional, a priori, and information. However, researchers have found it difficult to acquire or verify that prior information. This method is useful to establish the degree of sensitivity to various changes in coefficient constraints. The results from the method, however, are not fixed and are difficult to interpret (Yang and Land 2008). Therefore, recent researchers of APC methodology have studied the estimable function, which is not dependent on the variation of constraints.

Apart from using constrained GLM, Bayesian Ridge Estimation of Age-Period-Cohort Models can also be considered to solve collinearity problem using ridge priors (Duzan and Shariff 2015, Xu and Powers 2016). Ridge estimator has been reported as appropriate technique to find the desirable parameters for age, period and birth cohort under investigation.

Another alternative technique suggested by Yang et.al (2004) is Intrinsic Estimator (IE) model. This model has been proven to yields trustworthy estimates of age, period and birth cohort on mortality and morbidity, which determine the unique coefficients. Through empirical analysis utilising the IE model, Yang et al. (2004) concluded that it can be a useful alternative to conventional methods of APC. For any fixed number of periods, the IE model is not only unbiased but has a smaller variance compared to other conventional model. Luo (2013) assessed the validity and application scope of IE theoretically where constraint depends on the number of age groups, periods, and birth cohorts' categories has and non-trivial inferences for estimations.

2.3 Modelling

To study age effect across period and birth cohort; period effect across age and birth cohort; and birth cohort effect across age and period on mortality due to TB and HIV from 1997 to 2015 in South Africa, Age, Period and Cohort analysis was undertaken. The usual APC model is an additive model for the log rate of mortality. Let $i(1, \dots, I)$ index the age groups, where age group 1 includes 15-17 years old, age group 2 includes 18-20 years old, and so on ; $j(1, \dots, J)$ index 3 year period, with period 1 as 1997-1999, period 2 as 2000-2002, and so on; and $k(1, \dots, K)$ index cohort. Assuming that age and period are divided into equally spaced intervals, we have, $I=17$, $J=7$ and $K=23$, that is, $K = (I + J) - 1$ (Appendix I). The following assumptions were decided upon during the model construction.

The classical Lee-Carter (LC) method for modeling TB and HIV-related mortality rates is given by

$$\mathbf{ln}(R_{ijz}) = \mathbf{a}_i + \mathbf{b}_i \mathbf{t}_j + \varepsilon_{ijz} \quad (1)$$

Where, R_{ijz} denotes the matrix of age-specific mortality rates for age i at time period j and for a cause-specific z , a_i represents the mean age-specific mortality, b_i is the rate of change in TB and HIV-related mortality at age i , t_j indicates the TB and HIV-related mortality index at period j and ε_{ijz} is the residual term at age i and period j (Lee and Carter (1992)).

The Lee-Carter model assumes that residual term at age i and period j are distributed normally with mean 0 and variance σ^2 , while a change of mortality remain fixed over time for all defined age-groups and overall trend (t_j) is fixed for all age at a given period. Our model parameter for mean age-specific mortality rates (a_i) over time period j is then estimated as follows:

$$\hat{\mathbf{a}} = \frac{1}{j} \sum_{t=1}^j \mathbf{ln}(R_{ijz}) \quad (2)$$

According to Lee and Carter (1992), the age and period parameters on the right side of (1) cannot be observed, therefore ordinary least squares (OLS) approach cannot be used to estimate the parameters of b_i and t_j . Since the OLS approach is not plausible, Lee and Carter proposed a singular value decomposition (SVD) matrix to obtain the least squares estimates by transforming age-specific mortality vector into a mortality index scalar where the parameters are subjected to the following constraints:

$$\sum_{i=1}^I \mathbf{b}_i^2 = \mathbf{1}, \quad \sum_{j=1}^J \mathbf{t}_j = \mathbf{0} \quad (3)$$

From equation (1) and (2), b_i and t_j parameters are estimated by creating a matrix

$$\mathbf{A}_{ij} = (\mathbf{ln}R_{ij}) - \hat{\mathbf{a}}_i \quad (4)$$

Which can be broken down into three matrices; age component (A), singular values (X) and period component (T) by applying SVD which takes a form

$$\mathbf{SVD}(\mathbf{A}_{ij}) = \mathbf{AXT}' \quad (5)$$

By applying constraints of b_i and t_j (3), the estimates of b_i and t_j are therefore given by

$$\hat{\mathbf{b}}_i = \frac{1}{\sum_j \mathbf{a}_{i1}^2} \cdot (\mathbf{a}_{1,1} \ \mathbf{a}_{2,1} \ \mathbf{a}_{3,1} \ \dots \dots \ \mathbf{a}_{i,1}) \quad (6)$$

and

$$\hat{t}_j = \sum_i \alpha_{i,1}^2 \cdot X_1 \cdot (t_{1,1} \ t_{2,1} \ t_{3,1} \ \dots \ t_{j,1}). \quad (7)$$

The estimates of log of cause-specific mortality are therefore estimated by

$$\ln \hat{R}_{ijz} = \hat{a}_i + \hat{b}_i \hat{t}_j = \hat{a}_i + \hat{A}_{ij} \quad (8)$$

Although LC model can yield accurate average mortality rates for countries with available vital registrations over time, it has some limitations. Based on the assumption that overall mortality trend remains fixed over all age groups for a given period, this result to an over estimation of the modelled mortality. Furthermore, LC model limit the mortality rate of each age given that the log of mortality rate could be a linear function of time, which eventually causes mortality rates of all ages to be 0. Since the estimates of LC are over-parametrized and is implausible to be applied to many country- and cause-specific mortality data which it was not designed for. Renshaw and Haberman (2006) introduced the cohort extension of the age-period model given by equation (1) is defined by

$$\log(R_{ij}) = \alpha_i + b_i t_j + b_i \gamma_{(j-i)} + \epsilon_{ij} \quad (9)$$

Here $\gamma = (\gamma_{1-i}, \dots, \gamma_j)$ represents cohort effects. Again, this model is over-parametrized; due to the non-identifiable problem ($j = i + k$). Here we have $2i - 1$ parameter for the age effects and $j - 1$ for the period effects.

Identification Problem

Suppose the effect of age, period and cohort are given by α_{iz} , β_{jz} and γ_{kz} for deaths due to diseases $z = 1$ (TB) and $z = 2$ (HIV). Assuming that the three factor has an additive effect on the log rate, one can model the natural log as;

$$\ln(R_{ijkz}) = \mu + \alpha_{iz} + \beta_{jz} + \gamma_{kz} + \epsilon_{ijkz}, \quad (10)$$

which can be fitted using linear model and ϵ_{ij} is the random error. However, direct interpretation of these effects is difficult because the model is over parametrized. There are two sources of identifiability to consider. The simpler one to account for is that which always occurs in models with factors: with an intercept in the model, we have one more level than is estimable and so a constraint is required. A typical solution is to impose a sum-to-zero

constraint. The more indirect form of non-identification transpires because three factors (age, period and cohort) are linearly dependent, and there is no solution to this problem. Instead, one must make assumptions if one wishes to directly interpret the parameters in equation (10). Further, these assumptions are uncheckable from the raw data alone.

In most cases, the number of deaths O_{ijkz} , are taken to be independent Poisson distribution with mean (λ_{ijkz}) depending on Age, Period, and Cohort effects. Model (10) fall into the family of Generalized Linear Models (GLMs), it takes a log-linear regression form using a log-link function and can be modelled as;

$$\ln(\lambda_{ijkz}) = \log(N_{ijkz}) + \alpha_{iz} + \beta_{jz} + \gamma_{kz} + \varepsilon_{ijkz} \quad (11)$$

Since Age-Period-Cohort analysis suffer from identification problem, some parameters are constraints to zero in order to avoid the problem.

$$\sum_{i=1}^I \alpha_{iz} = \sum_{j=1}^J \beta_{jz} = \sum_{k=1}^K \gamma_{kz} = 0 \quad (12)$$

This problem persists due to the exact linear relationship of the age, period and birth cohort, which makes it impractical to recognize the individual contributions time effects factors, (Holford 1983).

After the constraining parameters to zero (12), and when R_{ijkz} is normally distributed, model (10) can also be presented in a form of a matrix for a linear model as follows:

$$\mathbf{R} = \mathbf{Y}\mathbf{b} + \boldsymbol{\varepsilon} \quad (13)$$

where \mathbf{R} is a mortality rates column vector, \mathbf{Y} is the dummy variables design matrix and \mathbf{b} is a model parameter vector,

$$\mathbf{b} = (\mu, \alpha_1, \dots, \alpha_{a-1}, \beta_1, \dots, \beta_{p-1}, \dots, \gamma_1, \dots, \gamma_{a+p-2})^T, \quad (14)$$

where $\boldsymbol{\varepsilon}$ is a random errors vector with mean 0 and variance σ^2 where ordinary least squares method can be adopted to obtain model parameter estimates vector \mathbf{b} :

$$\hat{\mathbf{b}} = (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{R} \quad (15)$$

where the indexed T denotes the transpose and the indexed -1 represent the inverse. Since at least one of the columns of \mathbf{Y} can be rewritten as a function of the other columns, the design

matrix \mathbf{Y} is a rank deficient one (i.e. singular). Consequently, a regular inverse $(\mathbf{Y}^T\mathbf{Y})^{-1}$ does not exist and Eq.15 cannot be estimated without an additional constraint.

The right side of equation 14 gives lot of information that helps in understanding how singular design matrix incurs during estimation of model parameters. The following property is obtained when design Matrix \mathbf{Y} is less than full column rank,

$$\mathbf{Y}\mathbf{b}_0 = \mathbf{0} \quad (16)$$

Where \mathbf{b}_0 is a zero vector to so that the product of a vector and a design matrix becomes 0. In this case, our null space design matrix \mathbf{Y} contains an empty eigenvector \mathbf{b}_0 that aligns with eigenvalue of zero. Therefore, \mathbf{b}_0 is estimated by decomposing \mathbf{b} in Eq (14) into two components, which represent non-null and null spaces of the design matrix \mathbf{Y} as follows:

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{q}\mathbf{b}_0 \quad (17)$$

where \mathbf{b}_1 is the prediction of \mathbf{b} on the non-empty space of \mathbf{Y} , \mathbf{b}_0 is the prediction of \mathbf{b} on the empty space of our design matrix \mathbf{Y} , and \mathbf{q} is an unpredictable scalar which produces crucial underlying understanding of identity problem. as the scalar \mathbf{q} possesses any real number, the age, period and birth cohort effects also change (eq 17). Nonetheless, our column vector \mathbf{R} is not affected by various values of \mathbf{q} , as shown in the following equation:

$$\mathbf{R} = \mathbf{Y}\mathbf{b} + \mathbf{e} = \mathbf{Y}(\mathbf{b}_1 + \mathbf{q}\mathbf{b}_0) = \mathbf{Y}\mathbf{b}_1 + \mathbf{q}\mathbf{Y}\mathbf{b}_0 \quad (18)$$

where $\mathbf{q}\mathbf{Y}\mathbf{b}_0$ is equal zero since $\mathbf{Y}\mathbf{b}_0$ from equation (16) is also zero. This implies that the column vector \mathbf{R} is dependent to $\mathbf{Y}\mathbf{b}_1$. Therefore, \mathbf{R} can take ant non-negative value that represent various arrays of age, period and birth cohort effects depending on the value of \mathbf{q} . Identification problem defined in this case arises when the true values of \mathbf{b} cannot be classified in terms of three time-components factors. Yang et.al (2004) developed an intrinsic estimator (IE) models to solve the persisting problem of the APC models which models the mortality rates of a subspace \mathbf{Y} by adopting principal component analysis methods which reduces the dimensions of analysis by 1 dimension and eliminates eigenvector \mathbf{b}_0 . Suppose. $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{2(a+p)-4}]$ and $\mathbf{\Lambda} = [\lambda_1, \lambda_2, \dots, \lambda_{2(a+p)-4}]$ are eigenvector and eigenvalue of $\mathbf{Y}^T\mathbf{Y}$ respectively, our new PC space design matrix. in the non-null subspace is defined by:

$$W = Y \cdot V' \quad (19)$$

where V' has an interval of $2(a + p) - 4$ by $2(a + p) - 5$, consisting of eigenvector which are non-empty.

IE algorithms yields some advantages over other techniques of solving identity problem including CGLIM, using prior estimator and Bayesian approximation methods (Yang et.al 2004). These methods have been found to produce unbiased estimator depending on the number of categories of age and period as well as their width intervals. Though, there are no available literature to back up the theory behind the IE's and on the selection of definite constraints, this technique has been proven to produce plausible estimates. Moreover, IE are recommended to produce efficient estimates than CGLIM models (Yang et.al. 2004)

All coefficients were estimated using the log-linear Poisson regression model. The temporal variations of mortality by each cause were examined based on age-period, age-cohort, period-cohort and age-period-cohort. The logarithmic transformation of mortality rates allows the formulation of generalised linear model (GLM) presented by the linear equation above, where μ denotes adjusted mean deaths and ϵ_{ij} denotes random error. The robustness of the APC model with various combinations of age, period and cohort was evaluated using deviance, log-likelihood, Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). A model with lower AIC, BIC and deviance was considered the most suitable model.

2.4 Ethical considerations

Data for this study are publicly accessible and include no personal identifiers. Therefore, no ethical review of the study protocol consent was necessary. These are secondary data which are recorded by the Department of Home Affairs (DHA) and Statistics South Africa processes and compiles them for public use.

CHAPTER 3: DATA DESCRIPTION AND RESULTS

3.1 Introduction

Data handling and the statistical approach adopted to analyse data in this study were described. Results of the analysis of the TB and HIV-related mortality data for the period 1997-2015 are presented in this chapter. Summary statistics are given by means of frequency and percentages for each cause-specific mortality and age. Age-specific and adjusted mortality rates are reported per 100,000 residents of South Africa together with their corresponding mortality plots. Estimated risk ratios are also presented adjusting for age, period and birth cohort.

3.2 Data

To study HIV and TB mortality trends, official mortality data obtained from Statistics South Africa was used. These data provide demographic information such as gender, age, death year and place of residence. Mortality data in South Africa are registered with the Department of Home Affairs (DHA) using the death notification forms. Statistics South Africa processes and compiles the mortality records forms into annual publications (STATSSA 2015). The mortality due to HIV and TB data was analysed for the years 1997-2015. The underlying causes of death were extracted using the ICD-10 (International Classification of Diseases) which are developed by the World Health Organization (WHO) and which the United Nations (UN) states members agrees and adheres to (WHO 2010). TB codes included A15 to A19, U51 and U52 while HIV codes range from B20 to B24. For this study, adults between the ages of 15 and 64 were considered where the standardised age groups, 15-24, 25-44, and 45-64 were utilised.

Statistics South Africa compiles mid-population totals at country level as well as at provincial level from 2002 till to date. However, this study covered death data from 1997 to 2015.

Therefore, for the years 1997 to 2000, an exponential growth rate from census 1996 was used to estimate the mid-population. The formula used is as follows:

$$P_{t+n} = P_t e^{nr},$$

where n is the length of time, r is the rate of population growth, and P_t is the total population at time t . Mid-population estimates for the year 2001 was obtained from the 2001 census report. For each cause-specific mortality, rates for each age group were calculated as per 100,000 residents of South Africa, that is, the proportion of the population dying from either TB or HIV and population at risk.

3.3 Results

Table 1 presents overall mortality rates of adults aged 15-64 by sex, age and cause. In 2015, a total of 1,162,832 adults' deaths from TB and HIV were registered in South Africa. The highest proportion of deaths was among the age of 25 to 44 years, followed by 45 to 64 years, 63.2% and 29.3% respectively. Considering deaths due to TB and HIV and among age 15-64, TB cause-specific mortality was accountable for 910 731 (78.3%) deaths by the year 2015, while 252, 101 (21.7%) deaths were observed for HIV cause mortality.

Table 1: Distribution of deaths by cause-specific and age groups, 1997-2015

Age Groups	Number of deaths (%)		
	Overall [1, 162 832 (100%)]	TB [910 731 (78.3)]	HIV [252 101 (21.7)]
15-24	87 417 (7.5)	68 291 (7.5)	19 126 (7.6)
25-44	735 268 (63.2)	561 322 (61.6)	173 946 (69.0)
45-64	340 147 (29.3)	281 118 (30.9)	59 029 (28.4)

Table 2 shows age-specific mortality rates calculated as the number of deaths per 100,000 residents in South Africa for TB cause-specific and HIV cause-specific, for each year and three age groups. TB-specific mortality rates were higher for the age groups 25-44 and 45-64 compared to age group 15-24 years. Mortality rates increased rapidly with time, reached a peak in 2006 and then declined until 2015. A normal distribution pattern was observed for the three age groups under study in Figure 1.

Mortality rates due to HIV increases with time from 1997 to 2015 for all age groups, that is 15-24, 25-44, 45-64 years. Higher mortality rates were observed for age group 25-44 followed by age group 45-64, while age group 15-24 had lower rates compared to other age groups. Trends for mortality due to HIV are presented in Figure 1, where mortality trends are observed to increase and decrease within the periods and age groups.

Table 2: Specific mortality rates calculated for each cause and age groups for each year

Year of death	Population at Risk			TB						HIV					
				Observed number of deaths by age and period			Age specific deaths (per 100 000)			Observed number of deaths by age and period			Age specific deaths (per 100 000)		
	15-24	25-44	45-64	15-24	25-44	45-64	15-24	25-44	45-64	15-24	25-44	45-64	15-24	25-44	45-64
1997	8417230	11617730	5705371	1575	9745	6632	18.7	83.9	116.2	688	3883	855	8.2	33.4	15.0
1998	8604167	11874984	5831239	2174	13813	7962	25.3	116.3	136.5	778	4592	988	9.0	38.7	16.9
1999	8795291	12137980	5959909	2585	17746	9264	29.4	146.2	155.4	1018	6508	1309	11.6	53.6	22.0
2000	8990697	12406852	6091443	3365	23225	10700	37.4	187.2	175.7	966	6909	1447	10.7	55.7	23.8
2001	9190481	12681729	6225903	4016	29438	12285	43.7	232.1	197.3	850	6092	1275	9.2	48.0	20.5
2002	9394742	12962748	6363357	4634	35896	14275	49.3	276.9	224.3	879	7112	1451	9.4	54.9	22.8
2003	9703620	13128343	6524329	5032	40260	16283	51.9	306.7	249.6	932	7980	1715	9.6	60.8	26.3
2004	9991817	13288949	6677279	5075	42292	17013	50.8	318.2	254.8	1020	9062	2184	10.2	68.2	32.7
2005	10231229	13474486	6824885	5140	43809	18500	50.2	325.1	271.1	1053	9639	2550	10.3	71.5	37.4
2006	10386442	13716965	6968450	5179	44568	19801	49.9	324.9	284.2	1041	9627	2688	10.0	70.2	38.6
2007	10480119	13998671	7121076	5105	44402	19943	48.7	317.2	280.1	934	8818	2618	8.9	63.0	36.8
2008	10532875	14318582	7278399	4677	42266	20433	44.4	295.2	280.7	979	9680	3118	9.3	67.6	42.8
2009	10551264	14666060	7435300	4240	37752	19602	40.2	257.4	263.6	1068	11108	3937	10.1	75.7	53.0
2010	10527466	15057211	7593685	3866	33477	18517	36.7	222.3	243.8	1125	11506	4310	10.7	76.4	56.8
2011	10447446	15505680	7752424	3112	27714	16820	29.8	178.7	217.0	1065	10483	4348	10.2	67.6	56.1
2012	10318325	15980998	7908754	2706	23507	15191	26.2	147.1	192.1	1090	11482	4797	10.6	71.8	60.7
2013	10157010	16482193	8062446	2208	19292	13369	21.7	117.0	165.8	1287	13896	6359	12.7	84.3	78.9
2014	10000435	16983627	8219731	1916	17570	12852	19.2	103.5	156.4	1220	12876	6426	12.2	75.8	78.2
2015	9869209	17473681	8386321	1686	14550	11676	17.1	83.3	139.2	1132	12693	6654	11.5	72.6	79.3

Mortality rates were calculated as number of deaths/estimated population and expressed per 100,000 persons.

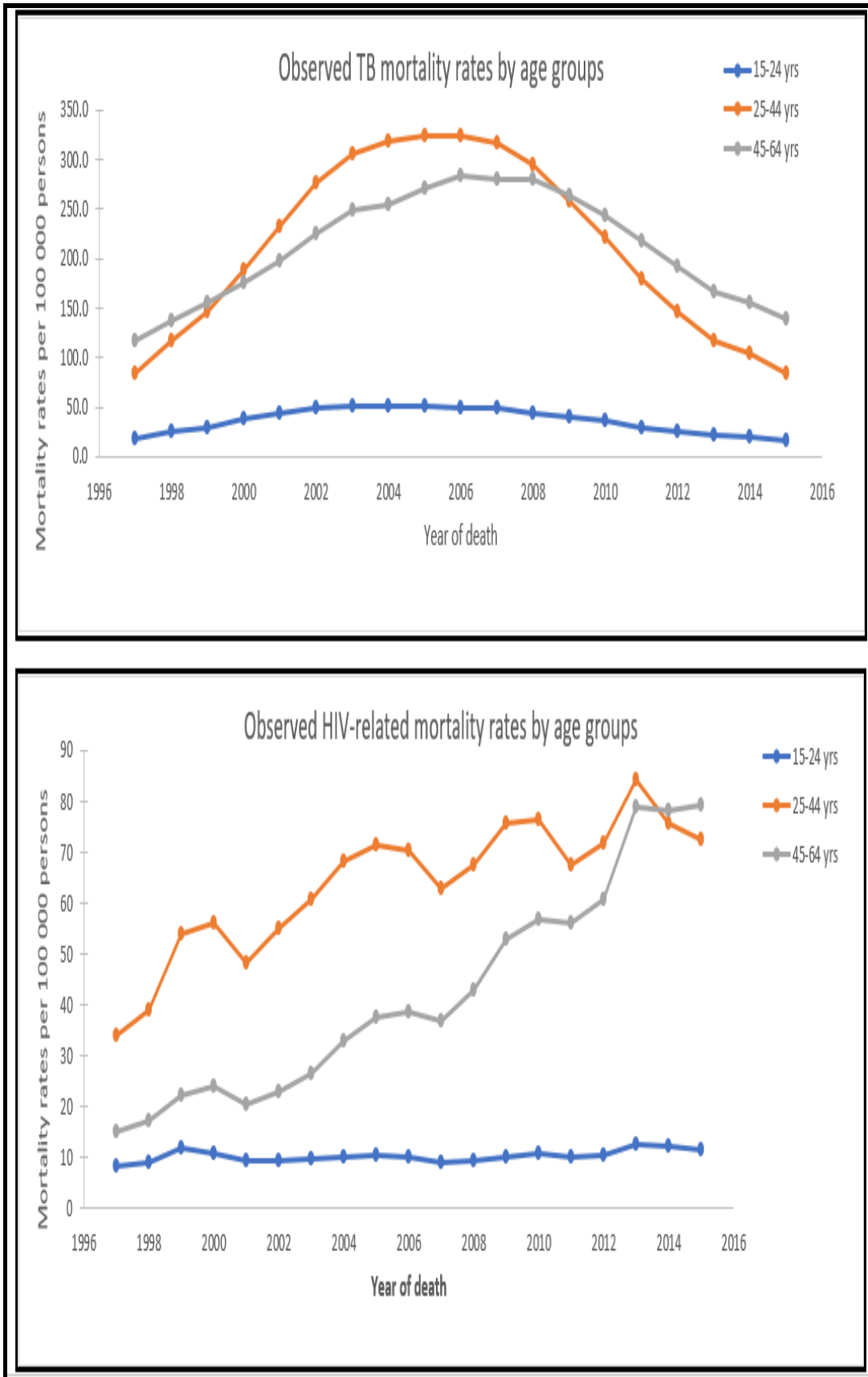


Figure 1: Observed TB and HIV-related mortality trends by age groups

3.4 HIV-related mortality data

In South Africa, vital data quality issues such as high proportion of misclassification of HIV/AIDS deaths, ill-defined causes, low-quality certification of underlying cause of death by health practitioner and headmen (rural areas), and a high proportion of deaths classified as undetermined unnatural causes have been identified (Groenewald 2005, Birnbaum 2012 and Pillay-van Wyk 2016). Over decades, completeness for death registration has been estimated using several indirect demographic and statistical techniques for example, the Brass Growth Balance, Generalised Growth Balance, Bennett–Horiuchi or Preston–Coale methods (Hill and Queiroz 2010, Silva 2013, Bennett and Horiuchi 1984 and Dorrington 1999). However, these methods are best used by national statistical offices, as they require the use of census data and an understanding of the different techniques. Adair and Lopez (2018) developed an empirical method to estimate death registration completeness, which utilizes data commonly available at the national and subnational level. This method has some advantages compared to other methods including better accuracy, improved timeliness, application to subnational areas and greater simplicity of use. To adjust for under-reporting and misclassification of HIV/AIDS death, our study has adopted the indirect descriptive statistics utilized by Fazito (2012) to identify and quantify misclassified and under-reported AIDS deaths in Brazil, and Groenewald (2005) to identify AIDS related deaths in South Africa. The adopted method for adjustment of HIV/AIDS related death has shown its significance in improving mortality information at country level, by both age groups and gender (Groenewald 2005).

Cause-of-death statistics are compiled using the International Classification of Diseases (ICD-10) which enables the identification of the exact cause of death. Statistics South Africa manually codes the causes to the 10th version of the ICD-10 and undertakes an automated selection of underlying causes of death according to ICD rules. According to ICD-10, HIV/AIDS codes range from B20 to B24. Young adults (15-64 years) mortality records of all deaths which occurred from 1997 to 2015 in South Africa with information on age, year of death, and underlying cause of death were analysed. Death rates for HIV/AIDS and AIDS-

related diseases were calculated for each of the three age groups (15-24, 25-44, and 45-64) for periods from 1997 to 2015 per 100 000 persons of South Africa.

For this study, 'misclassified' deaths are defined as deaths which were due to AIDS but not reported as such in the death certificates, while 'Under-reported' deaths are referred deaths which were not declared or registered in the civil registries. According to the ICD-10 underlying cause of death selection rules, a death may be considered as HIV/AIDS if there is reference to HIV/AIDS in any line of the death certificate, as well as mention of a condition that is considered a direct consequence of HIV/AIDS. The conditions that are assumed to be direct consequences of HIV/AIDS are: Kaposi's sarcoma, Burkitt's tumour, and any other malignant neoplasm of lymphoid, haematopoietic and related tissue, classifiable to C46 or C81-96; and any infectious diseases classifiable to A00-B19, B25-B49, B58-B64, B99 or J12-J18 (Centers for Disease Control and Prevention 2008).

To adjust for misclassification, potential 28 causes of death were investigated using age-specific rates trends plotted over time. AIDS-indicators following the same distinct age pattern typical of HIV/AIDS age-specific death rates trends plotted over time were selected for estimating excess mortality. Statistics South Africa (2016) indicated that 96% of completeness of death registration was achieved for 2015 period. Given that, we utilized 2015 deaths rate to calculate expected mortality (expected mortality was subtracted from observed mortality) and excess mortality from the causes which were identified as AIDS-indicators. Deaths records with no information on age were redistributed proportionally to all other age groups for the same year of death while deaths due to ill-defined causes were redistributed across all other natural causes of death. Expected numbers of AIDS deaths that were coded as ill-defined causes were estimated using the proportion of AIDS deaths among all-natural causes of deaths. To adjust for under-reporting, observed number of AIDS deaths were further adjusted for incompleteness. To do that, the ratio of overall observed death rate from our working data and overall estimated death rate from WHO was subtracted from 1. Furthermore, excess and

expected mortality due to ill-defined causes were scaled-up by the proportion of incompleteness.

Table 3 presents an observed number of AIDS deaths, excess mortality, adjustment for AIDS mortality to ill-defined and completeness, and the revised estimated number of AIDS deaths along with corresponding deaths rates per year. Death rates were used to plot the trend presented in Figure 2. Out of 28 AIDS-indicator deaths, only four conditions (Cryptococcosis, Kaposi's sarcoma, Pneumocystosis and Meningitis due to other and unspecified causes) were identified likely to be misclassified for AIDS deaths. There were a total of 11 968 deaths due to Kaposi's sarcoma, 24 166 deaths due to Cryptococcosis, 71 051 deaths due to Meningitis (from other and unspecified causes) and 37 294 deaths due to Pneumocystosis related to HIV/AIDS deaths from 1997 to 2015; a total of 144 479 AIDS deaths misclassified as other diseases. However, a total of 80 438 was redistributed to AIDS deaths due to over estimation for the period 1997 to 1999. A total of 33 358 deaths due to ill-defined causes were redistributed to AIDS, and 170 624 were added to the number of AIDS deaths to correct for completeness of the mortality in South Africa. Altogether, 284 420 deaths were recoded to AIDS between 1997 and 2015, representing 53% of all AIDS deaths. Of all deaths recorded to HIV/AIDS, 60% were due to incompleteness of deaths, 28% were due to conditions likely to be misclassified AIDS deaths while only 12% were due to ill-defined. Revised mortality has shown a concave down association between HIV/AIDS cause-specific mortality and rates over time with a peak during the period 2006.

Table 3: Summary of AIDS mortality Adjustments in the 15 to 64-year-old South African Population: 1997-2015

Year	Observed AIDS Mortality	Observed AIDS death rates	Excess Mortality	Adjustment for ill-defined	Adjustment for completeness	Revised number of AIDS deaths	Revised AIDS death rates
1997	5427	20.40	-1569	835	-1100	3594	13.51
1998	6358	23.34	-1078	1016	-93	6203	22.77
1999	8835	31.69	-417	1187	1154	10758	38.59
2000	9322	32.69	799	1228	3041	14389	50.46
2001	8217	28.27	2366	1103	5206	16892	58.12
2002	9442	31.88	4022	1247	7906	22616	76.35
2003	10627	35.21	5643	1419	10599	28288	93.74
2004	12266	39.93	7350	1623	13495	34733	113.06
2005	13242	42.38	8515	1751	15399	38908	124.53
2006	13356	42.15	9722	1827	17323	42227	133.26
2007	12370	38.52	9537	1745	16924	40577	126.35
2008	13777	42.34	8752	1872	15937	40339	123.96
2009	16118	48.86	7955	2174	15195	41442	125.63
2010	16949	50.64	6796	2239	13552	39536	118.12
2011	15912	46.81	4467	2102	9798	32279	94.96
2012	17376	50.29	3562	2264	8676	31879	92.27
2013	21562	61.38	2328	2683	7516	34088	97.05
2014	20533	57.52	1688	2539	6340	31100	87.12
2015	20499	56.54	0	2505	3758	26762	73.81
Total	252188		80438	33358	170624	536609	

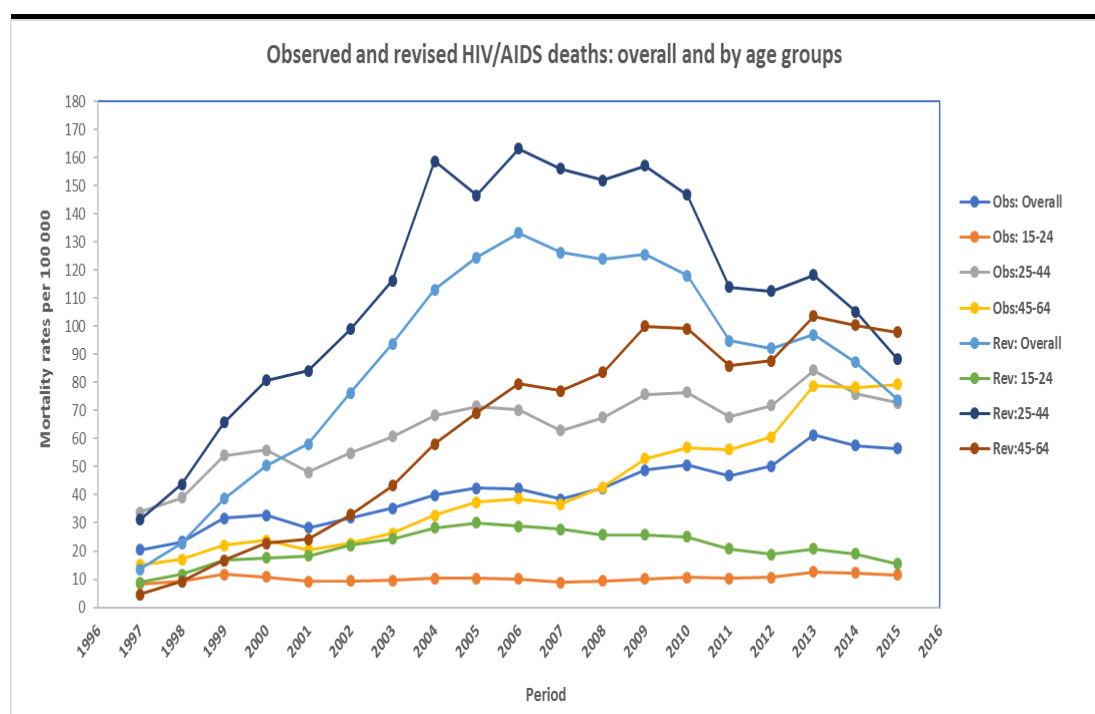


Figure 2: Observed and revised HIV/AIDS death rates for the 15 to 64-year-old South African Population

3.5 Analysis of the modelled data with reference to HIV/AIDS and TB related causes of death

The data covered HIV and TB cause-specific mortality for South Africa between 1997 and 2015. It was decided to restrict the age range of the dataset to people who were, at year of death, equal to or greater than 15 but less than 65 years of age. To highlight the possibility of including covariates into the analysis, the gender of patients and the cause-specific mortality was included when collapsing the dataset into unique records of age, period, and cohort. A three-year age, period and birth cohort intervals for 15-64 years, 1997-2015 and 1934-2000 were used respectively. This collapse led to $(64 - 15) \times (2015 - 1997)$ different age-period categories, each of which were further subdivided by date of birth into three categories. This gave a total of 238 observations, one for each triangular subset. However, since a completed dataset contains insufficient information, cause-specific mortality terms were included. To adjust for individual age populations, the Sprague multipliers method of estimation was employed. Calot and Sardon (2004) gave detailed information on how Sprague multipliers can be utilised to estimate the mid-population per single age. Therefore, the dataset contains (119×2) observations and population risk-time calculations were performed for single age using Sprague multipliers. Data used comprise 16 3-year and 1 2-year age groups ranging from 15-64 and 6 3-year and 1 single year period from 1997 to 2015. These yielded 22 successive 3-year and 1 single birth cohorts of which the oldest was born in 1934-1936 and the youngest cohort was born in 2000. A two-way table of modelled data is shown in the Appendix I estimated effects are based on the adjusted mortality data.

3.6 Data exploration by three-year age, period and birth cohort intervals

Figure 3 shows the percentage distribution of TB-related mortality by age group over a 3-year period interval and 3-year birth cohort. Higher proportion of deaths were observed for age group 25-44 -24 followed by age group 45-64 while age groups 15-24 had the lowest deaths

across all periods. Observed percentages showed that TB mortality for age group 45-64 increased with an increase of time while for other remaining age groups, mortality decreased with time period. Age group 15-24, 25-44 and 45-64 were born between the birth cohort 1973-2000, 1952-1990, and 1934-1972 respectively. For birth cohorts between 1952-1954 to 1958-1960, mortality was higher for age group 45-64 compared to age group 25-44 years, while higher mortality were observed for age group 25-44 birth cohort 1964-1984 compared to all other age groups.

Figure 4 shows proportions of HIV-related mortality by age group over a 3-year period interval and 3-year birth cohort. HIV-related mortality. Age group 25-44 had higher proportions of HIV-related mortality between 1997 and 2015 while age group 15 – 24 had the least deaths. HIV-related mortality decreased with period for age group 15-24 and 25-44 but increased with time for age group 45-64. Birth cohort 1952-1954 to 1961-1963, age group 45-64 had higher proportion of HIV-related mortality compared to same birth cohort age 25-44. Again, proportion of HIV-related death were higher among age group 25-44 compared to age group 15-24 birth cohort 1973-1975 to 1985-1987. The rest of the results are organised by the main research question in the study.

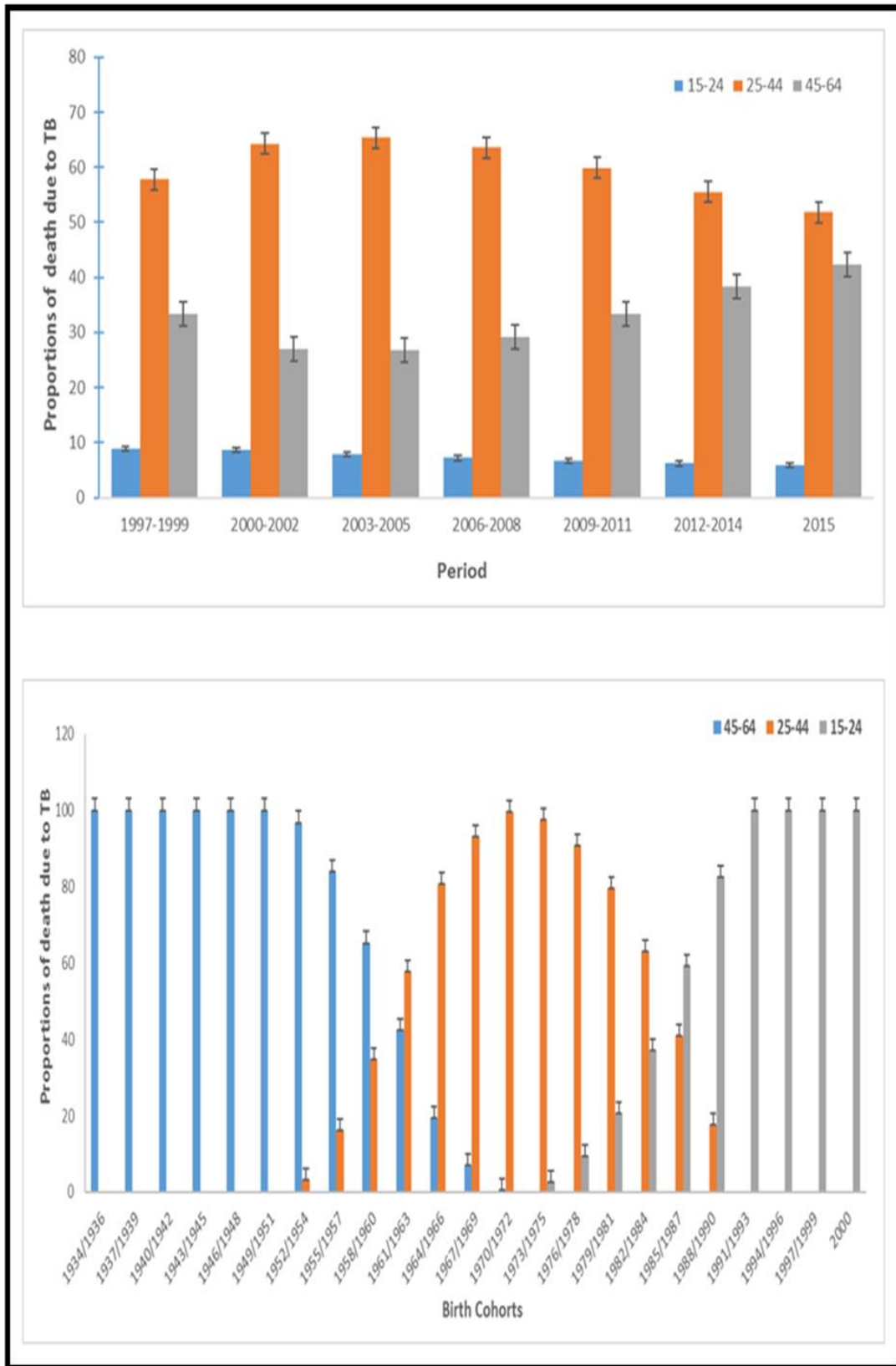


Figure 3: Percentage distribution of observed TB mortality by age groups over the 3-year period and 3-year birth cohorts

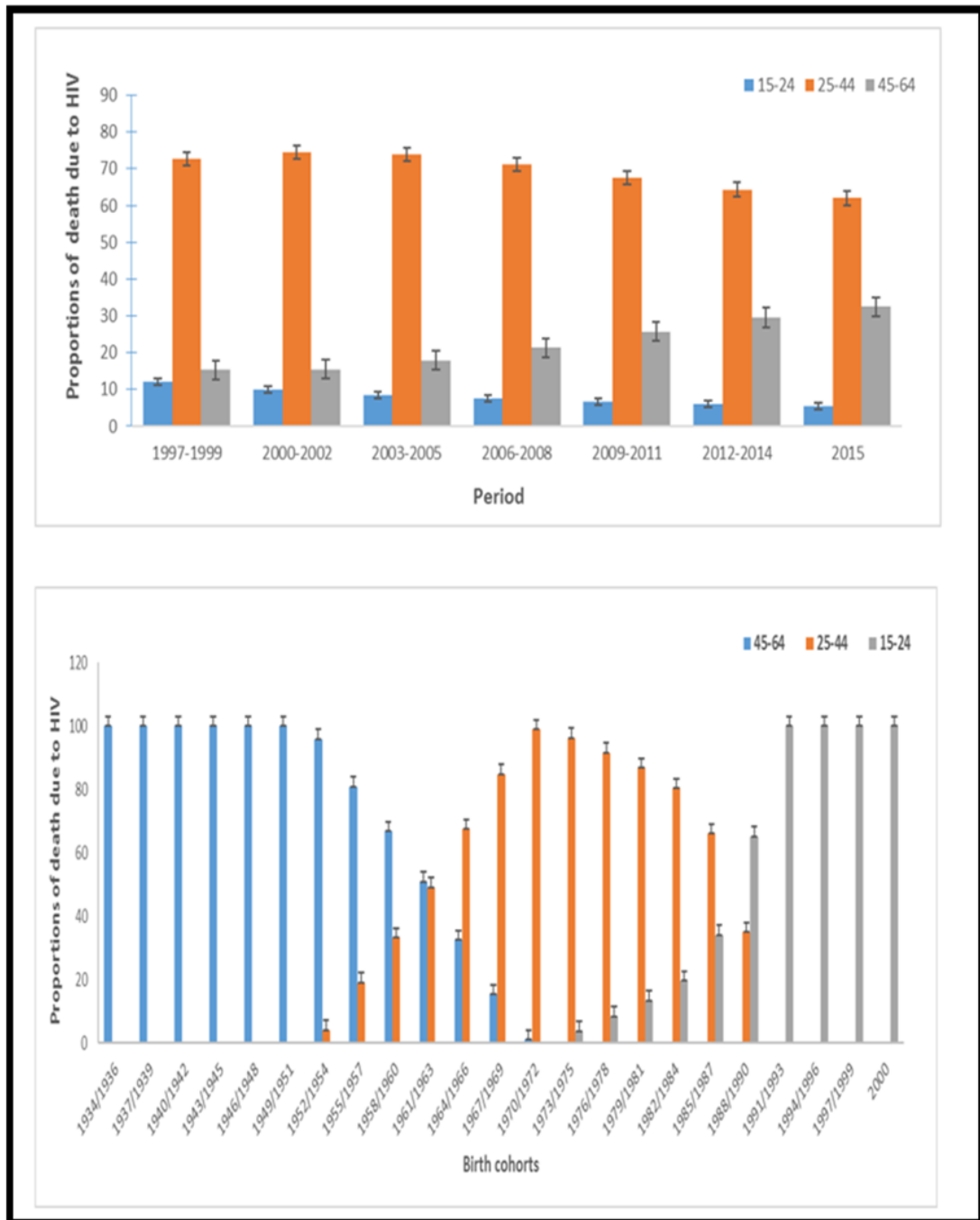


Figure 4: Percentage distribution of observed HIV-related mortality by age groups over a 3-year period and 3-year birth cohorts

3.7 Mortality relative risk (RR) for TB and HIV cause-specific.

To describe the age, period and birth cohort effects simultaneously, this study used the APC model (IE) to analyse TB and HIV-related mortality trends (1997-2015) of South African residents aged 15-64 years. To evaluate the robustness of the IE findings, I estimated the

relative risks using the the traditional CGLIM model that apply different constraints. Table 4 and 5 present the estimated relative risk (RR) and 95% confidence intervals for mortality due to TB and HIV respectively. The reference group for age, period and birth cohort relative risk is the mean influence of all ages, periods and birth cohorts combined (Keyes and Miech 2013). Age, period and birth cohorts' effects trends were plotted based on the relative mortality risk. Analysis for HIV-related mortality were conducted using the revised data.

Table 4 shows the estimated mortality relative risk for TB cause-specific mortality using the intrinsic estimator and Constrained Generalized linear models and depicted in Figure 5 and Figure 6 respectively. Mortality relative risk and 95% confidence intervals are reported. Our results show that age, period and cohort have a strong association with TB mortality. Controlling for period and birth cohort effects, age effect showed that the relative mortality increases as the age increases. Higher mortality was observed for individuals aged 27 years and above to 64 years. For the period effect, relative risk increased from 2000, reached a peak by the year period 2006-2008 and thereafter decline until 2015, holding age and birth cohort effects constant. The cohort effect reflects an aggregate effect of exposure in generations. Mortality relative risk was found higher for birth cohort 1934-1936 than birth cohort after 1934-1936. The results showed an upward trend in Tuberculosis mortality by age to around the age of 33-35 years, from which point the trend turned downward continuously to age group 63-64 years. There was a progressive upward trend in Tuberculosis mortality between 1997-2000 and 2006-2008, then declined thereafter to 2015 period. A downward trend between TB mortality relative risk and birth cohort was found (1937-2000).

Table 5 shows the estimated incidence rates for HIV cause-specific mortality using the intrinsic estimator and Constrained Generalized linear models and portrayed in Figure 7 and Figure 8 respectively. Age group 24-26 to 42-44 years were associated with higher mortality rates as compared to age groups younger than 24 or older than 44 years, controlling for period and birth cohort effects. Keeping age and birth cohort effects fixed, period effect increased significantly with time from 1997-1999 to 2009-2011 then declined significantly thereafter.

Cohort born in 1964-1966 were associated with higher HIV incidence rates as compared to cohorts born before 1964 or after 1966. An accelerating trend in Human Immunodeficiency Virus mortality with age, to age group 33-35 years, from which point the trend turned downwards until age group 57-59 years then increases again. There was a concave down association for HIV-related mortality during the study period (1997-2015). There was inverted V-shaped association between HIV-related mortality and birth cohort.

Estimates of age, period and cohort using CGLM approach depends on the choice of constraint, which can be by either age, period or cohort. Unfortunately, estimates of the CGLM approach completely depend on constraint chosen and consequently different results can be obtained when changing the identification restriction (Appendix III). Therefore, the constraints should be chosen carefully. For our study, estimates obtained from each of the three constraint are reported and similar significant results to those of the IE solution was yielded. Results are presented in table 4 and Table 5 and portrayed by Figure 5 and Figure 6; and Figure 7 and Figure 8, for TB and HIV cause- mortality respectively. The age and birth cohort effects are different for TB cause-specific mortality across the two models, while the period effects estimated by both models are basically similar. The age, period and birth cohort effects estimated by both IE and CGLM are virtually very similar for HIV cause-specific mortality.

Table 4: Age-period-cohort estimates for TB cause-specific mortality

TB Cause-Specific Mortality	IE			CGLIM		
				Constrained by Age group: 18-20 = 63-64		
	RR	95% CI		RR	95% CI	
Age						
15 - 17	0.12	0.11	0.12	1.00	-	-
18 - 20	0.29	0.27	0.30	2.22	2.08	2.36
21 - 23	0.58	0.55	0.60	4.10	3.82	4.39
24 - 26	0.95	0.92	0.99	6.17	5.72	6.65
27 - 29	1.28	1.22	1.33	7.52	6.89	8.21
30 - 32	1.45	1.40	1.51	7.81	7.07	8.62
33 - 35	1.51	1.46	1.57	7.42	6.64	8.29
36 - 38	1.50	1.45	1.56	6.71	5.93	7.59
39 - 41	1.47	1.42	1.52	6.00	5.25	6.86
42 - 44	1.42	1.37	1.48	5.29	4.57	6.13
45 - 47	1.40	1.34	1.46	4.73	4.03	5.55
48 - 50	1.36	1.29	1.44	4.21	3.53	5.02
51 - 53	1.34	1.26	1.43	3.78	3.13	4.57
54 - 56	1.31	1.22	1.41	3.37	2.75	4.12
57 - 59	1.30	1.18	1.43	3.05	2.46	3.78
60 - 62	1.24	1.12	1.38	2.66	2.17	3.26
63 - 64	1.14	0.97	1.34	2.22	2.08	2.36
Period						
1997 - 1999	0.60	0.56	0.64	1.00	-	-
2000 - 2002	1.10	1.08	1.12	2.01	1.87	2.16
2003 - 2005	1.50	1.46	1.53	3.00	2.77	3.25
2006 - 2008	1.58	1.55	1.61	3.47	3.20	3.77
2009 - 2011	1.25	1.23	1.27	3.01	2.76	3.29
2012 - 2014	0.81	0.79	0.84	2.14	1.93	2.38
2015	0.63	0.61	0.66	1.84	1.63	2.06
Birth cohort						
1934 - 1936	1.62	1.39	1.90	1.00	-	-
1937 - 1939	1.34	1.04	1.73	0.75	0.56	1.01
1940 - 1942	1.30	1.04	1.61	0.66	0.51	0.86
1943 - 1945	1.02	0.87	1.19	0.48	0.39	0.58
1946 - 1948	0.97	0.87	1.09	0.41	0.35	0.49
1949 - 1951	1.12	1.01	1.25	0.44	0.36	0.52
1952 - 1954	1.16	1.07	1.26	0.41	0.35	0.48
1955 - 1957	1.14	1.06	1.23	0.37	0.32	0.42
1958 - 1960	1.20	1.13	1.28	0.35	0.31	0.40
1961 - 1963	1.23	1.16	1.30	0.33	0.29	0.37
1964 - 1966	1.30	1.23	1.37	0.32	0.28	0.35
1967 - 1969	1.23	1.17	1.30	0.27	0.25	0.31
1970 - 1972	1.27	1.20	1.34	0.26	0.23	0.29
1973 - 1975	1.22	1.17	1.28	0.23	0.21	0.25
1976 - 1978	1.13	1.08	1.18	0.19	0.18	0.21
1979 - 1981	0.94	0.91	0.97	0.14	0.13	0.16
1982 - 1984	0.80	0.77	0.82	0.11	0.10	0.12
1985 - 1987	0.68	0.65	0.70	0.09	0.08	0.09
1988 - 1990	0.56	0.54	0.59	0.07	0.06	0.07
1991 - 1993	0.53	0.50	0.56	0.06	0.05	0.06
1994 - 1996	0.60	0.55	0.66	0.06	0.05	0.07
1997 - 1999	0.76	0.64	0.90	0.07	0.05	0.08
2000	0.85	0.79	0.92	0.07	0.06	0.08

Mortality relative risk (exp(b)) and 95% confidence intervals are reported.

Table 5: Age-period-cohort estimates for HIV cause-specific mortality

HIV Cause-Specific Mortality	IE			CGLIM		
	Constrained by Age group: 27-29 = 63-64					
	RR	95% CI		RR	95% CI	
Age						
15 - 17	0.47	0.42	0.54	1.00	-	-
18 - 20	0.52	0.50	0.55	1.09	0.95	1.25
21 - 23	0.70	0.67	0.74	1.44	1.24	1.68
24 - 26	1.31	1.27	1.36	2.65	2.26	3.11
27 - 29	1.07	1.04	1.11	2.14	1.81	2.52
30 - 32	1.66	1.61	1.71	3.25	2.73	3.87
33 - 35	1.61	1.56	1.66	3.10	2.59	3.72
36 - 38	1.53	1.48	1.58	2.91	2.41	3.51
39 - 41	1.44	1.40	1.49	2.70	2.22	3.28
42 - 44	1.40	1.33	1.48	2.59	2.10	3.18
45 - 47	0.81	0.74	0.88	1.47	1.17	1.84
48 - 50	0.78	0.74	0.83	1.40	1.12	1.75
51 - 53	0.78	0.75	0.82	1.38	1.10	1.73
54 - 56	0.82	0.77	0.87	1.43	1.12	1.81
57 - 59	0.89	0.84	0.95	1.52	1.19	1.95
60 - 62	1.05	0.98	1.13	1.77	1.38	2.27
63 - 64	1.29	1.13	1.47	2.14	1.81	2.52
Period						
1997 - 1999	0.36	0.34	0.38	1.00	-	-
2000 - 2002	0.67	0.66	0.69	1.91	1.80	2.03
2003 - 2005	1.15	1.13	1.17	3.32	3.12	3.53
2006 - 2008	1.41	1.38	1.44	4.12	3.85	4.42
2009 - 2011	1.45	1.42	1.47	4.30	3.99	4.63
2012 - 2014	1.37	1.34	1.40	4.13	3.78	4.51
2015	1.29	1.23	1.35	3.95	3.54	4.41
Birth cohort						
1934 - 1936	0.39	0.35	0.45	1.00	-	-
1937 - 1939	0.32	0.26	0.40	0.81	0.64	1.03
1940 - 1942	0.61	0.43	0.87	1.51	1.01	2.25
1943 - 1945	0.89	0.74	1.08	2.17	1.73	2.72
1946 - 1948	1.09	0.98	1.20	2.60	2.31	2.92
1949 - 1951	1.30	1.21	1.40	3.06	2.73	3.43
1952 - 1954	1.48	1.37	1.60	3.43	3.07	3.84
1955 - 1957	1.61	1.51	1.72	3.68	3.35	4.03
1958 - 1960	1.90	1.75	2.05	4.26	3.88	4.67
1961 - 1963	2.08	1.94	2.22	4.60	4.27	4.95
1964 - 1966	2.22	2.10	2.34	4.83	4.56	5.12
1967 - 1969	2.09	1.98	2.22	4.49	4.24	4.75
1970 - 1972	1.93	1.83	2.03	4.07	3.85	4.30
1973 - 1975	1.74	1.67	1.82	3.63	3.45	3.82
1976 - 1978	1.58	1.51	1.65	3.24	3.05	3.44
1979 - 1981	1.29	1.24	1.35	2.61	2.45	2.79
1982 - 1984	1.05	1.00	1.11	2.09	1.93	2.27
1985 - 1987	0.85	0.81	0.88	1.66	1.52	1.81
1988 - 1990	0.70	0.67	0.72	1.34	1.21	1.49
1991 - 1993	0.61	0.57	0.65	1.16	1.01	1.32
1994 - 1996	0.54	0.50	0.58	1.01	0.86	1.17
1997 - 1999	0.44	0.40	0.48	0.80	0.67	0.96
2000	0.38	0.34	0.43	0.69	0.56	0.86

Mortality relative risk (exp(b)) and 95% confidence intervals are reported.

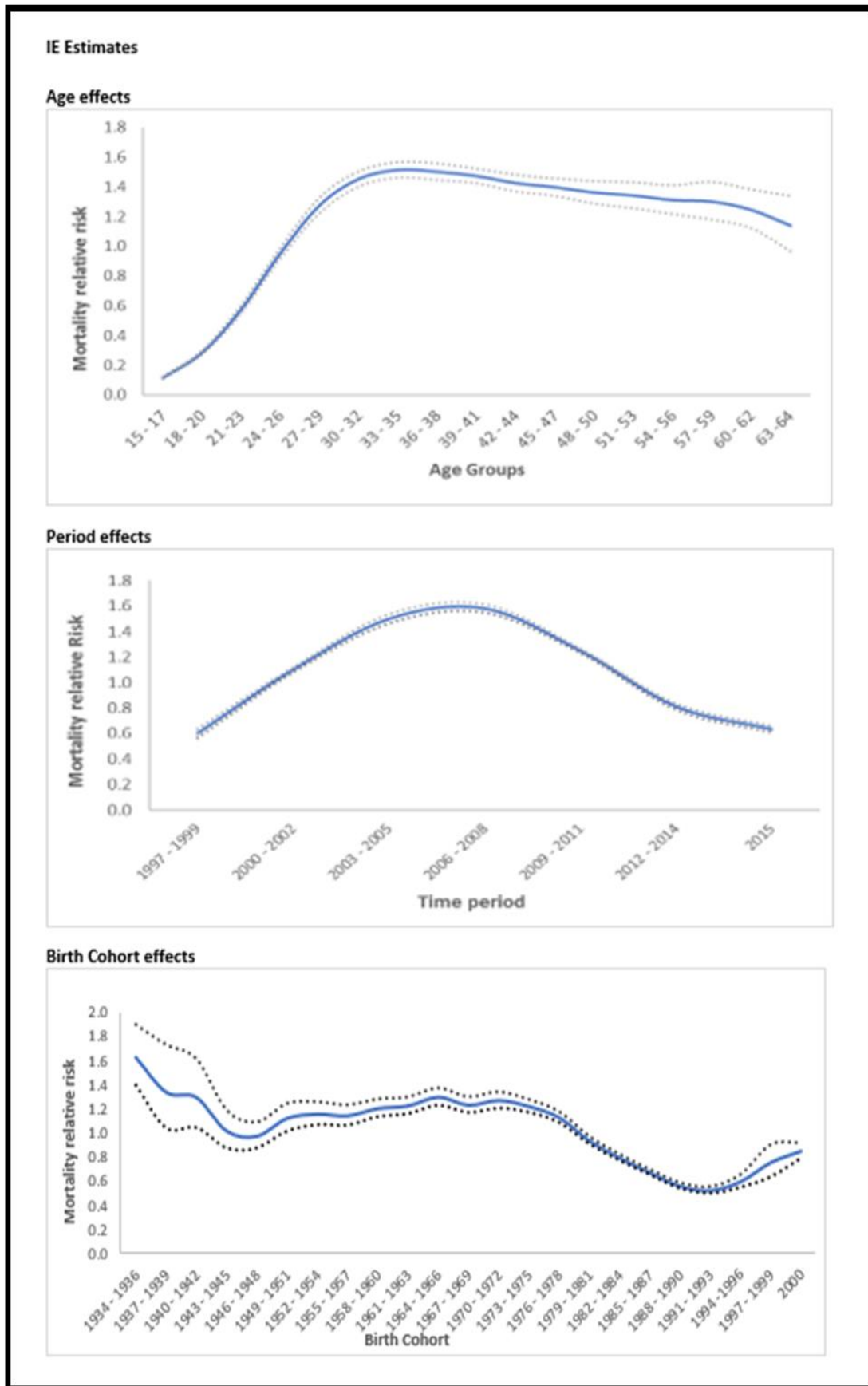
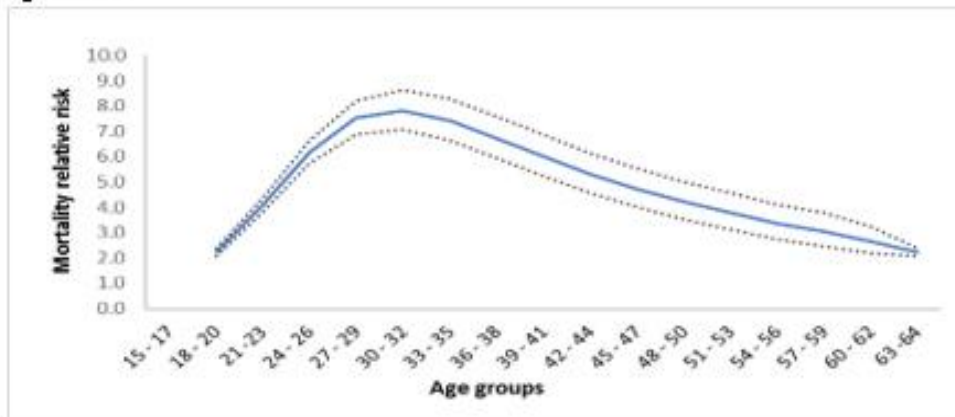


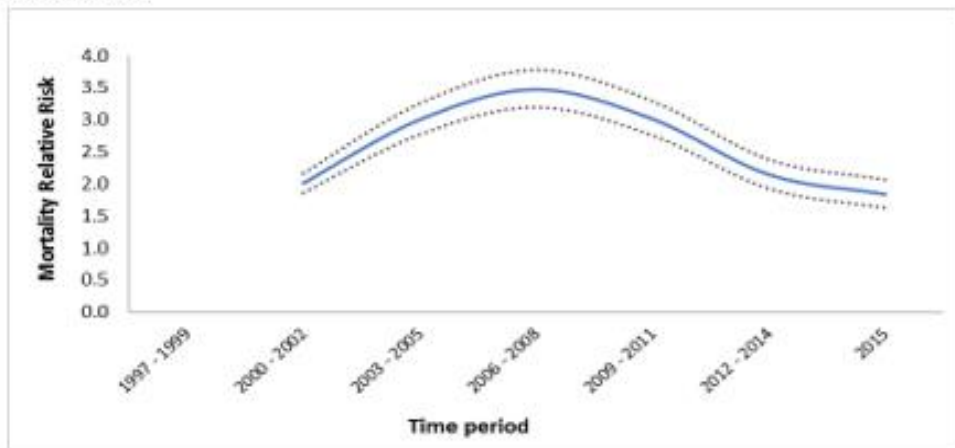
Figure 5: Age, Period and cohort effects of TB mortality. Notes: dotted lines represent 95% confidence intervals.

CGLIM Estimates

Age effects



Period effects



Birth cohort effects

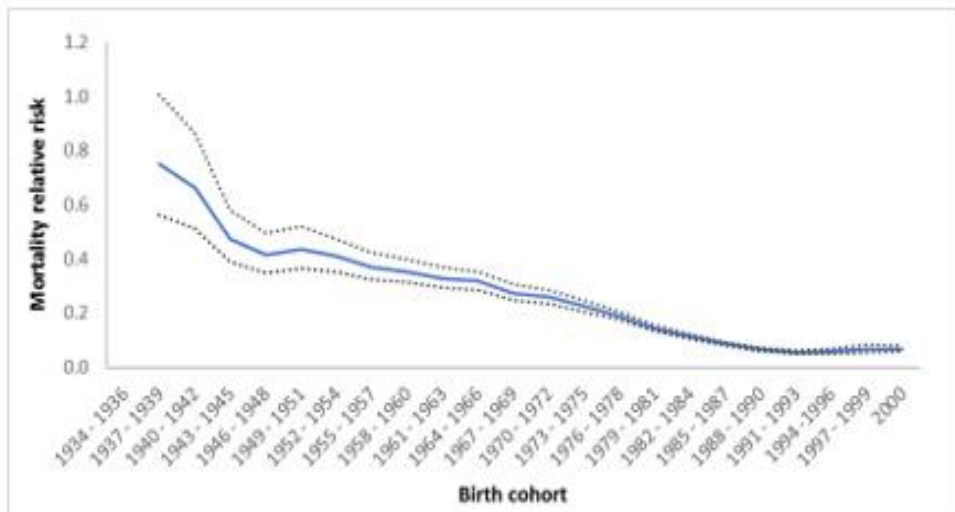


Figure 6: Age, Period and cohort effects of TB mortality. Notes: dotted lines represent 95% confidence intervals.

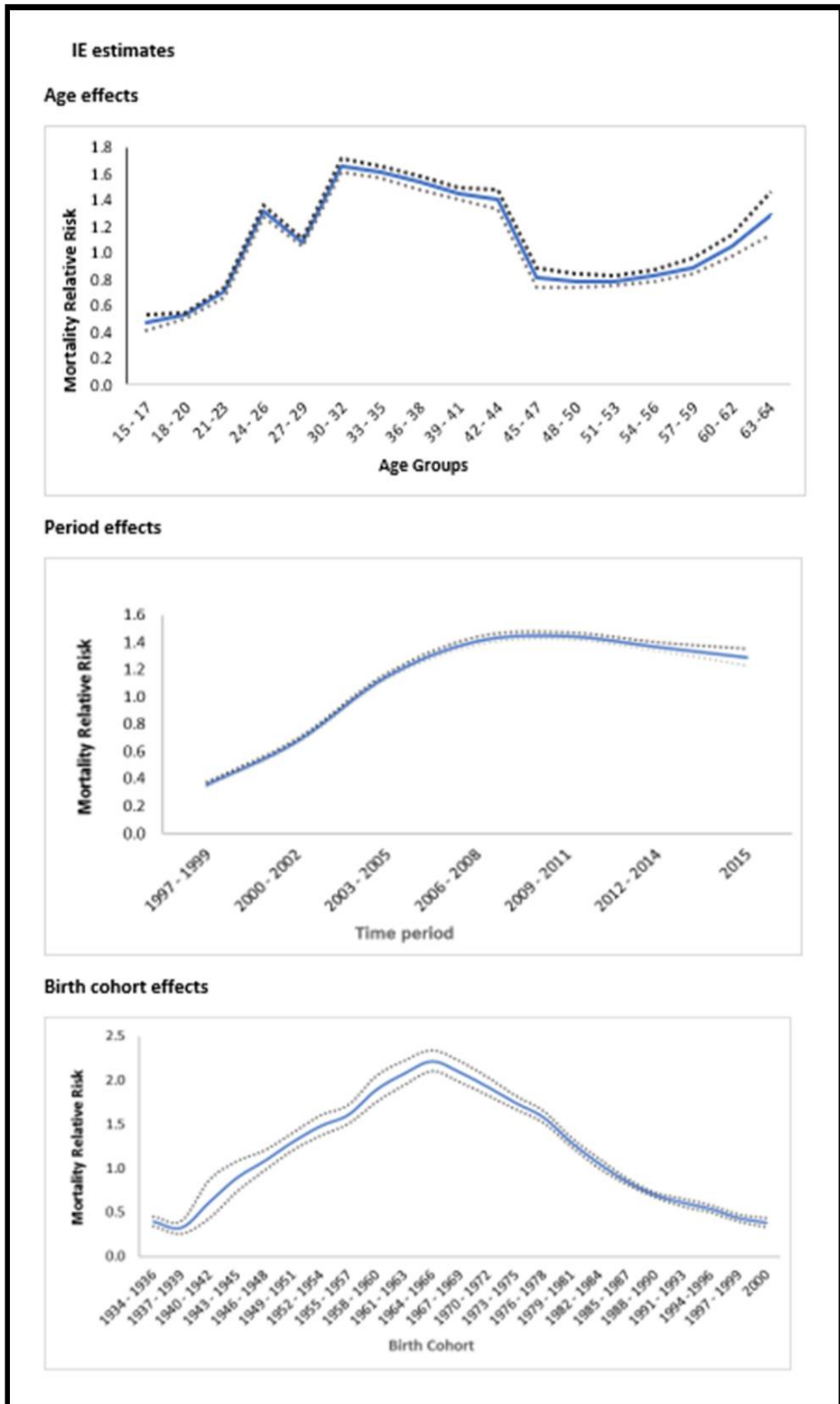


Figure 7: Age, period and cohort effects of HIV-related mortality. Notes: dotted lines present 95% confidence intervals.

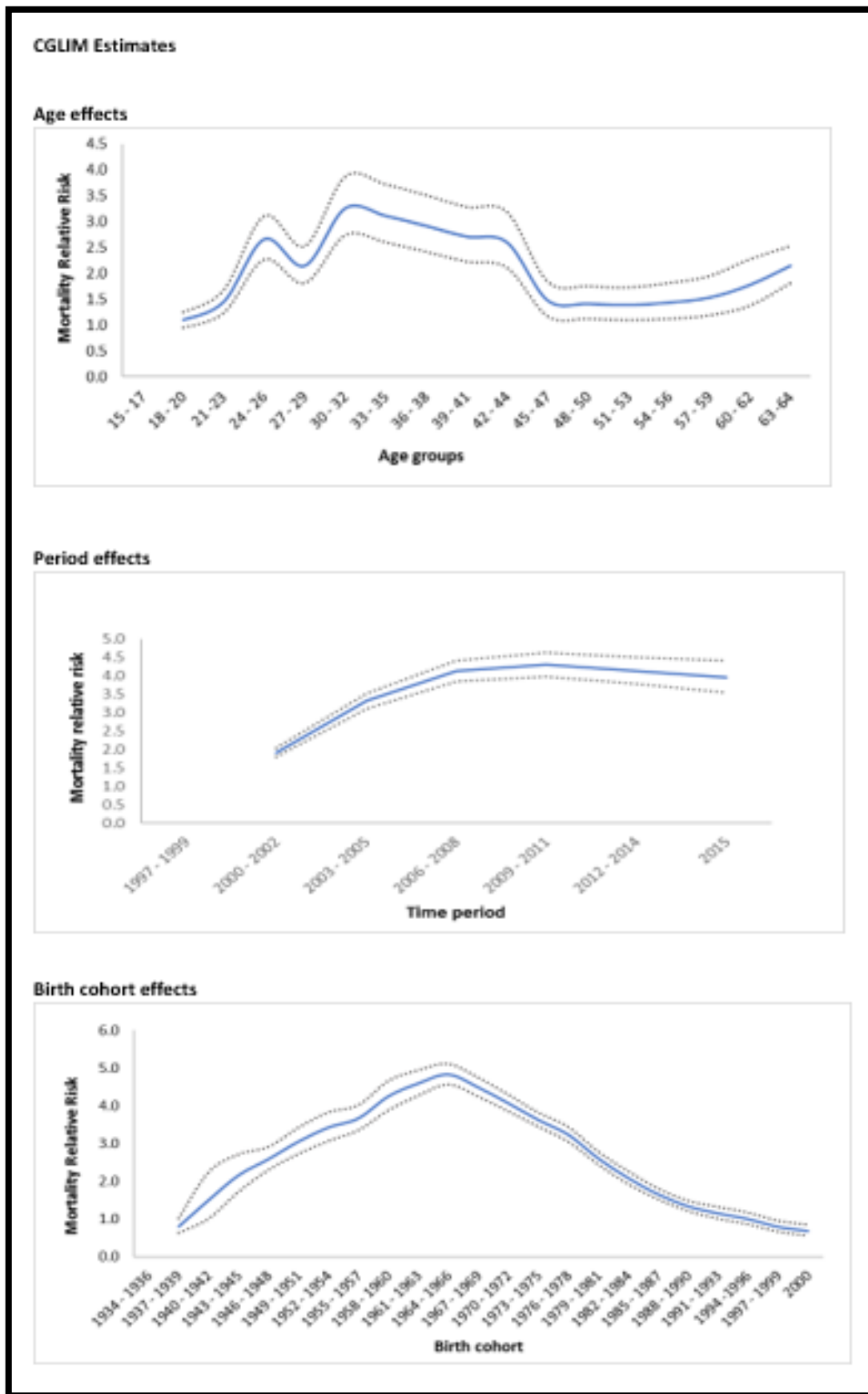


Figure 8: Age, period and cohort effects of HIV-related mortality. Notes: dotted lines present 95% confidence intervals.

Table 6 shows the model selection statistics for example deviance and AIC were used to choose the best model of two factor and full three factor APC models. Based on the measures of estimates, none of the three components of the APC models should be eliminated from the model specification. The full APC model fit the data significantly better than the other three models. Outputs from the fitted models are shown in Appendix II.

Table 6: Summary of the goodness of fit

Cause-Specific Mortality	Age and Period (AP)	Age and Cohort (AC)	Period and Cohort (PC)	Age, Period and Cohort (APC)
TB				
Log-likelihood	-365127	-395163	-390506	-354507
Deviance	727230	787302	777988	705991
BIC	725339	785411	781070	704124
AIC	2173	2353	781032	2110
HIV				
Log-likelihood	-98680	-96518	-108390	-92250
Deviance	194811	190486	214229	181951
BIC	192920	188595	216838	180084
AIC	588	575	216800	549

Lower AIC and Deviance indicates best model in each cause-specific mortality.

CHAPTER 4: DISCUSSION AND CONCLUSION

4.1 Introduction

This chapter discusses the key findings from the analysis in this study and how it varies from or is similar to findings from other studies. The strengths and limitations of the study have also been discussed. Future work has been suggested and conclusions are also provided in this chapter.

4.2 Discussions

In South Africa, young adult mortality is important to provide a reliable population estimates and projections that underlie planning in any sector especially with the impact of HIV/AIDS, TB, injuries and emerging non-communicable diseases. This study set out to simultaneously estimate effects of age, period and birth cohort on mortality due to TB and HIV among young adults (ages 15-64 years) in South Africa between 1997 and 2015. An Age, Period and Cohort Model based on Poisson Distributions for the deaths counts was used. Our study also adjusted HIV/AIDS deaths for misclassification and under-reporting using descriptive statistics.

Findings confirmed the substantial misclassification of HIV/AIDS deaths in South Africa's vital registration system reported in the literature. Our study identified four AIDS related conditions which are source of HIV misclassification and revealed a downward trend in HIV death rates over time, with a peak during the year period 2006. A concave down association between age and TB mortality, with a peak at 33-35 years. There was a concave down relationship between TB cause-specific mortality in the studies mortality data period: 1997 and 2015. There was a downward trend between TB mortality and the effect of birth cohort from 1934 to 2000. There was an inverse flatter U-shaped association between age and HIV-related mortality and was more pronounced at 30-32 years. The estimated relative risks showed an inverse U-shaped relationship between HIV-related mortality and effect of period

from 1997 to 2015. An inverted V-shape relationship between birth cohort and HIV-related mortality was estimated.

Results from age, and period analysis are comparable to official statistics report, global TB and HIV-related mortality report and other similar research findings in South Africa, however, differs from official statistics of South Africa according to period. Our study showed that age groups 36-38 were more at risk of TB mortality than all other age groups, which correspond to Kootbodien et.al (2018) findings. We reported a decline of TB death rates in South Africa since 2007 to 2015, which is in line with the 68.8% to 77.2% global reduction in TB mortality from 2007 to 2014 respectively. The similar trend for TB deaths over time was reported by the official statistics of South Africa which was also confirmed by Pillay-van Wyk et al. (2016). Furthermore, Day and gray (2015) reported a concave down relationship in TB mortality over the period 2001 to 2013 while an inverted u-shaped association between TB mortality and during the period 2000 to 2015 was reported by WHO (2018). Meanwhile, there are possible explanations for the decreasing TB death rates over time among young adults in South Africa. Such rapid decrease may be due to an increase in treatment success rate (83%) achieved in the year 2016 and reported by the National TB statistics for South Africa. The decline in the change rate over time represent improvements in the initiatives taken to mitigate the burden of TB mortality (Karim et.al 2009, World Mortality Report 2011 and Groenewald et.al 2017). A burden of TB mortality was noticed in the earlier cohorts compared to later cohorts. This may be due to the better living and health conditions of younger cohorts and their exposure at younger ages under these favorable circumstances (WHO 2018). This may also be related to advanced medical development and antibiotics usage during the 21st century. This study showed higher TB mortality among older birth cohorts; therefore, it is essential to further strengthen screening, immunization, and treatment for older cohorts at high risk.

Findings from this study are comparable to the National Burden of Diseases of South Africa and other studies conducted to assess and quantify the misclassification and under-reporting of HIV/AIDS deaths in South Africa (Pillay-van Wyk et.al 2016, Birnbaum, Murray and Lazano 2012). Among the four AIDS-related causes identified in this study, three are

document as HIV/AIDS ICD-10 codes by the national burden of disease list for the South Africa (Pillay-Van Wyk et.al 2014) while one was confirmed by Groenewald et.al (2005). This study has shown higher risk of HIV-related mortality among age groups 30-32, 33-35 and 36-38 years compared to all other age groups. The impact of HIV-related mortality among age group 30-34 years in South Africa is justified in a study conducted by De wet, Oluwaseyi and Odimegwu (2014) during 2001-2009 period. Furthermore, Herbest, Mafojane and Newell (2011) also suggest a huge impact of HIV-related mortality among the childbearing age (15-49 years). Pillay-van Wyk et.al (2016) reported a downward trend in HIV/AIDS deaths from 2007 while De Wet et.al (2014) showed that the probability of dying from HIV/AIDS decreased in 2007-2009, which is similar to the findings of this study. Factors like an introduction of antiretroviral (ARV) treatment programme at the public health sectors (WHO 2008) and an improvement towards completeness of vital registration with less misclassification over time (STATSSA 2016) may have accounted for decrease in HIV/AIDS Mortality. This study indicated higher HIV-related mortality risk for birth cohort 1946-1949 to 1982-1984 with even greater risk for birth cohort 1964-1966. The results in this study agree with the world mortality report by UN, where it is indicated that mortality risks due to HIV/AIDS compared to all causes of mortality rates were higher among 1950 to 1955 birth cohort. Furthermore, they found the highest AIDS mortality risk among 1970-1975 birth cohort compared to the earlier and later birth cohorts. MacLead, Majuba and Tipping (2018) confirmed a higher HIV/AIDS mortality among the older birth cohort of South Africa. Factors like time to exposure, weak immune system and lack of education may have contributed to greater risk of HIV-related mortality among older birth cohort (MacLead et.al 2018).

4.3 Limitations of the study

The findings and subsequent discussions are subject to limitations of our study. For example, one of the limitations of the study was that APC models do not test hypotheses about the effects of environmental or historical influences; instead, they organize data and provide useful mathematical formulae for summarizing mortality rates over time. Officially published

mortality data was used, which are based on underlying cause of death. Furthermore, mortality data are under-reported, misclassified and incomplete (Bradshaw et al. 2012), which could have affected and confounded the observed rates and their interpretation. Given these limitations, mortality estimates should be interpreted with caution.

4.4 Recommendations for Further Research

Misclassification of HIV/AIDS deaths and large number of ill-defined deaths is a problem in South Africa. Therefore, to estimate death rates which are reliable, it is necessary to adjust mortality for misclassification, ill-defined and under-reporting. This study only focused on the effect of age, period and cohort on TB and HIV-related mortality. However, one can assess the effect of adjusting other factors such as gender, race and socio-economic factors and consider the HIV\TB co-infected population. Future studies will include the top 10 leading causes of death with a longer period and adjusting with other demographic factors. How mortality trends vary at the provincial level could be looked at using the other APC approach, for example, Bayesian approach, nonparametric bounds and the hierarchical. Age-Period-Cohort (HAPC). This study used aggregated data for HIV and TB cause-specific mortality, however individual data can be used to assess the effect of age, period and birth cohort on mortality.

4.5 Conclusions

In conclusion, despite the limitations of official published mortality data, we still found a notable age and period effect on TB cause-specific mortality according to APC analysis, which were similar to the results of previous studies. However, in contrast with official report of South Africa, we found a different trend for HIV cause-specific death rates from revised mortality data. Age-Period-Cohort Model of HIV and TB mortality offers a more robust assessment of effect of age, period and birth cohort, which would not be possible using traditional Poisson regression model on the death counts separately. This provides more and

relevant useful information for effective monitoring and evaluation of public health policies and programmes targeting mortality reduction across age and period in South Africa.

REFERENCES

- AA, M.A. and Naing, N.N. 2012. Comparison between Negative Binomial and Poisson Death Rate Regression Analysis: AIDS Mortality Co-Infection Patients. *IOSR Journal of Mathematics*, 3, 34-38.
- Adair, T. and Lopez, A.D., 2018. Estimating the completeness of death registration: An empirical method. *PloS one*, 13(5), p.e0197047.
- Adeiza M.A., Abba A.A. and Okpapi J.U. 2014. HIV associated tuberculosis: A Sub-Saharan Africa perspective. *Sub-Saharan Afr J med*, 1, 1-14.
- Age-Period-Cohort Analysis available:
<https://www.mailman.columbia.edu/research/population-health-methods/age-period-cohort-analysis#Overview> [Accessed 02 February 2017].
- Aung, Z.Z., Saw, Y.M., Saw, T.N., Oo, N., Aye, H.N.N., Aung, S., Oo, H.N., Cho, S.M., Khaing, M., Kariya, T. and Yamamoto, E., 2019. Survival rate and mortality risk factors among TB–HIV co-infected patients at an HIV-specialist hospital in Myanmar: A 12-year retrospective follow-up study. *International Journal of Infectious Diseases*, 80, pp.10-15.
- Bennett, N.G. and Horiuchi, S., 1984. Mortality estimation from registered deaths in less developed countries. *Demography*, 21(2), pp.217-233.
- Bendavid, E., Holmes, C.B., Bhattacharya, J. and Miller, G., 2012. HIV development assistance and adult mortality in Africa. *Jama*, 307(19), pp.2060-2067.
- Biset Ayalew, M., 2017. Mortality and its predictors among HIV infected patients taking antiretroviral treatment in ethiopia: a systematic review. *AIDS research and treatment*, 2017(2017) 2017 :e5415298.

Bourne, D.E., Thompson, M., Brody, L.L., Cotton, M., Draper, B., Laubscher, R., Abdullah, M.F. and Myers, J.E., 2009. Emergence of a peak in early infant mortality due to HIV/AIDS in South Africa. *Aids*, 23(1), pp.101-106.

Bradshaw, D., Pillay-Van Wyk, V., Laubscher, R., Nojilana, B., Groenewald, P., Nannan, N. and Metcalf, C., 2010. Cause of death statistics for South Africa: Challenges and possibilities for improvement. South African MRC Burden of Disease Research Unit.

Bruno, G., Maule, M., Merletti, F., Novelli, G., Falorni, A., Iannilli, A., Iughetti, L., Altobelli, E., d'Annunzio, G., Piffer, S. and Pozzilli, P., 2010. Age-period-cohort analysis of 1990–2003 incidence time trends of childhood diabetes in Italy: the RIDI study. *Diabetes*, 59(9), pp.2281-2287.

Calot, G. and Sardon J.P. 2004. Methodology for the calculation of Eurostat's demographic indicators. Available: http://www.cedefop.europa.eu/files/Methodology_for_the_calculation_of_Eurostats_demographic_indicators.pdf [Accessed 12 November 2018].

Centers for Disease Control and Prevention. Appendix A: AIDS-Defining Conditions. *MMWR*. 2008; 57(RR10):9

Chang, Y.H., Li, C.Y., Tung, T.H., Tsai, J.J. and Lu, T.H., 2011. Age-period-cohort analysis of mortality from epilepsy in Taiwan, 1971–2005. *Seizure*, 20(3), pp.240-243

Curtin, L.R. and Klein, R.J. 1995. Direct standardization (age-adjusted death rates) (No. 6). Hyattsville, MD: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

Collet, D. (2015). Modelling survival data in medical research. Chapman and Hall/CRC.

da Silva Escada, R.O., Velasque, L., Ribeiro, S.R., Cardoso, S.W., Marins, L.M.S., Grinsztejn, E., da Silva Lourenço, M.C., Grinsztejn, B. and Veloso, V.G., 2017. Mortality in patients with HIV-1 and tuberculosis co-infection in Rio de Janeiro, Brazil-associated factors and causes of death. *BMC infectious diseases*, 17(1), p.373.

- Day, C. and Gray, A., 2014/15. Health and related indicators. South African health review, 2015(1), pp.213-344.
- De Wet, N., Oluwaseyi, S. and Odimegwu, C., 2014. Youth mortality due to HIV/AIDS in South Africa, 2001–2009: An analysis of the levels of mortality using life table techniques. African Journal of AIDS Research, 13(1), pp.13-20.
- Dorrington R.E., Bradshaw D., Laubscher R. and Nannan N. 2018. Rapid mortality surveillance report 2016. Cape Town: South African Medical Research Council.
- Dorrington, R.E., Bradshaw, D. and Wegner, T., 1999. Estimates of the level and shape of mortality rates in South Africa around 1985 and 1990 derived by applying indirect demographic techniques to reported deaths. Cape Town: South African Medical Research Council.
- Duzan, H. and Shariff, N.S.B.M. 2015. Ridge regression for solving the multicollinearity problem: review of methods and models. *Journal of Applied Sciences*, 15(3), 392-404.
- Fazito, E., Cuchi, P., Fat, D.M., Ghys, P.D., Pereira, M.G., Vasconcelos, A.M.N. and Pascom, A.R.P., 2012. Identifying and quantifying misclassified and under-reported AIDS deaths in Brazil: a retrospective analysis from 1985 to 2009. *Sex Transm Infect*, 88(Suppl 2), pp.i86-i94.
- Gaifer ZA. Risk factors for tuberculosis mortality in a tertiary care center in Oman, 2006–2016. *Int J Mycobacteriol* 2017;6:356-9.
- Gao, X., Wang, Z., Kong, C., Yang, F., Wang, Y. and Tan, X., 2017. Trends of esophageal cancer mortality in rural China from 1989 to 2013: an age-period-cohort analysis. *International journal of environmental research and public health*, 14(3), p.218..
- Gesesew, H., Tsehayneh, B., Massa, D., Gebremedhin, A., Kahsay, H. and Mwanri, L., 2016. Predictors of mortality in a cohort of tuberculosis/HIV co-infected patients in Southwest Ethiopia. *Infectious diseases of poverty*, 5(1), p.109.

UNAIDS. 2016. Global AIDS Response Progress Reporting (GARPR) 2016

Groenewald, P., Nannan, N., Bourne, D., Laubscher, R. and Bradshaw, D., 2005. Identifying deaths from AIDS in South Africa. *Aids*, 19(2), pp.193-201.

Groenewald, P. et al. 2017. Mortality trends in the City of Cape Town between 2001 and 2013: Reducing inequities in health. *South African Medical Journal*, 107(12), 1091-1098.

Herbst, A.J., Mafojane T. and Newell M. 2011. Verbal autopsy-based cause-specific mortality trends in rural KwaZulu-Natal, South Africa, 2000-2009. *Population Health Metrics*, 9, 47.

Hill, K. and Queiroz, B., 2010. Adjusting the general growth balance method for migration. *Revista Brasileira de Estudos de População*, 27(1), pp.7-20.

Holford, T.R. (1983). The Estimation of Age, Period and Cohort Effects for Vital Rates. *Biometrics*, 39, 311-24. 10.2307/2531004.

Jee Y.H. and Cho S. 2016. Age-period-cohort analysis of smoking prevalence among young adults in Korea. *Epidemiology and Health*. 38:e2016010. doi:10.4178/epih.e2016010.

Johnson, L.F., Dorrington, R.E., 2017. Thembisa version 3.2: A model for evaluating the impact of HIV/AIDS in South Africa. Centre for Infectious Disease Epidemiology and Research, University of Cape Town. Available:
http://www.thembisa.org/content/downloadPage/Thembisa2_4 [Accessed 26 October 2018].

Karim, S.S.A., Churchyard, G.J., Karim, Q.A. and Lawn, S.D., 2009. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *the Lancet*, 374(9693), pp.921-933.

Keyes K. and Li G. 2010. A Multiphase Method for Estimating Cohort Effects in Age-Period Contingency Table Data. *Annals of Epidemiology*, 20(10), 779-85. 10.1016/j.annepidem.2010.03.006.

- Keyes, K.M. and Miech, R., 2013. Age, period, and cohort effects in heavy episodic drinking in the US from 1985 to 2009. *Drug and alcohol dependence*, 132(1-2), pp.140-148.
- Kootbodien, T., Wilson, K., Tlotleng, N., Ntlebi, V., Made, F., Rees, D. and Naicker, N., 2018. Tuberculosis Mortality by Occupation in South Africa, 2011–2015. *International journal of environmental research and public health*, 15(12), p.2756.
- Lee, R.D. and Carter, L.R., 1992. Modeling and forecasting US mortality. *Journal of the American statistical association*, 87(419), pp.659-671.
- Lima, M.D.S., Martins-Melo, F.R., Heukelbach, J., Alencar, C.H., Boigny, R.N. and Ramos Júnior, A.N., 2016. Mortality related to tuberculosis-HIV/AIDS co-infection in Brazil, 2000-2011: epidemiological patterns and time trends. *Cadernos de saúde publica*, 32, p.e00026715.
- Luo, L. 2013. Assessing validity and application scope of the intrinsic estimator approach to the age-period-cohort problem. *Demography*, 50(6), 1945-1967.
- MacLeod, W., Majuba, P.P. and Tipping, B., 2018. Human immunodeficiency virus infection and older adults: A retrospective single-site cohort study from Johannesburg, South Africa. *Southern African journal of HIV medicine*, 19(1).
- Musenge, E., Vounatsou, P., Collinson, M., Tollman, S., & Kahn, K. (2013). The contribution of spatial analysis to understanding HIV/TB mortality in children: a structural equation modelling approach. *Global health action*, 6, 19266. doi:10.3402/gha.v6i0.19266
- Naing, N.N. 2000. Easy way to learn standardization: direct and indirect methods. *The Malaysian Journal of Medical Sciences: MJMS*, 7(1), 10.
- Namosha, E., Sartorius, B. and Tanser, F., 2013. Spatial clustering of all-cause and HIV-related mortality in a rural South African population (2000–2006). *PloS one*, 8(7), p.e69279.

Ngom, P. and Clark, S. 2003. Adult mortality in the era of HIV/AIDS: Sub-Saharan Africa. A paper presented in a workshop on HIV/AIDS and adult mortality in developing countries, 8-15, September 2003, New York.

Nguyen, D.T., Jenkins, H.E. and Graviss, E.A., 2018. Prognostic score to predict mortality during TB treatment in TB/HIV co-infected patients. *PloS one*, 13(4), p.e0196022.

Ocaña-Riola R., Mayoral-Cortés J.M. and Blanco-Reina E. 2013. Age-period-cohort effect on lung cancer mortality in southern Spain. *Eur J Cancer Prev*, 22(6), 549-57.

Onyango, D.O., Yuen, C.M., Cain, K.P., Ngari, F., Masini, E.O. and Borgdorff, M.W., 2017. Reduction of HIV-associated excess mortality by antiretroviral treatment among tuberculosis patients in Kenya. *PloS one*, 12(11), p.e0188235.

Osagie, A.M. and Adebukola, A.C. 2017. Fitting a Poisson Regression Model to Reported Deaths from HIV/AIDS in Nigeria. *International Journal of Statistics and Actuarial Science*, 1(5), 117-121. doi: 10.11648/j.ijzas.20170105.12

Pillay-van Wyk, V., Msemburi, W., Laubscher, R., Dorrington, R.E., Groenewald, P., Glass, T., Nojilana, B., Joubert, J.D., Matzopoulos, R., Prinsloo, M. and Nannan, N., 2016. Mortality trends and differentials in South Africa from 1997 to 2012: second National Burden of Disease Study. *The Lancet Global Health*, 4(9), pp.e642-e653.

Pillay-Van Wyk, V.; Laubscher, R.; Msemburi, W.; Dorrington, R.E.; Groenewald, P.; Vos, T.; Matzopoulos, R.; Prinsloo, M.; Nojilana, B.; Nannan, N.; et al. Second South African National Burden of Disease Study: Data Cleaning, Validation and SA NBD List; Burden of Disease Research Unit: Cape Town, South Africa, 2014.

Podlekareva, D.N., Panteleev, A.M., Grint, D., Post, F.A., Miro, J.M., Bruyand, M., Furrer, H., Obel, N., Girardi, E., Vasilenko, A. and Losso, M.H., 2014. Short-and long-term mortality and causes of death in HIV/tuberculosis patients in Europe. *European respiratory journal*, 43(1), pp.166-177.

Podlekareva, D.N., Efsen, A.M.W., Schultze, A., Post, F.A., Skrahina, A.M., Panteleev, A., Furrer, H., Miller, R.F., Losso, M.H., Toibaro, J. and Miro, J.M., 2016. Tuberculosis-related mortality in people living with HIV in Europe and Latin America: an international cohort study. *The lancet HIV*, 3(3), pp.e120-e131.

Queiroz, A.A.R.D., Berra, T.Z., Garcia, M.C.D.C., Popolin, M.P., Belchior, A.D.S., Yamamura, M., Santos, D.T.D., Arroyo, L.H. and Arcêncio, R.A., 2018. Spatial pattern and temporal trend of mortality due to tuberculosis. *Revista latino-americana de enfermagem*, 2018;26:e2992.

Renshaw, A.E. and Haberman, S., 2006. A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and economics*, 38(3), pp.556-570.

Rosenberg, P.S. and Anderson, W.F. 2011. Age-period-cohort models in cancer surveillance research: ready for prime time? *Cancer Epidemiol Biomarkers Prev.* 2011; 20:1263–1268.

Russo S., Flender, D. and Franssica da Silva, G. 2012. Poisson Regression Models for Count Data: Use in the Number of Deaths in the Santo Angelo (Brazil). *Journal of Basic & Applied Sciences*, 8, 266-269.

Sifuna, P., Otieno, L., Andagalu, B., Oyieko, J., Ogutu, B., Singoei, V., Owuoth, J., Ogwang, S., Cowden, J. and Otieno, W., 2018. A Spatiotemporal Analysis of HIV-Associated Mortality in Rural Western Kenya 2011–2015. *Journal of acquired immune deficiency syndromes* (1999), 78(5), p.483.

Simpfendorfer, C.A., Bonfil, R. and Latour, R.J., 2005. Mortality estimation. *FAO Fisheries Technical Paper*, 474, p.127.

Silva, R.M.A., 2013. Papers on Indirect Mortality Estimation & Analysis in Low-Resource Settings (Doctoral dissertation, UC Berkeley).

Statistics South Africa. Estimating completeness of adult mortality data at sub-national level, 2014: Findings from death notification / Statistics South Africa Pretoria. Stats SA Library Cataloguing-in-Publication (CIP) Data.

Statistics South Africa. Mortality and causes of death in South Africa, 2015: Findings from death notification / Statistics South Africa Pretoria. Stats SA Library Cataloguing-in-Publication (CIP) Data.

Statistics South Africa. Mortality and causes of death in South Africa, 2017: Findings from death notification / Statistics South Africa Pretoria. Stats SA Library Cataloguing-in-Publication (CIP) Data.

Takarinda, K.C., Sandy, C., Masuka, N., Hazangwe, P., Choto, R.C., Mutasa-Apollo, T., Nkomo, B., Sibanda, E., Mugurungi, O., Harries, A.D. and Siziba, N., 2017. Factors associated with mortality among patients on TB treatment in the southern region of Zimbabwe, 2013. *Tuberculosis research and treatment*, 2017.

Tetteh-Ahinakwa, L. and Oduro, F.T. 2017. Poisson Regression Model on Infant Mortality Occurrence and Incidence in the Kumasi District of Ghana. *International Journal of Statistics and Applications*, 7(2), 113-116.

United Nations, Department of Economic and Social Affairs, Population Division. 2012. World Mortality Report 2011 (United Nations publication).

Wang, Z., Bao, J., Yu, C., Wang, J. and Li, C., 2015. Secular trends of breast cancer in China, South Korea, Japan and the United States: application of the age-period-cohort analysis. *International journal of environmental research and public health*, 12(12), pp.15409-15418. World Health Organization ed. 2018. Global tuberculosis report 2018. World Health Organization.

Xu, M. and Powers, D.A. 2016. Bayesian ridge estimation of age-period-cohort models. In *Dynamic Demographic Analysis* (pp. 337-359). Springer, Cham.

Yang, Y., Schulhofer-Wohl, S., Fu, W.J. and Land, K.C., 2008. The intrinsic estimator for age-period-cohort analysis: what it is and how to use it. *American Journal of Sociology*, 113(6), pp.1697-1736.

Yang, Y. and Land, K.C. 2008. Age-period-cohort analysis of repeated cross-section surveys: fixed or random effects? *Sociological Methods & Research*, 36(3), 297-326.

Yang, Y., Fu, W.J. and Land, K.C. 2004. A Methodological Comparison of Age-Period-Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models. *Sociological Methodology*, 34, 75-110. doi:10.1111/j.0081-1750.2004.00148.

APPENDIX I

Two-way table of number of deaths that wer modelled by age-period-cohort models

		Time Period						
Age		1997-	2000-		2006-	2009-	2012-	
Groups		1999	2002	2003-2005	2008	2011	2014	2015
Mortality								
due to TB	15-17	672	1,086	1,242	1,281	1,312	1,187	373
	18-20	1,923	3,193	3,872	3,465	2,953	2,326	635
	21-23	4,119	6,992	8,800	8,674	6,669	4,586	1,221
	24-26	7,121	12,781	15,666	15,602	12,639	8,480	2,154
	27-29	8,821	17,514	23,327	21,787	18,200	12,855	3,309
	30-32	9,482	19,398	27,279	27,547	20,890	15,546	4,324
	33-35	9,752	18,175	27,173	28,317	23,366	16,512	4,654
	36-38	8,677	17,217	23,639	26,736	22,650	17,013	4,522
	39-41	7,861	14,536	22,209	22,674	20,332	16,104	4,702
	42-44	6,678	12,490	18,087	21,193	17,505	14,441	4,167
	45-47	6,009	10,099	15,292	17,584	16,225	12,451	3,889
	48-50	5,294	8,881	12,070	14,466	13,520	11,700	3,191
	51-53	4,024	7,069	10,138	11,509	11,540	9,884	3,287
	54-56	3,572	5,001	8,060	9,687	9,197	8,410	2,619
	57-59	3,758	4,318	5,580	7,315	7,768	7,048	2,242
	60-62	2,827	4,085	4,309	5,127	6,259	6,070	2,001
	63-64	1,526	1,980	2,796	2,913	3,025	3,431	1,101
Mortality								
due to HIV	15-17	547	951	1,115	1,118	1,064	845	216
	18-20	1,427	2,679	3,296	2,954	2,341	1,598	404
	21-23	2,917	5,625	7,276	7,213	5,128	2,925	720
	24-26	4,966	10,224	12,861	12,734	9,447	5,244	1,178
	27-29	6,229	14,090	19,075	17,848	13,576	7,882	1,715
	30-32	6,674	15,592	22,335	22,512	15,513	9,310	2,256
	33-35	7,067	14,686	22,345	23,183	17,345	10,002	2,449
	36-38	6,451	14,118	19,517	22,057	16,897	10,297	2,406
	39-41	6,034	12,036	18,537	18,852	15,527	10,061	2,562
	42-44	5,326	10,573	15,251	17,726	13,323	9,035	2,330
	45-47	4,935	8,724	13,096	14,937	12,677	8,068	2,283
	48-50	4,525	7,823	10,560	12,463	10,696	7,828	1,850
	51-53	3,545	6,315	9,029	10,124	9,330	6,901	2,052
	54-56	3,231	4,577	7,286	8,590	7,597	5,994	1,696

57-59	3,494	4,038	5,112	6,637	6,548	5,277	1,539
60-62	2,674	3,880	4,038	4,715	5,394	4,649	1,451
63-64	1,454	1,903	2,675	2,711	2,697	2,695	805

Populations

15-17	8505345	9140314	9625088	9565414	9088974	8547570	2769263
18-20	7,856,078	8,442,573	9,216,711	9,619,084	9,530,812	9,083,351	2,909,211
21-23	7,020,790	7,544,929	8,458,570	9,293,025	9,731,003	9,597,122	3,113,392
24-26	6,541,068	7,029,393	7,567,258	8,437,211	9,349,416	9,756,526	3,278,991
27-29	6,128,401	6,585,922	6,862,404	7,553,212	8,563,231	9,427,061	3,291,478
30-32	5,737,708	6,166,060	6,571,654	6,889,570	7,397,208	8,486,409	3,078,434
33-35	5,370,592	5,771,535	6,087,349	6,266,258	6,626,480	7,457,888	2,704,563
36-38	5,003,300	5,376,822	5,525,614	5,714,209	6,111,948	6,466,291	2,274,393
39-41	4,597,070	4,940,266	5,148,025	5,255,319	5,442,051	5,821,785	2,059,244
42-44	4,184,274	4,496,651	4,755,770	4,840,350	4,913,999	5,278,585	1,863,918
45-47	3,775,625	4,057,499	4,291,788	4,444,793	4,606,729	4,765,128	1,636,837
48-50	3,336,297	3,585,369	3,829,345	4,024,392	4,213,928	4,364,166	1,492,451
51-53	2,887,112	3,102,650	3,364,333	3,583,194	3,772,769	4,012,266	1,388,761
54-56	2,430,940	2,612,424	2,874,918	3,135,948	3,364,192	3,591,582	1,245,196
57-59	2,053,131	2,206,408	2,425,609	2,681,565	2,935,742	3,160,310	1,101,274
60-62	1,800,295	1,934,696	2,044,636	2,221,614	2,478,536	2,729,426	962,685
63-64	1,062,573	1,141,897	1,195,863	1,276,425	1,409,515	1,568,052	559,117

APPENDIX II

Estimates of the fitted models using APC

TB Cause-Specific Mortality	AP			AC			PC			APC		
	Coef.	95% CI		Coef.	95% CI		Coef.	95% CI		Coef.	95% CI	
Age												
_spA1_intercept	-7.30	-7.30	-7.29	-7.32	-7.32	-7.31	-	-	-	-7.19	-7.20	-7.19
_spA2	0.61	0.61	0.61	0.48	0.47	0.48	-	-	-	0.54	0.54	0.54
_spA3	0.61	0.61	0.61	0.38	0.38	0.39	-	-	-	0.40	0.40	0.41
_spA4	-0.22	-0.22	-0.21	-0.19	-0.20	-0.19	-	-	-	-0.20	-0.20	-0.19
_spA5	-0.05	-0.06	-0.05	-0.03	-0.03	-0.03	-	-	-	-0.02	-0.03	-0.02
_spA6	-0.02	-0.02	-0.01	-0.01	-0.02	-0.01	-	-	-	-0.01	-0.01	-0.01
Period												
_spP1	-0.03	-0.04	-0.03	-	-	-	0.00	0.00	0.00	0.00	0.00	0.00
_spP2	0.31	0.31	0.32	-	-	-	0.06	0.06	0.07	-0.05	-0.05	-0.04
_spP3	-0.05	-0.05	-0.04	-	-	-	0.47	0.47	0.47	0.29	0.29	0.29
_spP4	-0.08	-0.08	-0.08	-	-	-	0.07	0.06	0.07	-0.08	-0.08	-0.08
_spP5	-0.08	-0.09	-0.08	-	-	-	0.05	0.04	0.05	-	-	-
Cohort												
_spC1	-	-	-	-0.20	-0.20	-0.19	0.30	0.30	0.31	0.14	0.14	0.15
_spC2	-	-	-	0.33	0.33	0.34	0.63	0.63	0.64	0.27	0.26	0.27
_spC3	-	-	-	0.14	0.14	0.14	0.13	0.13	0.14	0.04	0.04	0.05
_spC4	-	-	-	0.08	0.07	0.08	0.02	0.02	0.02	-0.02	-0.02	-0.01
_spC5	-	-	-	-0.01	-0.01	-0.01	-	-	-	-	-	-
HIV Cause-Specific Mortality												
Age												
_spA1_intercept	-8.96	-8.96	-8.95	-8.47	-8.48	-8.46	-	-	-	-8.75	-8.76	-8.74
_spA2	0.39	0.38	0.40	1.02	1.01	1.03	-	-	-	0.42	0.41	0.43
_spA3	0.80	0.79	0.81	0.49	0.48	0.50	-	-	-	0.49	0.48	0.49
_spA4	-0.19	0.20	-0.19	-0.23	-0.24	-0.23	-	-	-	-0.23	-0.24	-0.23
_spA5	-0.03	0.04	-0.03	-0.03	-0.04	-0.03	-	-	-	-0.03	-0.04	-0.03
_spA6	-0.01	-0.02	0.00	-0.01	-0.02	-0.01	-	-	-	-0.01	-0.02	0.00
Period												
_spP1	0.31	0.31	0.32	-	-	-	0.05	0.04	0.05	0.06	0.05	0.06
_spP2	-0.01	0.01	0.00	-	-	-	0.14	0.14	0.15	-0.04	-0.04	-0.04
_spP3	-0.04	0.04	-0.04	-	-	-	0.11	0.11	0.12	-0.02	-0.02	-0.01
_spP4	-0.05	0.05	-0.04	-	-	-	0.10	0.09	0.10	-0.04	-0.05	-0.04
_spP5	-0.02	0.03	-0.02	-	-	-	0.07	0.06	0.07	-0.04	-0.04	-0.04
Cohort												
_spC1	-	-	-	0.75	0.74	0.76	0.15	0.14	0.15	0.03	0.03	0.04
_spC2	-	-	-	0.47	0.46	0.48	0.85	0.84	0.86	0.48	0.47	0.49
_spC3	-	-	-	0.06	0.06	0.07	0.27	0.26	0.28	0.07	0.06	0.08
_spC4	-	-	-	0.03	0.03	0.04	0.04	0.03	0.04	-0.03	-0.03	-0.02
_spC5	-	-	-	-0.02	-0.03	-0.02	-	-	-	-	-	-

APPENDIX III

Age-period-cohort estimates for TB cause-specific mortality

TB Cause-Specific Mortality	Constrained Generalised Linear Model					
	Period 2003-2005 = 2006-2008			Cohort 1937-1938=2000		
	IR	95% CI		IR	95% CI	
Age						
15 - 17	1.00	-	-	1.00	-	-
18 - 20	2.59	2.51	2.67	2.52	2.44	2.60
21 - 23	5.51	5.33	5.69	5.21	5.05	5.37
24 - 26	9.61	9.27	9.96	8.85	8.56	9.14
27 - 29	13.60	13.07	14.14	12.17	11.74	12.62
30 - 32	16.33	15.63	17.05	14.22	13.66	14.81
33 - 35	17.97	17.12	18.85	15.22	14.55	15.93
36 - 38	18.81	17.84	19.84	15.51	14.74	16.31
39 - 41	19.49	18.39	20.66	15.62	14.77	16.53
42 - 44	19.90	18.67	21.20	15.52	14.58	16.51
45 - 47	20.61	19.23	22.08	15.63	14.60	16.73
48 - 50	21.22	19.69	22.87	15.66	14.54	16.86
51 - 53	22.10	20.39	23.95	15.86	14.64	17.19
54 - 56	22.77	20.89	24.82	15.90	14.58	17.34
57 - 59	23.66	21.58	25.94	16.07	14.65	17.63
60 - 62	23.45	21.26	25.88	15.50	14.04	17.10
63 - 64	28.36	25.53	31.50	18.23	16.41	20.25
Period						
1997 - 1999	1.00	-	-	1.00	-	-
2000 - 2002	1.75	1.73	1.77	1.79	1.77	1.82
2003 - 2005	2.26	2.22	2.30	2.39	2.35	2.43
2006 - 2008	2.26	2.22	2.30	2.46	2.40	2.51
2009 - 2011	1.68	1.64	1.73	1.88	1.83	1.93
2012 - 2014	1.03	1.00	1.07	1.18	1.14	1.23
2015	1.14	1.10	1.19	1.35	1.30	1.41
Cohort						
1934 - 1936	1.00	-	-	1.00	-	-
1937 - 1939	0.79	0.74	0.84	0.77	0.72	0.82
1940 - 1942	0.78	0.73	0.82	0.73	0.69	0.78
1943 - 1945	0.63	0.59	0.66	0.58	0.54	0.61
1946 - 1948	0.59	0.56	0.63	0.53	0.50	0.56
1949 - 1951	0.73	0.68	0.77	0.63	0.59	0.67
1952 - 1954	0.80	0.74	0.85	0.67	0.63	0.72
1955 - 1957	0.83	0.77	0.89	0.68	0.64	0.74
1958 - 1960	0.92	0.86	1.00	0.74	0.69	0.80
1961 - 1963	0.99	0.92	1.07	0.77	0.71	0.84
1964 - 1966	1.11	1.02	1.21	0.84	0.77	0.92
1967 - 1969	1.11	1.02	1.21	0.82	0.75	0.90
1970 - 1972	1.21	1.10	1.33	0.87	0.79	0.96
1973 - 1975	1.23	1.11	1.36	0.86	0.77	0.95
1976 - 1978	1.20	1.08	1.33	0.81	0.73	0.91
1979 - 1981	1.05	0.94	1.17	0.70	0.62	0.78
1982 - 1984	0.94	0.84	1.06	0.61	0.54	0.68
1985 - 1987	0.85	0.75	0.95	0.53	0.47	0.60
1988 - 1990	0.75	0.66	0.85	0.45	0.40	0.52
1991 - 1993	0.75	0.65	0.85	0.44	0.38	0.51
1994 - 1996	0.89	0.77	1.02	0.51	0.44	0.59
1997 - 1999	1.18	1.02	1.38	0.66	0.57	0.78
2000	1.41	1.16	1.73	0.77	0.72	0.82

Constrained by period and birth cohort

Age-period-cohort estimates for HIV cause-specific mortality

HIV Cause-Specific Mortality	Constrained Generalised Linear Model					
	Period: 2003-2005 = 2006-2008			Cohort: 1937-1938 = 2000		
	IR	95% CI		IR	95% CI	
Age						
15 - 17	1.00	-	-	1.00	-	-
18 - 20	3.04	2.84	3.26	2.71	2.54	2.90
21 - 23	7.59	7.06	8.16	6.04	5.66	6.43
24 - 26	14.15	13.08	15.30	10.03	9.41	10.69
27 - 29	19.85	18.19	21.66	12.55	11.74	13.41
30 - 32	24.20	21.97	26.67	13.64	12.71	14.64
33 - 35	26.15	23.47	29.13	13.14	12.18	14.17
36 - 38	26.80	23.78	30.20	12.00	11.06	13.03
39 - 41	26.46	23.19	30.18	10.57	9.66	11.55
42 - 44	26.62	23.05	30.75	9.48	8.60	10.45
45 - 47	26.02	22.24	30.45	8.26	7.44	9.18
48 - 50	25.89	21.84	30.70	7.33	6.54	8.21
51 - 53	25.78	21.46	30.99	6.51	5.76	7.36
54 - 56	26.27	21.56	32.00	5.91	5.18	6.74
57 - 59	26.38	21.36	32.59	5.29	4.60	6.10
60 - 62	27.46	21.91	34.42	4.91	4.24	5.70
63 - 64	31.71	24.89	40.41	5.06	4.31	5.94
Period						
1997 - 1999	1.00	-	-	1.00	-	-
2000 - 2002	1.12	1.09	1.14	1.25	1.23	1.28
2003 - 2005	1.33	1.28	1.38	1.67	1.63	1.72
2006 - 2008	1.33	1.28	1.38	1.88	1.81	1.94
2009 - 2011	1.52	1.43	1.61	2.40	2.30	2.51
2012 - 2014	1.71	1.59	1.84	3.04	2.88	3.20
2015	2.54	2.33	2.77	5.05	4.75	5.38
Cohort						
1934 - 1936	1.00	-	-	1.00	-	-
1937 - 1939	0.94	0.72	1.23	0.84	0.64	1.10
1940 - 1942	1.25	0.97	1.61	0.99	0.78	1.28
1943 - 1945	1.43	1.11	1.82	1.01	0.79	1.29
1946 - 1948	1.71	1.34	2.19	1.08	0.85	1.38
1949 - 1951	2.55	1.99	3.27	1.44	1.13	1.83
1952 - 1954	3.23	2.51	4.16	1.62	1.27	2.07
1955 - 1957	3.69	2.85	4.78	1.66	1.29	2.12
1958 - 1960	4.60	3.53	5.98	1.84	1.43	2.36
1961 - 1963	5.28	4.03	6.92	1.88	1.46	2.42
1964 - 1966	6.18	4.68	8.15	1.96	1.52	2.54
1967 - 1969	6.46	4.86	8.59	1.83	1.41	2.37
1970 - 1972	7.00	5.22	9.39	1.77	1.36	2.30
1973 - 1975	7.23	5.35	9.77	1.63	1.24	2.13
1976 - 1978	7.14	5.24	9.74	1.43	1.09	1.89
1979 - 1981	6.19	4.50	8.53	1.11	0.84	1.47
1982 - 1984	5.58	4.02	7.76	0.89	0.67	1.19
1985 - 1987	4.91	3.49	6.89	0.70	0.52	0.93
1988 - 1990	4.28	3.02	6.07	0.54	0.40	0.73
1991 - 1993	4.09	2.85	5.88	0.46	0.34	0.63
1994 - 1996	4.56	3.13	6.63	0.46	0.34	0.63
1997 - 1999	6.04	4.07	8.94	0.54	0.39	0.75
2000	10.45	6.81	16.04	0.84	0.64	1.10

Constrained by period and birth cohort

APPENDIX IV

Project code

```
/**intrinsic Estimator
/**TB cause-specific
apc_ie deaths if cause==1, age( age) period(period) cohort( Coh ) family(poisson) link(log)
exposure( population) eform
/**HIV cause-specific
apc_ie deaths if cause==2, age( age) period(period) cohort( Coh ) family(poisson) link(log)
exposure( population) eform

//Constrained GLIM
/**TB cause-specific//
apc_cglim deaths if cause==1,age(age) period( period ) cohort( Coh )agepfx("_A")
periodpfx("_P") cohortpfx("_C")family(poisson) link(log) exposure( population )
constraint("p2003=p2006")eform
drop _A* _P* _C*
apc_cglim deaths if cause==1,age(age) period( period ) cohort( Coh )agepfx("_A")
periodpfx("_P") cohortpfx("_C")family(poisson) link(log) exposure( population )
constraint("a63=a15")eform
drop _A* _P* _C*
apc_cglim deaths if cause==1,age(age) period( period ) cohort( Coh )agepfx("_A")
periodpfx("_P") cohortpfx("_C")family(poisson) link(log) exposure( population )
constraint("c2000=c1934")eform
drop _A* _P* _C*

/**HIV cause-specific//
apc_cglim deaths if cause==2,age(age) period( period ) cohort( Coh )agepfx("_A")
periodpfx("_P") cohortpfx("_C")family(poisson) link(log) exposure( population )
constraint("p2003=p2006")eform
drop _A* _P* _C*
apc_cglim deaths if cause==2,age(age) period( period ) cohort( Coh )agepfx("_A")
periodpfx("_P") cohortpfx("_C")family(poisson) link(log) exposure( population )
constraint("a63=a15")eform
drop _A* _P* _C*
apc_cglim deaths if cause==2,age(age) period( period ) cohort( Coh )agepfx("_A")
periodpfx("_P") cohortpfx("_C")family(poisson) link(log) exposure( population )
constraint("c2000=c1934")eform
```