Evaluating the adequacy of the method of using vital registration and census data in estimating adult mortality when applied sub-provincially

Chido Chinogurei

University of Cape Town

A dissertation submitted to the Faculty of Commerce of the University of Cape Town in partial fulfilment of the requirements for the degree of Master of Philosophy in Demography

September 2017

ABSTRACT

In developing countries, vital registration is the best source of death data that can be used to estimate adult mortality provided they are sufficiently complete. However, they are usually insufficient for estimating mortality sub-nationally due to incomplete registration. This research adapts a method used by Dorrington, Moultrie and Timæus at the provincial level to determine whether it is adequate for estimating adult mortality at the district municipality level in the year prior to the 2001 census. The method uses registration data adjusted for completeness of registration to scale (up or down) the deaths reported by households in the census by age group for each sex.

The process of correcting the registered deaths in the year prior to the 2001 census involves estimating intercensal completeness for each population group and each sex between 1996 and 2001 using the average of results from the GGB and the SEG+δ methods. Thereafter, the results are used to estimate the completeness in each of the years within the intercensal period. Thus, an estimate of completeness is obtained in the year prior to the 2001 census for correcting the registered deaths at the population group level. These registered deaths are then used to obtain population group specific adjustment factors to correct the deaths reported by households at the district level, and thereafter to estimate adult mortality rates.

Most districts in Kwa-Zulu-Natal have amongst the highest rates of adult mortality, while most districts in the Western Cape have amongst the lowest rates. Results show the Buffalo metropolitan municipality to have higher mortality than that expected for most of the district metropolitan municipalities for both sexes. The same is true for women in Mangaung metropolitan district. It is suspected that HIV prevalence had a significant impact on different levels of adult mortality in the districts, although some adults in the more urban provinces may have died in other provinces. At the provincial level, the method produces marginally higher estimates of adult mortality than the other sources. Provinces that reflect a higher level of mortality appear to deviate more from other research findings than those reflecting lower mortality. In conclusion, the method produces district estimates of ${}_{45}q_{15}$ that are consistent with provincial estimates from other sources and with estimates of HIV prevalence at the district level.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Firstly, I would want to thank God for his grace. I would also like to thank my wife and my son for their endurance and patience through the time of my studies. I also want to thank my supervisor Professor Rob Dorrington for his guidance and assistance throughout the completion of this thesis. Last but not least, I would also want to thank the Andrew Mellon Foundation for funding my stay throughout the completion of my studies.

1 INTRODUCTION

1.1 Background

Adult mortality estimates are an essential component in guiding health policies and interventions. Since the start of the HIV epidemic in South Africa, over half of the deaths in the population have occurred in the adult ages from as early as 1999 (Statistics South Africa, 2005). Although it is important to understand the overall mortality in a nation, gaining a better understanding of the heterogeneity in mortality is of greater significance as it gives a better understanding when targeting specific areas for health intervention, allocating resources and setting priorities where mortality is high (Ahmed & Hill, 2011). At a sub-national level unhealthy districts and healthy districts can be used to generate debate about the general health in the country at the time (Mahapatra, Shibuya, Lopez et al., 2007). Sub-national estimates are essential in identifying the heterogeneity in the population, not just referring to the geographical location but also the population groups within the whole population. However, there is a debate about whether heterogeneity in the mortality of population groups is purely biological or due to social outcomes (Moultrie & Dorrington, 2012)? None the less it is certain that heterogeneity exists and should be accounted for. However, the required data for the measurement of sub-national level estimates are often not readily available, particularly in developing countries and the world's poorest populations (Joubert, Rao, Bradshaw et al., 2012).

Although much progress has been made in estimating adult mortality at national levels in developing countries where the data are often incomplete or biased, no method currently exists that can be used to determine mortality accurately at a sub-national level (Ahmed & Hill, 2011). In most developing countries, death data used to obtain adult mortality estimates is sourced from the vital registration system and the census. The methods used to obtain mortality estimates can be either direct or indirect, depending on the nature of the data being used. Indirect approaches to estimating adult mortality nationally include the orphanhood and sibling history methods, whilst direct techniques include death distribution methods. However, use of these methods on a sub-national level is problematic for several reasons, most resulting from incomplete and missing information from the data sources.

There have been several attempts to produce provincial estimates of adult mortality (Statistics South Africa, 2000, Dorrington, Moultrie & Timæus, 2004, Dorrington, Timæus, Moultrie et al., 2004, Msemburi, Pillay-van Wyk, Dorrington et al., 2014, Udjo & Lalthapersad-Pillay, 2014). The ASSA provincial demographic models also project the rates of adult mortality (Actuarial Society of South Africa, 2011). Among these, some studies (Udjo & Lalthapersad-Pillay, 2014, Dorrington, Timæus, Moultrie et al., 2004, Dorrington, Moultrie & Timæus, 2004, Dorrington & Timæus, 2015) have used a combination of vital registration and census data to come up with the provincial mortality rates.

Although the completeness of the vital registration deaths in South Africa has improved in the past two decades, vital registration still has some weaknesses that make it difficult to produce reliable estimates sub-nationally. As far as registered deaths are concerned there are three major concerns: coverage, completeness and missing data (Mahapatra, Shibuya, Lopez et al., 2007). Coverage is often higher in urban areas than rural areas, even in some areas where there is 100 per cent coverage, the data often contains missing information. Most if not all developing countries have incomplete death registration, but the completeness of registered deaths in South Africa has been improving over the last decade. Because ignoring unregistered deaths will lead to overstating life expectancy among other things, death distribution methods have been used to estimate and thus correct for incompleteness of death registration. However, one of the inherent weaknesses of vital registration data lies in its failure to locate the usual place of residence of the deceased, which is essential if mortality is to be measured at sub-national level.

On the other hand, data on deaths reported by households in the census are also problematic for different reasons, which include: recall bias, event omission, household collapse and inaccurate age reporting (Joubert, Rao, Bradshaw et al., 2012). Recall bias leads to deaths being reported for a longer or shorter period before the census than intended which results in under- (over-) reporting of deaths for a fixed period. Death omission may result from a household collapsing after the death of an individual which leads to under-reporting of deaths. Conversely, deaths of, for example, young adults who recently moved out of the household to look for work, may be reported by more than one household. Elderly households are prone to under-report deaths in the household as there may be no one to report the death. All of this could mean that completeness of deaths may differ by age.

Due to these errors it is more reasonable to assume that the reporting of registered deaths is more likely to be constant for all age groups above a certain minimum age than the deaths reported by households are (Dorrington, Moultrie & Timæus, 2004), which is one of the major assumptions when applying death distribution methods. Dorrington and Timæus (Dorrington, Moultrie & Timæus, 2004, Dorrington & Timæus, 2015) have estimated provincial mortality in South Africa for 2001 and 2011 using the strengths in both sources of data to produce estimates of mortality at the provincial level. In a similar way, this study aims to assess the strength of the method at a district level.

1.2 Objectives of the research

This study mainly aims to assess the suitability of the adult mortality estimates obtained for each sex for the 52 municipal and metropolitan districts of South Africa in the year prior to the 2001 census by using the method of combining vital registration data and census data adopted from Dorrington, Moultrie and Timæus (2004).

However, in the process of producing the adult mortality estimates the following objectives must be satisfied.

1. Estimate national completeness of registered deaths for both sexes and also by population group for each sex in the intercensal period of 1996 to 2001.
2. Estimate completeness of registered deaths for each sex and population group for the year prior to the 2001 census.
3. Estimate age-specific household correction factors for each sex and population group in the year prior to the 2001 census.
4. Estimate the true (corrected) deaths reported by households in each district by sex from the 2001 census.
5. Estimate district adult mortality rates by sex from the corrected reported deaths by households in the 2001 census.
6. Evaluate the suitability of the method in obtaining mortality estimates for 2001 at district level.

1.3 Significance of the research

The research enables us to examine the extension of the method proposed for provincial estimates by Dorrington, Moultrie and Timæus (2004) to the district level. Based on the outcome of the study, the method will either fail to produce convincing adult mortality estimates or produce reasonable estimates of adult mortality at the district level.

If the method fails to produce reasonable estimates of adult mortality at district level, it is essential that the weaknesses in applying the method are discovered from this study for the benefit of future studies using a similar approach. A cross examination of the methodology and the data will guide future studies in addressing the methodology changes and (or) fundamental data needs that may ensure more reasonable estimates of adult mortality at the district level.

On the other hand, if the method is found suitable its will enable us to make use of the method elsewhere. Estimates of adult mortality at the district level will help inform health policies. The estimates may be used as a guide for improving health delivery at sub-national level in an effort to reduce the burden of adult mortality, specifically targeting diseases with a high prevalence in adults that lead to early death such as HIV/AIDS.

1.4 Structure of the research

The dissertation consists of five chapters. Chapter two reviews the sources of mortality data, death distribution methods, the level and trend of mortality in South Africa and reviews studies that obtained mortality estimates using a combination of vital registration and census data. In Chapter 3 the method used specifically in this dissertation is presented. Chapter 4 presents the results obtained from implementing the method. The last chapter reflects on the results and discusses the extent to which the district adult mortality estimates are accurate.

2 LITERATURE REVIEW

This chapter generally seeks to understand the limitations and strengths of tools designed to produce estimates of adult mortality at the sub-national level. It also aims to understand the efforts made by different studies in the area of estimating mortality at the sub-national level. The chapter begins by reviewing the potential sources of mortality data. In the same section we consider the extent to which the different sources of data are useful to estimating mortality at the sub-national level.

Thereafter the chapter reviews several methods that have been proposed for estimating adult mortality. This starts with a review of the death distribution methods in depth by looking at its foundation and establishing the requirements essential to obtaining reliable estimates. This is followed by a review of the orphanhood method. In the last part of this section studies that have produced estimates of provincial adult mortality are reviewed.

The final section of the chapter covers estimates of the level and trends of adult mortality to understand the general trajectory of mortality up until 2001. In particular it covers the level of mortality for different population groups and provinces from various studies and the level of HIV prevalence at the sub-national level.

2.1 Sources of mortality data

In developing countries, the sources of data necessary for estimating mortality rates are usually obtained from vital registration, censuses, surveys and surveillance sites. The validity of any of the estimates produced will depend on the data quality. However, in most developing countries these different sources have problems which limit their usefulness in producing accurate mortality estimates.

Household surveys such as the Demographic Health Survey (DHS) have been used extensively in estimating mortality in developing countries, particularly of child mortality. Questions relating to sibling and parental survival are used in the indirect estimation of adult mortality. Survey data relating to child mortality has, generally, produced better estimates compared to adult mortality (Timæus, Dorrington & Hill, 2013). This is because unlike child deaths, adult deaths are not concentrated in a narrow age group, thus bias in results is greater in estimating adult mortality. Apart from that, data relating to adult deaths are not specific to a unique respondent; questions relating to orphanhood and sibling survival may be answered by multiple individuals referring to a single death, unlike the mother in the case of a child death. Adult deaths, thus, may be over-reported or under-reported. Although data quality can be monitored, survey death data are not appropriate for estimating mortality sub-nationally because the sampling procedure in identifying households for the survey is taken to be nationally representative, which may not be representative of the sub-national units. Generally in South Africa, the use of surveys in adult mortality estimation has been limited. During the late 1990s the October Household Survey (OHS) was used in the estimation of mortality as there were a few sources to rely on at the time (Udjo, 2016).

Small area systems such as the Health and Demographic Surveillance System (HDSS) sites have been useful over the past decade in monitoring demographic events. They ensure platforms for further research into cause-specific mortality, for example, maternal mortality and HIV/AIDS; however, they represent a small geographical area, which is usually not representative of the national population (World Health Organisation, 2006). Nonetheless they are useful in building an understanding of how different socio-economic and health outcomes will lead to different levels of mortality. Furthermore HDSS can be used for investigating causal factors of mortality, which may be difficult to identify using other sources of death data.

Deaths recorded in hospitals can also be a useful source of data, particularly when assessing the causes of death. In some countries, like Botswana where it is said that more than 90 per cent of the deaths occur in health facilities (Negussie, Will, Bodika et al., 2009), this source of death data can be a substitute for registered deaths and may also provide information about causes of death. However, for most developing countries, a significant number of deaths occur outside health facilities. Most of these deaths occur particularly in the rural areas resulting in an uneven death coverage of the population. In addition to the low coverage of deaths in hospitals, the deaths may be selective according to the causes of death, which results in the clustering of certain causes of death in the hospitals more than other causes. Recorded deaths from hospitals are likely to be selective of the location, in other words, deaths will likely end up in the nearest hospital which may even be outside the district of residence. In addition, data on hospital deaths are not easy to access as the databases for the different hospitals are not openly accessible and often not updated regularly (Khoza, 2009).

Registered deaths are the best source of mortality data for comprehensive and continuous monitoring of public health programmes over time (Mahapatra, Shibuya, Lopez et al., 2007, Mathers & Boerma, 2010), yet, only 30 per cent of the world live in areas where registration is more than 90 per cent complete (Joubert, Rao, Bradshaw et al., 2012). In South Africa, it is a legal requirement for all people to register a death once it has occurred. If the death occurred in a hospital, a family member or the funeral undertaker can report it to the nearest local Home Affairs office, otherwise, if it occurred elsewhere, either the family member or funeral undertaker registers the death with the local municipality as well as the municipal Home Affairs office for acquisition of burial space (Khoza, 2015). The population register is also updated at the local Department of Home Affairs (DHA), taking account of the deaths that have a South African identity number from the South African-born population. Thereafter, the death notification forms are sent to Stats SA for the production of vital statistics (Khoza, 2015).

The framework proposed by Mahapatra, Shibuya and Lopez et al. (2007) for assessing the quality of general vital registration data is to assess the accuracy, relevance, comparability, timeliness and accessibility of the data. The timeliness of vital registration data in South Africa has improved considerably since 2004. One of the reasons for the reduction in time to release mortality data is faster processing of the death notification forms by Stats SA (Khoza, 2015). The delay in releasing deaths data has reduced to a lag of 11 months from a lag of seven years in 2003. The accuracy of the vital registration deaths is also affected by coverage issues, which hinder registration as the DHA may not be easily accessible in some areas. Urban areas are more likely to have higher coverage than rural areas. Another problem that affects the quality of death registration data is variables with missing responses, especially on key variables such as gender and age, which will affect the usefulness of the data. For example, a small proportion of death notification forms lose either the first or second page when the death notification forms are transferred to Stats SA (Khoza, 2015).

Death notification forms now include the usual place of residence of the deceased as well as the place where the death occurred (Statistics South Africa, 2014), whereas previously registered deaths recorded only the place of occurrence, which probably concentrated deaths at health facilities (Dorrington, Moultrie & Timæus, 2004). In many developing countries, the time it takes for the data to be published and the ease of access are also a concern. However, in South Africa, Stats SA releases reports periodically on their findings from captured mortality data from the vital registration deaths.

Despite the errors in vital registration data, the major concern, that of incompleteness of registration, can be corrected for adult ages using death distribution methods. Although deaths are liable to late registration, the date on which the death occurred is accurately recorded, implying the correspondence of deaths to the population in estimating mortality is relatively accurate for those deaths that have been reported.

Unlike registered deaths, data on deaths reported by households are not collected continuously as they depend on the census having taken place, usually every ten years. Census data for South Africa currently incorporates questions involving the survival of parents as well as deaths reported by households in the last year. South Africa has had three post-apartheid censuses, namely in 1996, 2001 and 2011, although the 1996 census did not include the question on deaths reported by households in the last year. Deaths reported by households (HHD) rely on the respondents accurately recalling the death of household members in the year before the census.

One of the strengths of HHD over registered deaths is that they are reported according to the usual place of residence, which is useful in obtaining sub-national estimates of mortality directly. However, one of the concerns with HHDs is reference period error, where deaths which occurred outside the reference period are reported as occurring in the period and vice-versa, thus leading to over- (under-) reporting of deaths, which affects the quality of estimates (Mathers & Boerma, 2010).

Not every death is of a person living in a household, that is to say HHDs may not constitute full coverage in the population, for example, deaths occurring in retirement homes are not included under HHDs, which is apparent in the White population (Dorrington, 2013c). However, in the 2011 census, Stats SA report that converted hostels were included in the final reported deaths (Statistics South Africa, 2012a, p.68), although their definition is not clear on the difference between institutions and converted hostels (Statistics South Africa, 2012b). In addition, some deaths in the household may be missed, for example, households dissolving when a key member has died, which leads to under-reporting of deaths. More so for small and young families at the time when HIV was at the peak before the introduction ARVs (anti-retrovirals): they would then have a greater risk of both infected parents dying, leaving the household isolated or with young orphans who may be taken into orphanages or fostered.

Another weakness of HHDs can arise from multiple reporting of deaths. Young adult deaths may be reported by more than one household due to the young adults having dual residences, which may be caused by migrating elsewhere in search for employment. Also the age at death of a household member may be misreported, particularly at older ages where it may be exaggerated (Dorrington, 2013b).

Nonetheless, census data is easily accessible. Most of the census data are openly available from the Stats SA interactive data websites; Nesstar and SuperWeb, and variables not available may be provided upon request to Stats SA.

2.2 Methods of estimating adult mortality

2.2.1 Death Distribution methods

Death distribution methods are used to estimate the completeness of death reporting in the population, which essentially allows one to correct for the incompleteness of death reporting in producing estimates of mortality. The principle behind death distribution methods were developed by Brass (1975) and Preston, Coale and Trussell et al (1980). Originally, these methods compared deaths with population counts at a single point in time on the assumption that the populations are stable. However, they were later generalised to make use of two censuses and thus to be of more practical use for populations which are not stable.

There are two approaches: the General Growth Balance (GGB) method and the Synthetic Extinct Generations (SEG) method. Both of these, however, require certain assumptions about the data to be satisfied. These assumptions are: that there is a constant under- (over-) count in all age groups in each of the censuses; that the population is closed to migration; that completeness of reporting of deaths is constant (at least for adult ages); and that the ages of the living and the deceased are reported to be reasonably accurate (Hill & Queiroz, 2010). Although both methods use the same data, in some instances they produce different results due to the nature of data errors and violation of assumptions. One could also consider the method of initially applying the GGB method to estimate the parameter δ to be used in the SEG method on the same input data to estimate completeness, which is referred to as GGB-SEG throughout this dissertation.

2.2.1.1 General Growth Balance Method

The GGB method was generalised from the Brass Growth Balance (BGB) by Hill (1987). The BGB was proposed by Brass on the basis that it is possible to estimate the completeness of reporting of deaths if population counts are available from a single census provided the population is stable. The GGB method is derived from the population balance equation assuming the population is closed to migration:

$$P_2 = P_1 + B_{1,2} - D_{1,2}$$

where P_i is the total population at time of census i, also $B_{1,2}$ and $D_{1,2}$ are the total births and deaths respectively in the intercensal period. This equation can be generalised to the open age interval (aged a and older) to give:

$$P_2(a+) = P_1(a+) + B_{1,2}(a+) - D_{1,2}(a+)$$

Dividing the entire equation by the total person-years of exposure, which is approximated by the geometric average population in the open age interval ($a+$) at the start and end of the intercensal period multiplied by the length of the interval, $t[P_1(a+)P_2(a+)]^{0.5}$, and dropping the time of the census for convenience:

$$r(a+) = b(a+) - d(a+) \qquad (2.1)$$

where: $r(a+)$ is the true growth rate of the population; $b(a+)$ is the true rate at which the population turns a and older; $d(a+)$ is the true death rate of the population in the open age interval.

The BGB uses a constant growth rate since it assumes the population is stable. The GGB, on the other hand, tries to account for the violation of that assumption in reality by making use of two census counts to estimate the age-group specific growth rates in the open age interval. However, since the censuses are prone to under- (over-) count and deaths are liable to under- (over-) reporting relative to the second census, the observed census counts can be expressed as $P_1^o(a+)$ and $P_2^o(a+)$, and the reported deaths as $D(a+)$. The observed and true values are related as follows:

$$P_i(a+) = \frac{1}{k_i} P_i^o(a+)$$

$$D(a+) = \frac{1}{c} D^o(a+)$$

where: k_i is the proportion counted from census i; c is the proportion of deaths registered in the intercensal period reported.

In estimating the total population turning a in the intercensal period, $B(a+)$, where $N_k(i, j)$ is the population aged between i and j in census k, under the assumption that there is constant under- (over-) count in all age groups, substitution results in

$$B(a+) = \frac{t}{5}\left[N_1(a-5,a)N_2(a,a+5)\right]^{0.5} = \frac{t}{5}\left\{\left[\frac{1}{k_1}N_1^o(a-5,a)\right]\left[\frac{1}{k_2}N_2^o(a,a+5)\right]\right\}^{0.5} = \frac{1}{\left[k_1k_2\right]^{0.5}}B^o(a+)$$

Also substituting for total person-years of exposure in the intercensal period gives:

$$t\left[P_1(a+)P_2(a+)\right]^{0.5} = \frac{1}{\left[k_1k_2\right]^{0.5}}\left\{t\left[P_1^o(a+)P_2^o(a+)\right]^{0.5}\right\}$$

Substituting the above equations into each component of the population balance equation gives:

$$r(a+) = \frac{1}{t}\ln\left[\frac{P_2(a+)}{P_1(a+)}\right] = \frac{1}{t}\ln\left[\frac{k_1}{k_2}\right] + r^o(a+) \tag{2.2}$$

$$b(a+) = \frac{B(a+)}{t\left[P_1(a+)P_2(a+)\right]^{0.5}} = \frac{\left[k_1k_2\right]^{-0.5}B^o(a+)}{\left[k_1k_2\right]^{-0.5}\left\{t\left[P_1^o(a+)P_2^o(a+)\right]^{0.5}\right\}} = b^o(a+)$$

$$d(a+) = \frac{D(a+)}{t\left[P_1(a+)P_2(a+)\right]^{0.5}} = \frac{\left[k_1k_2\right]^{0.5}}{c}\frac{D^o(a+)}{t\left[P_1^o(a+)P_2^o(a+)\right]^{0.5}} = \frac{\left[k_1k_2\right]^{0.5}}{c}d^o(a+)$$

The growth rate is estimated differently in equation 2.1 and equation 2.2 given equation 2.1 assumes a linear growth, whereas equation 2.2 assumes an exponential growth, although the difference is insignificant for small values of the growth rate for the open age interval (Hill, 1987). Substituting the equations above into the balancing equation gives:

$$r^o(a+) + \frac{1}{t}\ln\left[\frac{k_1}{k_2}\right] = b^o(a+) - \frac{\left[k_1k_2\right]^{0.5}}{c}d^o(a+)$$

$$b^o(a+) - r^o(a+) = \frac{1}{t}\ln\left[\frac{k_1}{k_2}\right] + \frac{\left[k_1k_2\right]^{0.5}}{c}d^o(a+)$$

A straight line can be fitted to the data points with $b^o(a+) - r^o(a+)$ on the y-axis and the observed death rate in the open age interval on the x-axis, and an intercept of $\ln(k_1/k_2)/t$ and a slope of $(k_1k_2)^{0.5}/c$. Once the estimates of the slope and intercept are obtained, values of the census under- (over-) count and under- (over-) reporting of deaths can be obtained. However, since there are two equations and three unknowns, the equation can only be solved with two unknowns. This is done by expressing the under- (over-) count of one population relative to the other, that is one of the k's is usually assumed to be the lower of the two with the other taking a value of unity, which is appropriate since there is usually no ideal census with the true count. Although there are many methods that can be used in fitting the straight line, such as the method of least squares, orthogonal regression, grouped mean and weighted means, the method of least squares and un-weighted methods are not recommended for fitting the line as they give too much weight to extreme data points (Bhat, 2002, United Nations Population Division, 2002, Marandu, 2011). Orthogonal regression to fit the straight line is recommended (Bhat, 2002, Dorrington, 2013b) since it produces a better fit in cases where age exaggeration exists, which is typical of data in developing countries, more so at the older ages.

2.2.1.2 Synthetic Extinct Generations method

The SEG method proposed by Bennett and Horiuchi (1981, 1984) is a generalisation of the method proposed by Preston et al. (1980) which was itself an extension of the method of extinct generations proposed by Vincent (1951). The method of extinct generations uses the idea that the population at any age and at any point in time can be estimated from the future deaths from that cohort until the last individual dies, which can be expressed as follows:

$$N_a(t) = \int_0^{\omega} D_{a+x}(t+x)dx \qquad (2.3)$$

where; $N_a(t)$ is the true population aged a last birthday at time t; $D_{a+x}(t+x)$ are the true deaths in the year $t+x$ aged $a+x$ last birthday; ω is the age of the last death in the cohort. However, applying this method would require one to wait for the cohort to die out. Additionally this method requires deaths data to be collected on the same cohort over time, which may be problematic due to migration and lack of follow-up. This would be simpler if the population is stationary, as the number of deaths at each age remains constant over time, although this is not so in reality.

Preston, Coale, Trussell and Weinstein (1980) generalised the method for populations that are close to being stable so that one uses the current numbers of deaths by age to estimate the future deaths in the population since the deaths at each age change by a constant rate each year over time. That is,

$$D_{a+x}(t+x) = D_a(t)\exp[rx]$$

where r is the annual growth rate of the population.

However, most developing countries have populations that deviate from stability. Thus Bennett and Horiuchi (1981, 1984) further generalised the method proposed by Preston et al. (1980) to non-stable populations by using age-specific growth rates estimated over an intercensal period to estimate future deaths:

$$D_{a+x}(t+x) = D_a(t)\exp\left[\int_a^x r_i di\right]$$

where r_i is the growth rate for age i.

The method of synthetic extinct generations as proposed by Bennett and Horiuchi uses the assumption that the current growth rates in the intercensal period apply in the future. However, one is not necessarily restricted to this assumption as it can be shown that the current population can be expressed in terms of the current deaths and the current growth rates in the intercensal period (Dorrington, 2013c, Preston, Heuveline & Guillot, 2001, p171-172):

$$N_a(t) = \int_a^{\omega} D_x(t)\exp\left[\int_a^x r_i di\right]dx$$

Bennett and Horiuchi (1981, p221) propose the following approximation to the above equation to apply to five-year age groups:

$$N_a(t) = N_{a+5}\exp[5_5 r_a] + {}_5D_a \exp[2.5_5 r_a]$$

To take into account the under- (over-) reporting of deaths, let the observed deaths between age a and $a+5$ be represented by ${}_5D_a^o$ and the completeness of reported deaths be c. This means that the product of true deaths and c gives the reported deaths, and this adjustment can be allowed for in the above equation which can be shown to be:

$$\hat{N}_a(t) = \hat{N}_{a+5}\exp[5_5 r_a] + {}_5D_a^o \exp[2.5_5 r_a] \qquad (2.4)$$

However, since the equation is applied iteratively from the open age interval to lower ages, Bennett and Horiuchi (1981) also proposed a way to estimate the population at the open age interval, A. Apart from using the deaths in the open age interval one needs the life expectancy in the open age interval (e_A) to account for the variation in the age at death above A, more so when the open interval is set to be relatively low:

$$\hat{N}_A = {}_{\omega}D_A^o\left\{\exp[r_A e_A] - \frac{(r_A e_A)^2}{6}\right\}$$

The estimate of life expectancy at the age of the open interval can be obtained using several methods. One could estimate it based on independent estimates from other research on the same population. An alternative would be applying the GGB method to the same data to derive a life table and hence estimate life expectancy at age A. One could also use the ratio of reported deaths, ${}_{30}D^{o}_{10} / {}_{20}D^{o}_{40}$, and compare it to an equivalent level in the Coale and Demeny model life tables, but this may not give suitable estimates in populations affected significantly with HIV/AIDS as the Coale and Demeny model life tables do not reflect AIDS mortality (Dorrington, 2013c). Alternatively, one may also start with an initial guess of the life expectancy and obtain the results from the SEG, and thereafter use the results to construct a life table, then iterate to a solution (Dorrington, 2013c). The significance of the error in estimating life expectancy decreases the higher the age of the interval (Bennett & Horiuchi, 1981).

Once the number of people in the population in the age groups have been estimated from the deaths, the annual average population in the intercensal period in each of the five-year age groups (${}_{5}\hat{N}_{a}$) can then be estimated:

$${}_{5}\hat{N}_{a} = \frac{2.5}{t}\left[\hat{N}_{a} + \hat{N}_{a+5}\right]$$

The annual average population estimated from the numbers counted in the two censuses in the intercensal period is:

$${}_{5}N_{a} = \left[{}_{5}N^{1}_{a}\,{}_{5}N^{2}_{a}\right]^{0.5}$$

where ${}_{5}N^{i}_{a}$ represents the population in census i aged between a and $a+5$. Therefore, completeness of reporting deaths can be estimated by the ratio of the cumulated population obtained from the observed deaths and the age distribution from both censuses assuming the level of death reporting is constant over the cumulated age groups:

$$c = \frac{{}_{A-a}\hat{N}_{a}}{{}_{A-a}N_{a}}$$

The completeness of death reporting is not determined in the open age interval as this interval is more sensitive to the error in estimating the life expectancy, e_{A}. One may also consider changing the open interval A to lower ages if there is evidence of significant age exaggeration at the older ages (greater than 60), which can be noticed by comparing visually the data points a against the starting age a of each age interval to a trend line. A gradually consistent decrease in completeness at the older ages would indicate age exaggeration.

Since the censuses may be under- (over-) counted, the growth rates used in obtaining equation 2.4 should be adjusted to take this into account. This adjustment was previously derived in the review of the GGB as $\ln(k_1/k_2)/t$, is referred to as δ by Bennett and Horiuchi (Dorrington, 2013c), and can be estimated iteratively by adding different values of δ to the observed growth rates until the plot of a against the starting age a of each age interval reflects the most level set of ratios (Dorrington, 2013c). When this adjustment is applied, we refer to the method as the SEG+δ, otherwise the SEG.

2.2.1.3 Problems of using Death Distribution Methods with registered deaths

For the errors that the methods were developed to overcome, the methods are reported to perform particularly well by researchers who have looked at DDMs response to various errors (Hill, 1987, Hill, You & Choi, 2009, Dorrington, Timæus & Moultrie, 2008). These errors include omission of deaths, changes in census coverage and omission of deaths occurring together with a change in census coverage (Hill, You & Choi, 2009).

Research (Dorrington, Timæus & Moultrie, 2008, Hill, You & Choi, 2009) suggests that if migration in the population is significant and has not been accounted for, the error in the estimate of completeness is higher than most of the typical errors one may encounter. However, it is suggested (Dorrington, 2013b, Murray, Rajaratnam, Marcus et al., 2010) that if the points used to fit the straight line to determine completeness for the GGB indicate migration effects[1] it may be useful to confine the fit to higher ages where the fit seems to be consistent with the straight line to help eliminate potential migration effects, although there are concerns that it may increase the possibility of error in estimating the completeness due to age exaggeration at older ages (Hill, You & Choi, 2009). Hill and colleagues suggest that when emigration is unaccounted for and there is an error in estimating increasing census coverage, the SEG method performs particularly better than the rest of the methods.

The percentage errors in the estimate of ${}_{45}q_{15}$ was higher than that for ${}_{25}q_{60}$ in the GGB, SEG and GGB-SEG methods when completeness was either increasing or decreasing with age (Hill, You & Choi, 2009), this may also support the indication of using a narrower age interval in estimating completeness.

International migration is quite difficult to account for accurately in most countries. Immigrants may not be willing to answer questions on the census because of the fear of their illegal status in the country and also the fear of xenophobia, resulting in

[1] Migration effects in the fitted straight line can be noticed by deviation from the straight line at ages usually lower than 35 since much of the migration in a population is usually concentrated around these ages unless the context of the population suggests otherwise.

the net migration likely being under-represented (Dorrington & Hill, 2013). Emigration, on the other hand, is even more difficult to account for as it requires the combination of information from receiving countries in their census which may not correspond with the census date being used (Dorrington & Hill, 2013).

Stats SA's assumptions for the population projections used to produce mid-year population estimates (Statistics South Africa, 2015) were that generally, the African and Indian[2] population groups experience immigration, with the White population experiencing significant emigration (Dorrington, Moultrie & Timæus, 2004), and the Coloured experiencing negligible net migration. Khoza (2015), using the DDMs, assumes a zero net migration between 1996 to 2001 and 2001 to 2006; however, between 2001 to 2006 the results seem to suggest unaccounted for migration, although Khoza is open to the possibility that the effect may be due to the use of the 2007 Community Survey rather than an actual census.

At the national level net-migration is less volatile and less significant compared to the sub-national level. For this reason DDM's are increasingly less reliable due to migration effects at the sub-national level. The method of combining the vital registration and HHD proposed by Dorrington and Timæus was proposed as a way around having to measure completeness at the sub-national level.

Dorrington, Moultrie and Timæus (2004) estimated the completeness of death registration in each of the population groups in the intercensal period 1996-2001 using the GGB. Migration in other population groups except for the White population is identified to be insignificant. However, emigration is estimated using the censuses from the main recipient countries of the South African-born White population. In spite of the effort to estimate net migration the results for the White population appeared implausible. The need to allow for migration in applying DDMs restricts their use as there is need to accurately estimate net migration.

Age exaggeration at old ages is common, which results in bias in the estimates of overall completeness. When using the GGB method to estimate completeness, age exaggeration can be noticed from the diagnostic plots by points at the older ages deviating from the fitted straight line; likewise by an inconsistent rise and drop in the completeness at the older ages in the SEG and SEG+δ methods (Dorrington, 2013b).

[2] The immigration in the Indian/Asian population group is mostly significant among men.

Generally, the GGB method performs better where there is age exaggeration than does the SEG and SEG+δ methods (Dorrington, 2013b, Hill, You & Choi, 2009).

Murray, Rajaratnam and Marcus (2010) suggest there is evidence of significant variation in the age ranges that produce the best results when comparing the DDMs. In their research, the ideal age range for the GGB, SEG and GGB-SEG methods is 40 to 70, 55 to 80 and 50 to 70 respectively. In addition, the SEG reflected a consistent pattern of error reduction with older age ranges as opposed to the other methods. Although not specified it may be that SEG+δ follows the same behaviour[3]. Furthermore, when patterns of age misreporting differ widely between the two censuses and registered deaths, the error increases in the GGB, SEG and GGB-SEG methods (Murray, Rajaratnam, Marcus et al., 2010).

It is worth mentioning that the data used in both studies (Hill, You & Choi, 2009, Murray, Rajaratnam, Marcus et al., 2010) did not take into account the impact of HIV/AIDS unlike Dorrington, Timæus and Moultrie (2008), thus some of the results may possibly not be generalisable to the South African context. In addition Murray, Rajaratnam and Marcus et al. (2010) used a different age pattern for migration in their scenarios, and some of the statistical assumptions underlying their modelling are not clear enough to allow replication of the full scenarios.

Both Hill, You and Choi (2009) and Dorrington, Moultrie and Timæus (2008) seem to agree that the SEG performs better in errors combining emigration and increasing census coverage. However, the authors' conclusions seem to differ as Hill, You and Choi (2009) conclude that GGB-SEG is the best method with an age range of 5+ to 65+ and Dorrington et al. (2008) concludes SEG+δ performs better in most of the scenarios on average than the GGB and GGB-SEG. However, results in Dorrington, Moultrie and Timæus (2004) are strengthened as the methods are applied to typical African population data in the context of South Africa with the HIV/AIDS epidemic.

2.2.2 The Orphanhood method

The orphanhood method uses questions on the survival status of parents to estimate adult mortality. One of the advantages of this method is that the questions are simple to answer and eliminate recall bias by not requiring recall of either the date or the age when the death occurred.

[3] Since the two methods follow the same general approach in terms of the general methodology, it may be that the SEG and SEG+δ both respond better to older age ranges.

The method uses the age of the respondent and the mean age of childbearing of the mother[4] (conception for the father) in estimating the exposure to the risk of dying of the parents through the probability of the parents' survival from that age to the time of the survey. Conditional survivorship of the parent is estimated as the probability of the mother of the respondent having had survived from the mean age of child bearing to the age of the respondent (l_{m+n}/l_m, where *m* is the mean age of childbearing and *n* is the age of the respondent). To obtain the estimate of this conditional probability a regression equation making use of the proportion of respondents in a particular age group with living mothers and coefficients of survivorship for women (men) is used (Timæus, 1992). Thus, the older the respondent is, for example, the longer the parent has survived.

Since the conditional survivorship refers to different probabilities of survival as well as different mortality experiences, it becomes necessary to standardise the conditional survivorship into one measure that can be comparable through time, for example, ${}_{35}p_{25}$, the probability of a person aged 25 dying before reaching age 60. The conditional survivorship in each age group is standardised using the two parameter relational logit model, which connects the estimated mortality experience to that of any standard life table (Murray, Ferguson, Lopez et al., 2003, p.166-168). In application (Timæus, 2013), the age pattern of mortality is determined by a life table that reflects the same age pattern of mortality as the population on which we seek to estimate mortality[5]. The time location of the estimate from the time of the census depends on the level of mortality and can be estimated from the proportion of surviving parents, the age of the respondents and the mean age at childbearing of the parents (Timæus, 2013).

However, this method usually produces more reasonable estimates using responses from female respondents rather than male respondents, as women are usually closer to their parents than men. In addition, the method produces some bias in mortality estimates as those parents without surviving children are not represented whereas those with more than one surviving child may be over-represented; however, the bias in the final estimates is small, except in populations affected by HIV/AIDS, especially before ARVs were accessible, due to infected adults dying before their parents

[4] The mean age at childbearing is calculated from the census information by dividing the age-weighted sum of births by the total number of births. The mean age of conception for the father is calculated by adding the difference between the median age of marriage for men and that of women.

[5] Moultrie Tom A & Ian M Timæus 2013. *Introduction to model life tables. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population: Available: http://demographicestimation.iussp.org/content/introduction-model-life-tables.

since they will no longer be able to report the survival status of their parents (Timæus, 2013). The adoption effect also presents a problem at younger ages where children may report their parents as surviving when in fact they are deceased but due to the nature of their adoption they may consider their guardian as the parent. However, Timæus (2013) tries to reduce this bias by using respondents aged 20 onwards, since the bias is likely to be greater when the respondent is still young. Although it is possible to apply the orphanhood method sub-nationally, one should do so with caution in the knowledge that parents do not always live in the same area as their children and thus the mortality experience may not be consistent with the location.

2.3 Estimating mortality sub-nationally

Sub-national mortality estimates are quite problematic to produce accurately if the data is prone to errors and bias. However, in South Africa there is growing research in the area as the accuracy of data sources continues to improve. Usually when one thinks of estimating completeness of reporting deaths, death distribution methods are used, although as pointed out earlier, these methods cannot be applied at a sub-national level. The reason being compared to international migration, inter-provincial migration is more dominant and difficult to measure or estimate accurately, which is one of the requirements when applying death distribution methods (Dorrington, 2013b).

Dorrington and colleagues (Dorrington, Timæus, Moultrie et al., 2004) in producing provincial estimates of mortality from the vital registration deaths for 1996 linked completeness of registered deaths for children to that of adults, assuming a linear relationship between the two. The relationship was established on the assumption that the mortality rate of adults between 65 and 70 (${}_5m_{65}$) did not differ significantly between the provinces at the time, thus any difference between the provinces can be attributed to the completeness of death registration. Although there is no justification for the linear relationship between completeness in adults and that of children, a weak correlation for both sexes between child completeness and mortality for older ages was obtained (Dorrington, Timæus, Moultrie et al., 2004).

The results for completeness indicated high levels of death registration in Gauteng, Northern Cape, Western Cape, North West and Free State (118%, 118%, 117%, 107% and 105%), and low levels of registration in Eastern Cape, Limpopo and Mpumalanga (57%, 60% and 85%). It is suggested that the completeness of reporting of deaths in some provinces appeared implausible, namely, the high levels in completeness in Free State and North West could possibly be due to the poor coverage in the census rather than registration of deaths from other provinces. However, the post-enumeration survey results reflect the highest undercount in Northern Cape, Kwa-Zulu-Natal and Limpopo (15.59%, 12.81% and 11.28%), with the lowest in Western Cape, Free State and North West (8.69%, 8.75% and 9.37%) (Statistics South Africa, 1998). This conflict in results was not interrogated further.

One of the limitations of this study is that the questions relating to deaths in the household were not asked in the 1996 census. So for subsequent censuses Dorrington and colleagues (Dorrington, Moultrie & Timæus, 2004, Dorrington & Timæus, 2015) used a different approach, which we seek to adopt in this study, in their estimation of provincial mortality estimates in 2001 and 2011. Unlike the prior study (Dorrington, Timæus, Moultrie et al., 2004), deaths reported by households are scaled using the results of completeness from registered deaths.

Generally, the studies (Dorrington, Moultrie & Timæus, 2004, Dorrington & Timæus, 2015, Pillay-Van Wyk, Laubscher, Msemburi et al., 2014) apply the DDMs initially to the registered deaths in the intercensal periods in estimating overall national completeness of death registration for each sex and each population group. Completeness was higher for the White population, followed by the Indian, Coloured and least in the African population, with females having higher values of completeness than males, possibly due to the fact that women are not as mobile as men and therefore less difficult to track than men (Dorrington, Moultrie & Timæus, 2004). Interestingly, the difference between the completeness of Indian men and women was quite large.

After this the completeness of death registration in the year preceding the census is estimated by fitting a trend to estimates of completeness obtained from using previous intercensal periods, since DDMs estimate the average completeness in the intercensal period and not the year before the census. The trend in completeness is fitted using a regression curve, which follows a known pattern of how completeness has been changing over time, for example, the logistic curve (Dorrington & Timæus, 2015, Pillay-Van Wyk, Laubscher, Msemburi et al., 2014). Alternatively, one could fit a trend to annual estimates of an indicator of mortality, for example, using m_{65+} from which the trend in completeness could be implied (Dorrington, Moultrie & Timæus, 2004, Dorrington, Timæus, Moultrie et al., 2004). In obtaining the trend in completeness from 2001 to 2010, a logistic trend was fitted to the completeness in the period 2001 to 2007 after applying DDMs to the intercensal period (Pillay-Van Wyk, Laubscher, Msemburi et al., 2014). The trend was assumed to reach an asymptote of 93 per cent by 2009 as no data was yet available at the time for 2011 registered deaths. However, the trend was verified for consistency by applying the DDMs to the intercensal period 2001-2011 using the estimate of registered deaths from the national population register for 2011 where registered deaths were not available (Dorrington, Bradshaw, Laubscher et al., 2015).

Once completeness is estimated in the year preceding the census, it is applied to estimate the true registered deaths for each sex, population group and the entire population (Dorrington & Timæus, 2015). However, the sum of the deaths corrected for completeness by population group are adjusted to match the national deaths corrected for completeness. This is done to ensure consistency in the true national deaths including those that were not reported since realistically it is expected that the national deaths are sub-grouped into the four population groups.

The deaths reported by households from the census are then divided by the true reported deaths to estimate adjustment factors for each sex, population group and age-groups. These adjustment factors indicate the under- (over-) reporting of deaths by households. The deaths reported by households in each province according to sex, population group and age-group are multiplied by these adjustment factors. However, the adjustment factors assume that there is no significant difference in the under-(over-) reporting of deaths by households in each of the provinces. The true deaths reported by households in each province are then divided by the estimate of the population to produce crude estimate of the mortality rates.

After estimating ${}_{50}q_{15}$ for each of the provinces using the deaths reported by households, the results were compared to corresponding results from the orphanhood. The results showed a strong correlation, indicating consistency between the two different estimates (Dorrington, Moultrie & Timæus, 2004).

2.4 Levels and trends of adult mortality in South Africa

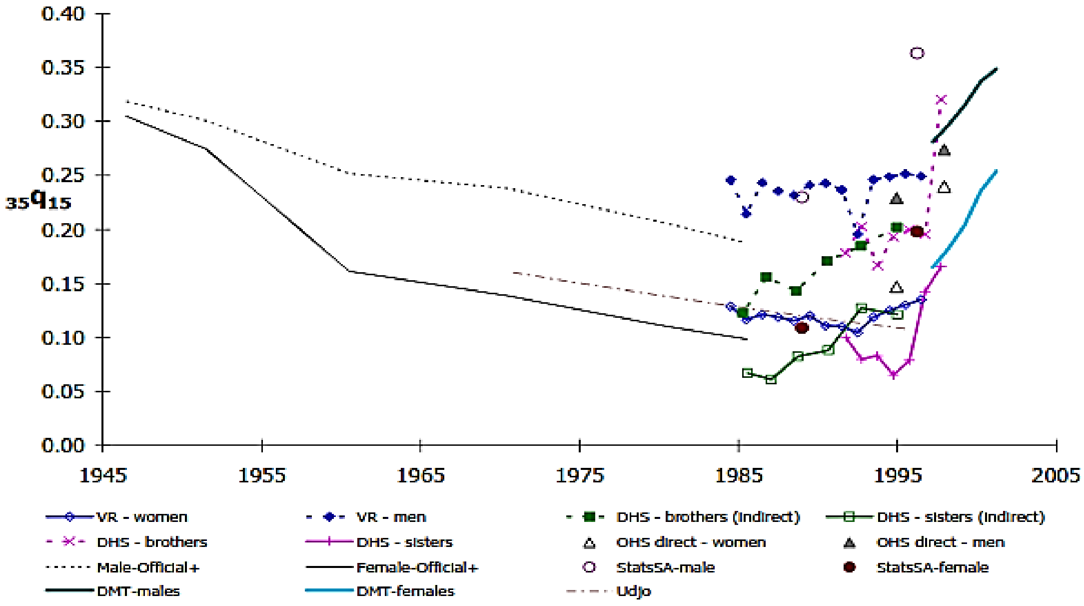
2.4.1 The national level and trend of adult mortality

A lot of uncertainty lies in the mortality estimates for the period 1983-93. During this period the usefulness of censuses or surveys from which mortality estimates could be obtained was limited. Another reason for the uncertainty in estimates is the fact that death registration in South Africa did not include the so-called TBVC (Transkei, Bophutswana, Venda and Ciskei) areas. The TBVC areas under the apartheid government were designated as self-governing countries and specifically for the residence of the African population groups. This among other things meant registration of deaths was done independently in each of these areas, although these areas were unable to carry out death registration. Because about 25 per cent of the total African population in South Africa resided in these areas it may be possible that the level of mortality estimated in the period excluding these areas could have under-estimated the true level of mortality. Also the official South African life tables for 1984-86 were not representative of the black African population (Dorrington, Bradshaw & Wegner, 1999). However, Dorrington, Bradshaw and Wegner (1999) estimated the mortality of the African population in order to produce a nationally representative life table. This was done by averaging the mortality estimates from two scenarios, namely, estimating the mortality of the African population on the assumption that all TBVC deaths were reported in South Africa and on the assumption that the mortality in TBVC areas reflected an equivalent mortality rate as that of Africans in South Africa[6].

From the late 1980s to the mid-1990s, adult mortality appears to have been constant to slightly decreasing, although the estimates from different authors vary on the extent of the decrease (Udjo, 2001, Dorrington, Moultrie & Timæus, 2004), as evidenced from most of the estimates in Figure 2.1 except for the indirect estimates obtained from the Demographic and Health Survey (DHS) in 1998. In 1985, the probability of a person aged 15 dying before age 60 ($_{45}q_{15}$) nationally for males and females was 38.4 and 25.4 per cent respectively (Bradshaw, Dorrington & Sitas, 1992), we expect the level of adult mortality 15 years later to be higher due to better completeness of recording deaths and also the contribution of the HIV epidemic that followed during the period.

[6] The first scenario estimates mortality using the total population of South Africa (including TBVC areas) and the deaths reported in South Africa, whereas the second scenario used the South African population excluding the TBVC.

Figure 2.1: The probability of a person aged 15 dying before reaching the age of 50 for South Africa up-to 2001.

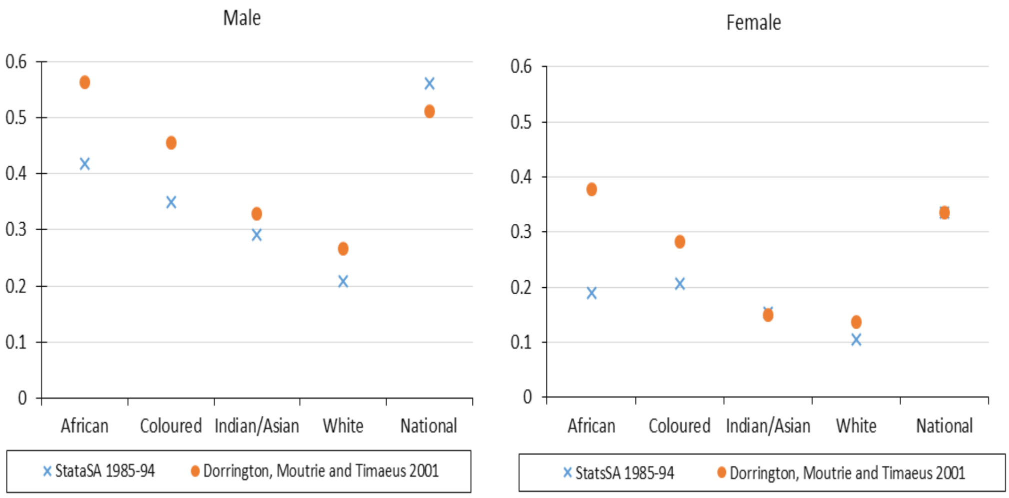


Source: Figure 6.1; page 65; Dorrington, Moultrie and Timæus (2004)
Notes: VR stands for vital registration deaths, DHS for Demographic Health Survey, OHS for October Household Survey and DMT for Dorrington, Moultrie and Timæus.

Udjo and colleagues (Udjo, 2016, 2001) also argue that female mortality in the period 1980-1990 generally declined for all population groups, with Africans having the highest mortality and the White population having the lowest mortality. Stats SA (2000) estimates from the 1985-94 abridged life table suggest a consistent ranking in the adult mortality for males with Africans having the highest mortality, followed by the Coloured, Indian/Asian and then the White population in that order; however, somewhat strangely, a different ranking is estimated for the female population. A slightly higher estimate was obtained for the Coloured population than the African population[7].

[7] The probability of a person aged 15 dying before age 60 for African females in 1985-1994 is estimated to be 19.1 deaths per 100 individuals, whereas that for the female coloured population is estimated to be 20.5 deaths per 100 individuals.

Figure 2.2: The probability of a 15 year old dying before reaching age 60 for different population groups in South Africa in 1985-94 and 2001.



Source: Table 1A-7B; page 14-17; Statistics South Africa (2000), Table 7.2; page 81; Dorrington, Moultrie and Timæus (2004)

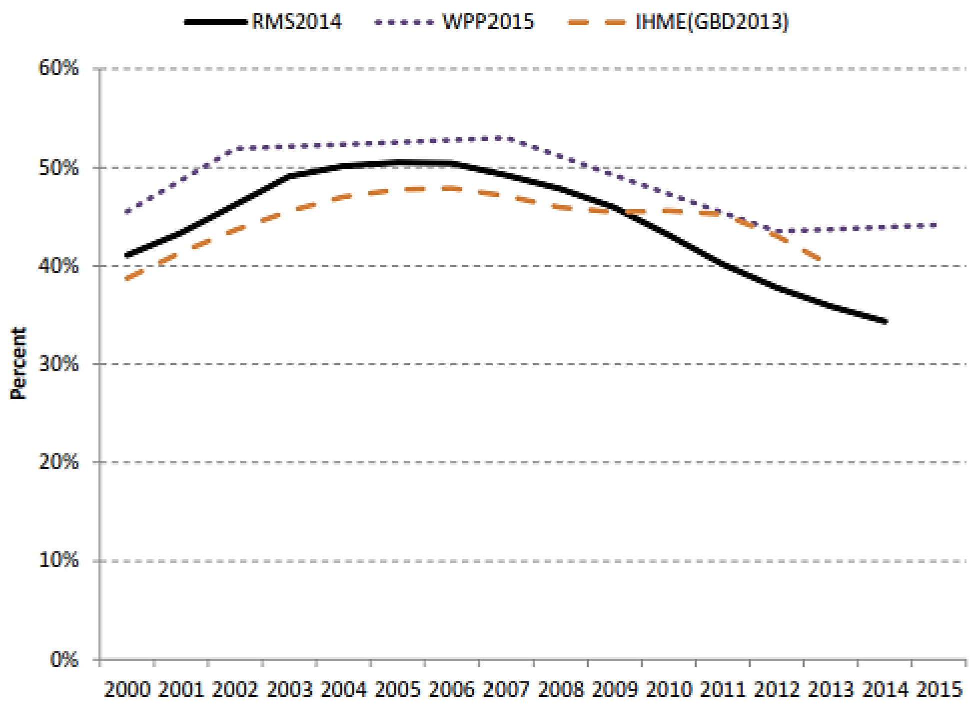
Timæus, Dorrington and Bradshaw et al. (2001) argue that the rapid increase in the mortality of women in the mid-1990s reflects the onset of the HIV/AIDS pandemic, which for the same period reflects an almost gentle increase for the men because of the pandemic and violence related incidences towards the end of apartheid. However, Timæus and Jasseh (2004) using the DHS, indirectly estimated mortality between the early 1990s and the mid-1990s to have increased slightly more for men than women. They estimate ${}_{45}q_{15}$ to have increased from 28.0 and 13.5 per cent in 1990 to 30.2 and 14.7 per cent in 1995 for males and females respectively, which indicates a slightly[8] greater increase for women rather than men. This supports the argument that the HIV/AIDS pandemic started around the early to mid-1990's, which is consistent with literature that suggests HIV pandemic in South Africa started late compared to other sub-Saharan countries in the north that had been affected by the pandemic (Timæus & Jasseh, 2004). Figure 2.2 shows how the mortality of the African and Coloured population increased more abruptly, unlike the Indian/Asian and White population.

For 1996, Stats SA (2000) estimates the probability of a person aged 15 dying before reaching age 60 was 56 per cent and 33.5 per cent respectively for men and

[8] The proportional increase in 45q15 for men is 7.86%, whereas it is 8.89% for women, implying a difference of around 1% between men and women at the time. Since the mortality of women is generally lower than that of men at adult ages, this increase in women's mortality could be a signal of the HIV/AIDS pandemic affecting the mortality of women earlier than that of men between 1990 and 1995 as mentioned in Timaeus, Dorrington and Bradshaw et al. (2001).

women. Dorrington, Moultrie and Timæus (2004) in their monograph argue that this was an over-estimate of mortality for the time on the basis of comparisons with other estimates. Between 1996 and 2001 mortality can be seen to have increased rapidly, mainly due to the HIV/AIDS pandemic (Udjo, 2006, Bah, 2016).

Figure 2.3: The combined probability of a person aged 15 dying before reaching age 60 for both men and women in South Africa up-to 2015.



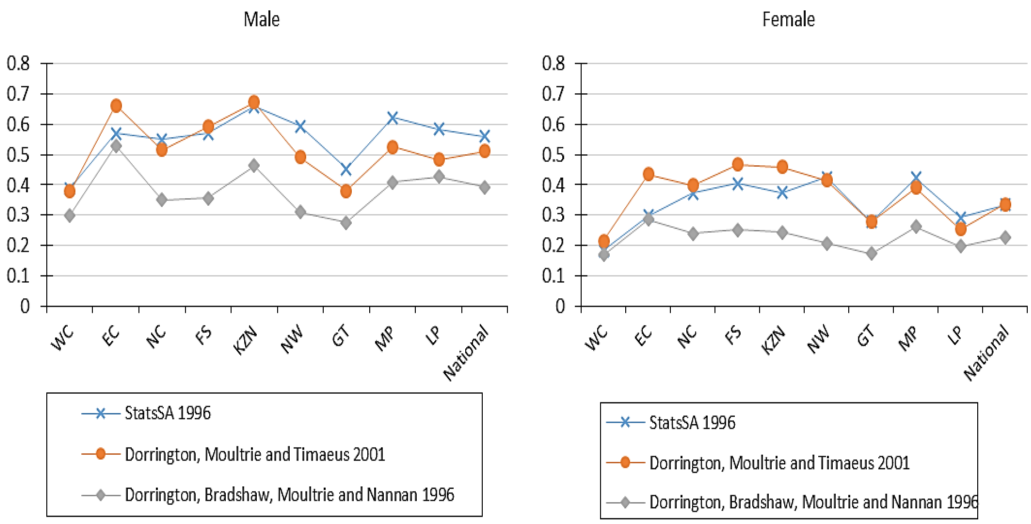
Source: Figure 6.1; page 65; Dorrington, Moultrie and Timæus (2004)
Notes: RMS2014 stands for Rapid Mortality Surveillance estimates in 2014, WPP2015 for World Population Prospects estimates in 2015 and IHME(GBD) for Institute of Health Metrics (Global Burden of Diseases) estimates in 2013.

Figure 2.3 shows the different estimates for $_{45}q_{15}$ from the Rapid Mortality Surveillance (RMS) report, the World Population Prospects (WPP) and the Institute of Health Metrics and Evaluation (IHME). All estimates reflect an overall consistent trend reaching a peak between 2006 and 2008, with a decrease in mortality after that. However, the estimates seem to reflect different levels from one another after 2007/2008. Dorrington, Bradshaw and Laubscher et al. (2015) suggest that the IHME estimates seem implausible as the decline in the mortality rates seems to stagnate from 2008-2011 at the time when anti–retroviral therapy (ART) was increasingly being provided. The WPP rates, however, suggest mortality to have been constant after 2011, unlike the other estimates.

2.4.2 The level of adult mortality in provinces of South Africa

The over-estimation of mortality from Stats SA in 1996 is apparent in Figure 2.4 and Figure 2.6 in all of the provinces when compared to estimates from Dorrington et al. (Dorrington, Timæus, Moultrie et al., 2004) for the same year. The estimates of male mortality seem to suggest a reasonably consistent increase in mortality from 1996 to 2001, with Kwa-Zulu-Natal, Eastern Cape and Limpopo reflecting the highest mortality and Western Cape, Gauteng and North West reflecting lower mortality. However, female mortality seems to reflect an inconsistent trend among the provinces from 1996 to 2001, especially in Kwa-Zulu-Natal and Free State, which seem to have the highest mortality in 2001, unlike in 1996.

Figure 2.4: The level of $_{45}q_{15}$ in the provinces among men and women in the years 1996 and 2001



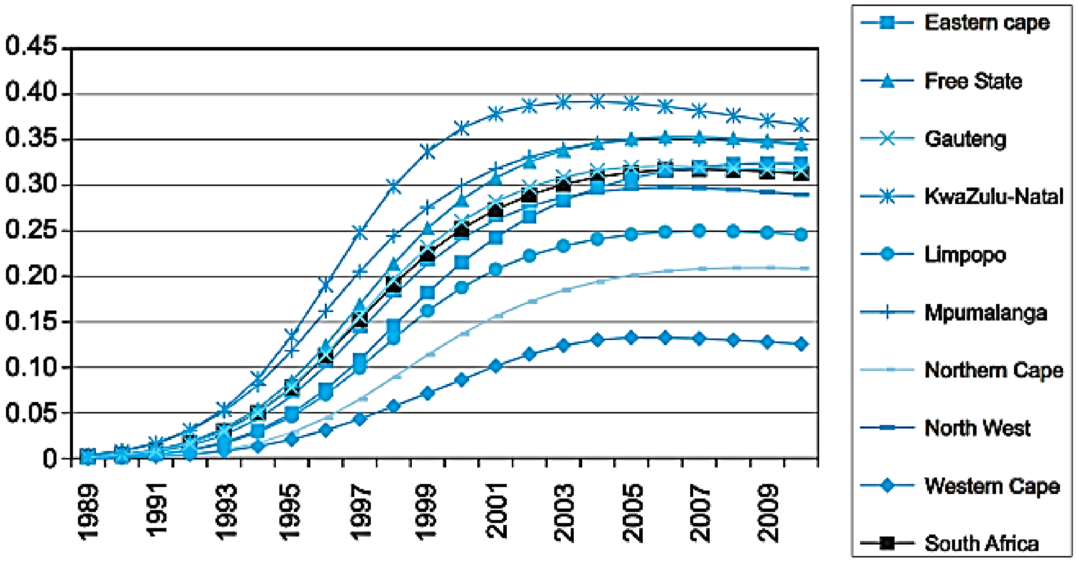
Source: Table 10A-18B; page 19-23; Statistics South Africa (2000), Table 7.3-7.4; page 82-83; Dorrington, Moultrie and Timæus (2004)

2.5 The impact of HIV/AIDS on adult mortality in South Africa

The rise in the level of adult mortality in South Africa is widely attributed by various authors (Hosegood & Timæus, 2005, Johnson, Dorrington & Matthews, 2007, Dorrington, Bourne, Bradshaw et al., 2001, Timæus, Dorrington, Bradshaw et al., 2001) to be as a result of HIV/AIDS. One of the limitations in deciding on the quantitative impact of HIV/AIDS on adult deaths is that mortality data on causes of death is often poor and diseases associated with HIV/AIDS are often the attributed cause of death in a number of cases (Hosegood, Vanneste & Timæus, 2004). Dorrington, Bourne, Bradshaw et al. (2001) estimate about 40 per cent of the deaths occurred between the ages of 15 and 49 in 2000. Although the peak ages where these deaths occur among men and women differ they are within the range of adulthood.

Although the national burden at this time was high, there are differences in the HIV burden at the lower sub-national levels. In Figure 2.5, it is clear that HIV prevalence is higher in Kwa-Zulu-Natal compared to all the other provinces after the onset of the pandemic. Mpumalanga, Free State and Gauteng were also reported by Dorrington, Bradshaw and Budlender (2002) to also have higher prevalence rates than the national.

Figure 2.5: Prevalence of HIV/AIDS among antenatal attenders

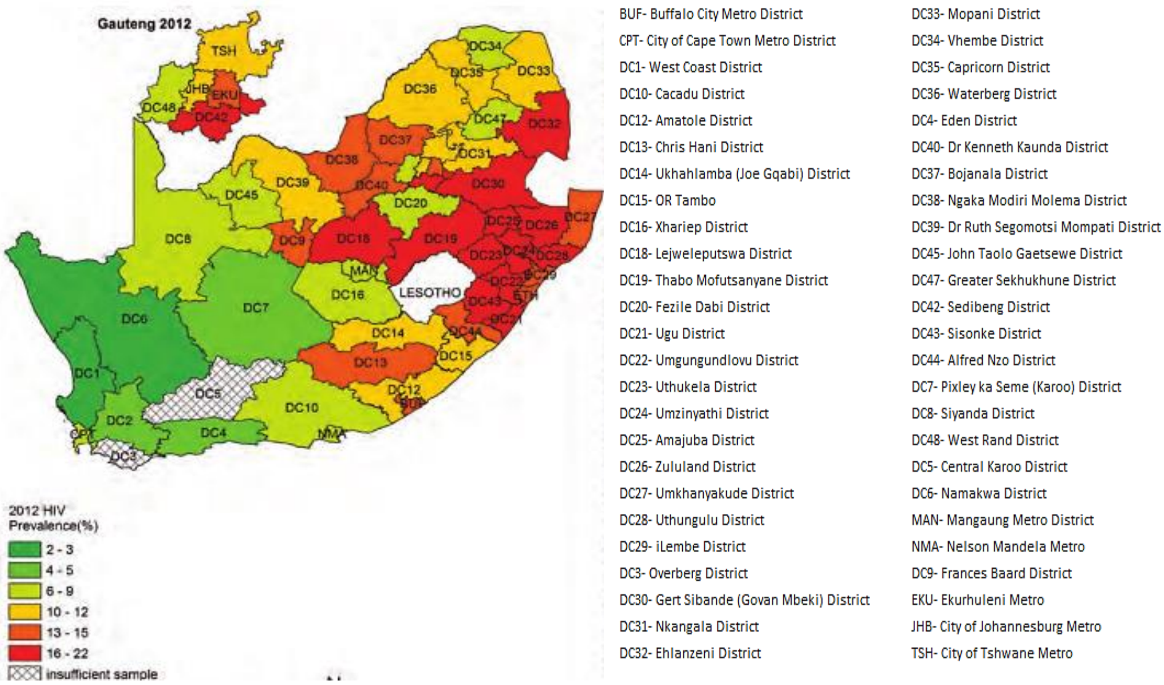


Source: Dorrington, Bradshaw and Budlender (2002, p.3)
Notes: The prevalence rates were projected from the ASSA model using data from women attending antenatal clinics and other demographic assumptions.

Figure 2.6 also shows the prevalence rates in 2012 at district level. Although the period was after the introduction of anti-retroviral treatment, the prevalence level of the districts may still help in understanding the adult mortality estimates 10 years prior to this time, because present prevalence rates are dependent on past prevalence rates, which in turn impacts mortality in the past.

HIV prevalence is generally lower in the western part of the country compared to the eastern part. Generally, we expect prevalence to be higher in the more urban districts of a province, especially the metropolitan districts since these are economic hubs where the risk of infection is higher. For example, prevalence is higher in Cape Town than the surrounding districts in Western Cape. However, there are exceptions to this generalisation. Nelson Mandala Bay Metropolitan in the Eastern Cape and Mangaung Metropolitan in the Free State show a lower prevalence compared to other neighbouring districts which may be due to the population group composition in the metros. For instance HIV prevalence is higher in the African population compared to the other population groups thus it may be expected that metropolitan districts with a dominant majority being the African population may reflect a higher prevalence level. Districts in Limpopo reflect a lower prevalence compared to the other districts in the eastern side of the country.

Figure 2.6: HIV prevalence level among the districts in 2012.



Source: Fig 3.3, Shisana, Rehle, Simbayi et al. (2014)
Notes: The prevalence rates were obtained from women attending antenatal clinics in 2012.

3 DATA AND METHODS

The main method used in this study for obtaining district estimates of adult mortality from the 2001 census is adapted from the methodology proposed by Dorrington, Moultrie and Timæus (2004).

3.1 Overview of the method

The method used to obtain district mortality estimates from the 2001 census is as follows. For each population group and each sex, the completeness of death registration is estimated for the intercensal period 1996 to 2001. Completeness of death registration of each of the population groups is obtained using the two census death distribution methods, the GGB and the SEG+δ methods, considering migration.

For the purpose of this dissertation it is assumed that change in completeness over time in the intercensal period of 1996-2001 was linear. Although this assumption is not tested this assumption is not expected to have a material impact on the results. On the assumption that completeness changed linearly over time such that the mortality rate for those aged 65 and older had a level trend over time, estimates of completeness for each of the years in the intercensal period are derived considering the intercensal estimate of completeness. As a result, the estimates of completeness of death registration in the year prior to the 2001 census are obtained for each population group and each sex.

On the assumption that completeness is constant for each population group, the latter estimate of completeness is then applied to the registered deaths in the year prior to the census to give the true number of deaths that occurred over that period by sex, population group and age. From the true numbers of deaths in the year before the 2001 census, age-specific adjustment factors are calculated for each population group to correct for the under (over) reporting of deaths in the census.

Subsequently, these adjustment factors are applied to correct the numbers of deaths reported by households in each district by sex, population group and age. The numbers of deaths in each district thus reflect estimates of the true numbers of deaths for each district. From the latter numbers, district adult mortality estimates of ${}_{45}q_{15}$ are obtained for each district and sex. The weighted average of the district adult mortality rates that lie in the same province will also be estimated for each of the nine provinces and each sex. This is done for comparison with other estimates since no estimates of adult mortality are available at district level.

To evaluate the suitability of the estimates derived, alternative estimates of adult mortality for each province and population group are obtained using the orphanhood method. The two sets of estimates are compared. However, it must be conceded that there is uncertainty as to the accuracy of the orphanhood for several reasons.

The mortality reflected by the orphanhood estimates may not be entirely representative of the province in which the children live as parents will not always live in the same province as their children. In addition the estimates of mortality rates are derived from a cohort measure whereas estimates from household data are a result of period estimate. Furthermore HIV/AIDS also has an impact on the estimate obtained from the orphanhood method. For instance if the mother is infected with HIV/AIDS there is a chance that the child will also be infected and may die hence the child will not be able to report the survival status of his/her parents. This phenomena is likely to result in understating mortality in the orphanhood method. Also a significant proportion of households do not have both parents present in the household, as a result some children do not know the survival status of their parents which likewise may result in understating mortality.

Estimates of provincial adult mortality obtained by Dorrington, Moultrie and Timæus (2004) are also compared with the estimate produced from the deaths reported by households.

Reasonability of the provincial adult mortality estimates derived from deaths reported by households is measured according to how close they fit to the alternative provincial estimates. Adequecy of the fit will be determined by how well the adult mortality rates in the nine provinces correlate to each other. In addition reasonabilty of the results will also depend on how consistent the district adult mortality estimates are with HIV prevelence in the districts at the time and also between the adult mortality for men and that of women.

3.2 Sources of data and software used

The numbers of registered deaths by sex, population group and age for the period from 1997 to 2001 were provided by Stats SA through the South African Medical Research Council (through personal communication with Professor R E Dorrington). The census data consists of population counts and the number of foreign-born immigrants from censuses for 1996 and 2001 are obtained from the Stats SA website[9]. The data for deaths reported by households from the 2001 census, according to 2011 boundaries, were provided by Stats SA by special request (personal communication, Professor R E Dorrington). The data are categorised by sex, population group and quinquennial age groups (with the 85+ open age interval). The counts for the 1996 and 2001 census by 2011 municipality and district boundaries are taken from the full census.

The main packages used for data processing and analysis are STATA[10] and Excel. STATA is used in the extraction and processing of registered deaths and census data. Excel workbooks on the death distribution methods (GGB and SEG+δ) (Dorrington, 2013b, 2013c) and orphanhood workbook (Timæus, 2013) are obtained from the Tools for Demographic Estimation website[11]. The Excel workbooks for the GGB and SEG+δ are used to estimate completeness of death registration in the intercensal period, whilst the orphanhood workbook is used to obtain alternative estimates of adult mortality at provincial level.

3.3 Method in detail

3.3.1 Data processing

Population numbers from the 1996 census had close to one per cent having no specified population group. These were distributed proportionally by age group to each of the population groups. This is done to improve the reliability of the age-specific growth rates to be obtained over the intercensal period for each of the population groups, especially for the smaller population groups. The numbers from the 1996 census with unknown age were ignored for the analysis as they contained less than 0.5 per cent of the overall population.

The numbers of registered deaths with missing population group were 83 per cent, 60 per cent and 25 per cent for the years 1997, 1998 and for 1999 to 2001 respectively. The numbers with unspecified sex were ignored in the analysis as they constituted less than 0.5 per cent of the total data in each of the years. Likewise, the registered deaths with unspecified age constituted less than one per cent of the total data in each of the years and were therefore ignored.

---

[9] http://interactive2.statssa.gov.za/webapi/jsf/login.xhtml Accessed [20/05/2016]

[10] http://www.stata.com/why-use-stata/

[11] http://demographicestimation.iussp.org/content/generalized-growth-balance-method
http://demographicestimation.iussp.org/content/synethetic-extinct-generations-methods
http://demographicestimation.iussp.org/content/indirect-estimation-orphanhood-multiple-inquiries

Incorporating registered deaths with a high proportion of unspecified population groups in any of the years within the intercensal period is problematic. The reason being, in order to estimate completeness of death registration annually the assumption of a linear trend in completeness is violated. This violation arises when completeness over the years with a high proportion of deaths having unspecified population group results in a significantly lower completeness estimate than any of the other years in the intercensal period. Thus, the numbers of registered deaths in 1997 and 1998 are replaced by estimates derived from back-projecting the average registered deaths between 1999 and 2001.

The numbers of registered deaths by age group in each of the population groups in 1997 and 1998 are obtained by back-projecting the average numbers of registered deaths between 1999 and 2001. This is done to ensure that completeness of registered deaths over the intercensal period is uniform and does not change abruptly between any of the census anniversary years. These numbers are estimated on the assumption that the deaths grew exponentially in the same manner as population growth in the intercensal period. This approach is preferred to that of calculating age-specific growth rates using registered deaths, mainly to limit the variability in the estimated deaths in each of the population groups as the numbers of deaths are relatively smaller than population counts. Also, ideally, population growth rates produce an even trend in the deaths if mortality rates are constant, which is more reasonable. The number of registered deaths occurring in the fraction of 1996 (that is between the date of the census and the end of the year) are estimated by extrapolating the deaths occurring in 1997 for each population group and sex on the assumption that the numbers of deaths are uniformly distributed over the whole period.

The data on deaths reported by households in the census does not specify the population group of the deceased. However, the census has information on the population groups of individuals resident in the household. It is assumed that the majority population group in each of the households corresponds to the population group of the deceased.

Issues connected with data cleaning and editing of the census data are likely to affect the number of deaths by age reported by households of the minor population groups. However, these will have limited impact as the deaths will be corrected for these problems using the sex-population group age-group specific factors. Although results from the Post Enumeration Survey (PES) are used to correct the census for under-coverage in the population counts, the correction for under-coverage may not be a true representation of the smaller sub-groups such as smaller districts and minor population groups because, the PES being a small sample provides only rough corrections for the smaller groups.

3.3.2 Estimating migration from the census

To estimate the completeness of death registration with limited bias, the net numbers of migrants are estimated. The method used in estimating net migration follows the approach from Tools for Demographic Estimation (Dorrington, 2013a). The following steps are used in producing the estimates of net numbers of migrants by sex, population group and age.

The numbers of foreign-born immigrants counted at the 1996 and 2001 census are extracted by sex, population group and age-group from the 10% sample of the 1996 and 2001 census. The Brass Standard life table is used to estimate survival factors to be used for estimating the deaths that occurred over the intercensal period. The Brass Standard life table is preferred over the AIDS Standard life table as we expect the AIDS standard may overstate mortality in the intercensal period between 1996 and 2001. Furthermore, the choice of the standard life table makes little difference to the final estimates of net migration for each of the population groups. The number of deaths aged between x and $x+5$, ${}_5^fD_x$ occurring over the intercensal period at the time of the 1996 census for each sex and population group are estimated as:

$$ {}_5^fD_{x-5} = \frac{1}{2}\left({}_5^fN_x^{2001} - {}_5S_{x-5}\,{}_5^fN_{x-5}^{1996}\right)\left(\frac{1}{{}_5S_{x-5}} - {}_5S_{x-5}\right) $$

where: ${}_5^fN_x^{1996}$ and ${}_5^fN_x^{2001}$ are the number of foreign-born immigrants aged between x and $x+5$ at the time of the 1996 and 2001 census; ${}_5S_x$ is the five-year intercensal survival factors respectively for the migrants aged between x and $x+5$ at the time of the 1996 census.

Similarly, the number of deaths in the open age interval are estimated as:

$$ {}_\infty^fD_x = \frac{1}{2}\left({}_\infty^fN_x^{2001} - {}_\infty S_{x-5}\,{}_\infty^fN_{x-5}^{1996}\right)\left(\frac{1}{{}_\infty S_{x-5}} - 1\right) $$

The net number of surviving foreign-born immigrants, ${}_5^fN_x$, are thus estimated by differencing the number of foreign-born immigrants in the second census from those who survived in the first census maintaining the same age cohort for each sex and population group.

$$ {}_5^fN_x = {}_5^fN_x^{2001} - {}_5^fN_{x-5}^{1996} + {}_5^fD_x $$

For the Coloured population group, net migration is assumed to be negligible. For the African and Indian (Asian) population group, the foreign-born immigrants are assumed to be approximately equal to the net number of migrants.

For the White population group, migration patterns are assumed to be an interaction of the South African-born emigrants and the foreign-born immigrants. The net numbers of White migrants are obtained by the difference between the numbers of foreign-born immigrants and the numbers of the South African-born emigrants. The net numbers of South African-born emigrants over the intercensal period are provided by Professor Dorrington (personal communication), although these do not make a big difference to the estimate of completeness of reporting of the registered deaths.

3.3.3 Estimating completeness

3.3.3.1 Estimating completeness in the intercensal period

The completeness of death registration is estimated for the intercensal period 1996 to 2001 for each population group by sex using the results from applying the GGB and SEG+δ methods. Estimates of completeness are derived using each of the six age ranges: 5 to 84, 15 to 84, 25 to 84, 5 to 75, 15 to 75 and 25 to 75. The final value of completeness is estimated from the median of these results. The choice of using the two methods and different age intervals is mainly to reduce variability in the final estimate among the two methods coupled with the desire to account for the possibility of falling completeness at extreme ages. Completeness of death registration in the intercensal period will consider net migration.

The results of completeness of death registration amongst each of the population groups in the intercensal period are checked to ensure they give a reasonable measure of completeness nationally by sex. This is done by ensuring that the numbers of deaths corrected for completeness by population group aggregate to give the expected numbers of deaths of the whole of South Africa. However, to accomplish this, we need to choose the most relevant expected national numbers of deaths. That is, are the expected national numbers derived from the overall national estimate of completeness or the aggregate of expected numbers of deaths from each of the population groups? The procedure to ensure this is as follows.

Using the result of completeness for each population group, the numbers of registered deaths are corrected for under- (over-) reporting to give the true numbers of deaths. The true numbers of deaths in each of the population groups are aggregated by sex to give the national numbers of deaths by sex and age, which will be referred to as the aggregated total numbers of deaths henceforth.

Ideally, we would expect the aggregate total numbers of deaths to be the same as the corrected national numbers of deaths for each age group; however, we do not expect to observe this as the numbers of registered deaths used before correcting for completeness are different for the national and the population groups. The national numbers of registered deaths consist of all the registered deaths, whereas the aggregated total from population groups did not consider registered deaths missing population group information and were made up of estimated deaths, which were not actually registered for the years 1997 and 1998. Apart from that, death distribution methods assume that as completeness of reporting deaths is constant by age, we expect the national estimate of completeness to differ from the estimate of completeness from aggregated deaths.

The corrected national numbers of deaths are obtained from the aggregate of corrected numbers from population groups rather than from the corrected numbers from the national alone. The choice for this decision is because completeness is not constant among the population groups and the proportions of the total deaths by age differ for each of the population groups.

To ensure the reasonability of the corrected national numbers of deaths, a final check to measure completeness is done to confirm that completeness is close to 100 per cent. To measure the completeness of the numbers of deaths we use GGB and the SEG+δ methods as used previously, however, considering results of the adjustments of the 1996 population counts in each of the population groups since each of the corrected population group numbers take account of the undercount observed in 1996 relative to 2001.

For estimating national completeness using the SEG+ δ, δ is set to zero to ensure that the population counts are not adjusted further, that also in the GGB *a* should measure to be approximately zero. This approach is preferred to using the 1996 national population counts before adjustment to consider the population group differences in under (over) count during between the two periods.

3.3.3.2 Estimating completeness in the year prior to the 2001 census

Once completeness is estimated for the intercensal period 1996 to 2001, for each of the population groups, the number of registered deaths in the intercensal period are sub-divided into annual partitions from one census anniversary to the next. For this purpose it is assumed that deaths in each regular calendar year occur uniformly over that year. It will be further assumed that completeness changed linearly over the intercensal period with the estimate of completeness throughout the intercensal period applying to the year in the middle of the intercensal period, that is, the year from 10 October 1998 to 9 October 1999.

For each population group, the numbers of deaths in the middle of the intercensal period are corrected for completeness. For the rest of the years, the numbers of deaths are corrected for completeness on the constraint that completeness changed linearly across the intercensal period and mortality for those 65 and older maintains a slope of zero across the intercensal period. The latter constraint is made on the assumption that mortality at old ages is relatively stable over short periods of time (Cutler & Meara, 2001). An Excel add-in, solver, is used to apply the constraints while iterating to an optimum solution of annual completeness for each of the annual periods.

From the estimate of completeness of death registration in the year prior to the 2001 census, the true numbers of deaths are estimated for each population group by sex. Estimates of ${}_{45}q_{15}$ will also be produced for each of the population groups, these will be compared against the estimates obtained from the orphanhood data.

3.3.4 Correcting deaths reported by households at district level in the 2001 census

For each population group and each sex, the estimated true numbers of deaths in the year prior to the 2001 census are divided by the corresponding numbers of reported deaths from the census in each of the quinquennial age groups to obtain age-specific household adjustment factors. The age-sex-population-group specific household adjustment factors are used to correct for under (over) reporting of deaths at the district level.

The numbers of deaths reported by households are categorised by sex, district, population group and age group. The age-group specific household adjustment factors are used to correct the district numbers of reported deaths from the census in each of the population groups by sex. On the assumption that the age-specific adjustment factors for each population group and sex apply equally to the 52 districts of South

Africa, the reported numbers of deaths are corrected for under- (over-) reporting. Upon estimating the true numbers of deaths in each population group in all the districts for both sexes, the numbers of deaths are aggregated from the population groups to give the estimate of the total numbers of deaths in each district by sex and age group.

3.3.5 Estimating adult mortality in population groups, districts and provinces

Estimates of ${}_{45}q_{15}$ for each sex are obtained as follows. Firstly, age-specific mortality rates are obtained by dividing the estimated number of deaths by the estimated exposed population at the time. The exposed population is obtained by interpolating from the 1996 and 2001 population numbers in each of the age groups. Thereafter, ${}_{45}q_{15}$ is obtained by calculating the exponent of the sum of the age-specific mortality rates between ages 15 and 60.

For each population group and each sex, estimates of ${}_{45}q_{15}$ are obtained from the estimates of the true numbers of deaths for each of the annual periods from one census anniversary to the next in the intercensal period. The estimates of ${}_{45}q_{15}$ obtained for the year prior to the 2001 census are compared to the orphanhood estimate for each population group on the assumption that the level of ${}_{45}q_{15}$ did not change significantly between the middle of the intercensal period and the year prior to the census. The comparison is used to determine how appropriate each of the population group specific adjustment factors are relative to the results produced from the orphanhood data.

Estimates of ${}_{45}q_{15}$ will be obtained from age-specific mortality rates in each district and each sex. The Brass relational logit model is also used to smooth the age-specific mortality for each district using the corresponding provincial standard derived from the ASSA2008 provincial model. This results in estimates of ${}_{45}q_{15}$ obtained before and after graduation for each district and each sex. Criteria for selecting the final estimate of ${}_{45}q_{15}$ is discussed below.

It is expected that the district estimates of ${}_{45}q_{15}$ within a province should be relatively close to the provincial estimate from the orphanhood method. The choice of the smoothed estimates of ${}_{45}q_{15}$ works in favour of some provinces more than others. To decide on this the sum of absolute difference between the district estimates of ${}_{45}q_{15}$ and the provincial estimate of ${}_{45}q_{15}$ for each province are compared between the graduated and ungraduated estimates of ${}_{45}q_{15}$. For each province, the least sum of absolute difference between the crude or graduated estimate of ${}_{45}q_{15}$ will determine the final district estimates of ${}_{45}q_{15}$ to be used further for that province.

The district estimates of ${}_{45}q_{15}$ within each province are weighted together to an average estimate of ${}_{45}q_{15}$ for the province. The provincial estimate of ${}_{45}q_{15}$ is obtained by summing the numbers of deaths in all the districts within that province by age-group for each sex. The provincial numbers of deaths along with the provincial population numbers are used to estimate ${}_{45}q_{15}$.

3.3.6 Estimating adult mortality using the orphanhood method

The orphanhood method can be used to estimate ${}_{45}q_{15}$ indirectly using questions of parental survival from the census. The purpose of the estimate of ${}_{45}q_{15}$ from the orphanhood data is to help in the evaluation of the estimates produced using the approach described above. Estimates of ${}_{45}q_{15}$ are to be obtained for each of the population groups and each of the provinces by sex. The orphanhood estimates refer to the time between the two censuses.

From the 1996 and 2001 censuses, responses to questions relating to the survival status of the parents are obtained. For 1996, the numbers of respondents who did not know the survival status of their parents are discarded from the analysis, on the assumption that they reflect the same survival distribution of parents as that implied by respondents who stated the survival status of their parents. However, for 2001, the data from the census had been edited by Stats SA, the survival status of parents of the respondents who did not know the survival status of their parents at the time of the census were imputed. This study assumes the differences in the editing of the responses to orphanhood questions between the two censuses will not have a significant impact on the estimates of mortality. Responses from both male and female respondents are used if there is negligible difference in the estimates produced from the survival status of either sex of respondents, otherwise only female responses to orphanhood questions are considered.

The standard life table to be used for the orphanhood workbooks is obtained from projecting the population group and provincial population to 1999 using the ASSA2008 model, thus extracting the full life table for both sexes. The standard life tables are chosen over the default standard life tables in the workbook to account for differences in the mortality distribution among population groups and provinces during the time of the HIV/AIDS pandemic.

The input data for the orphanhood workbooks, apart from the responses to questions about the survival status of parents and the standard life table include: the mean age at child-bearing of the mothers; and the mean age at child-conception of the fathers. For the women, the mean age at child-bearing of the women is obtained from the average of the mean age at their child-bearing from the 1996 and the 2001 censuses for each of the population groups and each of the provinces.

For the men, the mean age at child-conception of the father is obtained by adding the difference between the mean age at which sexual activity starts between men and women to the mean age at child-bearing of the mother. The difference between the mean age at which sexual activity starts between men and women is preferred to using the difference of ages at marriage between men and women as we suspect the latter difference to be biased because children are also born outside marriage, and thus using the differences of ages at marriage excludes children born outside marriages. The difference in the age at which sexual activity starts is derived from the ASSA 2008 full model. This approach is considered to the default setup of the orphanhood workbook of using the difference in the mean age at marriage as there is evidence that the mean age difference of marriage from the census may understate the mean age difference between father's age at conception and mother's age at child-bearing of their children since a significant number of births occur outside marriage.

To estimate the mean age at initiation to sexual activity in both sexes in the ASSA 2008 model, a new statistic is added for the mean age at which sexual activity starts in the output sheet for both sexes, which is linked to data on sexual activity assumptions. Once estimated for the initial year, the model is projected to 1999 to get the estimate desired. The difference is then used in place of the difference between mean age at child-conception of fathers and that at child-bearing of mothers.

For the synthetic cohort calculation of the orphanhood, the survivorship ratios are converted into measures of ${}_{45}q_{15}$ manually using the one parameter system of the relational logit model life table rather than the two-parameter fitted model automatically used in the orphanhood workbook.

4 RESULTS

This chapter reviews the results for all the districts in South Africa in the year 2001. Initially the quality of the input data is assessed and the necessary adjustments are made to the data to ensure it can be used.

After this the results obtained estimating completeness are shown. These results include the estimates of completeness of registered deaths by sex and population group in the intercensal period of 1996-2001 and in the year 2001. Then the estimates adult mortality obtained after correcting the registered deaths for under- (over-) registration are compared to those obtained from the orphanhood method. The results of the sex-population group age-specific factors for adjusting the deaths reported by households in the 2001 census are then shown.

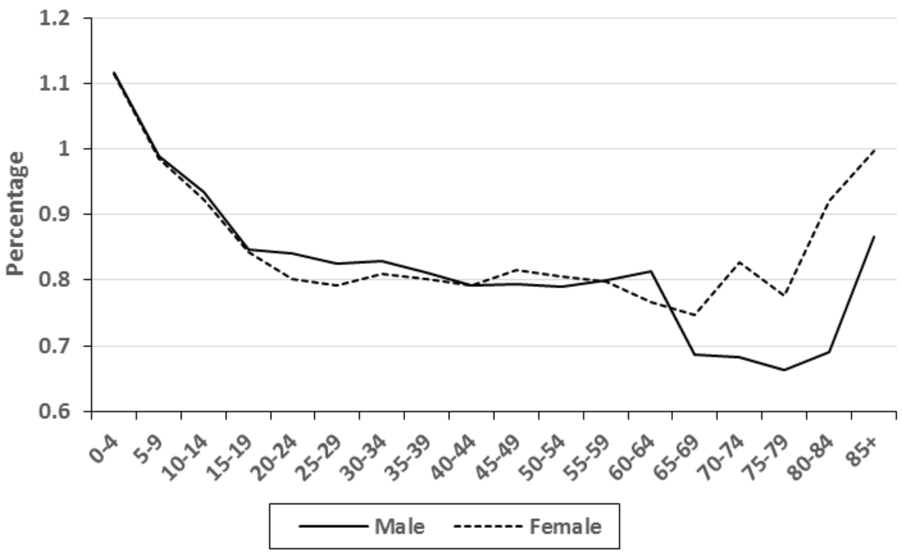
In the last sections of the chapter, the adult mortality estimates at district and provincial level are presented. The district estimates are compared to HIV prevalence from women attending antenatal clinics and ranked, whilst the provincial estimates are compared to the orphanhood adult mortality estimates.

4.1 Quality of data

4.1.1 Population counts

In the 1996 census, almost one per cent of the population did not specify their population group. Although this is a relatively small percentage, it can be seen from Figure 4.1 that the population with unspecified population group were higher at ages 0 to 14 and from age 80 onwards in women and 85 in men.

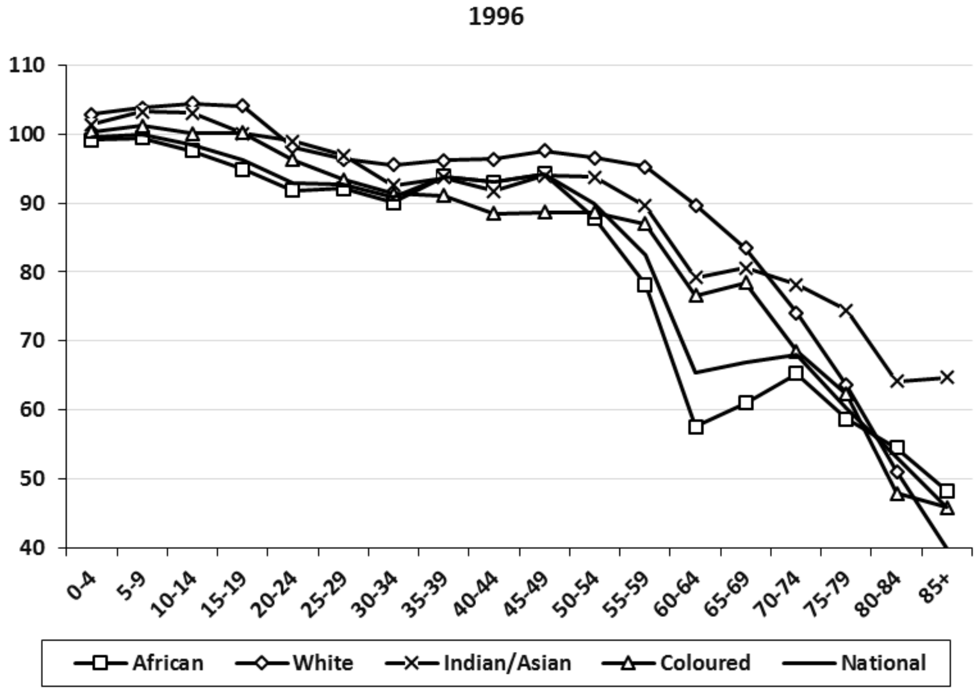
Figure 4.1: Percentage of the 1996 population with unspecified population group

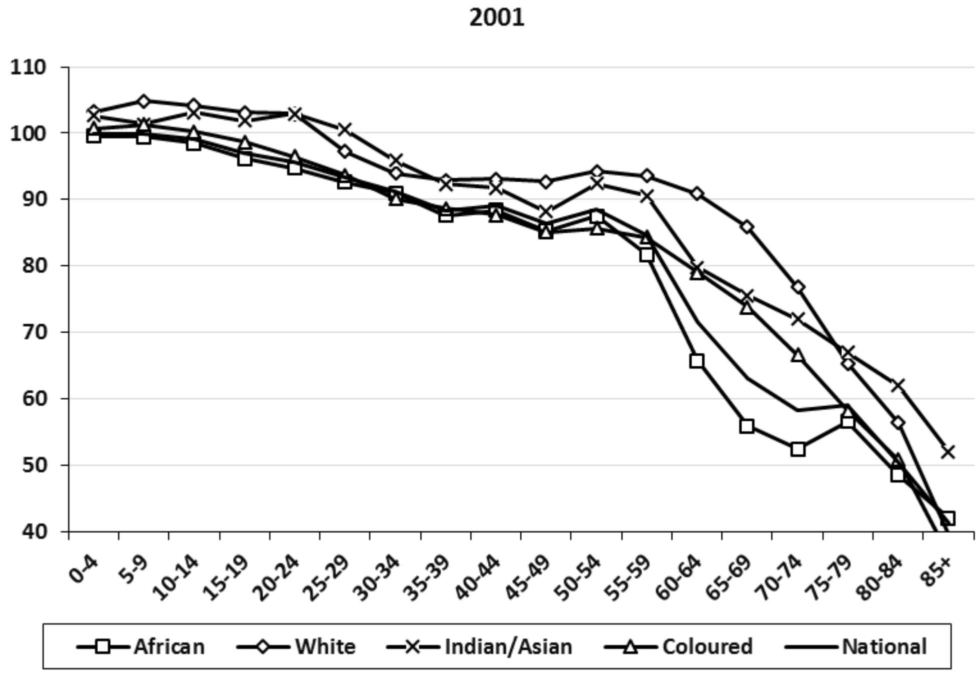


Since the proportions are different across age groups, the numbers with an unspecified population group are proportionately distributed to those with a known population group in each age group and each sex.

From the sex ratios of the population in Figure 4.2, we see that there is a noticeable dip at the middle ages (25 to 49), which increases in depth from 1996 to 2001. The dip initially is probably due to migration in men but deepens perhaps, because of the increased intensity of migration from 1996 to 2001. The steep fall in the sex ratio at the older ages in 1996 is perhaps due to a combination of women exaggerating their age to be above 60 and men to be above 65 to claim old age pensions (Dorrington, Moultrie & Timæus, 2004), hence, the dip becomes more spread out from 1996 to 2001.

Figure 4.2: Sex ratios of population counts by age for the national and all population groups combined from the 1996 and 2001 census





The age ratios[12] for the different population groups are shown in Figure 4.3, Figure 4.4, Figure 4.5 and Figure 4.6. Generally, the age ratios are close to 100 in most

[12] The age ratio for a given age-group is the ratio of twice the population in that age group to the sum of the population in each of the adjacent age groups.

of the age groups for all the population groups in both years. However, there is noticeable age exaggeration in the African population for men (age 65) and women (age 60) in 1996, which can also be traced to the same cohort in 2001. Age displacements are also apparent in the White men at the old ages (75+) in 2001 but this is not noticeable in 1996, which is not consistent with what we would expect. Also, displacement of ages of White women is evident at the young ages (15 to 24) in 2001.

Figure 4.3: Age ratios for the African population in the 1996 and 2001 census

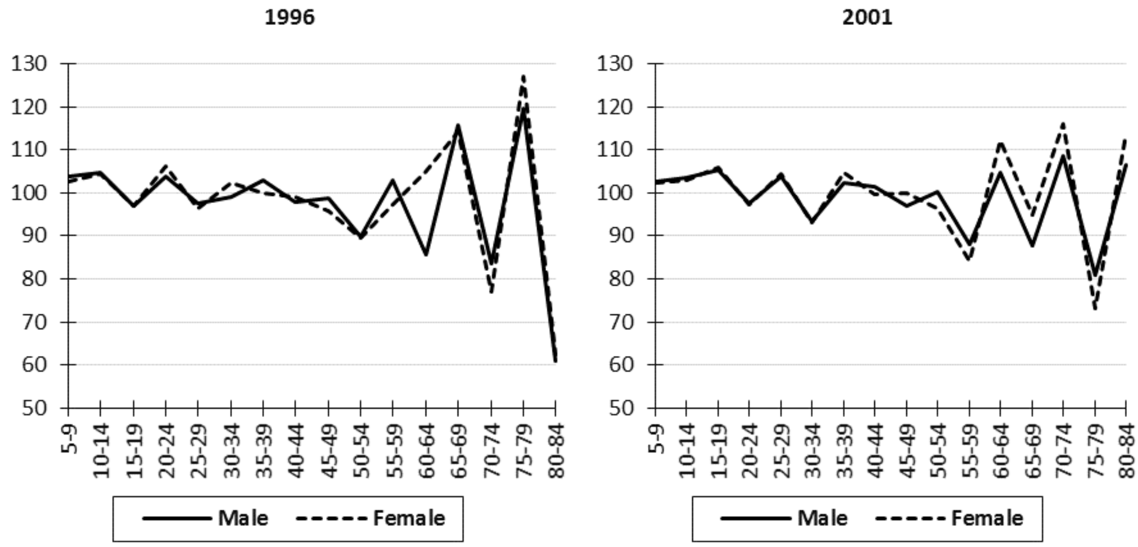


Figure 4.4: Age ratios for the White population in the 1996 and 2001 census

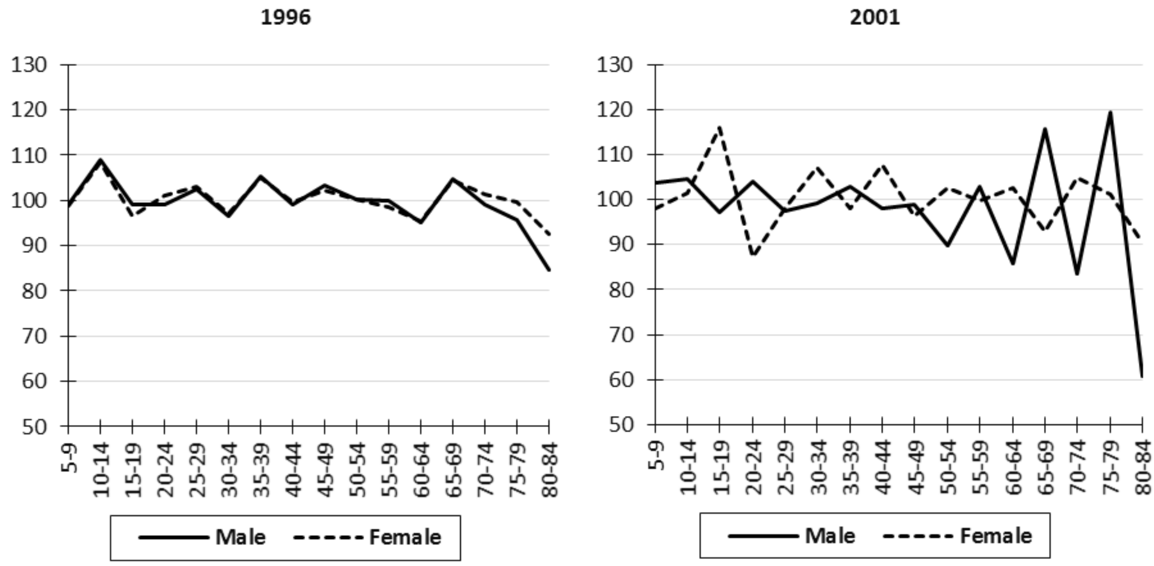
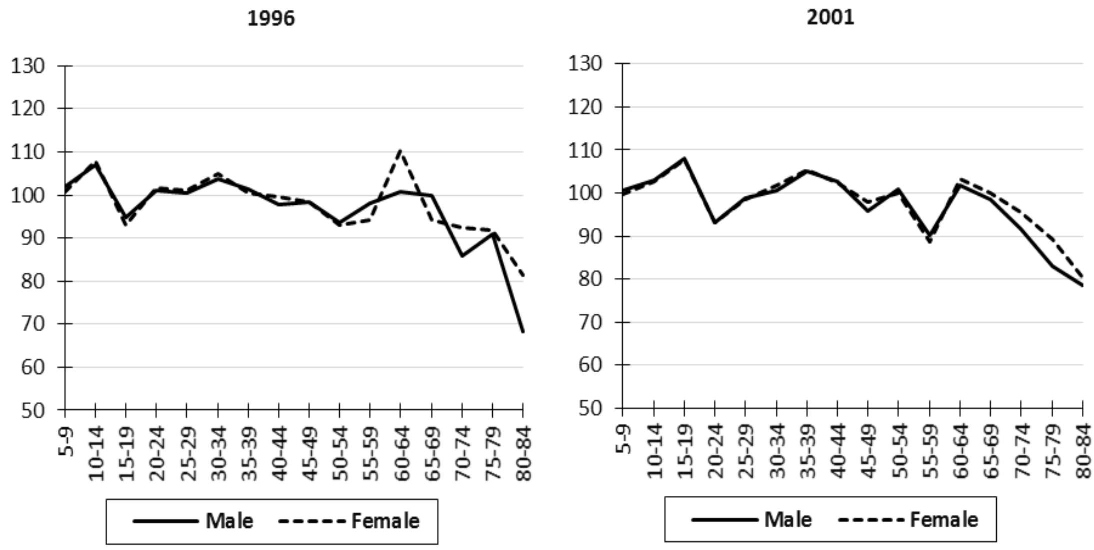


Figure 4.5: Age ratios for the Indian population in the 1996 and 2001 census

Figure 4.6: Age ratios for the Coloured population in the 1996 and 2001 census



4.1.2 Registered deaths

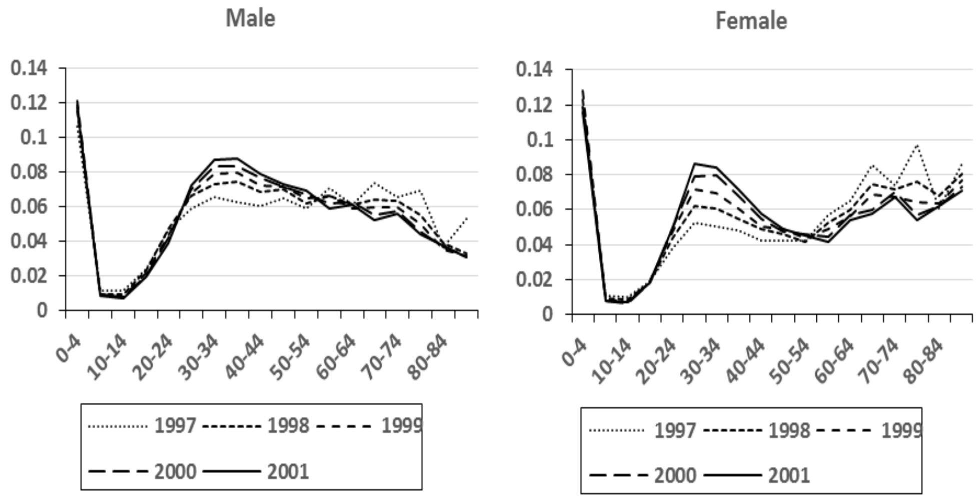
Table 4.1: Percentage of registered deaths with unknown population group in each of the years

Year	1997	1998	1999	2000	2001
Male	82.88%	60.08%	27.53%	22.91%	23.47%
Female	82.82%	60.14%	28.08%	23.59%	24.82%

The data on deaths for the years 1997 and 1998 were of no use since a large proportion had not specified their population group. Table 4.1 shows the percentages of registered deaths with the unspecified population group in each of the years. Although after estimating completeness of deaths registration for each population group, the overall expected number of deaths by age will change marginally by simply ignoring those who did not specify their population group. However, as mentioned in the previous chapter the method relies on the assumption that annual completeness within the intercensal period changed linearly, which will be violated if the registered deaths in 1997 and 1998 are under-registered, compared to the rest of the years in the intercensal period.

One of the essential requirements for the estimated registered deaths in the years 1997 and 1998 is that they have a similar distribution by age compared to the rest of the years in the intercensal period.

Figure 4.7: Proportions of African registered deaths by age for each sex

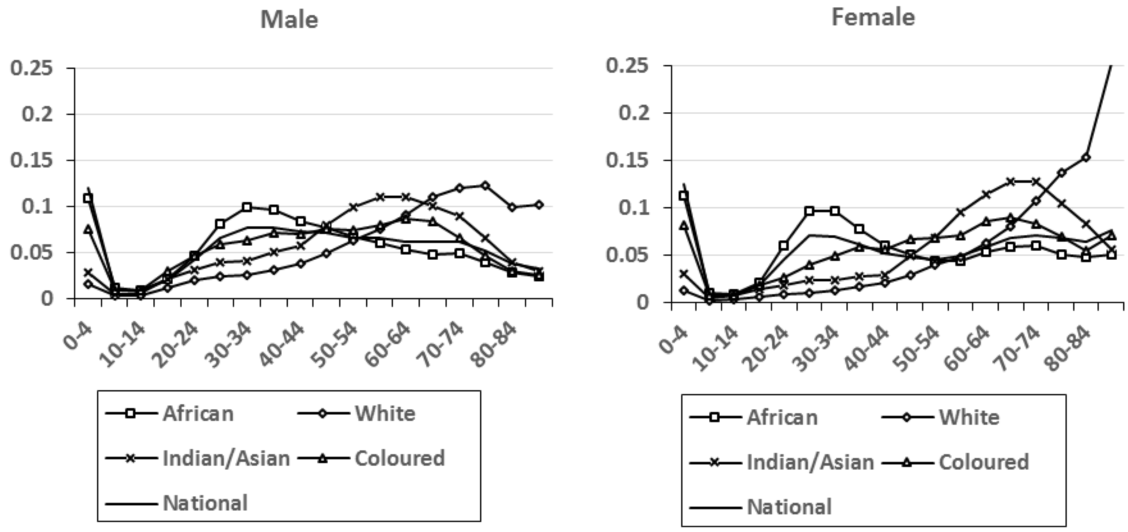


In Figure 4.7 it is evident that the distribution of deaths is fairly consistent across the years. An increase in the proportion of deaths in the adult ages over the years is attributed to the increase in HIV/AIDS related deaths. At the older ages, there is more pronounced age exaggeration, which is noticeable from the irregularity in the smoothness of the distribution, more so for the early years. This can be attributed to the use of population growth in estimating the deaths in 1997 and 1998, as we suspect there is more age exaggeration in the census than the registered deaths.

One of our major assumptions when estimating completeness of death registration in the intercensal period involves completeness of death registration being constant by age. We expect the distribution of registered deaths by age for each population group to be consistent with expectations of how mortality should be distributed by age.

In Figure 4.8, the major proportions of African deaths in the intercensal period occurred in children under the age of 5 and adults between the ages 25 to 44 in males and 20 to 44 in females mainly due to the HIV/AIDS pandemic, which is reasonable as women are more likely to be infected at a younger age than men since women generally are initiated to sexual activity earlier than men. On the other hand, for the White population, the bulk of the deaths occurred at the older ages, with women showing a higher proportion of old age deaths than their male counterparts since women generally live longer than men. For the Indians/Asians the bulk of the deaths are in the older ages (55 to 74). Coloured deaths are initially high for children and grow steadily as the age increases for both males and females. By and large, the distribution of deaths appears to conform to expectations.

Figure 4.8: Proportion of registered deaths by age in the intercensal period 1996-2001 by population group



4.1.3 Deaths reported by households in the census

From the distribution of household deaths (Figure 4.9) reported from the census, Africans have a distribution consistent with that of the registered deaths (Figure 4.8) in the intercensal period. Again, for the White numbers, the trend seems to be consistent with that of the registered deaths although there seems to be lower proportions of males and females at the extreme older ages than that observed in Figure 4.8, probably owing to the reporting of deaths in the census not covering retirement homes.

Figure 4.9: Proportion of household deaths by age reported in the 2001 census

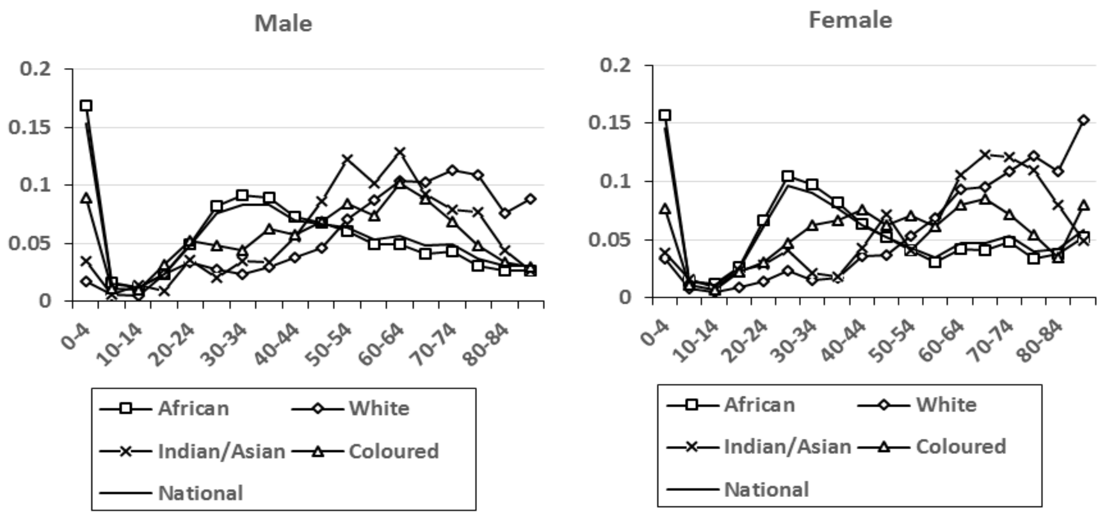
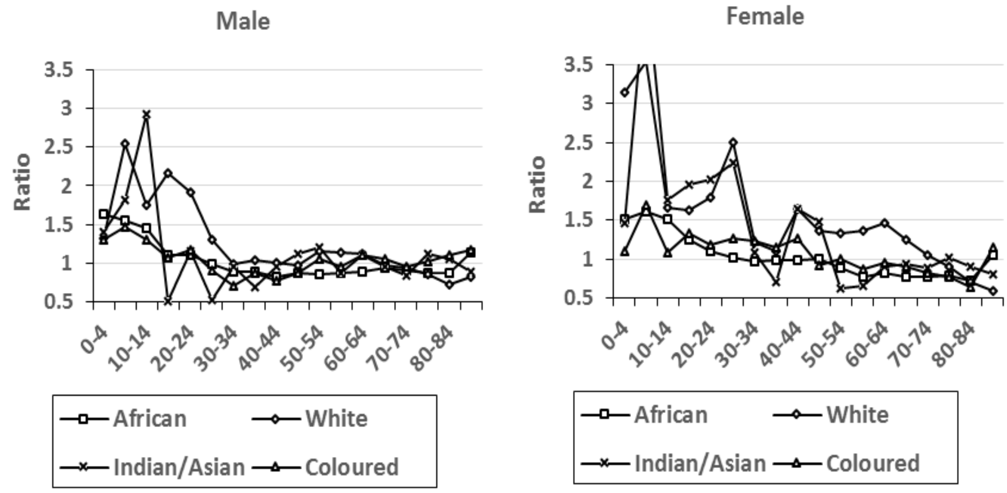


Figure 4.10 compares the ratio of deaths from the two sources between 10 October 2000 and 9 October 2001. At a glance, for males, the distribution of deaths across most age groups is similar from age 30 to the oldest ages. A wide and inconsistent disparity in the distribution of the Indian population at ages below 30 is noticeable, which arises because of age heaping in the household deaths (Figure 4.9), although it is not common for age heaping to occur between the ages 10 to 30, thus all age heaping can be attributed to data error. However, this does not have a significant impact on the quality of the final result as the household deaths will be adjusted accordingly using the true numbers of deaths for each population group derived from registered deaths corrected for completeness. For the other population groups, the ratios generally lie close to one except for the young ages (0 to 24), which possibly may be due to over-reporting of child deaths in the households, which is likely especially in the case where the mother and father of the deceased child are separated.

Figure 4.10: Ratio of the proportion of household deaths to registered deaths in the year prior to the 2001 census, by population group.



For the females, at a glance the ratios are not as close to one as those for males, although for most of the population groups there is a general over-reporting of deaths reported by households (0 to 19). The over-reporting is apparently higher and is more prolonged to age 34 in the Whites and Indians. We suspect it may be a case of data error for the Indians, but for Whites this is unexpected. Similarly, like the Indian males, the inconsistency in household deaths contributes to the irregular pattern between the ages 25 to 59. On the other hand, both sources of death data produce consistent distributions for African females in the ages 20 to 44 and Coloured females aged between 40 and 69.

In general, the distribution of household deaths differs from that of registered deaths in that registered deaths appear to show a smoother and more consistent shape by age-group compared to household deaths. This evidence supports the use of registered deaths (corrected for completeness) as a better indicator of the true level and distribution of overall deaths by sex, population-group and age-group. However, it is not clear why the disparity should be greater for White and Indian women than their male counterparts.

4.1.4 Migration data

The level of the net numbers of migrants was, as expected, higher among the African and White than the other population groups. The distribution by age of migration for males and females is similar for each of the two groups, although this is not so for the level of migration. The number of female migrants are almost half the net number of male migrants. The distribution of migrants by age for the population groups differs, with White emigrants and African immigrants mainly in the middle ages. Also of note is the African male emigration of those aged 40 to 60 which could be attributed to non-South African-born men returning home for retirement. The net numbers of African women migrants are mostly concentrated in the ages 15 to 49, whereas for the White women the net number of migrants are between the ages 15 and 30.

Figure 4.11: Net numbers of migrants in the intercensal period between 1996 and 2001

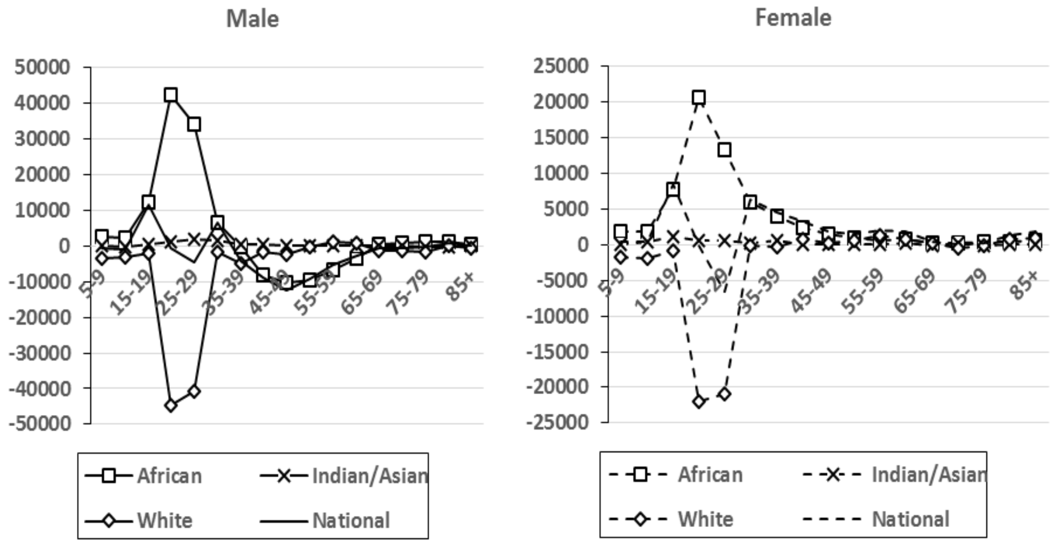
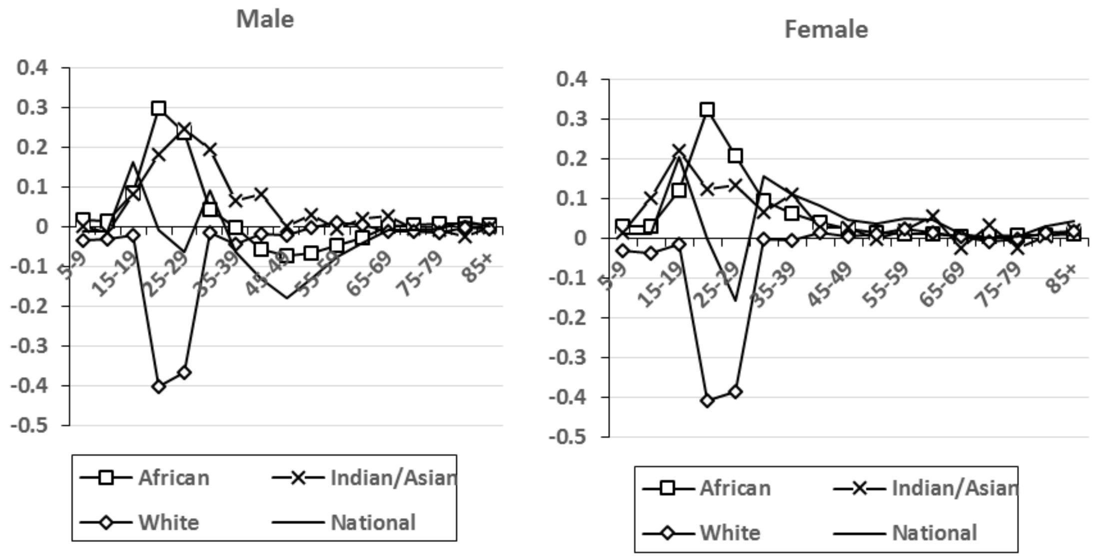


Figure 4.12 shows the distribution of migrants by age for each population group. The distribution of the net number of male migrants is concentrated between the ages 15 and 34 for all the population groups, although for females, the distribution of Indian net migrants seems to be concentrated in a wider age range, starting at the age of 10 to around 39.

Figure 4.12: Proportions of the net numbers of migrants between the 1996 and 2001 census by population group



4.2 Estimating completeness in the intercensal period 1996-2001

For each of the population groups, the GGB diagnostic plots in Figure 4.13, Figure 4.14, Figure 4.15 and Figure 4.16 generally indicate a relatively straight line when fitted to all age groups. Although there is noticeable curvature at the young ages, the fit is relatively close to the straight line. This curvature could be a result of problems with estimates of net migration, although it appears odd that this curvature is also in the Coloured population group. Perhaps the effect could be merely fluctuations in the estimate of completeness at young ages. By and large, the effect appears to be small and subtle. One can also notice traces of age exaggeration at the old ages in the Indian men and Coloured women.

Owing to the relatively straight line, there appears to be minimal age exaggeration when reporting deaths. Additionally, from the plot, the assumption of constant completeness of death registration across all age groups appears to certainly hold true to all the population groups. However, one should bear in mind that the GGB method is less sensitive to age exaggeration than the SEG method (Dorrington, 2013c).

Figure 4.13: GGB diagnostic plots for completeness of registered deaths for the African population group

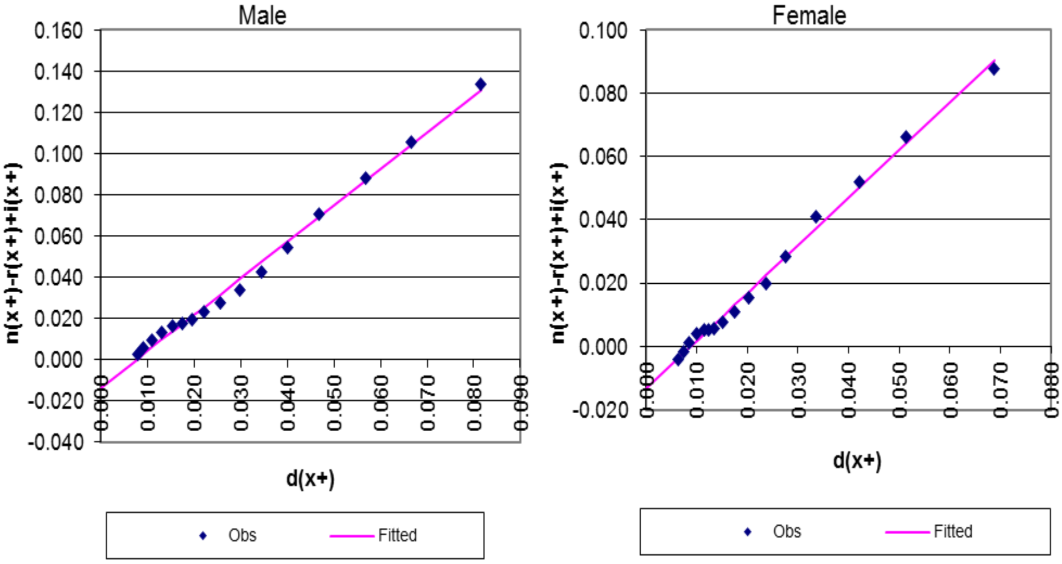


Figure 4.14: GGB diagnostic plots for completeness of registered deaths for White population group

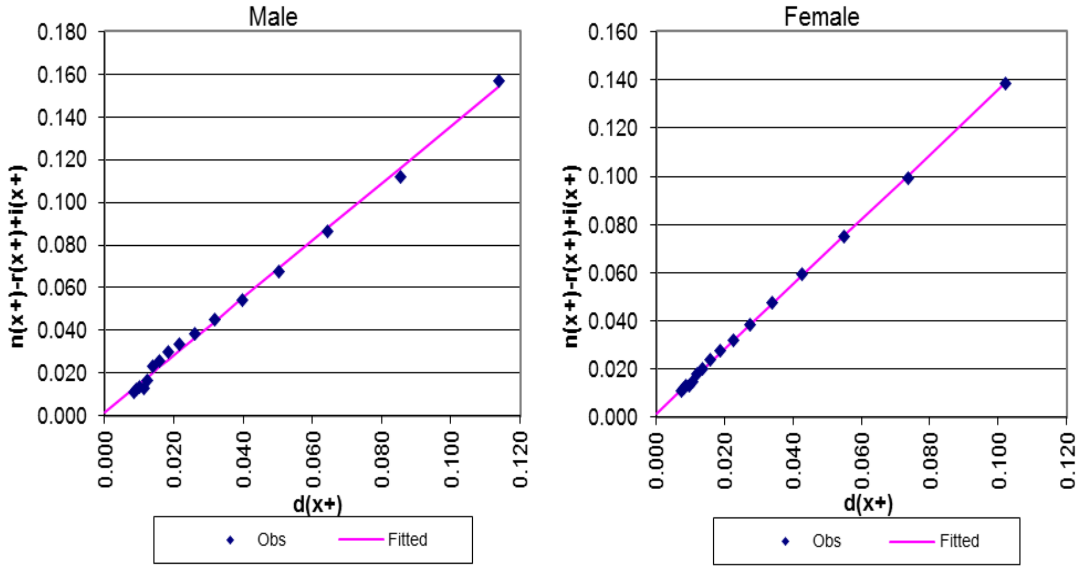


Figure 4.15: GGB diagnostic plots for completeness of registered deaths for Indian population group

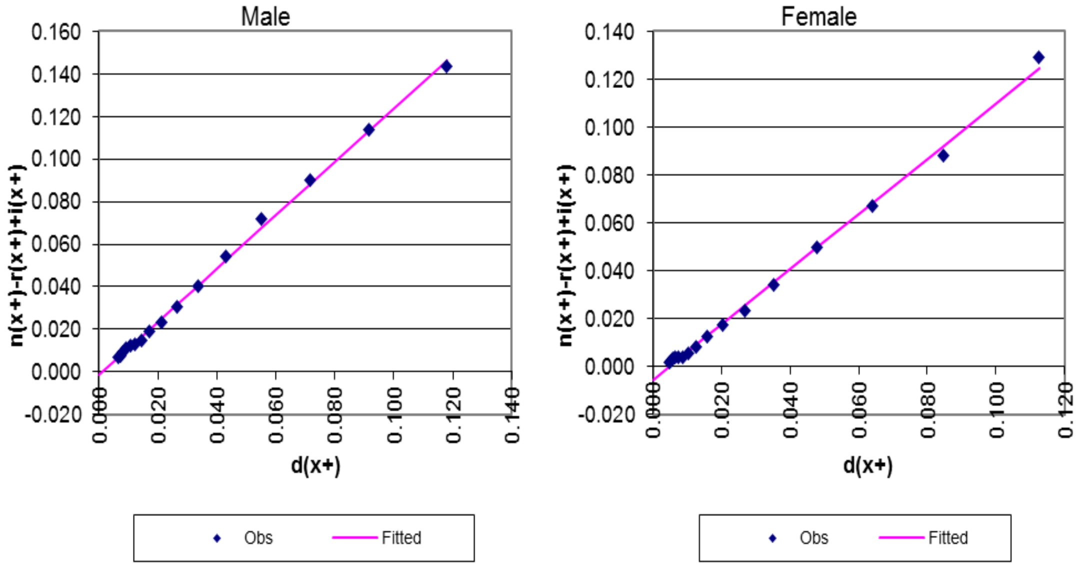
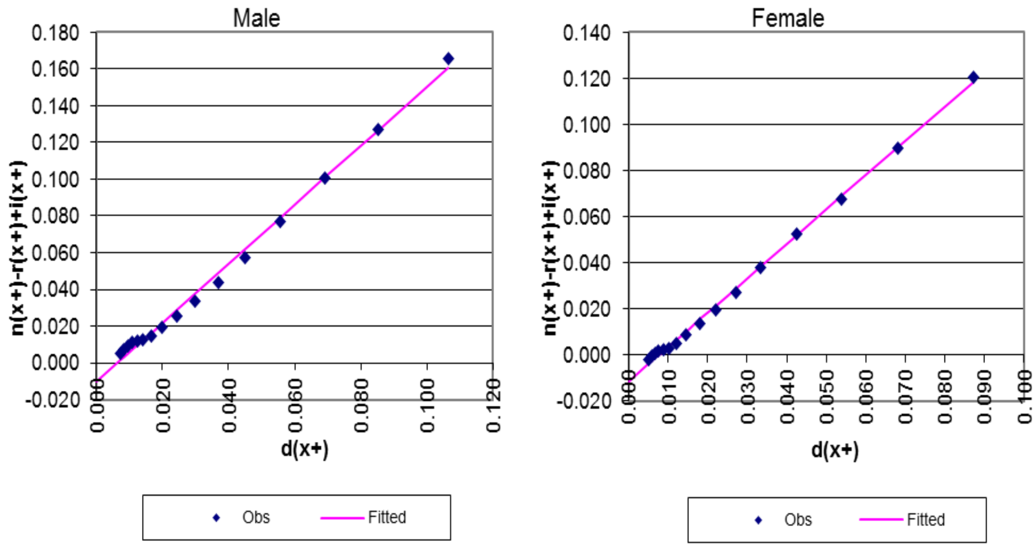


Figure 4.16: GGB diagnostic plots for completeness of registered deaths for Coloured population group

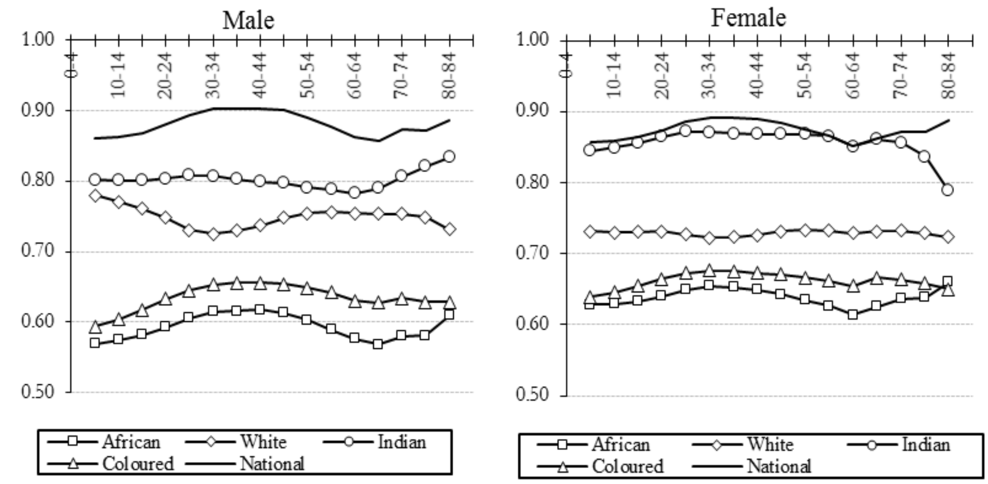


On the other hand, the SEG diagnostic plots shown in Figure 4.17 appear generally to be more level for the women than for men.

For the Indian population, completeness is relatively constant for both men and women at most of the ages except old ages. The rise and fall in completeness at the old ages may perhaps be due to either age exaggeration of the age at death or age exaggeration in the population counts relative to the registered deaths in the intercensal period. However, it is strange that completeness by sex for the population groups acts in an opposite pattern at the old ages.

For the Africans and the Coloureds, the dip in the childhood ages for females may indicate the extent of under-reporting of child deaths as there is a greater temptation to do so (since there is generally few incentives for registering a child death, unlike an adult death, which may be necessary for things such as wealth redistribution, etc.), more so in the rural areas. However, for males, the dip may be additionally lower due to problems in the census counts. This difference perhaps arises from mortality for men being generally greater than of women at all ages. Interestingly, the dip in African completeness from ages 50 to 79 may be an indication of reduced completeness of death reporting due to men retiring to their rural homes from the urban areas. However, you would expect the dip to remain low, yet it rises just after age 65 and 60 in males and females respectively.

Figure 4.17: SEG diagnostic plots for completeness of registered deaths in the intercensal period



For the White population group, there is evidence of constant completeness of death registration for the females. However, for the males, there is a dip between the ages of 15 to 44. Since this age range is characteristic of migration, perhaps this may have arisen from our estimate of net migration, although it should be noted that the effect of this fluctuation on overall completeness is relatively small.

Although there are indications of fluctuations in the completeness of reporting deaths across age groups, which is apparent from the SEG diagnostic plots, especially for the African and White male groups, the extent of the fluctuations seems to be relatively small, which is also strengthened by the diagnostic plots from the GGB method, indicating it can be fair to assume completeness is constant across all age groups in each of the population groups.

From the results of completeness in Table 4.2, completeness estimates from the GGB and SEG+δ methods were generally consistent over the different age ranges except in some instances where the age interval with ages 25 years or more gave slightly smaller estimates than the estimates in other age intervals. Estimates of completeness from the GGB were generally lower than those produced for the SEG+δ, which was more apparent in the African and Indian men than any other groups. This could be linked to a systematic bias of the two methods arising from unaccounted for immigration. However, the GGB produced higher estimates of completeness than the SEG+δ for Whites and Indian/Asian females when the upper age interval is 75.

Table 4.2: The completeness of death registration between 1996 and 2001 by sex among population groups and the whole of South Africa using the GGB and SEG+δ

	Age Interval	*African*		*White*		*Indian/Asian*		*Coloured*		*National*	
		GGB	*SEG*	*GGB*	*SEG*	*GGB*	*SEG*	*GGB*	*SEG*	*GGB*	*SEG*
Male											
	5-84	54%	59%	75%	76%	79%	81%	61%	63%	83%	86%
	15-84	54%	60%	75%	75%	79%	80%	60%	64%	83%	86%
	25-84	52%	61%	77%	75%	79%	81%	58%	64%	83%	87%
	5-75	55%	58%	76%	77%	78%	80%	63%	63%	85%	86%
	15-75	55%	59%	78%	75%	77%	80%	62%	64%	85%	86%
	25-75	52%	61%	80%	75%	77%	81%	60%	65%	84%	87%
	Migration	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Female											
	5-84	64%	63%	74%	73%	86%	86%	65%	66%	87%	87%
	15-84	64%	64%	74%	73%	85%	86%	65%	67%	87%	88%
	25-84	63%	65%	75%	73%	84%	87%	64%	67%	86%	88%
	5-75	62%	63%	74%	73%	90%	86%	66%	66%	86%	86%
	15-75	61%	64%	74%	73%	90%	87%	66%	67%	86%	88%
	25-75	60%	65%	75%	73%	89%	87%	64%	67%	83%	89%
	Migration	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes

Completeness in the whole of South Africa was higher than that of all the population groups since it also includes those deaths that were registered without specifying the population group.

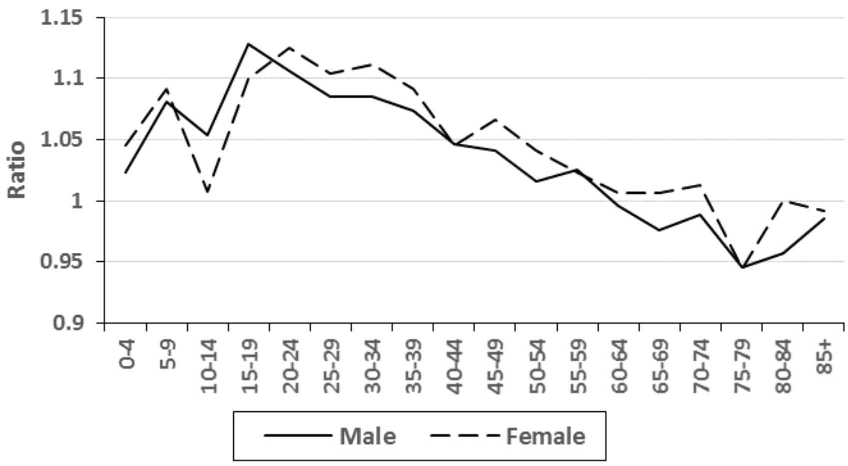
The final estimates of completeness are shown in Table 4.3. For the African and Indian population groups completeness of reporting deaths is 7 per cent higher in females than it is for males, unlike the other population groups and the whole of South Africa where the difference is quite small. Since a greater proportion of the African population live in the rural areas and completeness of reporting of the rural population is lower than that of the urban population, it is not surprising that it has the least complete death registration among the rest of the population groups.

Table 4.3: Final estimates of completeness in the intercensal period 1996-2001 for each population group and the population as a whole and each sex

	African	*White*	*Indian/Asian*	*Coloured*	*National*
Male	56.5%	75.5%	79.5%	63.0%	85.5%
Female	63.5%	73.5%	86.5%	66.0%	87.0%

The estimates of the true numbers of deaths by population group and for the entire population were calculated by dividing the number of registered deaths for each group by the corresponding estimates of completeness. The aggregate of the population group numbers of deaths are compared with the complete numbers calculated for the entire population. Figure 4.18 shows this comparison by age.

Figure 4.18: Ratio of the sum of expected deaths in the population groups to the number of deaths expected from the population as a whole by sex in the intercensal period 1996-2001



The age distribution of the ratio in both males and females is quite similar. Below the age of 64, the expected numbers of deaths from aggregating population group estimates appear greater than the numbers expected from all deaths combined, whereas the opposite is true for numbers in ages older than 65. The ratio peaks from the ages of 15 to 30 for males and 15 to 39 for females where the aggregated population group numbers are 10 per cent greater than the national numbers.

Because completeness of registering deaths is not the same for each population group by age and the proportion of the total number of registered deaths in each population group differ by age using the completeness estimate from the entire population would violate the assumption that completeness is constant by age. Therefore, the expected numbers of deaths from aggregating population groups are taken to be the true numbers rather than the numbers obtained from the result of completeness from the entire population. Table 4.4 shows the estimates of completeness obtained from the sum of expected deaths for each of the population groups, a final completeness estimate of 99 per cent and 98 per cent respectively for the males and females was obtained. Although the final estimate is not exactly 100 per cent, the estimates do not deviate markedly from being complete.

Table 4.4 Estimates of completeness derived from the sum of expected deaths from each of the population groups by sex

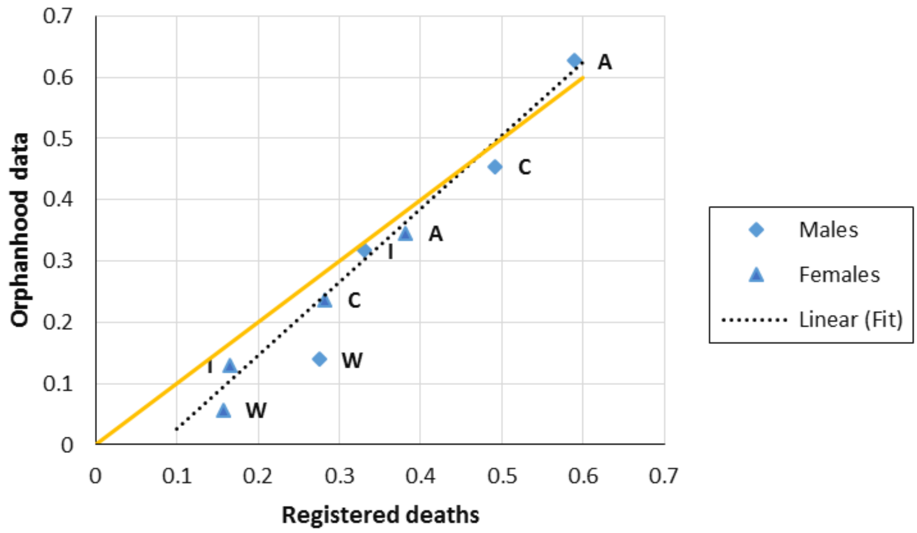
	GGB						*SEG*						*Final*
Age interval	*5-84*	*15-84*	*25-84*	*5-75*	*15-75*	*25-75*	*5-84*	*15-84*	*25-84*	*5-75*	*15-75*	*25-75*	
Male	98%	97%	95%	99%	99%	96%	99%	100%	101%	99%	100%	101%	99%
Female	100%	100%	99%	98%	98%	96%	97%	98%	99%	96%	97%	99%	98%

4.3 Comparing $_{45}q_{15}$ from registered deaths in the intercensal period to those from orphanhood data

Using the estimates of completeness from registered deaths for each population group and each sex, the deaths in the intercensal period were corrected for completeness to obtain the expected numbers of intercensal deaths from which estimates of $_{45}q_{15}$ were derived. These were then compared to estimates of $_{45}q_{15}$ from the orphanhood data. This is shown in Figure 4.19.

All the estimates of adult mortality for women from the orphanhood data are lower than those from the registered deaths. In contrast, for men, the estimates between the two approaches are quite close to each other, except for the Whites. Estimates of adult mortality from registered deaths for the Whites deviate from the orphanhood most. Perhaps the orphanhood estimates for the White population group are biased downwards or maybe estimates for completeness for the White population group were underestimated (the latter seems more likely for men considering outcomes from the SEG diagnostic plots). Above all, the estimates reflect a consistent ranking in population groups for both men and women between the registered deaths and the orphanhood data.

Figure 4.19: Comparison of $_{45}q_{15}$ derived from registered deaths after correcting for completeness with that derived from the orphanhood data for each population groups and each sex

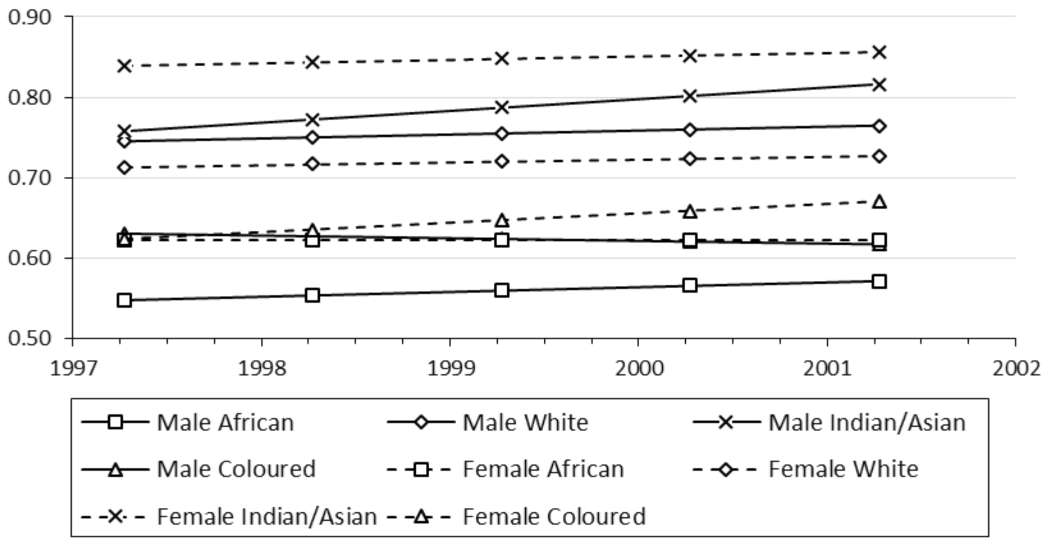


Note: The letter A represents Africans, W represents Whites, I represents Indians/Asians and C represents Coloureds, the solid dash line is a 45 degree line.

4.4 Estimating annual completeness in the intercensal period

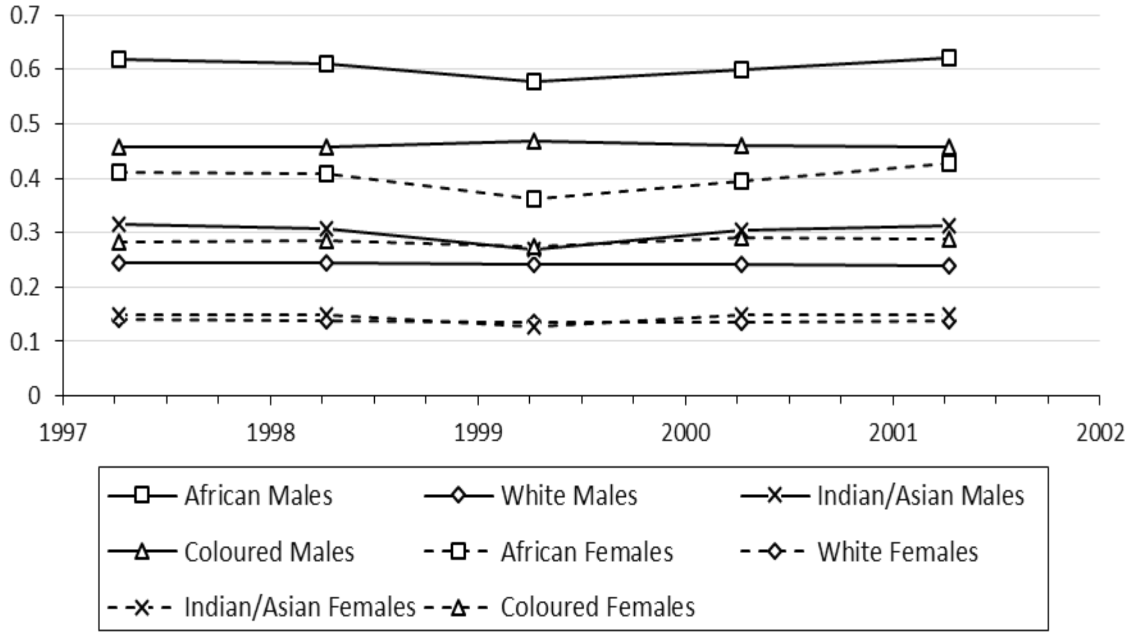
Using the estimates of completeness in the intercensal period assumed to apply six months before the census, annual census anniversary estimates of completeness of death registration are shown in Figure 4.20. Across the intercensal period completeness generally increased for most of the population groups.

Figure 4.20: Annual estimates of completeness for each population group and each sex over the intercensal period



The numbers of expected deaths for each population group and each sex between two consecutive census anniversary years is obtained by dividing the numbers of registered deaths in that period by the corresponding annual estimate of completeness. The expected numbers of deaths between two consecutive census anniversary years were then used to estimate adult mortality for each population group and sex, the trend is shown in Figure 4.21.

Figure 4.21: Estimates of $_{45}q_{15}$ for each year between the censuses of population groups by sex

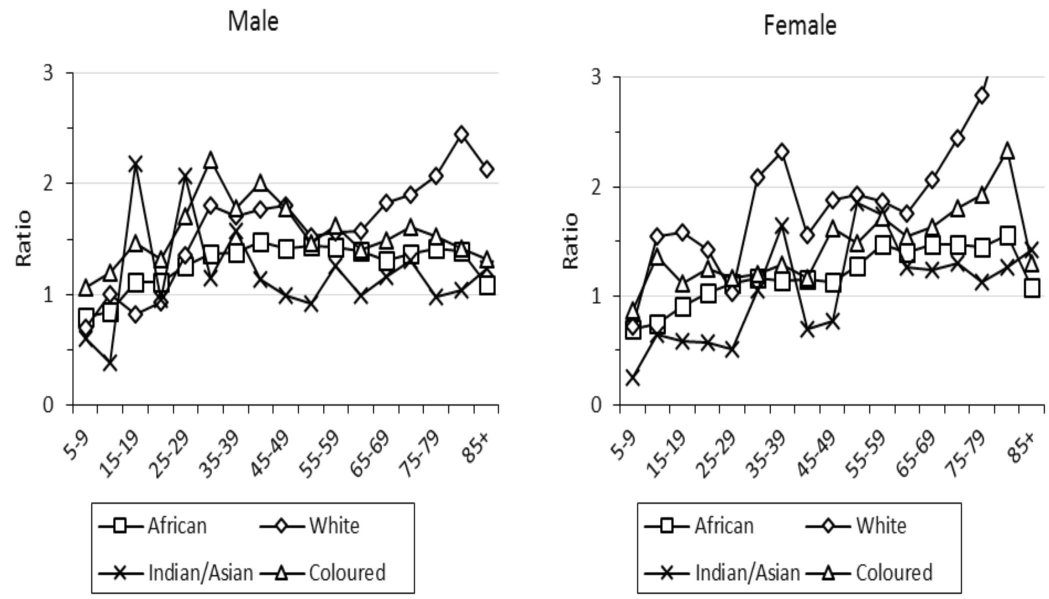


For the African population group, adult mortality increases in the period 1999 to 2001, whereas for the other population groups the increase is very small to non-existent. The increase in African mortality is higher for the females than for the males. We would have expected the mortality to reflect an upward trend throughout the entire intercensal period. However, the observed increases from the year 1999 are almost certainly due to the impact of HIV/AIDS. As expected, African mortality is higher than that of other population groups with the White population having the lowest mortality.

4.5 Adjusting deaths reported by households at district level

Sex population group age-specific adjustment factors for correcting the household deaths, calculated from the ratio of the expected numbers of deaths in the year prior to the 2001 census to the deaths reported by households are shown in Figure 4.22.

Figure 4.22: Ratio of adjustment factors in the 2001 census by age for each population group and each sex



Adjustment factors reflect by and large an under-reporting of deaths in the households, with the exception being children aged between 5 and 14 showing over-reporting of deaths in the 2001 census. Although it is quite likely that deaths may be missed in the census, for the children, it could be explained by child deaths occurring to separated parents, however, we do not expect that reasoning to produce as significant an effect as that observed. Over-reporting often results from double-counting deaths in multiple households, which is expected more in adult than child deaths since adults are more remote and likely to stay in multiple households due to the nature of their employment, for example. Thus, it could be that registered deaths between ages 5 and 14 are less complete than estimated. This reasoning is evident from the SEG diagnostic plot of completeness shown previously in Figure 4.17 reflecting a lower level of completeness for children aged between 5 and 14 years.

The adjustments to the numbers of deaths of women reported by households were greater than those for men for the Coloured and White population, which becomes more apparent at older ages. For the Whites, it may be a case of higher numbers of women dying in retirement homes which are not recorded in the census, although it is doubtful that the same reasoning explains the similar observation apparent in the Coloureds.

For men, the Indian adjustment factors indicate no smooth pattern, which is indicative of the inconsistency of Indian reported deaths by households apparent in

Figure 4.9 and Figure 4.10 which could be result from the scarcity of Indian deaths compared to deaths in other population groups. After age 29, the Indian adjustment factors tend to get closer to one, reflecting some degree of consistency in reporting between the registered deaths and the reported deaths from the census. For the White males, the adjustment factors tend to rise at the older ages due to under-reporting of old age deaths by households in the census highlighted earlier. Observable peaks for the Coloured and White groups at the adult ages indicate under-reporting of deaths in the census. For Whites this may be attributable to some households being missed in the census due to households being inaccessible to census enumerators, although it is not certain if this reasoning also holds for the Coloureds. The peak starts in the younger ages for the Coloured group unlike the Whites. On the other hand, over-reporting of child deaths in the census is apparent from the African, White and Indian groups. Although Africans have generally consistent ratios across all age groups greater than 15, they reflect a general under-reporting of deaths from the census across these ages.

For the females, the adjustment factors for Africans are closer to one with Coloureds and Whites having greater adjustment factors at older ages. However, adjustment factors for the White population are about double those of the Coloureds at the open age interval. For the Whites, this is almost certainly because a high proportion of White women live and die in retirement homes. Apart from the old ages and children aged less than 10, there is a general under-reporting of deaths in the census from the White population, for example, between the ages 30 and 60 it appears that about half of the deaths were not reported in the census, which shortage could be due to disintegration of the smaller households (single or two-person households) at the death of an individual. A consistent over-reporting of deaths between the ages 5 and 30 is evident from the Indian females in Figure 4.22. Reported numbers of African deaths are consistent with the registered deaths below the age of 50, although there is some under-reporting of deaths at the older ages.

4.6 Adult mortality at district level between the years 2000 and 2001

The household deaths were used to estimate an incomplete life table (not accounting for child mortality) for each district and each sex after applying the adjustment factors. Results of ${}_{45}q_{15}$ are then computed for each district and sex.

After this the incomplete life table is graduated using the provincial standard. The standard life table for each province is derived from the ASSA 2008 AIDS and demographic projection model for the province after projecting to 2001. The provincial standards are shown in Appendix A.

Since the mortality of underlying districts is more likely to be similar to the corresponding provincial estimate than not (because of similarity of conditions), it is assumed that in choosing between the crude and the graduated estimate of ${}_{45}q_{15}$ the better estimates for the districts are the ones that deviate less from the provincial estimate. The orphanhood provincial estimate of ${}_{45}q_{15}$ is assumed to be the provincial estimate in obtaining the sum of absolute errors for each province. The sum of the absolute errors within each province are weighted by the population aged between 15 and 60 for each district. Essentially the choice of whether or not to graduate all or none of the district estimates within each province depends upon which of the two results in a lower sum of deviations. Table 4.5 shows the mean absolute percentage error of both estimates and the decision on which estimates were used further. From the table, generally, the district estimates of ${}_{45}q_{15}$ for men have less deviation from the provincial orphanhood estimates compared to those of women. Before graduating, the estimate of ${}_{45}q_{15}$ from Western Cape and North West for women deviates widely from the orphanhood estimate of ${}_{45}q_{15}$. Unexpectedly, for women, all the district estimates favoured graduation to reduce their deviation to the provincial orphanhood estimate.

Table 4.5: Mean absolute percentage error of crude and graduated district estimates of ${}_{45}q_{15}$ to those of the province derived from orphanhood data

	Male			Female		
Province	*Crude*	*Graduated*	*Decision*	*Ungraduated*	*Graduated*	*Decision*
WC	13.6%	25.2%	Crude	45.1%	19.4%	Graduated
NC	12.8%	11.2%	Graduated	33.2%	24.6%	Graduated
EC	23.8%	18.8%	Graduated	38.2%	34.6%	Graduated
FS	8.7%	9.3%	Crude	38.8%	32.2%	Graduated
KZN	8.1%	6.5%	Graduated	18.0%	14.4%	Graduated
MP	8.1%	12.9%	Crude	15.0%	13.5%	Graduated
LP	16.2%	23.2%	Crude	12.4%	11.7%	Graduated
NW	12.9%	8.4%	Graduated	43.9%	32.9%	Graduated
GT	19.5%	13.3%	Graduated	21.9%	21.8%	Graduated
Overall	13.9%	14.1%	Crude	29.3%	22.6%	Graduated

After deciding on using the crude or graduated life table in obtaining an estimate of ${}_{45}q_{15}$ for each district in that province, the resulting estimates are shown in Figure 4.23 and Table 4.6.

Generally, in the absence of exception circumstances such as particular epidemics (e.g. HIV) or high maternal mortality, men experience higher mortality than women. For men, the level of mortality is lower in districts around Western Cape and Gauteng as they are predominantly urban than the rest of the provinces. Most of the districts in Kwa-Zulu-Natal have a relatively higher adult mortality, as well as Eastern Cape to a lesser extent. On the other hand, women were observed to have the lowest level of adult mortality in districts in Western Cape and Limpopo, whereas districts within Kwa-Zulu-Natal, Eastern Cape and Free State had the highest level of adult mortality for women.

Table 4.6: Top 15 lowest and highest ranking districts for $_{45}q_{15}$ in South Africa in 2001

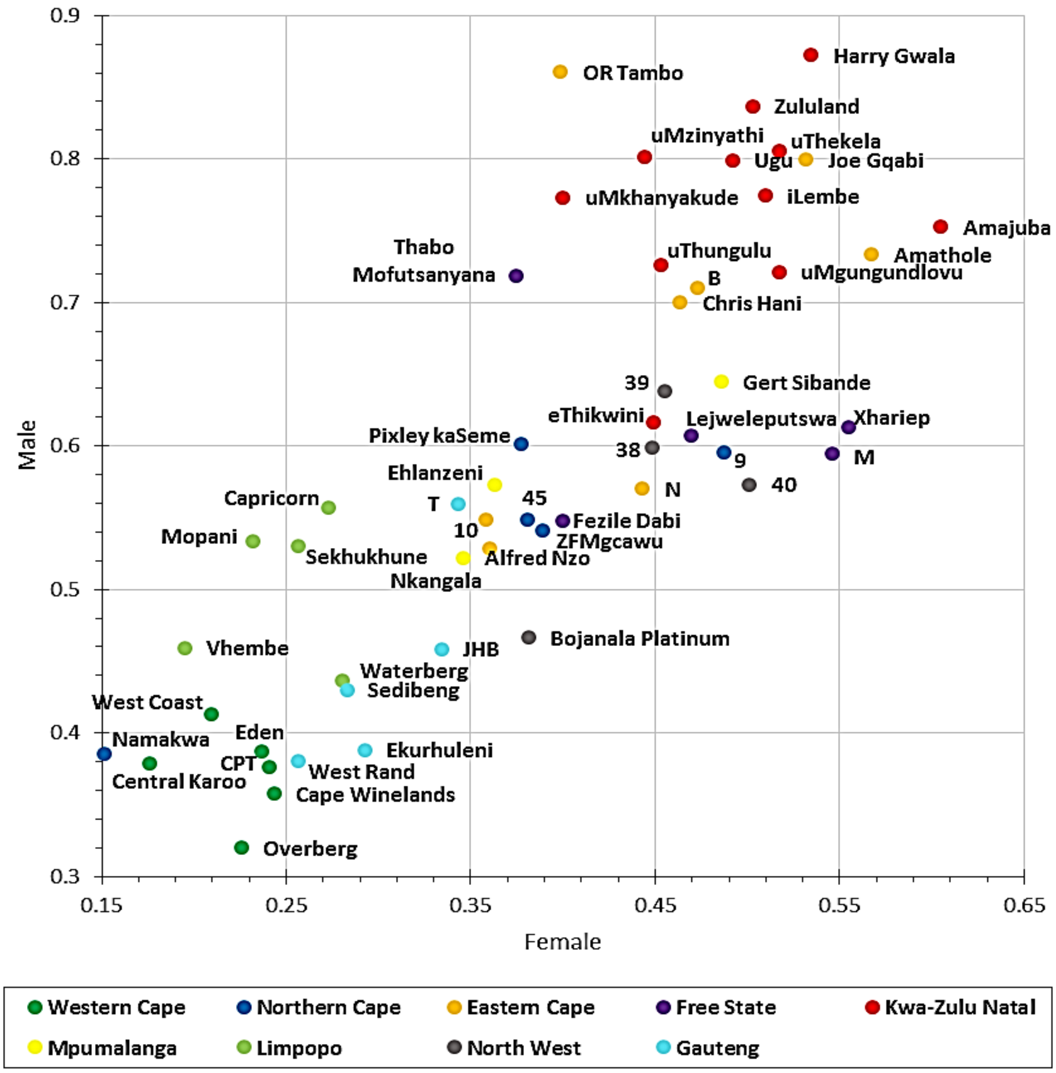
	Lowest mortality				*Highest mortality*			
	Male		*Female*		*Male*		*Female*	
Rank	*District*	*Province*	*District*	*Province*	*District*	*Province*	*District*	*Province*
1	Overberg	WC	Namakwa	NC	Harry Gwala	KZN	Amajuba	KZN
2	Cape Winelands	WC	Central Karoo	WC	OR Tambo	EC	Amathole	EC
3	CPT	WC	Vhembe	LP	Zululand	KZN	Xhariep	FS
4	Central Karoo	WC	West Coast	WC	uThekela	KZN	Mangaung Metropolitan	FS
5	West Rand	GT	Overberg	WC	uMzinyathi	KZN	Harry Gwala	KZN
6	Namakwa	NC	Mopani	LP	Joe Gqabi	EC	Joe Gqabi	EC
7	Eden	WC	Eden	WC	Ugu	KZN	uThekela	KZN
8	Ekurhuleni	GT	CPT	WC	iLembe	KZN	uMgungundlovu	KZN
9	West Coast	WC	Cape Winelands	WC	uMkhanyakude	KZN	iLembe	KZN
10	Sedibeng	GT	Sekhukhune	LP	Amajuba	KZN	Zululand	KZN
11	Waterberg	LP	West Rand	GT	Amathole	EC	Dr Kenneth Kaunda	NW
12	JHB	GT	Capricorn	LP	uThungulu	KZN	Ugu	KZN
13	Vhembe	LP	Waterberg	LP	uMgungundlovu	KZN	Francis Baard	NC
14	Bojanala Platinum	NW	Sedibeng	GT	Thabo Mofutsanyana	FS	Gert Sibande	MP
15	Nkangala	MP	Ekurhuleni	GT	Buffalo Metropolitan	EC	Buffalo Metropolitan	EC

There is a consistent trend in provinces where there are extremes in adult mortality among men and women. However, the districts by themselves do not rank consistently between men and women, with the exception of Eden, Buffalo Metropolitan and Joe Gqabi district.

All the districts in Western Cape are among the 15 districts with the lowest levels of $_{45}q_{15}$ among both sexes, the same goes for women in districts of Limpopo. At the opposite end, districts from Kwa-Zulu-Natal dominate the districts with the highest levels of $_{45}q_{15}$, as well as those in Free State and Eastern Cape to a lesser extent.

As can be seen for Figure 4.23, that for the majority of the districts the chances of men aged 15 not reaching the age of 60 are over 50 per cent provided the conditions in 2001 prevail in the future. For women, the chances are over 35 per cent, and although lower than that of the males, adult mortality was generally high. It is also visible that there is no consistency between the level of ${}_{45}q_{15}$ among districts of the same province except for those in Western Cape, Gauteng, Limpopo and Kwa-Zulu-Natal with a few exceptions. One of the exceptions is in Kwa-Zulu-Natal, the district of eThikwini reflects a relatively lower mortality for men compared to all the other districts in that province.

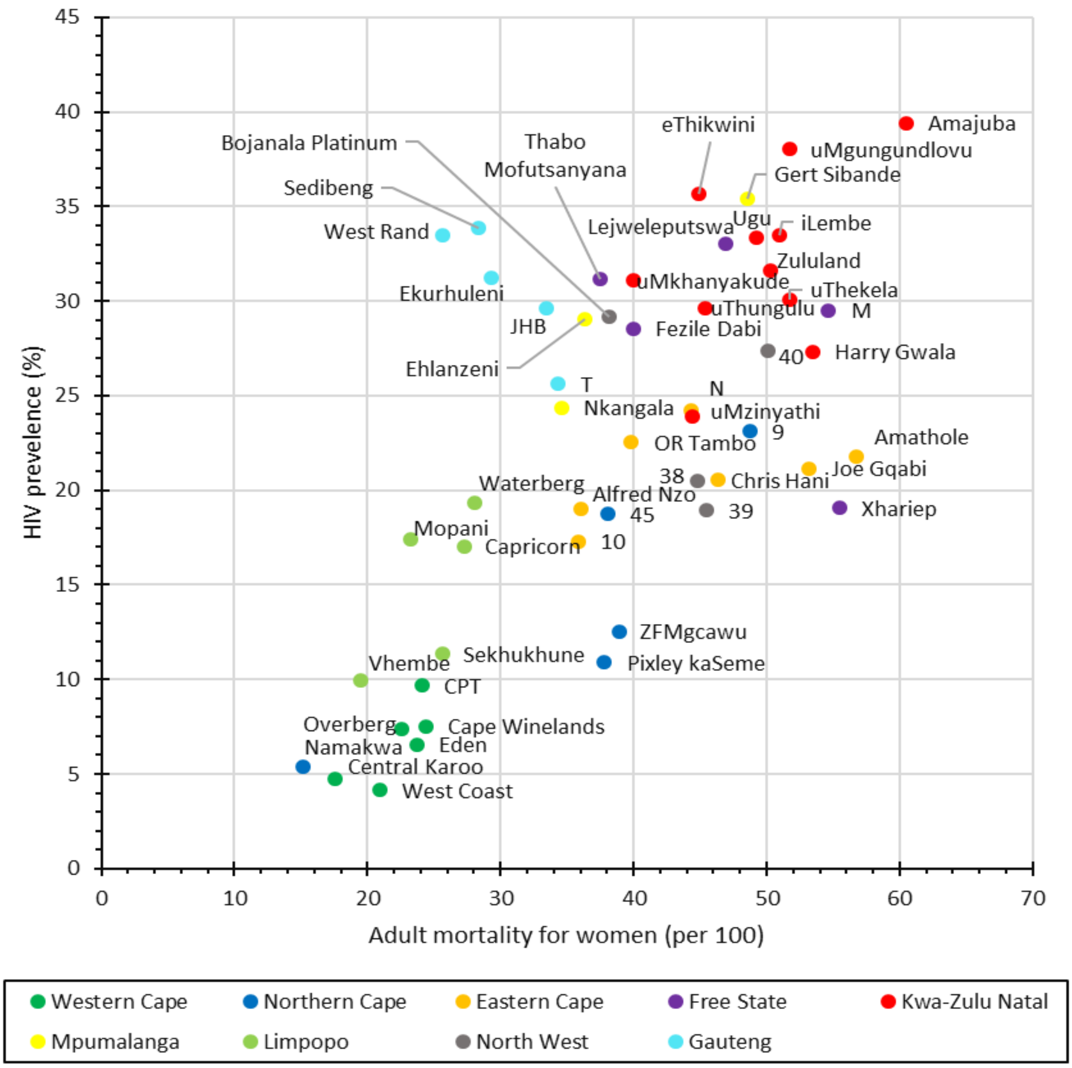
Figure 4.23: Estimates of ${}_{45}q_{15}$ among men and women by district



Note: CPT- City of Cape Town 9- Francis Baard 45- John Taolo Gaetsewe 10- Sarah Baartman N- Nelson Mandela Bay Metropolitan B- Buffalo Metropolitan M- Mangaung Metropolitan 38- Ngaka Modiri Molema 39- Dr Ruth Segomotsi Mompati 40- Dr Kenneth Kaunda T-City of Tshwane Metropolitan JHB- City of Johannesburg

In Figure 4.23 Gert Sibande has a higher adult mortality compared to the other provinces in Mpumalanga. Similarly, in Northern Cape, Namakwa has lower level of adult mortality than other districts in that province perhaps because the population in the district is small. In Eastern Cape, there is also a disparity between districts with higher mortality and a few reflecting lower mortality including Alfred Nzo and Sarah Baartman. There exists a wide difference in the ranking and the level of ${}_{45}q_{15}$ in OR Tambo between men and women, the level of mortality is extremely high for men whereas it is somewhat lower when compared to the other district.

Figure 4.24: Estimates of ${}_{45}q_{15}$ for women against the HIV prevalence from antenatal women by district in 2001



Note: CPT- City of Cape Town 9- Francis Baard 45- John Taolo Gaetsewe 10- Sarah Baartman N- Nelson Mandela Bay Metropolitan M- Mangaung Metropolitan 38- Ngaka Modiri Molema 39- Dr Ruth Segomotsi Mompati 40- Dr Kenneth Kaunda T-City of Tshwane Metropolitan JHB- City of Johannesburg

In Figure 4.24, the relationship between adult mortality and HIV prevalence of women attending antenatal clinics is apparent. However, it must be acknowledged that the HIV prevalence estimates are correlated with population group which is also correlated with the level of mortality which weakens the causal correlation to some extent. For example, Amajuba in Kwa-Zulu-Natal has the highest adult mortality for women and has the highest HIV prevalence among women attending antenatal clinics in the year 2001. Districts in Western Cape, Limpopo and Northern Cape experience low HIV prevalence, which may explain the low levels of adult mortality, the converse is true for districts in Kwa-Zulu-Natal. For districts in Gauteng, however, there is high HIV prevalence but the districts reflect lower adult mortality estimates. This may be attributed to people who may become infected not dying within that province.

4.7 Comparing estimates from deaths reported by households with those from orphanhood data at provincial level

After obtaining the provincial estimates of adult mortality from the district estimates weighted by the district population, the estimates of adult mortality were compared to the orphanhood estimate of adult mortality in Figure 4.25 and in Table 4.7.

Table 4.7: Comparison of the ranking of the provinces with respect to $_{45}q_{15}$ (lowest to highest)

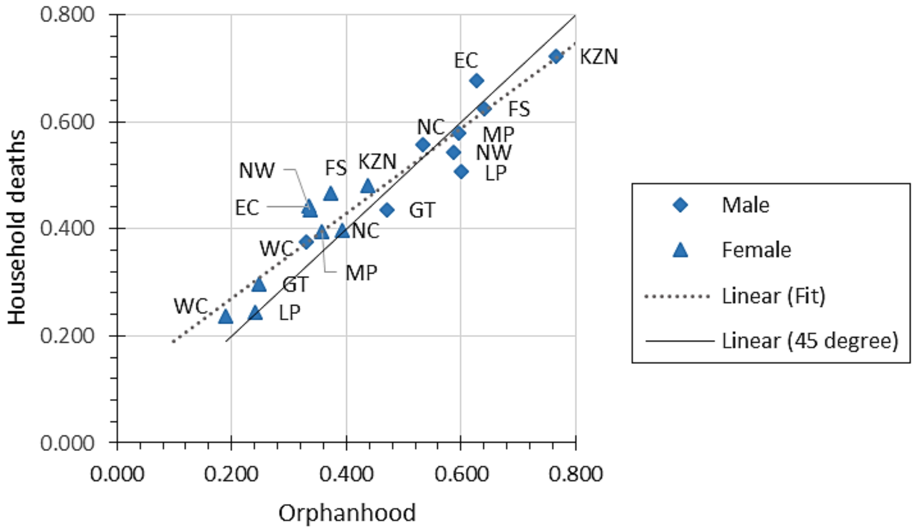
	Male		*Female*	
Rank	*Orphanhood*	*Household deaths*	*Orphanhood*	*Household deaths*
1	**Western Cape**	**Western Cape**	**Western Cape**	**Western Cape**
2	**Gauteng**	**Gauteng**	**Limpopo**	**Limpopo**
3	Northern Cape	Limpopo	**Gauteng**	**Gauteng**
4	**North West**	**North West**	Eastern Cape	Mpumalanga
5	*Mpumalanga*	Northern Cape	*North West*	Northern Cape
6	Limpopo	*Mpumalanga*	Mpumalanga	*North West*
7	*Eastern Cape*	*Free State*	*Free State*	Eastern Cape
8	*Free State*	*Eastern Cape*	Northern Cape	*Free State*
9	**KwaZulu-Natal**	**KwaZulu-Natal**	**KwaZulu-Natal**	**KwaZulu-Natal**

Note: Bold indicates consistent ranking, italics indicates one position difference in ranking

Estimates of adult mortality for men are more consistent between the two sources than the estimates of adult mortality for women. For example, women the Northern Cape rank eighth using estimates from the orphanhood whereas they rank fifth with estimates from deaths reported by households. On the other hand, for men the Northern Cape ranks third and fifth respectively. Certainly for the orphanhood, estimates of adult mortality will not always be consistent as parents do not always live in the same province as their children over the period for which the rates are being estimated.

As observed previously, in Figure 4.25 estimates of female mortality from deaths reported by households are higher than the estimates from orphanhood data, although this pattern is not that distinct in males. Although we suspect the adoption effect to bias the estimates of ${}_{45}q_{15}$ downwards for both sexes, the effect of absent fathers biases the estimates for men upwards, which counteracts, to the same extent, the bias introduced by the adoption effect. Thus, estimates of women from the orphanhood data are biased downwards. For men, by and large the household estimates are lower than the orphanhood estimates, although the pattern is not as distinct as that of women. This may perhaps be due to the absence of fathers having a greater impact on the bias of the orphanhood data than the adoption effect.

Figure 4.25: Comparison of provincial weighted average of district rates for ${}_{45}q_{15}$ from the households reported deaths against the provincial rates from the orphanhood method



From this comparison, it should be noted that the mortality reflected by the orphanhood estimates may not be entirely representative of the province in which the children live as parents will not always live in the same province as their children, which particularly holds true in provinces such as Western Cape and Gauteng (since the provinces are predominantly urban and attract people due to the opportunities of employment they offer).

Additionally, in converting to ${}_{45}q_{15}$ the orphanhood data uses a cohort measure to arrive at this estimate whereas the estimates derived from household data give a period estimate, and thus we would expect the estimates derived from the orphanhood data to be generally lower compared to those from the deaths reported by households, especially in this scenario where adult mortality appears to increase in the years leading to 2001. This does not hold for all provinces since in some provinces, specifically the more urbanised ones, adult mortality may not be increasing at the same level as provinces like Eastern Cape and Kwa-Zulu-Natal for several reasons, including better health facilities in the more urbanised provinces.

Lastly, the time reference points of the household deaths and the orphanhood estimates are not exactly the same, but this should have a marginal effect since the period is merely a year apart.

5 DISCUSSION

5.1 Introduction

This chapter evaluates the district estimates of adult mortality obtained in the previous chapter. Thereafter we assess into the limitations of this approach to obtaining adult mortality estimates at the district level. The chapter concludes on the reasonableness of the estimates of adult mortality for the districts using the approach proposed by Dorrington, Moultrie and Timæus (2004) and suggest areas for further research.

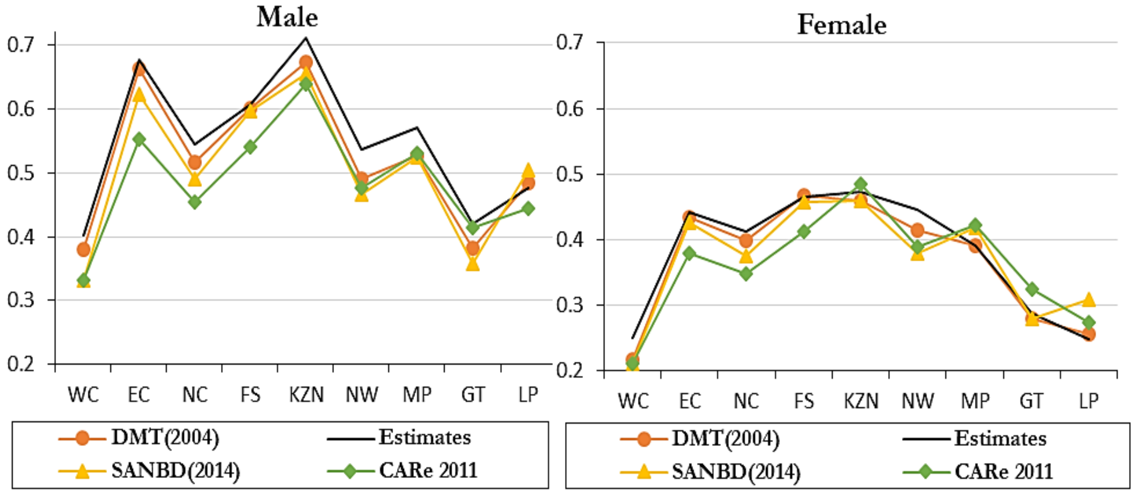
5.2 Discussion of results

Adult mortality in South Africa has been largely affected by the HIV/AIDS epidemic and to a lesser extent chronic diseases and accidents (Dorrington, Bourne, Bradshaw et al., 2001). These interact to cause differences in adult mortality among provinces and among districts. Although HIV prevalence was still rising over time in 2001, studies (Dorrington, Bourne, Bradshaw et al., 2001, Johnson, Dorrington & Matthews, 2007) show a significant number of deaths having occurred because of the HIV/AIDS pandemic at the time.

The predominantly urban provinces of Gauteng and Western Cape are shown in this study to reflect much lower mortality than the rest of the provinces, mainly because of population group composition and also better socio-economic standards than other provinces. Also, since these urban areas attract much of the employment, if one becomes sick or physically unable to work, for example, there may be selective out migration of this group, which is specific to the Africans, some of whom reside in dual residences between the urban and rural areas. Western Cape has a much lower mortality than Gauteng because of the lower HIV prevalence in Western Cape compared to Gauteng. Kwa-Zulu-Natal, Eastern Cape and Free State have the highest adult mortality rates associated with high HIV prevalence in these provinces. In addition to the high prevalence, those who were residing in the urban provinces are more likely to die in these provinces since they are dominantly rural.

The orphanhood method suggests that for the White population this study may have overstated ${}_{45}q_{15}$, although it is likely to be a problem with the orphanhood method of understating mortality. Apart from this most of deaths in the White population occur at the ages over 60, which have limited impact on ${}_{45}q_{15}$. The estimates of ${}_{45}q_{15}$ in this study for women appeared to be high when compared to that of the orphanhood method; however, comparing them with other sources, the estimates look reasonable and close to other estimates in Figure 5.1.

Figure 5.1: Estimates of $_{45}q_{15}$ obtained from this study compared against alternative estimates from other authors and models in 2001



From Figure 5.1, estimates from this study are generally higher than those from other studies and models apart from a few exceptions, one of which being women in the CARe 2011 model having higher estimates for Kwa-Zulu-Natal, Mpumalanga, Gauteng and Limpopo than this study. Provincial estimates of $_{45}q_{15}$ from this study are more consistent with those obtained by Dorrington, Moultrie and Timæus (2004). This should be since the same method was used in producing the estimates. However, estimates from this study appear higher in most of the cases than those of DMT (2004), more so for men. This could be explained by the estimates of completeness from this study being generally lower than those in the study of DMT (2004). For Africans, the estimates of completeness from this study are almost 9 per cent lower for men and 3 per cent lower for women than those from DMT (2004). Although the numbers of registered deaths used in this study were slightly higher due to late registered deaths and also DMT (2004) had to estimate registered deaths, we do not expect the registered deaths in the intercensal period to be significantly different as the method used in estimating registered deaths in the earlier years is not that significantly different.

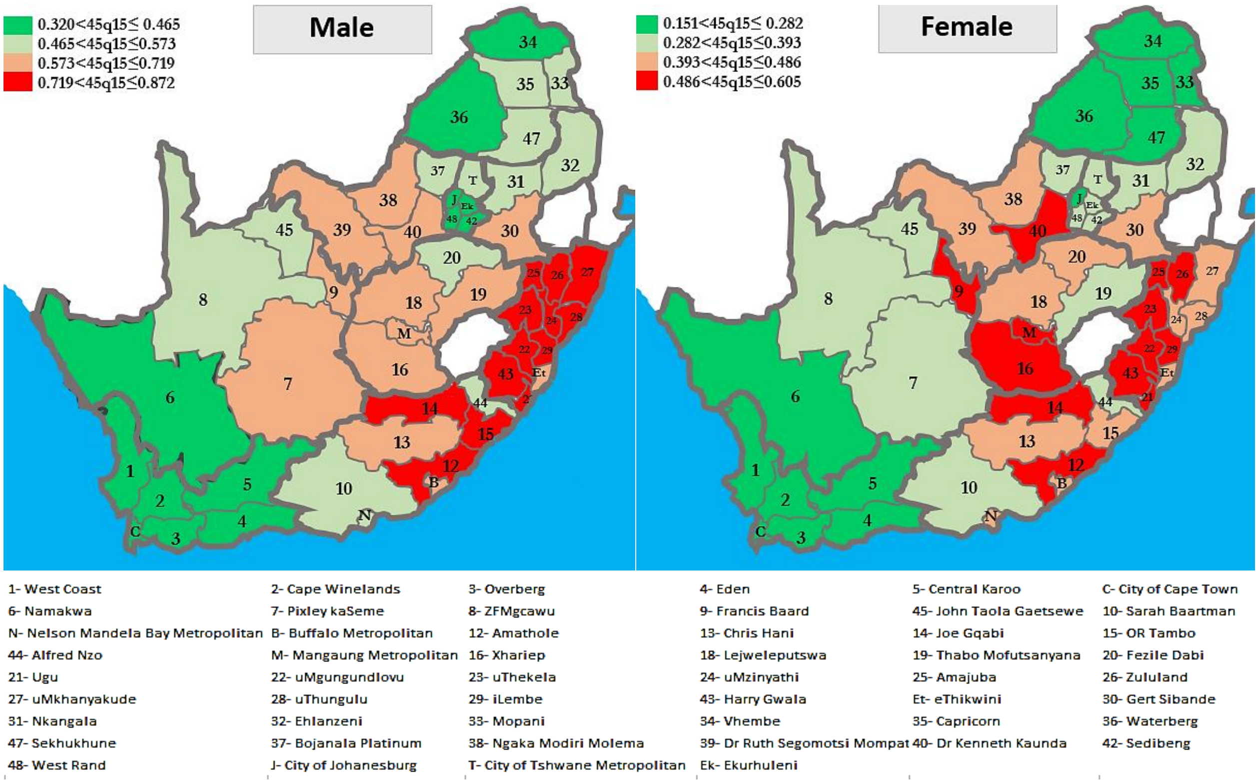
Estimates from the SANBD (South African National Burden of Diseases) (Msemburi, Pillay-van Wyk, Dorrington et al., 2014) generally lie in between those from the CARe 2011 model and those from this study in most of the provinces, which can be expected, as the method used in obtaining the estimates was also like the one from this study. On the other hand, the CARe 2011 model reflects lower mortality in most of the provinces among men and women, which we should expect since the model uses a different approach to the one used in this study. However, the estimates from the CARe 2011 model have a relatively consistent distribution with those obtained from this study.

From Figure 5.1, it is unlikely that Northern Cape has the same level of adult mortality as North West for women as suggested by estimates from DMT (2004) and SANBD (2014) since HIV prevalence in Northern Cape was lower than that in North West at the time. Similarly, considering HIV prevalence, it would be expected that Kwa-Zulu Natal has a higher adult mortality than Free State for women.

In districts around Eastern Cape and Kwa-Zulu-Natal, the results are consistent with the prevalence level observed in 2001 (Department of Health, 2003). However, the level of prevalence may not necessarily imply the AIDS deaths are occurring in that province, especially for districts within Gauteng province, where people may live there primarily for work and their main place of residence is elsewhere, once they fall ill or are unable to work they may migrate to their main place of residence.

There is some consensus as to where the true estimates lie for provinces with lower estimates of mortality. However, as mortality increases the disparities are more apparent among different authors. For provinces, such as the Western Cape, Mpumalanga and Limpopo there is less variation between the districts and the estimate of ${}_{45}q_{15}$ among different authors in the province. The method produces higher adult mortality estimates at provincial level in the more rural provinces compared to the demographic models. This could be attributed to the method taking into account of the deaths that occur in the rural areas, which are usually under-reported to a larger extent than the deaths occurring in the urban areas.

Figure 5.2: Estimates of $_{45}q_{15}$ for men and women in districts of South Africa



The distribution in the level of adult mortality among districts is not consistent when comparing between men and women. There seems to be much consistency for districts within low mortality provinces. Higher mortality areas are more scattered for women than men mainly because men are typically the bread winners and are more likely to influence the decision migration once they fall ill or are unable to work.

For women, Mangaung metropole and Xhariep in Free State province reflect higher mortality on average in the districts than men. The N1 highway to Gauteng and N6 highway to the Eastern Cape passes through Xhariep, the road links may have a rise in prostitution activities, which may affect the HIV prevalence in the area and this may also have an effect on Mangaung metropole since they have a shared border. Although metropolitan districts are usually associated with lower mortality due to better services and infrastructure it is quite odd that Mangaung metropolitan is the only metropolitan in the whole of South Africa which shows high mortality for women.

Joe Gqabi, Alfred Nzo, OR Tambo, Ugu, Thabo Mafutsanyana, Umzinyathi, Zululand and Umkhanyakude are among the top most disadvantaged rural districts in 2001 (Ntuli, Suleman, Barron et al., 2002), which is consistent with the results of this study as we notice high mortality in those districts. Although Central Karoo in Western

Cape and districts in Limpopo bordering Gauteng are also listed under the most disadvantaged district but there is no indication of high mortality in those areas, perhaps due to lower HIV prevalence in those provinces.

Although Mpumalanga has relatively light mortality in all the districts within it for both men and women, from the high prevalence in the province in 2001 we would have expected adult mortality to be high in some parts of the province. This is also expected for North West. However, these provinces lie close to the border of other countries and are mainly involved in mining and agricultural activities, which are likely to attract more migrant labour, whom once they fall sick may return to their countries of origin.

5.3 Limitations

Data quality of registered deaths was one of the limitations of this study. We were forced to estimate the registered deaths for each population group in the early years of the intercensal period because a high proportion of the registered deaths were not specified by population group. In producing the estimate population growth rates were used rather than growth rates obtained from deaths. Because population growth rates are smaller than death growth rates when back-projecting we may have overstated the registered deaths in the earlier years by not considering the rapid increase in AIDS deaths over the intercensal period but rather a gradual increase reflected by population growth rates. This has a direct implication of a higher estimate of completeness in the intercensal period than should be in reality, which will have the same impact on the estimate of completeness in the year prior to the 2001 census since the estimate of completeness in the year prior to the census is dependent on the intercensal estimate.

In this study, the intercensal completeness is assumed to be equivalent to completeness in the middle year, implying that the assumption that completeness increases linearly whilst old age mortality is constant is entirely dependent on the deaths at the old age being complete in the middle period. If the old-age mortality rate is biased this may affect the estimate of completeness in the final years.

Although the numbers of deaths reported by households in the census were corrected for completeness, the quality of the population counts at district and provincial level were not assessed. Prior to the release of census data, results from the Post Enumeration Survey are used to adjust for under- (over-) count in the census, however, it is not certain whether the population counts used in estimating adult mortality are accurate since this was not assessed in this study at the sub-national level. This has increased significance at the district level because of the small numbers.

Because death distribution methods cannot be used at the sub-national level to measure completeness due to the difficulty of accounting for inter-regional migration, to assess the reasonability of the corrected deaths at the provincial level, the orphanhood method was used as a measure of comparison to the estimates of ${}_{45}q_{15}$ obtained from each of the provinces. The orphanhood method at the provincial level is not the gold standard. For example, the mortality represented by the method may not be entirely representative for that province as it is not guaranteed that respondents will stay in the same province as their parents.

5.4 Conclusions

Based on the comparison of provincial estimates obtained from the orphanhood method with the provincial estimates obtained from this study, there is insufficient evidence to validate a conclusion on the reasonableness of the method. This is because the orphanhood estimate is not based on the location of the parent, but rather of the respondent, and thus may be an estimate representative of mortality in another provinces. Despite this the comparison of the two estimates indicates strength in the method as most provinces appear to be ranked consistently although the level of mortality is different between the two methods.

On the other hand, the method proposed by Dorrington, Moultrie and Timæus (2004) produces adult mortality rates at district level that rank consistently with HIV prevalence rates at district level, especially in districts with extreme levels of both adult mortality and HIV prevalence. Moreover, when the district estimates of ${}_{45}q_{15}$ are weighted to the provincial level, they are consistent with those obtained by Dorrington, Moultrie and Timæus (2004) and other studies. The method works well estimating mortality at sub-national levels, although the estimates are marginally higher. We suspect that the marginal differences may have arisen from boundary differences in population counts from the 2001 census. This study used 2011 census boundaries whereas Dorrington, Moultrie and Timæus used the 2001 census boundaries.

Overall, it is safe to say that the method produces reasonable results at the sub-provincial level after adjusting the deaths reported by households using data from registered deaths. Although most of the comparisons were at the provincial level due to limited estimates at the district level, the results strengthen the reasonability of the methodology.

On the contrary it is worth noting that the conclusion is limited in several aspects. Undoubtedly the use of adult mortality estimates from the orphanhood method at the provincial level for the purposes of comparison to the estimates obtained in this study should not be considered at face value as the orphanhood method has its own limitations. In addition, apart from death being a rare event, splitting the death counts in the districts by sex, population group and age group requires that the data quality not be compromised as the numbers are small. Because of editing of census data to account for the under- (over-) count in the population using the Post Enumeration Survey, the quality of data at lower levels of the population may be compromised.

Some of the limitations, however, are less severe and these include, among others, the household population group being the proxy for the population group of the reported household death. In addition, the national adjustment factors used may not be an adequate representation for households in all districts. For instance it may not be sufficient to use the same adjustment for a death reported in a rural district and one reported in a more urban district based only on the sex, population group and age group of the death. If the data were available from registered deaths, residency classification should also be included amongst the other factors (age, population group and sex) that justify adjusting the deaths reported by households in future studies. This is because residency classification is certainly one of the main reasons for differences in completeness of reporting deaths given in the literature, and with increased urbanisation population group divisions in completeness will be replaced predominantly by residency classification divisions. However, considering that this study took place only seven years into post-apartheid South Africa the differences in completeness of registration of deaths by population group still appear to be associated with residency divisions, for instance the majority of the African population residing in the rural areas. Thus excluding the residency classification may have limited impact in this study.

5.5 Areas of further research

Further research is needed to examine several areas. In this study, we were not able to investigate whether the old age mortality for those aged 60 and older used in estimating annual completeness is consistent with the level of old age mortality in the intercensal period, which would give insight on the level of completeness desired. Although we estimated old age mortality from correcting the deaths in the middle year, assuming completeness in the intercensal period applies to the middle year, if this assumption does not hold it may lead to biased estimates of old age mortality. Thus, the need of deriving the estimates of old age mortality independently.

As the most common knowledge we have for the districts relates to their degree of urbanisation, it would be useful to also investigate the implied completeness of death registration in each of the districts and assess whether the estimates are consistent with the degree of urbanisation.

In this study, there was no time to consider the correlation between adult mortality with proximity to health facilities, as well as socio-economic factors at the district level, which would allow us to directly identify the effect of other factors besides HIV prevalence on differences noted in adult mortality. This could help confirm other conditions that may have led to the estimated level of adult mortality at district level.

The lack of alternative estimates at the district level in the time before or after the year 2001 made it difficult to tell whether the level of mortality in the district was plausible. Doing a similar study would help in understanding the plausibility of the level of adult mortality at different time points and assessing the reasonableness of the method used to come up with the estimates. For example, by estimating district mortality in the year 2010, we can expect the adult mortality level in districts with high HIV prevalence in 2001 to reflect a lower level of adult mortality due to the introduction of ART.

Although we used HIV antenatal prevalence to assess the level of adult mortality in some provinces in this study, further investigation is needed to ascertain the extent to which this hypothesis explains adult deaths at district level. We would expect the distribution of deaths by age in the districts to be consistent with the ages where we normally expect HIV/AIDS deaths to occur.

REFERENCES

Actuarial Society of South Africa. 2011. *ASSA 2008 full model* [Online]. Available: http://www.actuarialsociety.org.za/Societyactivities/CommitteeActivities/DemographyEpidemiologyCommittee/Models.aspx [Accessed 20/01/2016].

Ahmed Saifuddin & Kenneth Hill. 2011. "Maternal mortality estimation at the subnational level: a model-based method with an application to Bangladesh" *Bulletin of the World Health Organization* **89** (1):12-21.

Bah Sulaiman 2016. *HIV/AIDS in the light of death registration data. In:* Zuberi, T., Sibanda, A. & Udjo, E. O. (eds.) *The Demography of South Africa.* Routledge:120-159.

Bennett N. G. & S. Horiuchi. 1981. "Estimating the Completeness of Death Registration in a Closed Population" *Population Index* **47** (2):202-221.

Bennett Neil G & Shiro Horiuchi. 1984. "Mortality estimation from registered deaths in less developed countries" *Demography* **21** (2):217-233.

Bhat PN Mari. 2002. "General growth balance method: a reformulation for populations open to migration" *Population Studies* **56** (1):23-34.

Bradshaw Debbie, Rob E Dorrington & Freddy Sitas. 1992. "The Level of Mortality in South-Africa in 1985 - What Does It Tell Us About Health" *South African Medical Journal* **82** (4):237-240.

Cutler David & Ellen Meara. "Changes in the Age Distribution of Mortality over the 20th Century". Perspectives on the Economics of Aging, 2001. National Bureau of Economic Research Available from: http://www.nber.org/papers/w8556 [Accessed 20/03/2016]

Department of Health. 2003. *National HIV and syphilis antenatal sero-prevelence survey in South Africa 2002.* Pretoria.

Dorrington R E, D Bradshaw, R Laubscher & N Nannan. 2015. *Rapid mortality surveillence report 2014.* Cape Town South Africa: South African Medical Research Council. Available from: http://www.mrc.ac.za/bod/reports.htm [Accessed 25/03/2016].

Dorrington Rob 2013a. *Estimation of migration from census data. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population Available: http://demographicestimation.iussp.org/content/estimation-migration-census-data [Accessed 25/01/2016].

Dorrington Rob, David Bourne, Debbie Bradshaw, Ria Laubscher & I M Timæus. 2001. *The impact of HIV/AIDS on adult mortality in South Africa.* Cape Town: Burden of Disease Research Unit, Medical Research Council. Available from: http://www.mrc.ac.za/bod/ [Accessed 20/01/2016].

Dorrington Rob, Debbie Bradshaw & Debbie Budlender 2002. *HIV/AIDS profile in the provinces of South Africa: Indicators for 2002*, Centre for Actuarial Research.

Dorrington Rob E 2013b. *The generalized growth balance method. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population Available: http://demographicestimation.iussp.org/content/generalized-growth-balance-method [Accessed 20/01/2016].

Dorrington Rob E 2013c. *Synthetic extinct generations method. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population: Available:

http://demographicestimation.iussp.org/content/synthetic-extinct-generations-methods [Accessed 20/01/2016].
Dorrington Rob E, Debbie Bradshaw & Trevor Wegner. 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*. Available from: www.mrc.ac.za/bod/ [Accessed 20/01/2016].
Dorrington Rob E & Ian M Timæus. 2015. "A method for estimating sub-national mortality rates by using incomplete civil registration to correct census data on recent deaths". Paper presented at the African Population Conference, Johannesburg, South Africa.
Dorrington Rob E, Ian M Timæus & Tom A Moultrie. 2008. "Death distribution methods for estimating adult mortality: Sensitivity analysis with simulated data errors, revisited". Paper presented at the Population Association of America New Orleans, LA.
Dorrington Rob E., Tom A. Moultrie & Ian M Timæus. 2004. *Estimation of mortality using the South African Census 2001 data*. Cape Town: University of Cape Town. Available from: http://unstats.un.org/unsd/censuskb20/KnowledgebaseArticle10546.aspx [Accessed 19/01/2016].
Dorrington Rob & Kenneth Hill 2013. *Introduction to migration analysis*. *In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population.
Dorrington Rob, Ian M Timæus, Tom Moultrie & Nadine Nannan. 2004. "Estimates of provincial fertility and mortality in South Africa, 1985—1996" *Southern African Journal of Demography*:25-57.
Hill K. 1987. "Estimating Census and Death Registration Completeness" *Asian and Pacific Population Forum* **1** (3):8-13: 23-24.
Hill Kenneth & Bernardo Queiroz. 2010. "Adjusting the general growth balance method for migration" *Revista Brasileira de Estudos de População* **27** (1):7-20.
Hill Kenneth, Danzhen You & Yoonjoung Choi. 2009. "Death distribution methods for estimating adult mortality: sensitivity analysis with simulated data errors" *Demographic Research* **21** (9):235-254.
Hosegood V. & I. M. Timæus. 2005. "The impact of adult mortality on the living arrangements of older people in rural South Africa" *Ageing and Society* **25**:431-444.
Hosegood Victoria, Anna-Maria Vanneste & Ian M Timæus. 2004. "Levels and causes of adult mortality in rural South Africa: the impact of AIDS" *AIDS* **18** (4):663-671.
Johnson L. F., R. E. Dorrington & A. P. Matthews. 2007. "An investigation into the extent of uncertainty surrounding estimates of the impact of HIV/AIDS in South Africa" *South African Journal of Science* **103** (3-4):135-140.
Joubert J., C. Rao, D. Bradshaw, R. E. Dorrington, T. Vos & A. D. Lopez. 2012. "Characteristics, availability and uses of vital registration and other mortality data sources in post-democracy South Africa" *Global Health Action* **5**:1-19.
Khoza Christine. 2009. *The extent of misclassification of causes of death: a case study of Emalahleni Municipality, Mpumalanga Province, South Africa*. PhD, Australian National University.
Khoza Christine 2015. *An Assessment of Quality of Death Registration Data for Mortality Estimation in South Africa*. *In:* Odimegwu, C. O. & Kekovole, J. (eds.) *Social Demography of South Africa: Advances and Emerging Issues*. New York: Routledge:34.

Mahapatra Prasanta, Kenji Shibuya, Alan D Lopez, Francesca Coullare, Francis C Notzon, Chalapati Rao & Simon Szreter. 2007. "Civil registration systems and vital statistics: successes and missed opportunities" *The Lancet* **370** (9599):1653-1663.

Marandu Simon H. 2011. *Full life tables for South Africa from vital registration data, 2006-2008.* MPhil in Demography, University of Cape Town.

Mathers Colin & Ties Boerma. 2010. "Mortality measurement matters: improving data collection and estimation methods for child and adult mortality" *PLoS Med* **7** (4):e1000265.

Moultrie T. A. & R. E. Dorrington. 2012. "Used for ill; used for good: a century of collecting data on race in South Africa" *Ethnic and Racial Studies* **35** (8):1447-1465.

Moultrie Tom A & Ian M Timæus 2013. *Introduction to model life tables. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population: Available: http://demographicestimation.iussp.org/content/introduction-model-life-tables.

Msemburi W, V Pillay-van Wyk, R E Dorrington, I Neethling, N Nannan, P Groenewald, R Laubscher, J Joubert, R Matzopoulos, E Nicol, B Nojilana, M Prinsloo, N Sithole, N Somdyala & D Bradshaw. 2014. *Second national burden of diseases study for South Africa: Cause of death profile for South Africa, 1997–2010.* Cape Town: South African Medical Research Council. Available from: http://www.mrc.ac.za/bod/SouthAfrica2010.pdf [Accessed 19/01/2016].

Murray C. J. L., B. D. Ferguson, A. D. Lopez, M. Guillot, J. A. Salomon & O. Ahmad. 2003. "Modified Logit Life Table System: Principles, Empirical Validation, and Application" *Population Studies* **57** (2):165-182.

Murray Christopher J L, Julie Knoll Rajaratnam, Jacob Marcus, Thomas Laakso & Alan D Lopez. 2010. "What can we conclude from death registration? Improved methods for evaluating completeness" *PLoS Med* **7** (4):1549-1676.

Negussie Taffa, Julie C Will, Stephane Bodika, Laura Packel, Diemo Motlapele, Ellen Stein, Thierry H Roels, Gail Kennedy & El-Halabi Shenaaz. 2009. "Validation of AIDS-related mortality in Botswana" *Journal of the International AIDS Society* **12** (1):24.

Ntuli Antoinette, Fatima Suleman, Peter Barron & David McCoy. 2002. *South African Review 2001.* Available from: http://www.hst.org.za/sahr/ [Accessed 28/10/2016].

Pillay-Van Wyk Victoria., Ria. Laubscher, William. Msemburi, Rob E. Dorrington, Pam Groenewall, Theo Vos, Richard Matzopoulos, Megan Prinsloo, Beatrice Nojilana, Nadine Nannan, Nontuthuzelo Somdyala, Nomfuneko Sithole, Ian Neethling, Edward Nicol, Anastasia Rossouw, Jané D. Joubert & Debbie Bradshaw 2014. *Second South African National Burden of Disease Study: Data cleaning, validation and SA NBD List.* Cape Town: South African Medical Research Council. Available from: http://www.mrc.ac.za/bod/reports.htm [Accessed 19/01/2016].

Preston Samuel, Ansley J Coale, James Trussell & Maxine Weinstein. 1980. "Estimating the completeness of reporting of adult deaths in populations that are approximately stable" *Population Index*:179-202.

Preston Samuel H, Patrick Heuveline & M Guillot 2001. *Demography: Measuring and Modelling Population Processes,* USA, Blackwell Publishing.

Shisana O, T Rehle, L C Simbayi, K Zuma, S Jooste, N Zungu, D Labadarios, D Onoya, A Davids, N Wabiri, J van Zyl, N Mbellle & S Ramlagan. 2014. *South African National HIV Prvelence, Incidence and Behaviour Survey*. Cape Town: Human Sciences Research Council.

Statistics South Africa. 1998. *The people of South Africa Population Census, 1996: Calculating the the undercount in Census '96*. Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/255/.../3941 [Accessed 23/02/2016].

Statistics South Africa. 2000. *South African life tables, 1985-1994 and 1996*. Pretoria South Africa: StatsSA.

Statistics South Africa. 2005. *Mortality and causes of death in South Africa, 1997-2003, Findings from death notification*. Pretoria South Africa: StatsSA. Available from: https://www.statssa.gov.za/publications/P03093/P03093.pdf [Accessed 20/01/2016].

Statistics South Africa. 2012a. *Census 2011*. Pretoria: Statistics South Africa. Available from: http://www.statssa.gov.za/publications/P03014/P030142011.pdf [Accessed 19/01/2016].

Statistics South Africa. 2012b. *Census 2011 Metadata*. Available from: http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Metadata.pdf [Accessed 19/01/2016].

Statistics South Africa. 2014. *Mortality and causes of death in South Africa, 2013: Findings from death notifications* Pretoria Available from: www.statssa.gov.za/publications/P03093/ [Accessed 20/01/2016].

Statistics South Africa. 2015. *Mid-year population estimates 2015*. Pretoria: Statistics South Africa. Available from: https://www.statssa.gov.sa/publications/P0302/P03022015.pdf [Accessed 13/4/2016].

Timæus Ian M. 1992. "Estimation of adult mortality from paternal orphanhood: a reassessment and a new approach" *Population Bulletin of the United Nations* **33** (2):47-63.

Timæus Ian M 2013. *Indirect estimation of adult mortality from orphanhood. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for demographic estimation*. Paris: International Union for the Scientific Study of Population Available: http://demographicestimation.iussp.org/content/indirect-estimation-orphanhood-multiple-inquiries.html.

Timæus Ian M, Rob Dorrington & Kenneth Hill 2013. *Introduction to adult mortality analysis. In:* Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. & Zaba, B. (eds.) *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population Available: http://demographicestimation.iussp.org/content/introduction-adult-mortality-analysis.

Timæus Ian M, Rob Dorrington, Debbie Bradshaw, Nadine Nannan & David Bourne. 2001. "Adult mortality in South Africa, 1980–2000: From apartheid to AIDS". Paper presented at the annual meeting of the Population Association of America, March, New York. 31.

Timæus Ian M & Momodou Jasseh. 2004. "Adult Mortality in Sub-Saharan Africa: Evidence from Demographic and Health Surveys" *Demography* **41** (4):757-772.

Udjo Eric O. 2001. "A Four "Race" Model of Mortality in South Africa". Fourth African Census Analysis Project (ACAP) Workshop, Dakar, Senegal, January.

Udjo Eric O. 2006. "Estimation of mortality from vital registrations in South Africa" *Current HIV Research* **4** (4):469-474.

Udjo Eric O 2016. *An examination of recent census and survey data on mortality within the context of HIV/AIDS*. *In:* Udjo, E. O., Zuberi, T., Sibanda, A. & Udjo, E. (eds.) *The Demography of South Africa.* Routledge:90-119.
Udjo Eric O. & Pinky Lalthapersad-Pillay. 2014. "Estimating maternal mortality and causes in South Africa: National and provincial levels" *Midwifery* **30** (5):512-518.
United Nations Population Division. 2002. *Methods of estimating adult mortality*. New York: . Available from: www.un.org/esa/population/publications/adultmort/Complete.pdf [Accessed 19/01/2016].
Vincent Paul. 1951. "La mortalité des vieillards" *Population (french edition)*:181-204.
World Health Organisation. 2006. *Demographic surveillance sites and emerging challenges in international health* [Online]. Available: http://www.who.int/bulletin/volumes/84/3/editorial20306html/en/ [Accessed 29/02/2016].

APPENDICES

Appendix A: Incomplete provincial life table standard for the year 2001 derived from the ASSA 2008 projection model by sex

	Western Cape	*Northern Cape*	*Eastern Cape*	*Free State*	*Kwa-Zulu Natal*	*Mpumalanga*	*Limpopo*	*North West*	*Gauteng*
Male									
5	10000	10000	10000	10000	10000	10000	10000	10000	10000
10	9975	9970	9961	9960	9958	9963	9971	9969	9971
15	9947	9936	9915	9919	9916	9926	9941	9938	9942
20	9871	9838	9791	9807	9799	9825	9858	9849	9861
25	9698	9623	9522	9548	9529	9593	9677	9642	9666
30	9413	9277	9110	9119	9055	9192	9396	9276	9305
35	9057	8847	8630	8593	8460	8684	9049	8803	8826
40	8656	8377	8111	8015	7817	8125	8657	8274	8293
45	8210	7869	7569	7423	7179	7549	8219	7726	7738
50	7685	7285	6979	6806	6539	6947	7725	7147	7149
55	7036	6589	6304	6132	5860	6286	7126	6498	6486
60	6268	5795	5558	5401	5138	5561	6414	5771	5748
65	5370	4898	4731	4603	4359	4759	5568	4950	4916
70	4309	3876	3811	3717	3506	3857	4555	4009	3956
75	3158	2811	2866	2801	2634	2904	3427	3001	2925
80	2077	1855	2015	1968	1861	2016	2328	2048	1958
85	1241	1144	1368	1330	1290	1326	1432	1294	1205
Female									
5	10000	10000	10000	10000	10000	10000	10000	10000	10000
10	9982	9979	9973	9972	9970	9973	9979	9978	9979
15	9965	9957	9945	9946	9944	9950	9960	9958	9960
20	9901	9880	9845	9853	9844	9864	9898	9883	9891
25	9741	9693	9611	9606	9548	9614	9732	9665	9679
30	9507	9425	9280	9222	9087	9217	9462	9323	9342
35	9258	9150	8940	8829	8632	8816	9185	8979	8998
40	8995	8862	8608	8457	8210	8435	8906	8649	8662
45	8715	8563	8284	8116	7842	8091	8634	8341	8344
50	8379	8221	7930	7770	7488	7750	8353	8018	8014
55	7953	7798	7513	7375	7094	7364	8016	7637	7626
60	7391	7251	6993	6882	6607	6881	7557	7147	7124
65	6625	6516	6318	6234	5970	6246	6953	6492	6446
70	5663	5595	5494	5431	5187	5454	6163	5662	5586
75	4535	4519	4545	4490	4281	4518	5175	4666	4557
80	3291	3345	3506	3455	3291	3476	4022	3542	3408
85	2059	2189	2463	2409	2305	2408	2799	2393	2252