Evaluating the plausibility of the method of using both the civil registration and census data in estimating adult mortality at district level in South Africa, circa 2011

Nicola Marozva
University of Cape Town

Thesis submitted for the degree of Master of Philosophy in Demography
to the Faculty of Commerce

February 2018

ABSTRACT

The challenge in estimating mortality, both at national and sub-national levels, in developing countries such as South Africa is that neither of the death data sources (vital registration and census) are one hundred percent complete, that is, vital registration data is prone to incompleteness and deaths reported by household are subject to over- or under-reporting which may vary by age. Also, apart from issues with data sources, there is no method that estimates mortality accurately at subnational level and the methods for estimating the level of completeness of reporting of deaths cannot be applied at sub-national level (due to issues with migration). Thus, measuring mortality rates at sub-national level is a challenge.

This research seeks to employ a method used by Dorrington, Moultrie and Timæus (2004) that makes use of both data sources in combination so as to overcome the weakness and makes use of the strength of each data source.

To estimate the level of completeness in the year prior to the 2011 Census (to correct the number of deaths registered), first, the Death Distribution Methods (Synthetic Extinct Generations +delta and General Growth Balance method) are used to estimate the level of completeness of the vital registration deaths for the intercensal period 2001-2011 by population group. Thereafter, the level of completeness for each of the years in the intercensal period is estimated by fitting a logistic curve to the level of completeness for the intercensal period of 1996-2001 and 2001-2011 (derived by both Chinogurei (2017) and Richman (2017)). Thus, the number of deaths registered in the year prior to the 2011 census are then corrected for either under- or over-reporting using the estimates of completeness to obtain the true number of deaths by population group and age group for each sex.

The corrected true numbers of registered deaths are then used to determine the age-specific correction factors by population group for correcting the household reported deaths at district level and thereafter estimates of mortality at district level are determined.

Comparison of estimates derived in this study to estimates by other studies indicated that the method produces plausible estimates at district level, thus, findings in this research strengthens the reasonability of the method.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to express my deep and sincere gratitude to Professor Rob Dorrington for his guidance, tireless efforts, immense knowledge and the time spent supervising and supporting this research with invaluable feedback as well as for his patience and encouragement at times when challenges were hindering my progress.

My deepest gratitude also goes to my angel mother, Angeline Ndhlovu Marozva for her guidance, love and support throughout my life. All that I am or ever hope to be, I owe to her.

I wish to thank the Andrew Mellon Foundation for without their support, my dream to pursue this MPhil would not have been possible.

Above all, I would like to thank my Lord, my God for taking me this far. Indeed, He is Jehovah-Jireh.

1 INTRODUCTION

1.1 Background

South Africa is one of the countries in the world that has been most affected by HIV/AIDS. Adult mortality increased sharply from 2000 to 2005/6 due to the HIV/AIDS epidemic, but this has improved dramatically due to the introduction of ART programme (deaths due to HIV/AIDS sharply increased from 1997 to 2006 and steadily declined thereafter whereas deaths from non-communicable diseases remained roughly constant from 1997 to 2010) (Dorrington, Bradshaw, Laubscher *et al.* 2016; Pillay-van Wyk, Dorrington and Bradshaw 2017). The ART programme has seen the country taking measures (such as prevention of mother-to-child transmission, care and support and ARV treatment for those infected) in mitigating the disease. Following this, it is of great importance to understand the burden of disease in populations through measuring the level and trend of mortality in a nation (as these estimates are a cornerstone of public health) (Mathers and Boerma 2010).

Due to the fact that there are disparities within geographical areas in the same country, for example, substantial discrepancies between urban and rural areas, it is desirable to estimate mortality at a sub-provincial level as it helps in better allocation of resources, identifying areas that should be prioritised and the evaluation of effectiveness of intervention programmes as well as decision making. However, the major challenge in estimating mortality at sub-provincial level in developing countries is that the data required are usually not readily available (Joubert, Rao, Bradshaw *et al.* 2012) and where it is available, the sources of data, that is, the census and vital registration data are incomplete and sometimes inaccurate, and as a result, information from these sources without adjustments is of limited use in estimating mortality (Dorrington 2013c).

Vital registration data are prone to three major generic problems, namely, coverage, incompleteness, omission of events (Joubert, Rao, Bradshaw *et al.* 2012). Though Death Distribution methods are used to estimate the level of completeness of reporting (as estimating mortality using incomplete vital registration death data leads to lighter mortality), the vital registration at sub-national level might not accurately state the region of permanent residence prior to death, as highlighted by Dorrington, Moultrie and Timæus (2004), which makes it difficult to estimate mortality at the sub-national level. In South Africa, completeness of registration of deaths has improved over the past years, however, use of the vital registration death data to estimate mortality at both sub-national and sub-provincial level is still a challenge due to the reason mentioned previously.

On the other hand, census data also suffers from its own inherent problems. Dorrington, Timaeus and Gregson (2006) point out that deaths reported by households suffer from problems associated with under-reporting, over-reporting, age misreporting, recall bias and reference period errors. Under-reporting may occur in instances where a household collapses after the death of a breadwinner or in elderly households where there might be no one left to report the death, whereas over-reporting may occur in instances where deaths of young adults who migrate from the areas they grew up in to urban areas to look for work are reported more than once as they may be regarded as members of more than one household (Dorrington 2013b). Also, apart from these errors, Bogue (1965) highlights that census data are often inaccurate, as respondents might choose not to disclose certain information while enumerators might not allow respondents sufficient time to answer all the census questions.

In developed countries, mortality rates can be estimated directly using either deaths captured by the country's vital registration system (since they have fully functioning vital registration systems) or household deaths reported in census data which are often accurate in the information provided, which makes estimation of mortality a trivial exercise (Setel, Macfarlane, Szreter *et al.* 2007). However, in developing countries, indirect methods are often used to estimate mortality due to the lack of accurate and complete data sources. These methods include death distribution methods, sibling history and orphanhood method. Use of these methods at sub-national level is problematic as neither of the data sources are hundred percent complete and they contain missing information. Applying orphanhood method and sibling history method at sub-provincial level is problematic in the sense that the sample size at sub-provincial level is too small to produce reliable estimates. In addition, parents and siblings may not have lived where the children are surveyed, thus the mortality experience may not be consistent with the location.

In South Africa, several studies have been carried out that estimate adult mortality at sub-national level. Most of these have used both data sources in combination to overcome the weakness and makes use of the strength of each data source. These include studies by Dorrington, Moultrie and Timæus (2004); Dorrington, Timæus, Moultrie *et al.* (2004); Udjo and Lalthapersad-Pillay (2014) and Dorrington and Timæus

(2015; 2017). In addition, a mathematical model (CARe_3.2 provincial demographic model) has been developed that projects mortality at a provincial level. However, there has been only one study, by Chinogurei (2017), focusing on estimating adult mortality at a sub-provincial level, possibly due to the fact that there is no method that can estimate mortality accurately at sub-provincial level. Chinogurei (2017) investigated if the method used by Dorrington, Moultrie and Timæus (2004) to estimate mortality at a provincial level could be used to produce reasonable estimates at district level in the year prior to the 2001 census. This study seeks to update and strengthen the research findings of the study carried out by Chinogurei (2017) through investigating if the same method produces reasonable estimates at district level in South Africa for a different time period (that is, the year prior to 2011 census).

1.2 Research Question

Does the method proposed by Dorrington and Timæus (2015) produce plausible adult mortality estimates at district level in South Africa circa 2011?

1.3 Objectives

The main objective of this study is to assess whether the method devised by Dorrington and Timæus (2015) produces plausible adult mortality estimates for the year prior to the 2011 census at district level. To estimate adult mortality at district level the following specific objectives must be achieved:

1. Assess and evaluate the quality of data to be used in estimating mortality in South Africa;
2. Determine the level and extent of intercensal completeness of vital registration deaths (national and by population group for each sex) in the intercensal period 2001-2011;
3. Estimate the trend in completeness of registration of deaths over time so as to determine completeness in the year prior to the 2011 census by each sex and population group;
4. Estimate age-sex specific household correction factors by population group and apply them to the number of deaths reported by households in each district to produce estimates of the correct number of deaths reported by households in each district;

5. Estimate adult mortality rates (${}_{45}q_{15}$) at district level (directly and thereafter graduate the rates obtained) for each sex using the true number of deaths, that is deaths corrected for either under- or over-reporting

1.4 Importance of the study

This research aims to investigate whether the method proposed by Dorrington and Timæus (2015) produces reasonable estimates at a district level. If the method works, it can be adopted to estimate mortality at district level at different time points, which is useful since much research conducted so far on mortality focused only on estimating mortality at national or province level, or certain geographical units. Accurate mortality estimates at lower geographical areas are essential for evaluating development programmes, effectiveness of intervention programmes, as well as decision-making.

However, if the method does not produce plausible estimates, weaknesses and limitations of the method (through interrogation of the method and data used) will be identified so that they can be addressed in future studies that will employ a similar method.

1.5 Summary of the study

The rest of this study is divided into four chapters. Chapter 2 reviews the literature on data sources for estimating adult mortality in South Africa (as well as errors that these data sources are prone to) and desirable methods for estimating adult mortality (particularly the ones that would be applied in this study). It also presents the literature on past studies that have been carried out in estimating mortality as well as the level and trend of adult mortality in South Africa. Chapter 3 gives a detailed description of the method (devised by Dorrington and Timæus (2015)) employed in the study. Chapter 4 reviews the estimates of adult mortality derived at district level in South Africa and Chapter 5 discusses the reasonableness of the results obtained and presents the overall conclusion of the study.

2 LITERATURE REVIEW

This chapter consists of four sections. The first section outlines data sources for estimating mortality in South Africa as well as errors to which these data sources are prone. Section 2.2 focuses on past studies that have been carried out in estimating mortality in South Africa. The third section then covers analysis of the level and trend of adult mortality in South Africa at both national and subnational level and the final section discusses methods of estimating adult mortality, particularly the ones that are applied in this study.

2.1 Data sources for estimating mortality in South Africa

The best and most reliable source of data for estimating the level of mortality is a sufficiently complete registration system that also records the age, sex and cause of death of the deceased (Gakidou, Hogan and Lopez 2004; Mathers and Boerma 2010; Setel, Macfarlane, Szreter *et al.* 2007; United Nations 2014). However, in developing countries this is not the case as vital registration data suffers from incompleteness. Joubert, Rao, Bradshaw *et al.* (2012) highlight that only two per cent of developing countries have death registration systems that are complete, and most of these countries do not document the cause of death. Mathers, Ma Fat, Inoue *et al.* (2005) carried out a study, which assessed death registration completeness and quality for 115 countries of the World Health Organisation (WHO). They noted that the level of completeness of registered deaths varied from 100 per cent (in several European countries) to less than 10 per cent (in some African countries). Phillips, Lozano, Naghavi *et al.* (2014) carried out a similar study whereby they applied a composite index to access the performance of VR systems for computing reliable mortality estimates in 148 countries for the period 1980 - 2012. They also concluded that the best performing systems were in European countries, whereas systems in African countries did not perform well with only a few countries showing improvement in the civil registration system over the past 50 years. Thus, mortality estimates generated from civil registration remain problematic in developing countries due to lack of confidence in the level of completeness of registered deaths (Murray, Rajaratnam, Marcus *et al.* 2010).

In South Africa, the death registration is administered by the Department of Home Affairs (DHA), and is guided by the Births and Deaths Registration Act (Act No. 51 of 1992) which stipulates that the registration of deaths is a legal requirement (Stats

SA 2014). Whenever a death occurs it should be registered either at the local Home Affairs office (if it occurred at a hospital) or at both the local municipal office and Municipal Home Affairs office (if it occurred at home) by relatives of the deceased, and an abridged death certificate is issued upon registration. From the death notification forms, the National Department of Home Affairs office updates the National Population Register, which captures only deaths of those with a South African identity number or registered birth.

Statistics South Africa (Stats SA) receives all the death notification forms (that is, for both non-South African citizens and South African citizens) from the National Department of Home Affairs. It then captures manually both socio-demographic and health data as well as causes of deaths from these death notification forms (Stats SA 2012).

In the 1990s the South African registration system suffered from problems of timeliness. It has significantly improved since 2004 (Khoza 2015), with only approximately 15 per cent of the death notification forms (from the deaths that had occurred during the year) being captured annually by Stats SA during the AIDS epidemic era when more up-to date mortality statistics were needed (Joubert, Rao, Bradshaw *et al.* 2012). This saw the South African Medical Research Council (SA MRC) in collaboration with University of Cape Town, establishing the Rapid Mortality Surveillance (RMS) system, which monitors mortality data received electronically from the DHA. This system is similar to the Stats SA approach with the difference being that Stats SA captures all deaths that are notified to DHA, whereas the SA MRC captures only deaths of individuals who are recorded on the National Population Register (NPR) by DHA, that is, deaths of those individuals with South African birth certificates or identity numbers (Joubert, Rao, Bradshaw *et al.* 2012). Also, the SA MRC receives regular monthly updates, thus allowing it to produce mortality estimates at least a year ahead of those produced by Stats SA (who, also have to capture the causes of death from the death notification forms).

Bah (1998b) argues that the political history, governance and the civil society, in general, influences the quality and nature of vital statistics in a country. This argument is borne out by South Africa's political history, as it has contributed immensely to the incompleteness of the vital registration system. From 1867 onwards, certain laws governing both birth and death registrations were enacted at a subnational level. In 1923, the Births, Deaths and Marriages Registration Act was enacted and under this Act the registration of rural black Africans was not obligatory (Bah 1998b; Joubert, Rao, Bradshaw *et al.* 2012). Kok (2003) notes that this meant that approximately 86 per cent of black Africans had no legal requirement to register their deaths, which led to the vital registration being substantially incomplete. However, during the 1990s, there were changes that led to an improvement of the civil registration system. In 1992, the Birth, Deaths and Registration Act of 1992, which stated that registration was to be compulsory in South Africa, replaced the previous Act, although information on the population group (race) of the deceased was no longer captured. In 1998, a revision of the 1992 Act was made to reintroduce the collection of information on population group for deaths. The 1992 and 1998 Acts led to an improvement in the statistics, between 1992 and 2000. For the period 1996 to 2000, the completeness of death registration improved from 85 percent to 90 per cent (Dorrington, Bourne, Bradshaw *et al.* 2001).

Even though improvements have been noted so far, the vital registration still has its weaknesses and it suffers from incompleteness, as deaths are not fully registered. The death notification forms are often found to be inaccurate and incomplete as a result of omissions and content errors (Khoza 2015; Stats SA 2005). Thus, for example, medical practitioners might not report accurately the required information, or the death might not be reported at all, leading to inaccuracy of the information received from the death notification forms. Also, the death notification forms often misreport the place of residence as the place where the deaths were registered (Dorrington, Moultrie and Timæus 2004), which makes it a challenge to estimate mortality at sub-national level. However, Stats SA (2014) indicates that the death notification forms now include place of permanent residence of the decedent, which hopefully, will improve the estimation of mortality at subnational level.

Other available sources, apart from the vital registration system, that provide data on mortality include censuses and surveys. In the absence of a sufficiently complete registration system in developing countries like South Africa, surveys and censuses are the major sources of information on mortality (Mathers and Boerma 2010; Udjo 2016). Gakidou, Hogan and Lopez (2004:715) point out that, "household surveys provide the only way, at present, to get timely information on the impact of health problems, such as HIV/AIDS in countries without elaborate monitoring systems". They argue that in most countries the methods available for estimating the level of completeness might be subjective and thus require considerable judgment. Timæus (1991) also highlights that owing to ineffective systems for data collection on deaths, most developing countries have resorted to surveys and *ad hoc* inquires.

In South Africa, existing household surveys include censuses, demographic and health surveys, general household surveys and the National Income Dynamics Study. Surveys like censuses give the number of deaths that have occurred in the household in the year prior to the survey date. Thus, for example, in the 2011 census questionnaire, respondents were asked if in the household there were any deaths in the past 12 months, that is, from 10 October 2010 to 9 October 2011. Also, surveys gather useful information that can be utilised to estimate mortality indirectly, for instance, responses to questions relating to the death of a family member (either siblings or parents) could be used in estimating adult mortality through the siblings' history method or the orphanhood method (Joubert, Rao, Bradshaw *et al.* 2012).

However, the data on deaths reported by households in censuses or surveys are prone to errors such as, under- or over-reporting, reference period errors (dating errors) and age misreporting (Timæus 2013b). Under- or over-reporting errors are largely because identifying a suitable respondent in a household who can provide reliable information about adult deaths may be difficult. Also, young adults who leave their usual place of residence to work elsewhere might be reported more than once or not at all thereby leading to under- or multiple-reporting (Dorrington 2013b). Apart from that, households might disintegrate and in cases of single-person households, deaths of such individuals will be unreported. Timæus (1991) attributes underreporting of deaths to either enumerators laxity, that is, excluding the questions that ask about household deaths and leaving the question blank, or omissions that might be due to respondent's unwillingness to talk about it.

To improve measurement of mortality in countries where vital registration suffers from incompleteness, censuses should be carried out at regular time intervals with appropriate mortality questions and demographic and health surveys should be conducted regularly, asking about full birth history as well as sibling survival histories (Mathers and Boerma 2010). Also, Timæus (1991) suggests that improvements can be achieved if there is better questionnaire design, field procedures, training of interviewers as well as multi-round surveys. He further highlights that multi-round surveys and surveillance systems reduce problems encountered in a single-round survey, as multi-round surveys are likely to eliminate omissions and reference period errors. This is because inquires on the second visit can be made of a different person from the one found on the first visit, and also enumerators would be able to identify if there are missing households (in case of disintegration of the household) from the ones that were there at the start of the survey.

2.2 History of South African Life Tables

Official life tables and mortality estimates can be traced as far back as 1920s[1]. These life tables were published as South African Life Tables (SALT) for each and every three year period around each census that occurred between 1920 and 1986, but these were only available for the Whites, Coloured and Indians (Bah 1998a; Dorrington, Bradshaw and Wegner 1999). Dorrington, Moultrie and Timæus (2004) note that until 1979, black Africans were not eligible to be included as part of the vital registration system and even after that people who lived in Transkei, Bophuthatswana, Venda and Ciskei were excluded as they were deemed to be citizens of "independent homelands", that is, they were not considered as South Africa citizens. This, as a result, made it difficult for demographers to produce life tables for the population as a whole over that period.

Van Eeden and Van Tonder (1975) produced abridged life tables for all the population groups except black Africans from the life tables computed by the Department of Statistics (which today is known as Statistics South Africa) for the period 1921-1970.

Dorrington, Moultrie and Timæus (2004) record that the last set of SALTs to be produced were for the years 1984 to 1986 for Whites, Coloureds and Indians, since, between 1991 and 1997, the death certificates did not capture information on a decedent's race. As a result, no life tables were produced for this period, however, Stats SA produced abridged life tables by population group for the period 1985 to 1994, using data from the 1996 census on the survival of parents as well as children ever born (Dorrington, Timæus, Moultrie *et al.* 2004). Stats SA also produced abridged life tables for 1996 for the provinces using registered deaths and the 1996 census population, and the Brass Growth Balance method was applied to the whole age range to adjust for incompleteness of registration (Stats SA 2000). However, application of the Brass Growth Balance method to estimate the level of completeness for all ages is not ideal as the population is not stable and completeness of reporting of deaths tends to differ for children compared to adults. This violates two of the fundamental assumptions of the

[1] The life tables of the Indians and Coloured population only started later, that is 1945 and 1936 respectively.

method. Dorrington, Timæus, Moultrie *et al.* (2004) further point out that the level of completeness produced was lower than the one produced by Dorrington, Bourne, Bradshaw *et al.* (2001), which leads to lack of confidence in the estimates produced.

Dorrington, Bradshaw and Wegner (1999) were the first to derive life tables for black Africans using registered deaths. They applied indirect estimation techniques to registered deaths for the period of 1984 to 1986 and the mid-1985 population estimates. A weighted average of the black Africans life tables combined with the official SALT for the Whites, Coloured and Indians was used to produce a national life table.

Dorrington, Moultrie and Timæus (2004) also attempted to produce full life tables by province as well as by population group. To estimate adult mortality, they used population estimates from the 1996 and 2001 censuses and the registered deaths, adjusted for incompleteness using the Generalized Growth Balance method together with childhood mortality[2] estimated directly from household deaths and census counts for children under five years of age.

2.3 Adult mortality estimates in South Africa

2.3.1 Level and trend of adult mortality in South Africa at national level.

Mortality estimation in South Africa has been complicated, particularly before 1994, largely due to a lack of a fully functioning vital registration system under the Apartheid government of the time (Dorrington, Moultrie and Timæus 2004). The new government identified the problem with the mortality statistics (Dorrington, Bradshaw and Wegner 1999) and efforts were made to collect data through the use of censuses and surveys as well as improving the vital registration system. This saw Statistics South Africa (Stats SA) carrying out surveys such as the 1998 and 2003 Demographic Health Surveys, the 2001 and 2011 censuses, and the 2007 and the 2016 Community Surveys.

Bradshaw and Timæus (2006) classify the trends in mortality into three time periods, that is, before 1980, when mortality was improving each year, between 1980 and 1995 when mortality was stagnant and thereafter, the HIV/AIDS epidemic era when mortality was rising. Dorrington, Moultrie and Timæus (2004) carried out an analysis of trends in adult mortality as documented by different sources in South Africa and produced the trends as shown in Figure 2.1. The "official+"[3] series, for both males

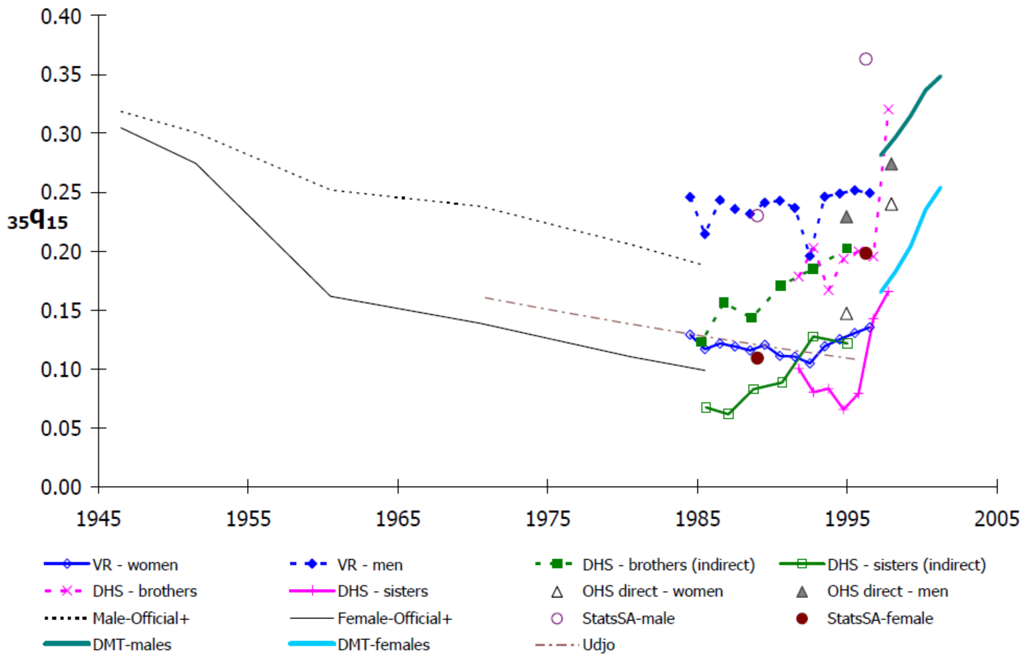
[2] The level of completeness of the registered deaths for the children under five years of age could not be assessed hence direct estimation of childhood mortality was done.

[3] The "official+" estimates for the period 1946 to 1985 were obtained by weighting the Whites, Coloureds, Indians estimates from SALT as well as black African estimates derived by Sadie (1988)

and female, shows that mortality declined sharply from 1945 to around 1960, particularly for females, as shown by the estimates of mortality ($_{35}q_{15}$) decreasing from 30 per cent to about 15 percent (thus, a 50 percent decline), and further declining to around ten per cent in 1985 for females.

The female mortality estimates are also consistent with the estimates of Udjo (1998) for the period 1970 to 1985, as both are declining at the same rate although the level is different.

Figure 2.1 Trends in the probability of dying for a person aged between 15 and 50 in South Africa for the period 1945-2005.



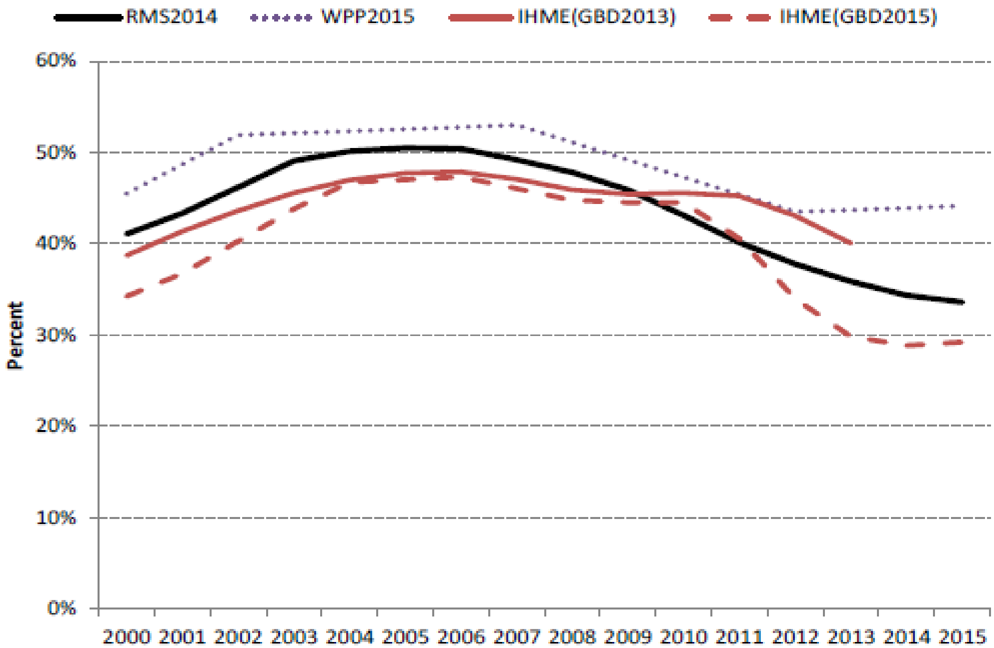
Source: Dorrington, Moultrie and Timæus (2004); Figure 1.1; page 65

Dorrington, Timæus, Moultrie *et al.* (2004) also observed that, during late 1980s to mid-1990s both estimates of male and female mortality from these various sources showed gradually decreasing or constant mortality, as confirmed by consistent estimates derived by Udjo (1998), Stats SA and vital registration data (after adjusting for incompleteness). However, estimates from the DHS data differ, probably indicating that estimates from DHS data are not reliable for adult mortality. This is also supported by UN Population Division (2009) as they estimated adult mortality rate ($_{35}q_{15}$) for females and males, for the period 1985-1990 to be around 12 per cent.

A rapid increase in both male and female mortality from the mid-1990s is also noticeable (in Figure 2.1), which is attributed to the HIV/AIDS epidemic by many authors (Bah 2016; Pillay-van Wyk, Laubscher, Msemburi *et al.* 2014; Timæus and Jasseh 2004; Udjo 2006). Dorrington, Moultrie and Timæus (2004) show that female mortality (${}_{35}q_{15}$) between 1995 and 2005 increased rapidly from around 17 per cent in 1997 to 26 per cent in 2003 and male mortality increased from 27 per cent in 1999 to 35 per cent in 2000. This shows that the mortality effects of the HIV pandemic in South Africa started around the mid-1990s, and Bradshaw and Timæus (2006) indicate that by 2001, 51 percent of the deaths were due to HIV/AIDS, which shows how severe the epidemic was.

Dorrington, Bradshaw, Laubscher *et al.* (2016) compared estimates for ${}_{45}q_{15}$ from the Rapid Mortality Surveillance (RMS) system, the World Population Prospects (WPP2015), and Institute of Health Metrics and Evaluation (IHME(GBD2013)) as shown in Figure 2.2. These estimates show that adult mortality reached its peak between 2006 and 2008, after which, mortality declined.

Figure 2.2 Combined mortality estimates (${}_{45}q_{15}$) for both men and women in South Africa for the period 2000-2015.



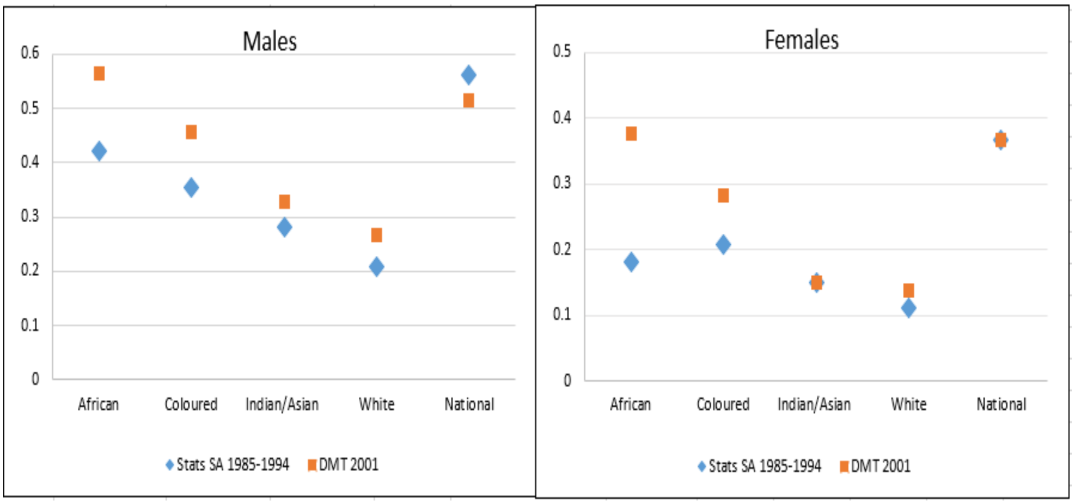
Source: Figure 24; page 20; Dorrington, Bradshaw, Laubscher *et al.* (2016)
Notes: WPP- World Population Prospects, HME(GBD)- Institute of Health Metrics (Global Burden of Diseases), RMS- Rapid Mortality Surveillance estimates

However, differences in trend were observed around 2008, with estimates from RMS and WPP showing a consistent declining trend, which is not the case with estimates from IHME. The IHME estimates exhibit an implausible trend, with no decline in mortality rates from 2008-2011 followed by a rapid decline, and this is not consistent with the fact that during that period, mortality should have been on the decline since starting from 2005 ARVs were extensively provided (Dorrington, Bradshaw, Laubscher *et al.* 2016).

Further analysis of the estimates by population group, in South Africa indicated that during the period 1980-1996, the White population had the lowest mortality followed closely by the Indian and then the Coloured population groups and the black Africans had the highest mortality Udjo (2016). However, estimates from Stats SA (2000), indicate an odd ranking for the female population, with the Coloured population having a slightly higher mortality than the African population, which shows weakness in estimates derived by Stats SA (2000).

Comparison of mortality estimates derived by Stats SA (2000) to those of Dorrington, Moultrie and Timæus (2004) by population group indicate that African and Coloured mortality sharply increased in contrast to that for the Indian/Asian (female estimates remained the same) and White population as shown in Figure 2.3.

Figure 2.3 Adult mortality rate ($_{45}q_{15}$) in South Africa in 1985-1994 and 2001



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

Richman (2017) estimated the trend and level of old age mortality, that is, from age 70 and above, for the period 1985-2011. He used the Death Distribution Methods to correct the registered deaths for incompleteness and the Near Extinct Generation (NEG) methods to estimate the population from the deaths, and hence old age mortality. However, he observed that mortality rates produced using the NEG methods were biased, and thus adapted the NEG method (NEG-GAM) to produce a generalised additive version of the model, which was then used to re-estimate the population by projecting future deaths of nearly extinct cohorts to produce less biased estimates of mortality.

Richman (2017) compared his estimates to other studies conducted in South Africa and the Human Mortality Database and found that the estimates were consistent for the White and Indian population groups, while the estimates for the Africans and the Coloured population groups increased too slowly with age at older ages. He attributed this to age exaggeration in the death data. Also, the population estimates of those aged 70+ obtained by summing the cohort deaths were consistent with 2011 census population estimates (Richman 2017). He concluded that the mortality rates of those aged 70-79 improved over time for the Indian, Coloured and White population groups and deteriorated slightly for Africans.

2.4 Mortality estimates at a sub-national level in South Africa

Estimation of mortality rates at district level is challenging in developing countries because:

- The vital registration system does not accurately capture the place of residence prior to death, as place of death is often mistaken as place of residence thus violating the principle of correspondence[4]. Thus, mortality estimates, calculated by dividing the observed number of deaths in a province by the population that is at risk of dying in that province would be biased. (Dorrington 2013b; Dorrington, Timæus, Moultrie *et al.* 2004).
- At the subnational level, inter-regional migration is both more significant and more difficult to measure, and thus, application of Death Distribution methods to correct reported deaths for incompleteness at subnational level might produce unsatisfactory results (Dorrington 2013a, b).

Dorrington, Timæus, Moultrie *et al.* (2004) estimated both provincial child and adult mortality for the period 1985 to 1996. Estimates of child mortality were derived using the 1996 census data as well as estimates from the 1998 Demographic and Health Survey, whereas adult mortality estimates were derived from the 1996 population estimates, registered deaths as well as completeness of reported child deaths. The level of completeness of reported child deaths in each province was obtained by dividing the

[4] The principle of correspondence states that the event should be included in the denominator if, had the event occurred, it would have been included in the numerator and vice versa

number of reported child deaths by the expected number of deaths[5]. Due to reasons outlined above (vital registration system's inaccurate capturing of place of residence prior to death and the complications caused by inter-regional migration), completeness of adult mortality was estimated based on the assumption that it is linearly related to completeness of reported child deaths (Dorrington, Timæus, Moultrie *et al.* 2004).

A constraint on the study by Dorrington, Timæus, Moultrie *et al.* (2004) was that the 1996 census did not include questions on deaths of household members in the year prior to the census, which were included in the subsequent censuses. As a result, Dorrington, Moultrie and Timæus (2004) used a different approach to estimate mortality at a provincial level for the year prior to the 2001 census. This method makes use of both the registered deaths and household deaths in combination, to overcome the weaknesses, and make use of the strengths, of each data source in estimating mortality at a sub national level.

Dorrington, Moultrie and Timæus (2004) used Death Distribution methods to estimate the level of completeness at the national level by sex and population group, so as to correct the vital registration data for underreporting. However, the level of completeness obtained is the average for the whole intercensal period (1996 to 2001 census). Thus to estimate completeness in the year prior to 2001 census, a trend was fitted to the level of completeness based on the assumption that the level of completeness changed linearly over time such that the trend in the mortality rate of those aged 65 and above over time had a slope of zero (Dorrington, Moultrie and Timæus 2004).

The level of completeness for the year prior to the 2001 census was then used to estimate the true number of reported deaths for each population group and the population as a whole for each sex. It was observed that the deaths corrected for completeness for the population as a whole did not match the sum of the number of deaths corrected for completeness by population group. Thus, the corrected total numbers of deaths by population group were adjusted to ensure consistency between the two estimates.

The age-sex specific adjustment factors (denoting the level of either under- or over- reporting of the number of deaths reported by households) for each sex and population group were produced by dividing the corrected registered deaths in the year

[5] The expected number of deaths were obtained by subtracting the number of children still alive from the expected total number of children born. (Age specific fertility rates were used to calculate the expected total number of children born)

prior to the 2001 census by reported household deaths. The estimated adjustment factors were then applied to the number of reported household deaths in each province based on the assumption that the age-sex specific adjustment factors for each population group were the same for all provinces. Mortality rates were then derived by dividing the corrected reported household deaths by population estimates from the 2001 census and the reasonableness was checked by comparing the mortality estimates to estimates derived from the orphanhood data.

Similarly, Msemburi, Pillay-van Wyk, Dorrington *et al.* (2014) produced mortality rates at a provincial level. To estimate completeness of registered deaths for the period between the 2001 census and 2007 Community Survey, the same approach to that used by Dorrington, Moultrie and Timæus (2004) was used. However, to obtain the level of completeness of registered deaths for each of the years between 2001 and 2011 censuses, a logistic curve was fitted to the level of completeness obtained for the mid-intercensal periods of the 1996-2001 census, the 2001 census and the 2007 Community Survey, and the period between the 2001 and 2011 censuses (Pillay-van Wyk, Laubscher, Msemburi *et al.* 2014). A key assumption was that the level of completeness by 2009 and beyond remained level at approximately 93 per cent.

Dorrington and Timæus (2015; 2017) also estimated adult mortality at a provincial level for the years prior to the 2001 and 2011 census respectively. They used the same procedure to that used by Dorrington, Moultrie and Timæus (2004), though a trend in completeness over time was fitted using the logistic curve.

Chinogurei (2017) assessed whether the method employed by Dorrington, Moultrie and Timæus (2004) and Dorrington and Timæus (2015; 2017) at provincial level could be employed at district level. He used the same approach to estimate mortality rates at district level assuming that the errors in the reporting of the numbers of deaths at each age for each population group are the same for all districts.

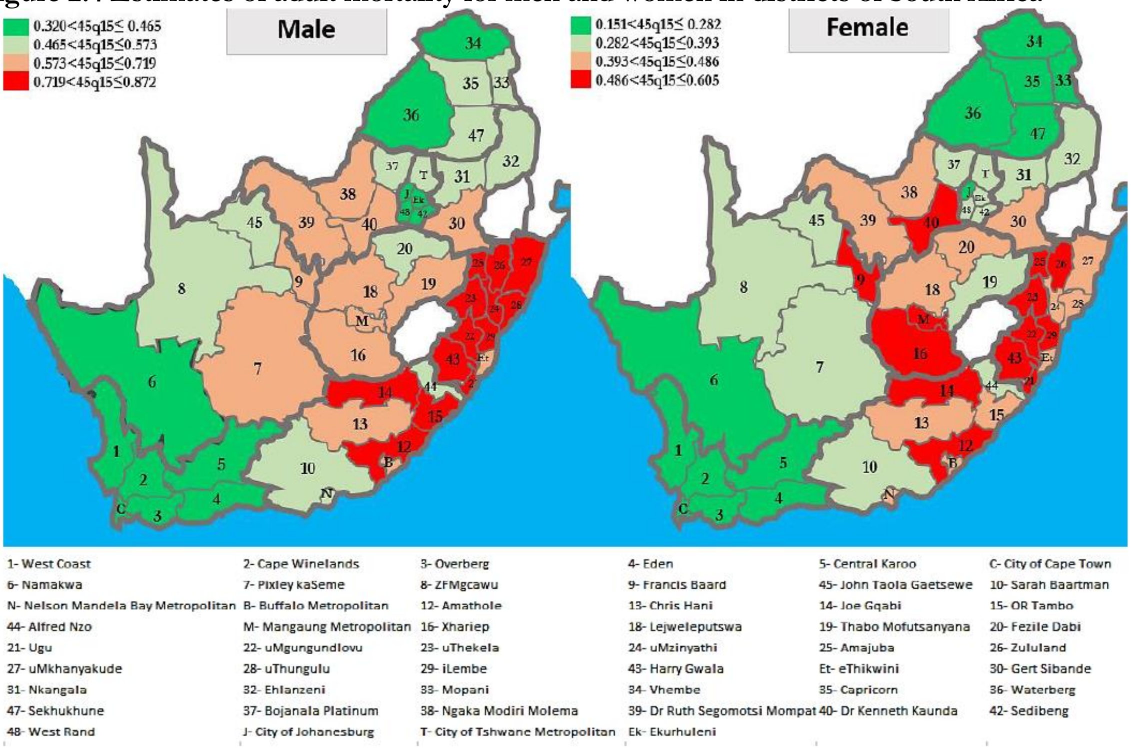
He assessed the reasonableness of the estimates derived by checking if the average estimates of adult mortality for the province (produced by weighting together district estimates of adult mortality within each province) agreed with estimates for the province produced by Dorrington, Moultrie and Timæus (2004), Msemburi, Pillay-van Wyk, Dorrington *et al.* (2014), orphanhood estimates and the CARe 2011 model[6]. The weighted district level estimates to provincial level mortality estimates were found to be

[6] Estimates from the CARe2011 model and the orphanhood data were derived by Chinogurei (2017) to obtain alternative estimates of adult mortality at provincial level

consistent (although they were marginally higher) with estimates derived by Dorrington, Moultrie and Timæus (2004) and Msemburi, Pillay-van Wyk, Dorrington *et al.* (2014). Estimates from the CARe2011 model and the orphanhood data were lower than his estimates although the ranking of the provinces was consistent, thus indicating strength in the method.

Chinogurei (2017) also considered the correlation between HIV prevalence and adult mortality and observed, that, by and large, districts (with the exception of districts that lie in Gauteng province) with high HIV prevalence have high mortality and vice versa (which was consistent with his expectations) as shown in Figure 2.4.

Figure 2.4 Estimates of adult mortality for men and women in districts of South Africa



Source: Bangha and Simelane (2013); Figure 5.2; page 73

Thus, from his findings, he concluded that although the method produces marginally higher estimates (most likely due to boundary differences in population counts from the 2001 census) to a large extent it produces reasonable estimates, which lent confidence in the method (Chinogurei 2017). However, he pointed out that his conclusion, to some extent, was limited due to the data (as splitting the death data in the districts by sex, population group and age group requires that the data quality not be compromised as the numbers are small) and the assumptions made in his study.

2.4.1 The level of completeness of death registration in South Africa

From the study by Dorrington, Timæus, Moultrie *et al.* (2004), it was observed that some of the provinces had a level of completeness of death registration exceeding 100 per cent. These included Gauteng, the Northern Cape, the Western Cape, the North West and Free State with the following percentages respectively 118, 118, 117, 107 and 105 per cent. Dorrington, Timæus, Moultrie *et al.* (2004) point out that this could be due to mis-recording of place of residence prior to death, that is, deaths from other province being recorded as deaths in Gauteng, the Northern Cape and the Western Cape. They further argue that a more likely explanation of the results for the Free State and the North West was under-coverage of census. The provinces with the lowest levels of registration were found to be Eastern Cape, Kwa-Zulu-Natal, Mpumalanga and Limpopo, that is, the poor and mostly rural provinces.

Estimates from Pillay-van Wyk, Laubscher, Msemburi *et al.* (2014) indicate that the Western Cape and Gauteng had the highest level of completeness for the period 1997 to 2005 (which is consistent with Dorrington, Timæus, Moultrie *et al.* (2004) results) whereas for 2005 to 2006, Limpopo, Free State and Western Cape had the highest level of completeness

Dorrington, Moultrie and Timæus (2004) observed that the White population group had the highest level of completeness[7] (77 and 78.5 per cent for males and females respectively) compared to other population groups whereas the black African population had the lowest level of completeness (63.9 and 66.6 per cent for males and females respectively). It was noted that for all population groups, females had a higher level of completeness than males. Of particular interest was the huge difference between Indian males and females, although no explanation for this could be found.

2.4.2 The level of adult mortality at sub-national level in South Africa.

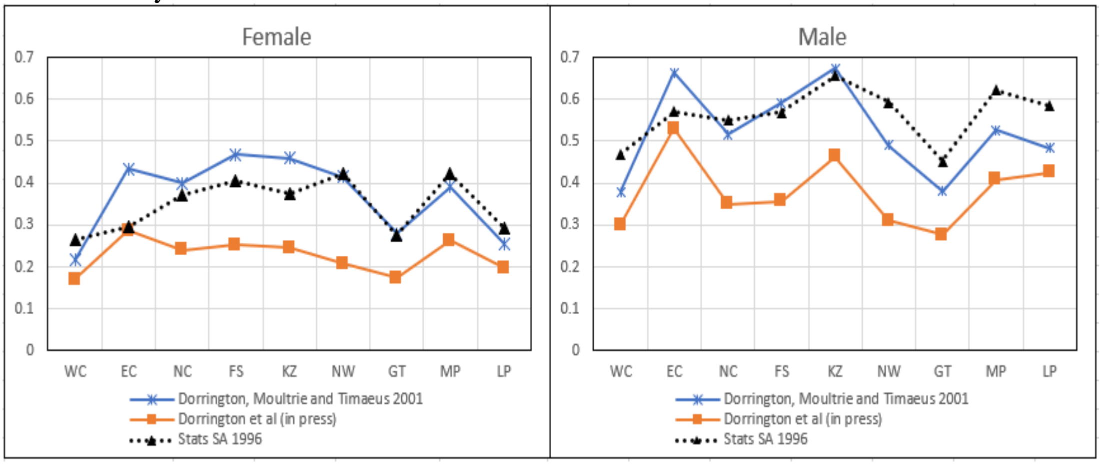
From the above-mentioned studies, the estimates derived are as shown in Figure 2.5. Stats SA mortality estimates for both male and female, in 1996, were higher than estimates derived by Dorrington, Timæus, Moultrie *et al.* (2004) in all provinces. Dorrington, Timæus, Moultrie *et al.* (2004) argue that Stats SA overestimated mortality as, in general, adult mortality for the population as a whole, in 1996, was lighter.

In support of their argument, they compared their estimates to estimates derived from 1996 Census orphanhood data. They observed that these two sets of estimates were consistent albeit that the orphanhood estimates were lower (Dorrington, Timæus,

[7] Note that this measure of completeness includes the requirement that the population group of the death is also recorded

Moultrie *et al.* 2004). However, comparison of their estimates to estimates from the ASSA 2000 model of the Actuarial Society of South Africa were inconsistent. They pointed out that this could be because the provincial mortality estimates from Actuarial Society of South Africa were obtained by weighting national mortality rates for each population group by the proportion of each population group by age in each province, which, as a result was not accurate for some of the provinces.

Figure 2.5 Comparison of adult mortality ($_{45}q_{15}$) in South Africa by provinces in the years 1996 and 2001



Source: Prepared using estimates from Stats SA (2000); Table 10A-18B; page 19-23, Dorrington, Moultrie and Timæus (2004); Figure 6.9 and 6.10; page 73

The male mortality estimates derived by Dorrington, Timæus, Moultrie *et al.* (2004) and Dorrington, Moultrie and Timæus (2004) (as shown in Figure 2.5) indicate a consistent increase in mortality levels from 1996 to 2001, with provinces such as Kwa-Zulu-Natal, Eastern Cape and Limpopo having high mortality levels while provinces with low mortality were Western Cape and Gauteng. However, looking at the estimates of female mortality, inconsistency is apparent within the provinces, that is, in 1996 the Eastern Cape and Mpumalanga were provinces with highest mortality whereas in 2001 Kwa-Zulu-Natal and the Free State were the provinces with highest mortality.

Figure 2.6 shows estimates of adult mortality, $_{45}q_{15}$, from Dorrington, Bradshaw, Laubscher *et al.* (2016)These estimates are consistent over all the years from 2000 to 2012 with KwaZulu-Natal, Eastern Cape and Free State being the provinces with highest mortality and the Western Cape and Gauteng being the provinces with lowest mortality. These estimates also indicate that mortality rates increased from the year 2000 to year 2005 in all provinces and declined from around 2005 to 2010 except for the Northern Cape.

Figure 2.6 The level of adult mortality ($_{45}q_{15}$) in South Africa's provinces.

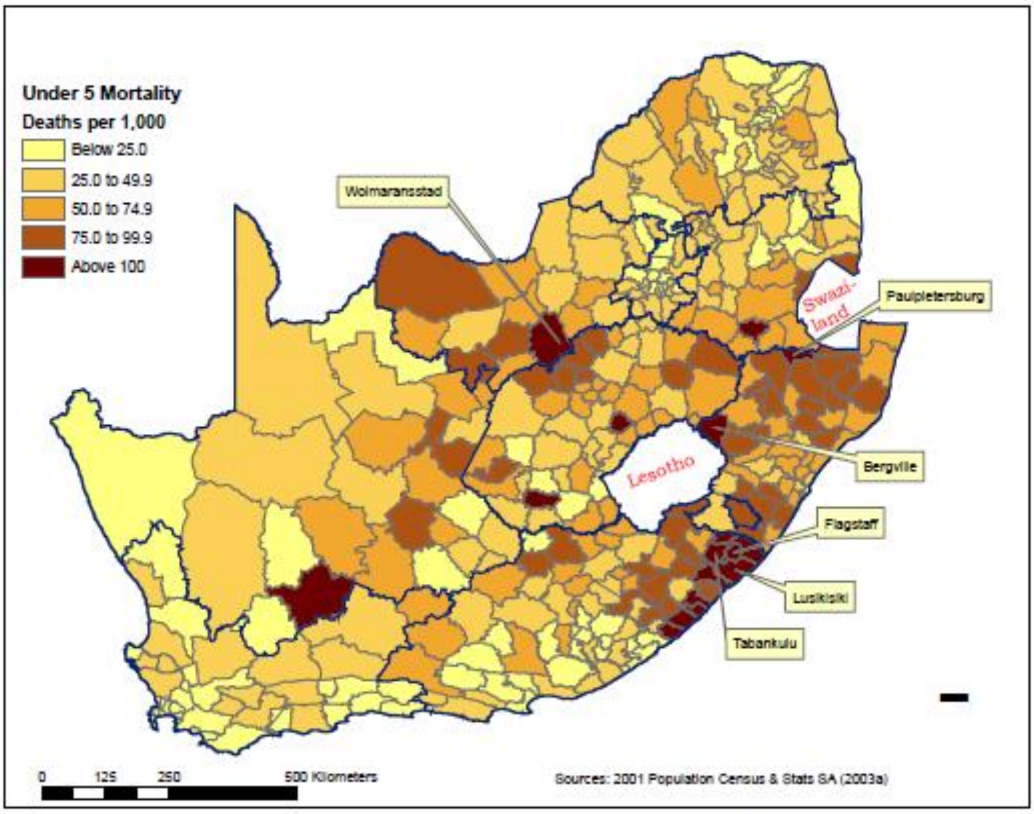
Source: Prepared using estimates from Dorrington, Bradshaw, Laubscher *et al.* (2016); Table 6; page 20

Chinogurei (2017) observed that the estimates of male mortality were lower in districts around the Western Cape and Gauteng and higher in districts that lie in KwaZulu-Natal and the Eastern Cape. This was in line with what might be expected, as generally one would expect high mortality levels in districts that lie in provinces with high mortality and low mortality in districts that are in provinces with low mortality. For women, it was observed that districts in the Western Cape and Limpopo had lower levels of mortality, whereas districts that lie in KwaZulu-Natal, Eastern Cape and Free State provinces had the highest levels of adult mortality.

Bangha and Simelane (2013) analysed under-five mortality by magisterial districts in South Africa for the year prior to the 2001 census. They argued that most studies of mortality in South Africa concentrate on the provincial level differentials, which is not ideal, as wide disparities exist within provinces. Thus, they aimed at estimating childhood mortality within the 354 magisterial districts to indicate mortality concentration within South Africa's geographical localities. They used the Brass children ever born/children surviving technique, and the GIS tool for spatial analysis. They observed that Eastern Cape and KwaZulu-Natal provinces had the high child mortality and most of the magisterial districts that had highest mortality lay in those provinces, however, even within those provinces, there were magisterial districts that had low mortality (as shown in Figure 2.7). In addition, in provinces with low mortality, Western

Cape, Gauteng and Limpopo, it was observed that there were magisterial districts with high mortality thus indicating disparities within the same provinces.

Figure 2.7 Under- five mortality estimates in South Africa



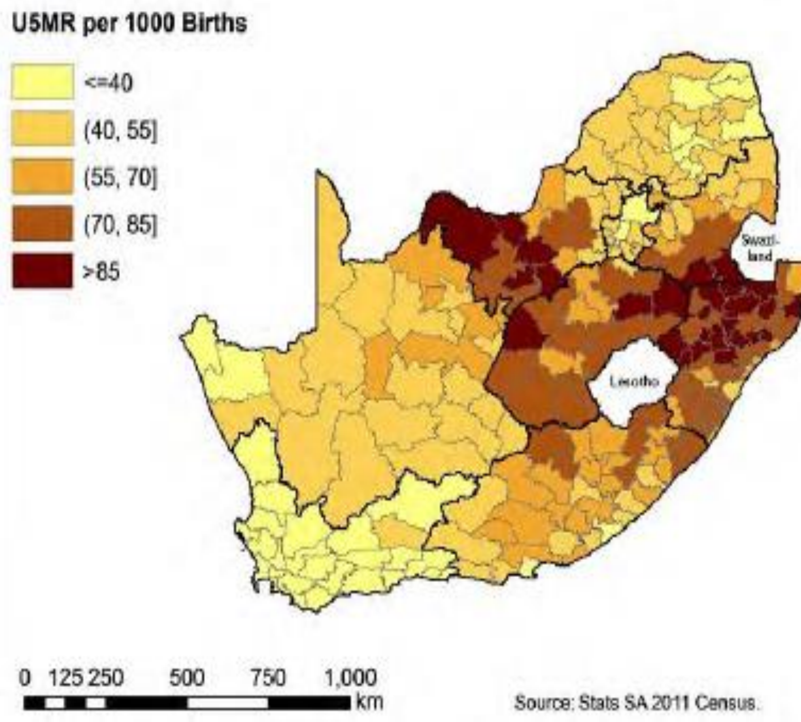
Source: Bangha and Simelane (2013); Figure 2; page 12

However, it should be noted that the 2001 census data do not produce reliable childhood mortality estimates as pointed out by Dorrington, Moultrie and Timæus (2004). Thus, estimates of childhood mortality derived by these studies, particularly by Bangha and Simelane (2013) should be interpreted with care as problems with the data leads to biased estimates. In addition, apart from data problems, estimating mortality at smaller geographical areas leads to random fluctuations, which could distort the picture.

Zewdie (2014) also derived child mortality estimates in South Africa at both provincial and municipal level using the 2011 census data. He employed direct synthetic cohort methods with Bayesian spatial smoothing to estimate child mortality at municipal level. From his findings, he observed that child mortality estimates in 2011 at the provincial level were lower in the Western Cape and Gauteng and high in provinces like KwaZulu-Natal and the Free State (having mortality 2.7 times that of Western Cape). The ranking of the provinces by Zewdie (2014) was consistent with the ranking of the provinces for adult mortality derived by studies discussed above.

Figure 2.8**Error! Reference source not found.** illustrates child mortality estimates by district level derived by Zewdie (2014). He observed that districts that lie in KwaZulu-Natal province and the North-West province had high mortality as compared to districts that lie in provinces like Gauteng, the Western Cape and Limpopo. He also concluded that there are spatial differences within districts that lie in the same province.

Figure 2.8 District child mortality estimates in South Africa



Source: Zewdie (2014); Figure 4-8; page 56

2.5 Methods of estimating adult mortality

In most developing countries, census coverage is incomplete as not all people are enumerated, and it is common for a certain percentage of recent deaths experienced by a population to remain unreported and for the vital registration to suffer from incompleteness. Methods have been developed to adjust for these errors for estimating mortality, so as to be able to produce reasonable and reliable mortality estimates (Bennett and Horiuchi 1981). These methods are categorised into three groups, that is, Death Distribution Methods (DDMs), census survival methods and methods based on the survival of relatives.

2.5.1 Death Distribution Methods

Death Distribution Methods are used for estimating the level of completeness of reported deaths, thus, allowing one to obtain estimates of the number of deaths corrected for incompleteness of reporting. The major approaches are the General

Growth Balance method (GGB) and the Synthetic Extinct Generations method (SEG). In order to make use of these methods certain assumptions have to be met and these assumptions, as set out by Dorrington (2013b, 2013c) and United Nations (2002) are that:

1. the population is closed to migration (however at some point, this assumption can be relaxed if sufficiently accurate estimates of migration are available),
2. completeness of the recording of deaths is constant, at least above a certain age, (usually age 15),
3. completeness of the population count is constant, at least above a certain age (usually age 15), and
4. the ages of the living and the dead are reported without error.

2.5.1.1 General Growth Balance Method (GGB)

The GGB method, proposed by Hill (1987), is an extension of the Brass Growth Balance (BGB) method originally formulated by Brass (1975) and is described in Dorrington (2013b) and in Hill K (2017). The BGB method estimates the level of under-reporting of registered deaths relative to the population count, making use of the demographic balance equation on the assumption that the population is stable.

The GGB method was formulated through generalising the BGB method such that it could be used without having to assume that the population is stable. The GGB method estimates the level of completeness of reporting of registered deaths through comparison of mortality estimates of those aged x and older to estimates derived from the population balance equation connecting a population aged x to x+5 at the beginning of the period to a population aged $x+t$ to $x+t$+5 at the end of the period, say t years later. To gain clarity of the concepts behind the method, we begin by evaluating the demographic balance equation, thus:

$$P_2 = P_1 + B - D + M$$

where P_1 and P_2 represent the true population numbers at the beginning and end of the period, at say, times t_1 and t_2 respectively, and B, D and M represent the true numbers of births, deaths and net number of migrants during the period t_1 and t_2 (United Nations 2002). Assuming that migration is insignificant, the demographic balance equation will reduce to:

$$P_2 = P_1 + B - D$$

This equation can be expressed as:

$$N(0) + P_1(0+) - P_2(0+) = D(0+) \qquad 2.1$$

where $P_1(0+)$ and $P_2(0+)$ represent true population of those aged 0 and over within the period and $D(0+)$ represent the number of deaths of those aged 0 and over within the period and $N(0)$ represent the total number of births during this period (United Nations 2002).

Generalising equation **Error! Reference source not found.** to all ages above a certain age *x* we obtain:

$$N(x) - P_2(x+) - P_1(x+) = D(x+) \qquad 2.2$$

where $P_1(x+)$ and $P_2(x+)$ represent true populations of those aged *x* and over at time t_1 and t_2 respectively, $N(x)$ represent number of people turning exact age *x* in the period and $D(x+)$ represent number of deaths aged *x* and over, in the period.

Assuming that the person-years lived within the period by persons aged *x* and over, during the period t_1 to t_2, can be derived using the geometric mean of $P_1(x+)$ and $P_2(x+)$ ((United Nations 2002), we have:

$$\boldsymbol{PYL(x+) = (t_2 - t_1) \times \sqrt{P_2(x+) \times P_1(x+)}} \qquad 2.3$$

Dividing equation **Error! Reference source not found.** by $PYL(x+)$ reduces the equation to:

$$\boldsymbol{\frac{N(x)}{PYL(x+)} - r(x+) = \frac{D(x+)}{PYL(x+)}} \qquad 2.4$$

where $r(x+)$ is the growth rate of the population aged *x* and older.

In cases where the first and second censuses have different coverage, we can let k_1 and k_2 denote the level of coverage of the enumeration of these two censuses respectively, and where death registration is incomplete, we let *c* be the level of completeness of registered deaths. Then:

$$P_1(x+) = \frac{P_1^*(x+)}{k_1} \qquad (a)$$

$$P_2(x+) = \frac{P_2^*(x+)}{k_2} \qquad (b)$$

$$D(x+) = \frac{D^*(x+)}{c} \qquad \text{(c)}$$

where, P_1^*, P_2^* and D^*represent observed values (enumerated population and registered deaths) of true values $P_1(x+)$, $P_2(x+)$ and $D(x+)$.

Based on the assumption that the population grows exponentially, the true annual growth rate of a population would be:

$$r(x+) = \frac{1}{t_2 - t_1} \ln\left[\frac{P_2(x+)}{P_1(x+)}\right] \qquad 2.5$$

Substituting *a* and *b* into equation **Error! Reference source not found.** gives:

$$r(x+) = r^*(x+) + \frac{1}{t_2 - t_1} \ln\left[\frac{k_1}{k_2}\right] \qquad 2.6$$

where $r^*(x+)$ is the growth rate of the observed values for people aged *x* and over (United Nations 2002).

Also, substituting *a* and *b* into equation **Error! Reference source not found.** we obtain:

$$PYL(x+) = \frac{PYL^*(x+)}{[k_1 k_2]^{0.5}} \qquad 2.7$$

Similarly, the number of people turning exact age *x* would be given by:

$$N(x) = \frac{N^*(x+)}{k_1 k_2} \qquad 2.8$$

Substituting equation **Error! Reference source not found.**, **Error! Reference source not found.** and **Error! Reference source not found.** into equation **Error! Reference source not found.**:

$$\frac{N*(x)}{PYL(x+)} - r^*(x+) = \frac{1}{t_2 - t_1} \ln\left[\frac{k_1}{k_2}\right] + \frac{(k_1 k_2)^{0.5}}{c} \frac{D^*(x+)}{PYL(x+)}$$

and rewriting the above equation in a simple form gives:

$$n^*(x) - r^*(x+) = a + b.d \times (x+) \qquad 2.9$$

where

$n^*(x) = \frac{N^*(x)}{PYL(x+)}$, $a = \frac{1}{t_2 - t_1}\ln\left[\frac{k_1}{k_2}\right]$, $b = \frac{k_1 k_2}{c}$, and $d^*(x+) = \frac{D^*(x+)}{PYL(x+)}$.

Thus, equation **Error! Reference source not found.** gives a straight line with the intercept *a* and slope *b* and the *y*-points being equal to $n^*(x) - r^*(x+)$ and the *x*-points being $d^*(x+)$. Hence, fitting a straight line to all (*x,y*) points for all age groups and inspecting the plotted points (so as to decide which age range will be suitable) provides estimates of the level of completeness of deaths and population counts. In order to obtain estimates of k_1, k_2 and *c,* we assume that the greater of either k_1 or k_2 is equal to 1 and thus,

if $\frac{k_1}{k_2} < 1$, then assume $k_2 = 1$ and hence:

$k_1 = e^{a(t_2 - t_1)}$ and $c = \frac{e^{a(t_2 - t_1)}}{b}$ and if:

$\frac{k_1}{k_2} > 1$, then assume $k_1 = 1$ and hence:

$k_2 = e^{-a(t_2 - t_1)}$ and $c = \frac{e^{-a(t_2 - t_1)}}{b}$

Individual judgement is exercised on which type of method to use when fitting the straight line described above, resulting in researchers using different methods (Bhat 2002) and thus producing different estimates. These methods include least squares regression (LSR), weighted trimmed means, grouped mean and orthogonal sum of squares regression. Bhat (2002) points out that each method has its own weakness and recommends use of the orthogonal regression method as he argues that it fits an accurate straight line when dealing with data that suffer from age exaggeration, which is usually the case with developing countries data. He further points out that the use of the unweighted least squares method is not recommended as the method gives too much weight to outliers.

Bhat (2002) and others have reformulated the GGB method to apply it to populations that are not closed to migration and the balancing equation generalises to:

$n^*(x+) - r^*(x+) + nm^*(x+) = a + bd^*(x)$ (Hill K 2017).

where $nm^{*}(x+)$ is the observed partial net migration rate within the intercensal period. Thus if data on migration are available, this method can still be applied.

2.5.1.2 Synthetic Extinct Generations (SEG) Method

The SEG method originates from the extinct generations method originally formulated by Vincent (1951). The idea behind Vincent's method was that one could estimate the population in a cohort at a particular time, say time *t*, aged *a* last birthday, using the number of deaths in that same cohort until the last person dies, that is:

$$N(z,t) = D(z,t) + D(z+1,t+1) + D(z+2,t+2) + D(z+3,t+3) + ... + D(\varpi, t+\varpi - z)$$

and simplifying the above equation gives:

$$N(z,t) = \sum_{y=0}^{\varpi - z} D(z+y, t+y)$$

where:

$N(z,t)$ represents the population at time *t* aged *z* last birthday,

$D(z+y,t+y)$ represents the deaths in year $t+y$ of people who were aged $z+y$ last birthday at the start of the year $t+y$ and ϖ represents the age at death of the last survivor in the same cohort.

However, the method is not particularly useful as it requires one to have to wait for the last person in the cohort to die, which makes it to be impractical to use (Bennett and Horiuchi 1984). Preston, Coale, Trussell *et al.* (1980) highlighted that if the population can be assumed to be stable, then the number of deaths in future could be predicted from current deaths, that is:

$$D(z+y,t+y) = D(z+y)e^{ry} \qquad 2.10$$

where *r* is the constant annual growth rate of the population.

Thus, the deaths of those aged *z*, in any year, would differ from the deaths in the preceding years by a factor of e^{r} and ultimately from the deaths in year *t* by a factor of e^{ry}.

Hence, substituting equation **Error! Reference source not found.** into the equation above gives:

$$N(z,t) = \int_{0}^{\varpi - z} D(z+y,t)e^{ry}dy \qquad 2.11$$

Bennett and Horiuchi (1981, 1984) formulated the Synthetic Extinct Generations (SEG) method by generalising the relationship identified by Preston, Coale, Trussell *et al.* (1980) so as to estimate the level of completeness of reported deaths and hence estimate mortality using registered deaths corrected for incompleteness (Hill K 2017). Thus, they generalised equation **Error! Reference source not found.** to

$$N(z,t) = \int_{0}^{\varpi - z} D(z+y,t) e^{\int_{0}^{y} r(z+v,t)dz} dy \qquad 2.12$$

where $r(z+v,t)$ is the age specific growth rate of the population aged $z+v$ at time t_0 Transforming $z+y$ into h implies that $N(z,t)$ would be equivalent to:

$N(z,t) = \int_{z}^{\varpi} D(h,t) e^{\int_{z}^{h} r(v,t)} dh$, implying that the number of people alive at a particular point in time, *t,* is equivalent to the number of people who die in that cohort from time *t* until the last person dies.

However, Dorrington (2013c) argues that describing the method as the Synthetic Extinct Generations is misleading as the method is not an extension of Vincent's proposed idea of extinct generations as is the case with the Preston and Coale method. Generally, the method does not rely on the idea of projecting deaths into the future, but on a reconstruction of the past (Dorrington 2013c). Thus, it is important to note that equation **Error! Reference source not found.** does not imply that future deaths grow at current growth rates.

In essence, the Synthetic Extinct Generations method can be used to determine the completeness of death registration to improve the accuracy of mortality estimates. If registered deaths are incomplete, the level of completeness can be obtained from the ratio of population estimates obtained from the registered deaths relative to the population used to determine the denominator.

Bennett and Horiuchi (1981) propose that this method be applied to data in five-year age groups for two consecutive censuses that are *t* years apart, that is, $P_1(x,5)$ and $P_2(x,5)$, and to age specific deaths registered in the intercensal period, $D(x,5)$. Using five-year age groups, equation **Error! Reference source not found.** becomes:

$$N(z,t) = N(z+5)e^{5 {}_5r_z} + {}_5D_z e^{2.5 {}_5r_z}$$

where ${}_5D_z$ denotes the true deaths in the age group z to z+*5*, implying that the relationship between the observed deaths (denoted by ${}_5D_z^*$) and the true number of deaths would be:

$${}_5D_z = \frac{{}_5D_z^*}{c}$$

where *c* is the level of completeness. To adjust for under- or over-reporting of deaths we substitute ${}_5D_z^*$ for ${}_5D_z$, and the above equation becomes:

$$\hat{N}(z) = \hat{N}(z+5)e^{5\,{}_5r_z} + {}_5D_z^* e^{2.5\,{}_5r_z}$$

The above equation is iterative in that in order to estimate the value of *N* (z)*,* one needs to know the value of *N*(*z*+5). This implies that the value of the oldest age group, $\hat{N}(Z)$ *(Z* being the lower age of the open-ended age interval) must be estimated independently. Bennett and Horiuchi (1981) proposed that this could be estimated using:

$$\hat{N}(Z+) = D(Z+)[e^{r_Z + e_Z} - (r_Z e_Z)^{2/6}]$$

The age-specific growth rate (${}_5r_z$) can be estimated from the two-census population as:

$${}_5r_z = \frac{1}{t}\ln\frac{P_2(z,5)}{P_1(z,5)}$$

The life expectancy at the lower age of the open interval, denoted as e_Z, is required apart from the deaths. The life expectancy can be obtained from either:

- independent research estimates based on the same population, or
- the GGB method applied to the same population data, from which a life table is derived and the life expectancy at age *Z* estimated, or
- by trial and error, using a model life table fitted to the rates produced using the SEG method.

Following this, the values of the *N*(*z*) can be generated for all age groups. The average number of people reaching exact age *z*, in a *t*-year interval, in each of the age groups, can be estimated as:

$\hat{N}(z) = 0.2t\sqrt{P_1(z,5) * P_2(z,5)}$ (Dorrington 2013c; United Nations 2002)

The average annual number of people in conventional five-year age groups can be computed as

$${}_5\hat{N}_z = 2.5(\hat{N}(z) + \hat{N}(z+5))$$

and the years of life lived by the population between t_1 and t_2, can be estimated as:

$${}_5N_z = (t_2 - t_1)[{}_5N_z(t_1) * {}_5N_z(t_2)]^{0.5}$$

where ${}_5N_z(t_1)$ and ${}_5N_z(t_2)$ are estimates of the population aged between z and $z+5$ at time t_1 and t_2 respectively.

Completeness of reporting of deaths (c) can then be estimated from each $c(z)$, where $c(z)$ is the ratio of the average annual population estimated from the reported deaths to the average annual population from the censuses, that is,

$c(z) = \frac{{}_5\hat{N}_z}{{}_5N_z}$, (where ${}_5\hat{N}_z$ is estimated on basis of reported deaths), based on the assumption that reporting of deaths is constant over all age groups.

To obtain the level of completeness of deaths, c, either the mean or median value of $c(z)$ values, has to be calculated, and thus, the level of overall completeness will be obtained. The level of completeness is then used to adjust the registered deaths for incompleteness, and mortality rates can then be obtained from the adjusted registered deaths as:

$${}_5m_z = \frac{[\frac{{}_5D_z^*}{c}]}{t\sqrt{P_1(z,5) * P_2(z,5)}}.$$

From the above equation, the adult mortality rate (${}_{45}q_{15}$) can be estimated based on the assumption that the force of mortality, $u(y)$, does not change over each age group interval and also assuming that $5{}_5m_z = \int_0^5 u(z+t)dt$, that is, ${}_5p_z = -5{}_5m_z$, and hence:

$${}_{45}q_{15} = 1 - e^{-\int_0^{45} \mu(15+t)dt}.$$

2.5.1.3 Allowing for differential census coverage

Bennett and Horiuchi (1984) observe that the age specific growth rates estimated from two consecutive censuses will lead to biased mortality estimates if these censuses have different coverage. They, thus, suggest adjusting the age-specific growth rates obtained so as to take into account differential coverage of censuses. They suggested adding a delta (δ) to the computed growth rate (${}_5r_z$) to obtain the corrected growth rate, (${}_5\hat{r}_z$), thus:

${}_5\hat{r}_z = {}_5r_z + \delta$ and the value of delta is such that series of $c(z)$ values are roughly constant. The optimum value of delta gives:

$\frac{k_1}{k_2} = e^{-\delta \cdot t} = 1 - \delta \cdot t$, where k_1 and k_2 are defined previously and t is the length of the intercensal period.

2.5.2 Sensitivity of Death Distribution Methods to data errors

The death distribution methods work quite well for errors that they are designed to overcome, that is, digit preference, non-registration of deaths, differential census coverage (Dorrington, Timæus and Moultrie 2008; Hill, You and Choi 2009). However, questions arise as to how accurate the mortality estimates derived using these methods are if the assumptions required do not hold and where there are errors in census enumeration as well as in reported deaths (particularly in developing countries).

Hill and Choi (2004) and later Dorrington, Timæus and Moultrie (2008) and Hill, You and Choi (2009) attempted to provide answers to such questions by considering how DDMs respond in the presence of diverse errors and cases where the assumptions on which these methods are based do not hold. Hill and Choi (2004) and Hill, You and Choi (2009) simulated different error scenarios, applied them to DDMs, and tried to measure how ${}_{45}q_{15}$ estimated by these DDMs differ. However, Hill and Choi (2004) did not employ the SEG+delta (SEG+ δ) approach (described in the previous section), which takes into account differential census coverage (Dorrington, Timæus and Moultrie 2008). They concluded that the best approach was to use the GGB method to estimate the level of one census coverage relative to the other and then correct the censuses before applying the SEG method to estimates of the level of completeness (Hill and Choi 2004; Hill, Choi and Timæus 2005).

Dorrington, Timæus and Moultrie (2008) investigated whether they would reach similar conclusions to Hill and Choi (2004) if they applied the SEG+ δ to similar data sets and various errors used by Hill and Choi (2004). Furthermore, they explored the response of the methods using same simulated error cases based on the dataset of an hypothetical African country that experienced high HIV mortality (Dorrington, Timæus and Moultrie 2008), thus making the study more relevant to African countries, in particular, South Africa with high HIV prevalence.

All studies (Dorrington, Timæus and Moultrie 2008; Hill and Choi 2004; Hill, You and Choi 2009) concluded that all the DDMs show bias in instances where migration is present and is not taken into account. This error produced a more biased estimate of completeness when compared to other typical errors. Hill, You and Choi (2009) argue that the SEG method was less affected by this error when compared to other DDMs. Migration effects can be seen by points that deviate from the fitted straight line (when applying the GGB method), mostly for ages 15 to 30. To remove the influence of these points on the estimates one can confine the age range on which to fit the line to ages 30 or 35 and upwards depending on which age range will fit a better line (Dorrington, Timæus and Moultrie 2008). Hill, You and Choi (2009) used age ranges 30 to 65+ for both the GGB and SEG methods and this reduced the RMSE by approximately 40 per cent.

Both Dorrington, Timæus and Moultrie (2008) and Hill, You and Choi (2009) came to the conclusion that the GGB method is a better method when dealing with the error of age misreporting. Age misreporting usually occurs at older ages and its effects can be seen from the diagnostic plots of the GGB method if the points, at those older ages, diverge from the straight line fitted whereas on the SEG method they are seen by a rising and falling level of completeness (with age) at the older ages. (Dorrington 2013b).

In general, Dorrington, Timæus and Moultrie (2008) concluded that the SEG+ δ method performs better (out of 23 error scenarios the method worked well in 15 scenarios) than the other methods. However, for the African female population, the method performed well in only 7 scenarios, though it performed much better for the male population. Hill, You and Choi (2009), on the other hand, concluded that the GGB+SEG method performed better than the other methods. In this research, the SEG+ δ method is to be used as Dorrington, Timæus and Moultrie (2008) concluded it was the best when applied to African populations making it more valid to use it for South Africa.

Hill K (2017) points out that what has not been addressed in systematic evaluations is a GGB-SEG combination is using different age trims for the two methods. He proposes estimating census completeness change by applying GGB to the wide age range 5 to 65 years, adjusting the populations accordingly, and thereafter applying SEG, estimating coverage for the age range 50 to 70 years. He suggests that this approach should maximize the effectiveness of GGB for estimating completeness change, while minimizing SEG errors from migration.

2.5.3 Orphanhood method

The idea of using questions on the survival status of parents to produce mortality estimates can be traced from the time of Lotka, in 1939, who used the life table to derive the total number of orphans (Luy 2009). Henry (1960) proposed the idea of calculating adult mortality from the reported number of orphans. Brass and Hill (1973) further developed the idea along with several other authors (Hill and Trussell 1977; Timæus 1992; Timæus and Nunn 1997). The concept behind the method is that for a mother to give birth she has to be alive at the time of birth, and hence the mother's exposure to the risk of mortality is the time from when she gave birth up to the age of the child at the time of the survey, say age *y* years. Assuming that mortality of children is independent of that of their mothers, and that the mother is aged *z* when she gives birth, it follows that the probability of a mother surviving (based on the proportion of children with mother alive at the time of the survey) from the time of giving birth to the time of the survey (that is, from between ages *z* to *z+y*) is given by $\frac{l_{z+y}}{z}$. In this method, *z*, is given by a convenient age closer to the mean age of child bearing (or conception, for the father) for each age group and *y* represents the age closest to the midpoint of each age group.

The conditional survivorship probability of the chosen suitable age in each age group are derived by employing a regression model (Timæus 1992). Coefficients which convert the proportion of children whose mothers are reported dead into life table survivorship probabilities were developed by Brass and Hill (1973). They developed these coefficients based on model life tables on the assumption that mortality schedules of the populations where the orphanhood method can be employed had the same shape regardless of the mortality level. Hill and Trussell (1977) later developed another set of coefficients, which were then followed by coefficients produced by Timæus (1992). These sets of estimates were similar for maternal orphanhood for both studies, and were better than the ones developed by Brass and Hill (1973).

It is worth noting that the conditional survivorship probabilities derived refer to different mortality experiences and measures of likelihood of survival for different ages and time periods, hence there is a need for a common measure which can be used to compare this estimates over time (Timæus 2013a). To convert these probabilities into this measure, an appropriate standard life tables is chosen and the Brass relational logit model is used, namely,

$$\lambda(l(x) = \alpha + \beta * \lambda(l^{s}(x)$$

where $l(x)$ and $l^s(x)$ is the proportion of those that survive up to age x in the fitted as well as in the standard life table and alpha and beta represent the level and the shape of mortality respectively (Murray, Ferguson, Lopez *et al.* 2003).

This method often works better using female rather than male respondents due to the fact that it is usually the case that daughters are more connected to their parent's than sons.

The method produces biases in mortality estimates as parents with more than one surviving child are over-represented and parents with no surviving children are not represented at all. In addition, the method assumes that the mortality of the parents is independent of the number of children, which it is not. Selection biases arising from this are usually small, however, in populations affected by generalized HIV epidemics, it is likely to be more severe, especially at the time when ARV's were not available. Timæus and Nunn (1997) developed a method that estimates adjustment factors to correct for this bias.

Also, the orphanhood method suffers from adoption effects which lead to an underestimate of mortality, as orphaned children (usually those aged less than 20) might misreport their guardians as their parents whilst in actual fact their parents are deceased thus leading to lower mortality estimates as they would be enumerated as though their parents are alive. However, Timæus (2013a) suggests that by using respondents aged above 20 years, one is able to reduce this bias, since the bias is likely to be greatest for young children.

2.5.4 Other Methods

Schmertmann and Gonzaga (2018) developed a Bayesian regression model for small-area mortality schedules that attempts to address simultaneously the problems of small local samples and underreporting of deaths. They combined a relational model for mortality schedules with probabilistic prior information on death registration coverage derived from demographic estimation techniques, such as Death Distribution Methods, and from field audits by public health experts. They then tested the model on small area data from Brazil, showing notable differences in the age patterns of mortality between adjacent small areas. They concluded that incorporating external estimates of vital registration coverage through priors improves small-area mortality estimates by accounting for underregistration and automatically producing measures of uncertainty.

2.6 Conclusion

It is evident from the literature review that there is little research that estimates both child and adult mortality at sub-provincial level in South Africa as sub-provincial mortality estimates are quite problematic to produce accurately if the data are prone to errors and bias. The Death Distribution Methods and the orphanhood method have limitations when used to estimate mortality at sub-provincial level, as the province to which the death is attributed may not be the province of residence. Also, migration not accounted for in DDMs may make the estimates unreliable.

On the other hand, unless the sample is very large (for instance, if the questions were part of the census), the orphanhood method may involve sample sizes at sub-provincial level that are too small to produce reliable estimates.On the other hand, using

For these reasons, it is difficult to produce reliable mortality estimates at sub provincial level from either the vital registration or census data alone. However, Dorrington and Timæus (2015) and Dorrington, Moultrie and Timæus (2004) have produced reasonable adult mortality estimates at provincial level using an approach that uses both data sources.

Chinogurei's investigation of the method at district level for the year prior to the 2001 census concluded it produced reasonable estimates of adult mortality. He reached the conclusion that the method, to a large extent, produces reasonable results though there were limitations in his conclusion. Thus, this research aims to investigate if the similar method produces plausible estimates when applied to data for 2011 so as to strengthen the conclusion reached by Chinogurei (2017).

3 DATA AND METHODS

3.1 Overview

The outline of the method used here to obtain adult mortality estimates at district level from the 2011 census is as follows. First, the data to be used are assessed and evaluated for errors. The level of completeness of vital registration deaths for the intercensal period 2001-2011 by population group for each sex is determined using the SEG+delta method. Thereafter, the level of completeness for each of the years in the intercensal period is estimated by fitting a logistic curve to the level of completeness for the intercensal period of 1996-2001 and 2001- 2011 (derived by both Chinogurei (2017) and Richman (2017)) so as to determine the level of completeness in the year prior to the 2011 census. The number of deaths registered in the year prior to the 2011 census are then corrected for either under- or over-reporting using the estimates of completeness so as to obtain an estimate of the true number of deaths by population group and age group for each sex.

Thereafter, the age-sex specific correction factors for correcting deaths reported by household at district level are estimated by dividing the true number of registered deaths by the number of household deaths in the year prior to the 2011 census. Deaths reported by households by population group and age group for each sex are then corrected for either under- or over-reporting by applying the age-sex specific correction factors obtained on the assumption that the correction factors at district level do not differ from those at national level. District adult mortality for each sex is then estimated using the corrected number of deaths reported by households and person-years lived.

Finally, the estimates derived are evaluated to assess whether the method employed produces plausible estimates. This is done by, first, comparing the weighted average of the estimates of the district level mortality to alternative estimates for each province that would be derived using the orphanhood method and to estimates derived by Dorrington and Timæus (2015) at provincial level. Second, these estimates are compared to Chinogurei (2017) estimates of district level mortality in 2001 and, third, they are compared to the ranking of childhood and adult mortality estimates derived by Zewdie (2014) and Chinogurei (2017) respectively.

3.2 Sources and assessment of quality of data

Census population counts and net foreign-born immigrants by population group, age group for each sex for the years 2001 and 2011 were obtained from SuperWEB2 via the Stats SA website[8]. The number of deaths reported by households for each district in the year prior to the 2011 census were provided by Stats SA through a special request (personal communication with Professor R. E. Dorrington). Registered deaths by population group and age group were provided by Stats SA through the South African Medical Research Council (personal communication with Professor R. E. Dorrington).

STATA 14 (StataCorp 2015) was used to extract and process the vital registration deaths and census data. GGB and SEG+delta excel workbooks (Dorrington 2013b, c) for applying death distribution methods were used to estimate the level of completeness for each population group at the national level and the orphanhood workbook (Timæus 2013a) was used to estimate the level of mortality at provincial level.

3.2.1 Census data

Figure 3.1**Error! Reference source not found.** shows the age distribution of the population from the 2001 and 2011 census estimates by single years and by grouped ages. The population distributions by single ages indicate occurrences of age heaping and digit preference. The graph of data by grouped ages smooths the data, as age heaping and digit preference are no longer evident.

[8] http://interactive2.statssa.gov.za/webapi/jsf/dataCatalogueExplorer.xh

Figure 3.1 Population age distribution from the 2001 and 2011 census

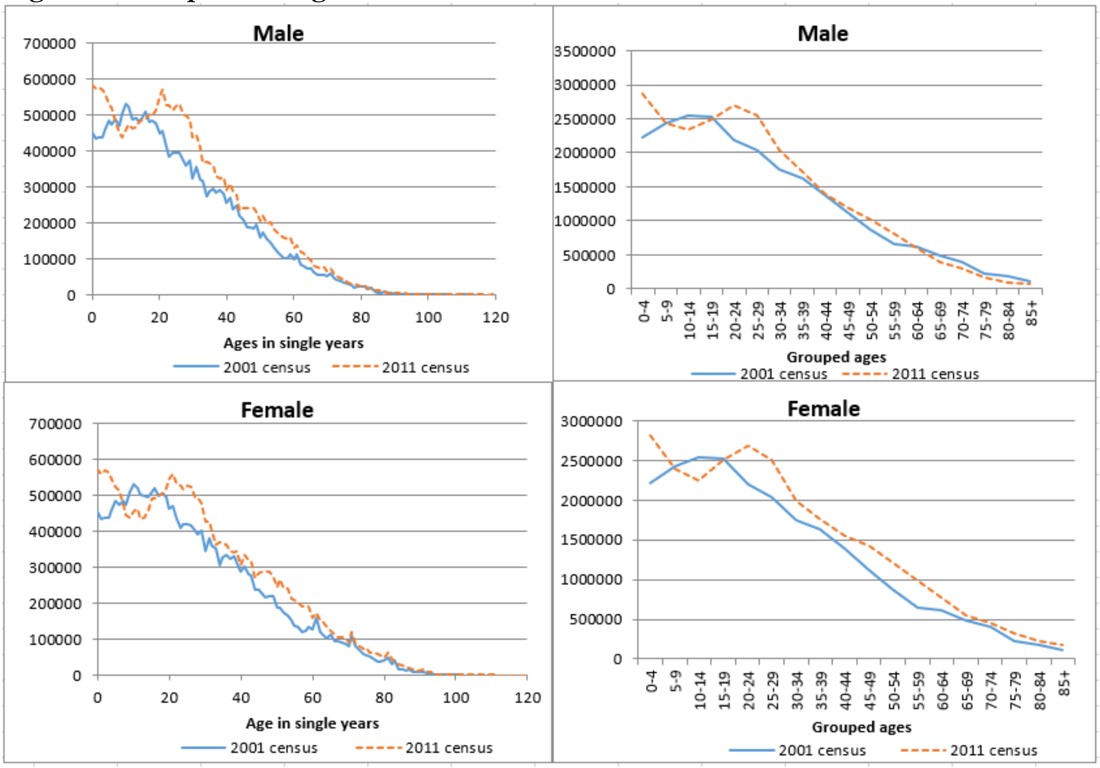
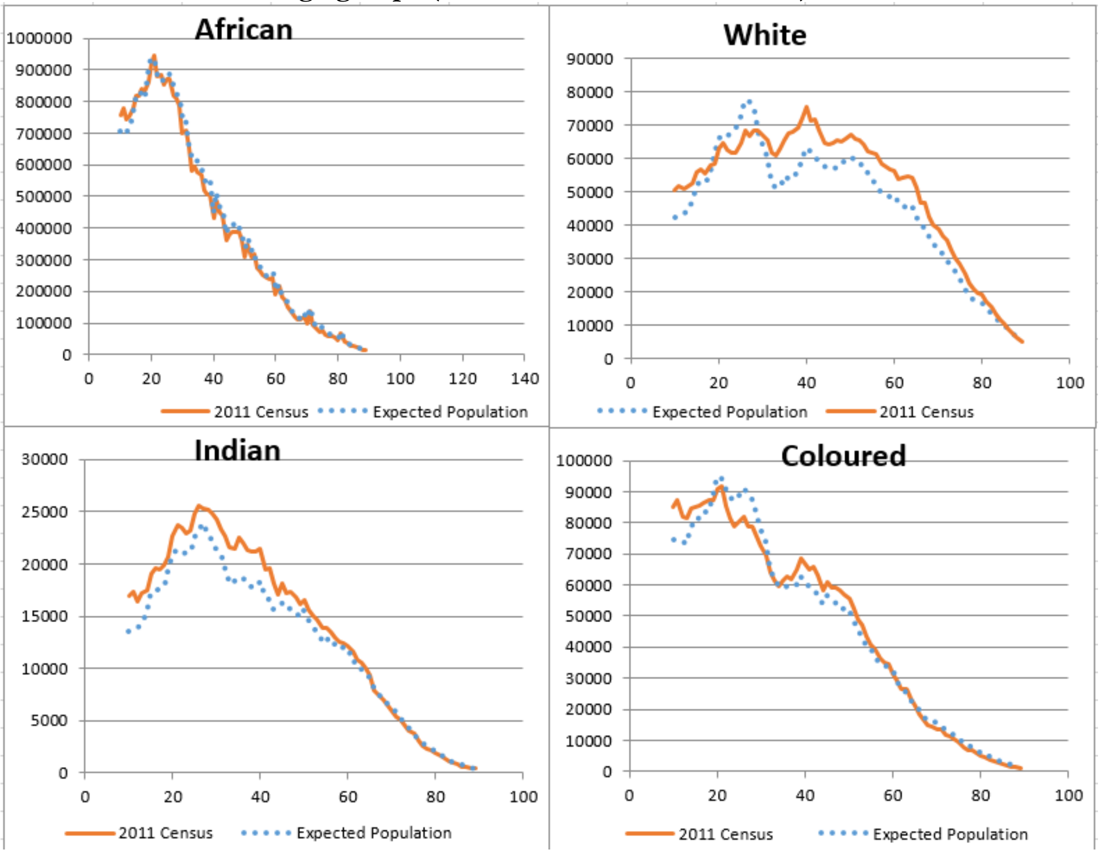


Figure 3.2 shows 2011 census population estimates and the 2001 census projected to 2011 by population group, both sexes combined. To project the 2001 census estimates to 2011, survival ratios estimated from the life tables in the CARe2011_160901 model (allowing the same rate of migration as the CARe2011_160901 model) for each sex, were used. Comparison of these figures indicate that the 2001 census substantially undercounted children aged five years and under for all population groups. For the White and the Coloured populations, it can be observed that the expected population between ages 20 and 30 is higher than the 2011 census population which could be explained by labour or education emigration. On the other hand, the 2011 census appears to overestimate those in age range 30 to 69 and 30 to 49 for the White and Coloured population respectively. For the Indian population, the 2011 census estimates are higher than the projected estimates possibly due to migration. Otherwise, in general, the expected population and the 2011 census population are broadly similar in pattern and very close over most of the age range.

Figure 3.2 Projected population from 2001 census vs 2011 census population by individual age groups (males and females combined)

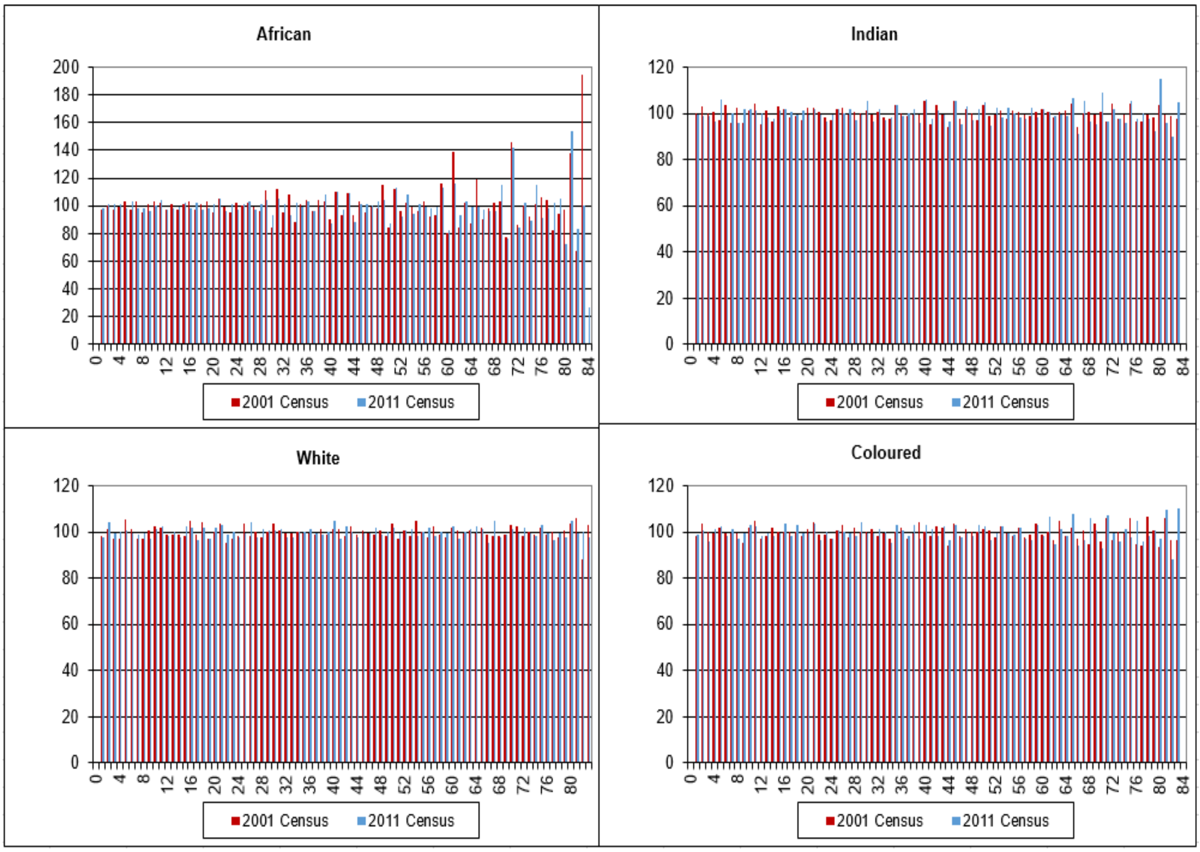


3.2.1.1 Age Ratios

Calculated measures, such as age ratios indicate possible age misreporting in the age-specific data. The age ratio should be fairly close to 100, based on the assumption that population numbers change linearly over three consecutive age groups. Age ratios that vary much from 100 constitute a warning of errors such as age misreporting and digit preference.

Figure 3.3 shows the age ratios for each population group by age for both the 2001 and 2011 censuses. From this it can be observed that these age ratios are close to 100 for ages below 80 for most population groups although, for the African population there is an indication of worse age heaping than the other population groups, particularly for ages 30 and above (which is also shown in the same cohort in 2011). In other population groups (excluding the white population) age heaping is apparent at much older ages, that is, ages 60 and above. In general, we observe that the age ratios are close to 100 for the younger ages.

Figure 3.3 Age ratios for male and female by population group in the 2001 and 2011 Census

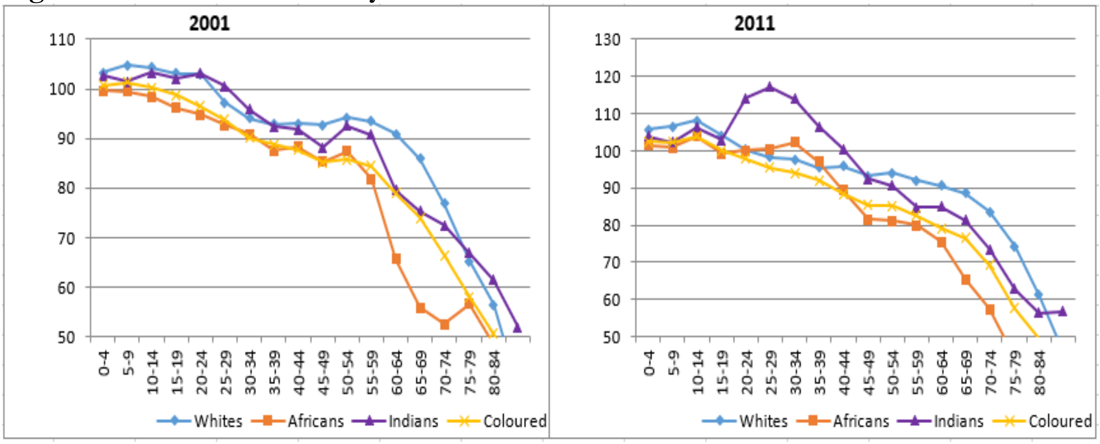


3.2.1.2 Sex Ratios

In the absence of significant sex-specific differences in migration we expect the sex ratio to start slightly above 100 and decline slowly between birth and late middle age (45-60) and thereafter we expect it to fall more rapidly at the older ages as male mortality begins to greatly exceed female mortality. Sex ratios that deviate from this pattern could indicate significant differences by sex in migration or mortality at certain ages or if not, it could indicate that the data suffer from errors.

From Figure 3.4 it can be seen that for the White population there is a noticeable dip between ages 20 to 49 in 2001 which is also apparent in ages 25 to 59 in 2011 which is probably due to undercount of white men. Also, the sex ratio of the Indian population in 2011 indicates that there is an excess of Indian males relative to females between the ages 15 and 49 (which is an unexpected pattern) which is probably due to young adult male immigration. The ratios for those aged 25 to 39 for the African population indicate that there are more males than females suggesting male labour in-migration or possibly under-enumeration of females. For all population groups, cohort effects are apparent, particularly at ages 10-14 in 2011 (that is, 0-4 in 2001) as a peak is noticed in the sex ratios of those aged 10-14 in 2011 this could be due to an undercount of female children in the population. Also, between ages 10-14 in 2001 (that is 20-24 in 2011) and ages 30-34 in 2001 (that is, 40-44 in 2011) the sex ratios indicate that there are more males than females in 2011 compared to 2001, this could be due to hidden migration. It can also be observed that at ages 60 to 64 there are more males than females which we do not expect, which could indicate the possibility of errors in the data.

Figure 3.4 Sex ratios for the years 2001 and 2011



Comparison of the 2001 shifted sex ratio (to match 2011 sex ratios) to 2011 sex ratio for the population as a whole, which are shown in Figure A. 1 (Appendix A) gives further evidence of undercount of female children relative to male children in 2001. The comparison also indicates that for those aged between 20 and 44 there are more males than females in 2011 compared to 2001 which could be explained by male migration.

3.3 Data on registered deaths

Table 3.1 indicates the distribution of deaths by population group in the intercensal period 2001-2011. Approximately 25 per cent of the deaths had unspecified population group. No imputations were done since DDMs were used to obtain the level of completeness of registration of deaths as well as completeness of recording of the population group, thus allowing the true number of deaths by population group to be estimated. This was done based on the assumption that completeness of registration of deaths and of population group are constant with age. Note that the proportion of deaths with unspecified population group was level for each of the age groups (Figure A.1, Appendix A) which validates the latter assumption.

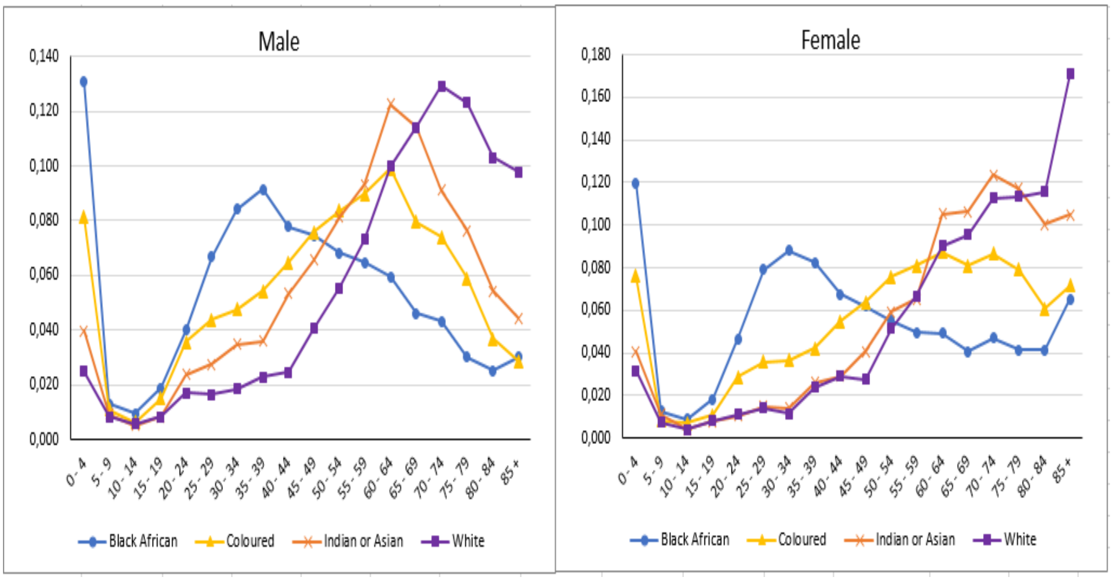
Table 3.1 The distribution of registered deaths by population group in the intercensal period 2001-2011

Population group	Frequency	Percentage
Black African	3 968 122	62.8
Coloured	289 139	4.6
Indians	83 434	1.3
White	401 388	6.4
Other	8 524	0.1
Unspecified	1 565 053	24.8

3.3.1.1 Registered deaths by population group

Figure 3.5 shows the proportion of deaths by age for each population group for males and females. It can be observed that for the African population, the age group with the highest number of deaths are children under the age of 5, with a hump for adults aged around 20 to 44 for females and around 25 to 44 for males. This is to be expected as these age ranges are mostly associated with high mortality risks in developing countries (due inter alia to HIV/AIDS epidemic), thus high proportions of deaths are observed at these ages. Also, of interest, is the observation that for women the age range at which most deaths occur is lower than that for males which is probably attributable to the impact of the HIV/AIDS pandemic, as women become sexually active at an earlier age than men (as a result, they are likely to be infected at earlier ages) and also the probability of transmission from men to women is higher than from women to men.

Figure 3.5 Distribution of registered deaths by population group in the period 2001-2011

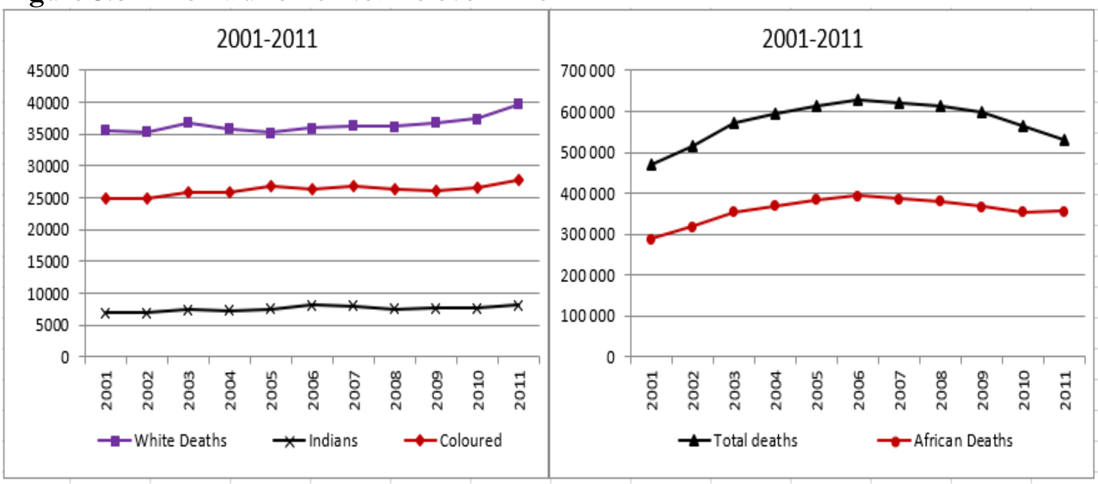


In contrast to the distribution of the African population, the distribution of deaths by age for the White population indicates that the majority of deaths are observed to be at older ages, that is, around 70 to 85+. For the Indian population they are observed to be at ages 60 to 74 and 60 to 79 for males and females, respectively. This result, indicates that the distribution of deaths by age differs significantly for each population group, thus applying DDMs to the population as a whole would violate the assumption that the reporting of deaths is constant with age (since completeness differs for each population group).

3.3.1.2 Distribution of deaths over time

From Figure 3.6 it can be seen that for all population groups combined, total deaths gradually increased from 2001 to the peak around 2006 (largely due to HIV/AIDS, with some population growth) and thereafter the deaths started to decrease. The most likely explanation for the decrease is the significant increase in provision of ARV treatment from around 2005, however, this could also be due to changes in the level of completeness of registered deaths (personal communication with Professor R. E. Dorrington) or the exclusion of late registration of deaths (although, for the data set used in this research the registration of deaths is sufficiently delayed thus reducing the effects of this to insignificant).

Figure 3.6 Distribution of deaths over time



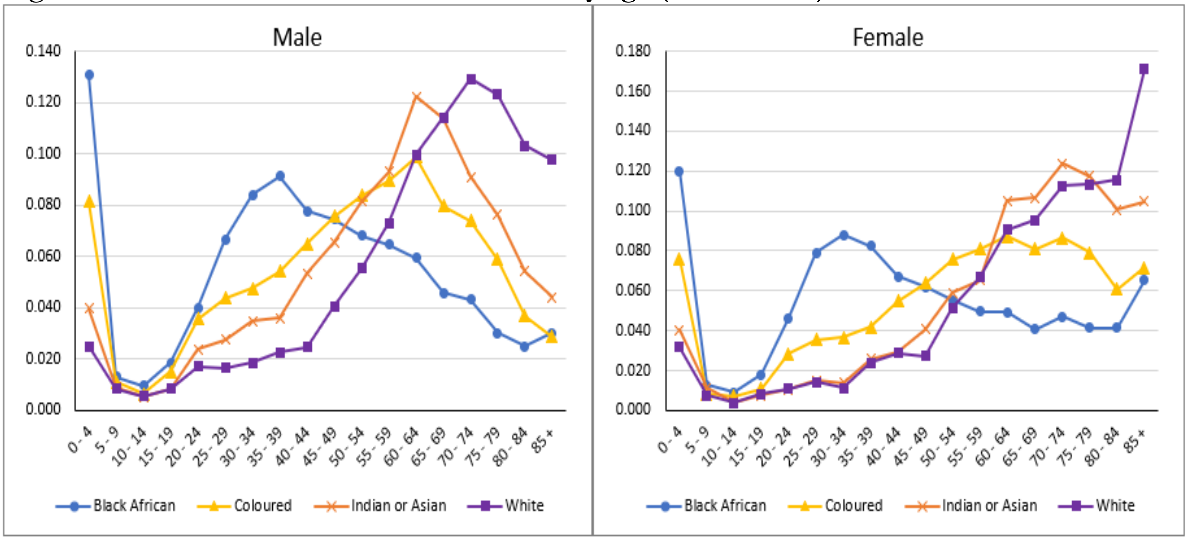
A different pattern is observed when looking at the number of deaths over time by population group for the non-African population groups (Figure 3.6). Although for the African population, deaths started levelling off round about 2010, for the White and Coloured population an uptick is observed from 2010 possibly because there was improvement in the completeness of reporting of the population group of the deaths (personal communication with Professor R. E. Dorrington).

For the African population, deaths also increased and started levelling off around about 2006 whereas for other population groups deaths were approximately level. This could be explained by fact that the prevalence of HIV is much higher for the African population, and also accessibility of ARVs to the African population was a challenge hence the increment in the number of deaths prior to 2006.

3.3.2 Household deaths from the 2011 census

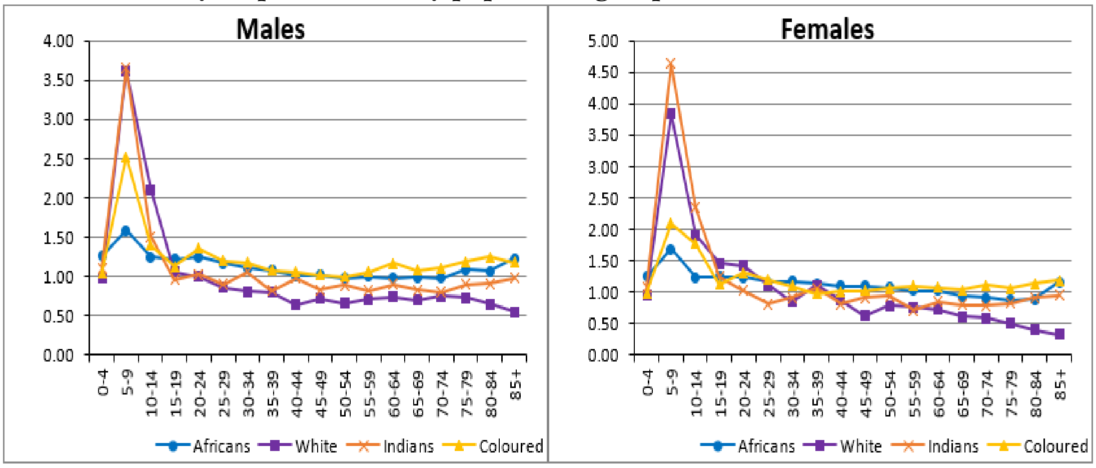
Figure 3.7 shows the distribution of deaths by age for each population group by sex reported by household from the 2011 census. The distribution of household deaths for all population groups is similar to that of the registered deaths in the intercensal period.

Figure 3.7 Distribution of household deaths by age (Census 2011)



Looking at the ratio of the proportion of household deaths to registered deaths (Figure 3.8) it can be seen that for all population groups, between ages 5 and 14 the ratios are significantly above 1. This might be explained by multiple reporting of child deaths by households due to extended households (as some of the children might live at more than one place for part of the time leading to their deaths being reported by more than one household (Dorrington, Moultrie and Timæus 2004)). For the White population the pronounced deviations below 1 are observed at older ages which is probably due to the fact that the census does not include retirement homes, whereas for the African and Coloured populations at ages 25 and above the ratios deviated only slightly from 1. In general, we do not expect these errors to distort the quality of our estimates as the age specific correction factors (obtained from the registered death data corrected for incompleteness) would be used to adjust household deaths for either under-or over-reporting.

Figure 3.8 Ratio of the proportion of household deaths to vital registration deaths in the year prior to 2011 by population group



3.4 Estimation of migration

In cases where net in-migration is ignored, the death distribution methods tend to underestimate the level of completeness and hence the number of deaths over the intercensal period would also be underestimated, and vice versa. Thus, adjustments to these methods should be made to allow for net in-migration when it is significant and this can be done if the estimates of net in migrants are available and sufficiently accurate.

In this research, we consider the net difference in the numbers of foreign-born between censuses to estimate net migration over the intercensal period (2001-2011 censuses). The outline of the method used to estimate migration is as follows:

First, we consider estimation of foreign-born immigrants who are still alive at the end of the intercensal period using:

$$Net\,M' = {}^{f}P_{x+10,t+10} - S_x \cdot {}^{f}P_{x,t} \text{ (which gives the number of surviving migrants)}$$

where:

$Net\ M'$ is the net number of foreign-born immigrants,

${}^{f}P_{x,t}$ and ${}^{f}P_{x+10,t+10}$ are the net number of foreign-born immigrants aged x to x+4 last birthday and x+10 to x+14 last birthday at the time of the 2001 and 2011 censuses, respectively.

S_x is the ten-year survival factor for migrants aged between x and x+5 at the time of the 2001 census.

The survival factors for the population as a whole were obtained by comparing the number of males (females) aged between x and x+5 (obtained by projecting the population using the CARe2011_160901 model) in 2001 to males (females) aged between x+10 and x+15 in 2011. It was then assumed that that the survival factors for the population as a whole apply to each of the population groups.

Second, we consider estimating the number of migrants alive at the start of the period (that is the 2001 census in this case) who will migrate during the intercensal period and this is done by applying reverse survival ratios (the reciprocal of forward survival ratios) to the number of foreign-born immigrants aged between x+10 and x+15 in the 2011 census, thus obtaining the number of those that would have been aged between x and x+5 years old at the 2001 census and then subtracting the actual number of those aged x to x+5 at the 2001 census from the resulting estimate, i.e.:

$$Net\,M'' = \frac{1}{S_x}\,{}^{f}P_{x+10,t+10} - {}^{f}P_{x,t}$$

However, the first approach tends to underestimate the number of migrants since it does not take into consideration the number of migrants who died during the intercensal period and the second approach tends to overestimate the number of migrants since it includes the number of migrants who would have died during the intercensal period. Thus, combining the two estimates (Average Survival Ratio Method) we obtain the correction factor at each age group, as it gives:

$Net\,\bar{M} = \frac{1+S_x}{2S_x} * Net\,\mathrm{M}'$, an estimate of the numbers who migrated.

The above method was used to estimate the net number of foreign-born immigrants for all population groups. The net number of African and Indian population group migrants was assumed to be roughly equal to the number of foreign born immigrants whereas for the Coloured population the net number of migrants was assumed to be insignificant. For the White population, the net number of migrants was estimated by deducting the estimates of net South African-born emigrants from estimates of the net number of foreign-born immigrants. Estimates of the net number of South African-born White emigrants (less those who returned) were obtained by from the online database (OECD. Stat)[9] by looking at the inflows and outflows of South Africans counted in 16 overseas countries recorded by the OECD[10] (as shown in Appendix B) receiving South Africans over the intercensal period and these were all

[9] The data were obtained from http://stats.oecd.org/ and it was assumed that these flows are for the year starting 1 July of the year in question).

[10] OECD represents Organisation for Economic Co-operation and Development.

assumed to be White. The figures obtained were for all ages in total and they were apportioned according to Stats SA (2016) percentage distribution of emigrants by age group.

3.5 Estimation of the level of completeness

3.5.1 Derivation of the level of completeness in the intercensal period by population group (2001-2011)

The level of completeness of reporting of deaths for each population group are derived using Death Distribution Methods (DDMs) that is, the GGB method and the SEG+δ method. This was done by first collating the number of registered deaths, net migration over the intercensal period and population data as reported by the 2001 and 2011 censuses. The GGB method was applied to the data over different age ranges so as to decide the age range over which the line was to be fitted. The age range used was determined by first setting the lower age to 5 and the upper to *A*-1 (*A* is the age of the open interval), and if a poor fit was observed then the age range was narrowed by decreasing *A* in 5-year steps.

For the SEG+δ method, the same age range used in the GGB method was used to estimate delta. The SEG diagnostic plots were investigated for any evidence of differential completeness of registered deaths with age and the level of completeness was estimated using ratios in the age range 25-64 so as to avoid the fluctuations at older ages and a drop of estimates at younger ages (as these age groups showed a departure from the average level of completeness).

However, for the African population it was decided to investigate whether the assumption that completeness is level by age for the African population as a whole hold since the proportions of the population urban and non-urban differ significantly by age and the level of completeness of registration of deaths in the rural areas is expected to be lower than it is in the urban areas. To investigate this, the level of completeness for the African population was estimated by applying DDMs separately to urban and non-urban populations, that is, treating African urban and African non-urban as two separate population groups. The paragraphs below indicate how data on urban and non-urban areas were obtained and how migration was estimated.

To obtain 2011 urban/non-urban population estimates it was decided first to classify municipal and metropolitan districts as either urban or non-urban and then aggregate the urban and non-urban district populations. Classification of districts was done by looking at the proportion of population in each district, districts that had at least 70 per cent or more of their individuals geocoded as urban were classified as urban (70 per cent was considered to be sufficiently urban as people would be near enough to services to register the deaths) and the remaining districts were classified as non-urban districts. The 2001 population estimates for these urban and non-urban districts (according to 2011 municipal district boundaries) were obtained from Super Web2, and were similarly aggregated to give urban/rural populations for 2001.

Registered deaths for the period 2001-2011 for the African population by district and sex were provided by South Africa Medical Research Council (through personal communication with Professor R. E. Dorrington).

Net internal migration and net immigration of foreigners were considered for estimation of net migration during the intercensal period, that is, foreign-born immigrants as well as local migrants. To estimate net immigration of foreigners, place of birth data were considered whereas for net internal migration (between urban areas and rural areas) data from the 2011 census on place of residence at previous census (that is, 2001 census) was used.

District boundaries changed during the intercensal period and are not the same as those in 2001, thus, to obtain the number of non-urban and urban foreign-born population estimates in 2001, municipal districts in 2001 were mapped on to either urban or non-urban municipal districts in 2011 and were thus classified accordingly. In all cases but two, mapping identification of municipal districts on to either rural or urban districts was clear. For these two districts, that is, Amatole (split into Buffalo City and Amatole in 2011) and Metsweding (absorbed into the Tshwane Metropolitan Municipality), further interrogation at municipality level was done. Most municipalities in the Amatole district mapped to rural municipalities in 2011, thus it was classified as a rural district and the Metsweding district was classified as urban since it lies in the Gauteng province which is essentially urban. Applying similar reasoning to 2001 census population counts produced population estimates that are consistent with those obtained from SuperWeb2 by 2011 census boundaries, thus, giving confidence to foreign-born estimates obtained.

To estimate intercensal internal migration, the 2011 census was used, as it identifies people who moved since the 2001 census at municipality level and these were grouped into municipal districts. The net number of internal migrants was obtained by differencing flows from rural to urban areas and vice versa. These were then added to the net immigration estimate of foreigners. Together the net in-migration into urban and into rural areas summed to the net immigration of the African population as a whole

3.6 Estimating annual completeness of registration of deaths in the year prior to the 2011 census

Annual estimates of the level of completeness of death registration for the White and Indian population groups during the intercensal period are derived by fitting a logistic curve to the level of completeness obtained for the intercensal periods of 1996 to 2001 and 2001 to 2011. The level of completeness for the intercensal period 1996 to 2001 is set to be consistent with the estimates derived by both Chinogurei (2017) adjusted to be relative to the population at 2011) and Richman (2017), each given the same weight. This is done on the assumption that the estimates of completeness for an intercensal period applies to the year in the middle of the intercensal period. For the Coloured population it was assumed that completeness was constant (at an average of 68 per cent) throughout the intercensal period since the level of completeness for the Coloured population obtained in the 2001-2011 intercensal period is similar to the level of completeness obtained by Richman (2017) in the 1996-2001 intercensal period.

Separate logistic curves of the form $C_x = \frac{L}{1 + Ae^{-k(x-x_0)}}$ were fitted to the estimates of completeness (using an Excel add-in, Solver) for the White and Indian populations, with the estimates for males and females combined into one data set (since the estimates of completeness of reporting for males and females are reasonably close to each other and there are no reasons for expecting a difference).

To derive annual estimates of completeness for the African population it was assumed that the Coloured, Indian and White deaths are completely reported and the estimates of completeness derived using DDMs are the estimates of completeness of reporting of population group of the deaths. For the African population group, it was assumed that completeness of reporting of population group and reporting of deaths are both incomplete. Following these assumptions, the true number of deaths of Africans were estimated to be the difference between true national number of deaths (obtained by scaling up vital registration deaths using national estimates of completeness obtained by Richman (2017)) and the aggregated true number of deaths for the White, Coloured and Indians (obtained by scaling up vital registration deaths using annual estimates of completeness of reporting of deaths). Annual estimates of completeness of registration of the African population were then calculated as a ratio of the vital registration to the estimated true number of African deaths the estimates produced.

3.7 Age-sex specific correction factors for correcting deaths reported by household at district level.

Age-sex specific correction factors for each population group, age and sex are estimated as the ratio of registered deaths in the year prior to the 2011 census, corrected for under-registration, to the number of deaths reported by household. These age-sex specific correction factors are to be used to correct the reported household deaths (for each population group) for under- or over-reporting of deaths at district level, based on the assumption that the age specific correction factors for each population group at the national level do not differ at a district level. The corrected household deaths for each population group are aggregated to obtain the total number of deaths in each district by sex.

3.8 **Adult mortality at district level in the year prior to the 2011 census**

Estimates of ${}_{45}q_{15}$ are obtained from age-specific mortality rates in each district and each sex. The age-specific mortality rates are obtained by dividing the number of deaths by estimated population at risk of dying in the year prior to the 2011 census. An abridged life table is estimated from the age specific mortality rates and thereafter estimates of ${}_{45}q_{15}$ are obtained using the equation: $1-\frac{l_{60}}{l_{15}}$.

In addition, the Brass relational logit model is used to smooth the age-specific mortality for each district using as standard the life table derived from the mortality rates from the CARe 3.2 (demographic projection model) for each province for 2011. Two estimates of mortality rates for each district are thus obtained, that is, crude/unsmoothed mortality estimates and graduated mortality estimates using corresponding provincial standards. To determine the final estimate to be used further for each province, the crude and graduated mortality estimates by province are compared to the provincial estimates of ${}_{45}q_{15}$ from orphanhood data that are derived using the orphanhood method. The estimate that gives the least sum of mean absolute percentage error is used.

The adult mortality estimates at provincial level are obtained by average weighting the district adult mortality estimates that lie in the same province to an estimate of $_{45}q_{15}$ for the province.

3.9 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 from the orphanhood data is to evaluate the estimates produced using the approach described above, understanding that the orphanhood estimates are not a 'gold standard' since they refer to a different period (the period between the two censuses) and not all parents live in the province of the respondent. Estimates of $_{45}q_{15}$ are to be obtained for each of the population groups and each of the provinces by sex.

From the 2001 and 2011 censuses, responses to questions relating to the survival status of the parents are obtained. The numbers of respondents who did not know the survival status of their parents are not included in the analysis, on the assumption that they reflect the same survival of parents as that implied by respondents who stated the survival status of their parents.

The standard life table used was obtained from the CARe 3.2 model (male and female life tables were averaged to give the final standard life table), which was chosen over the default standard life tables in the workbook to account for differences in mortality distributions between provinces and the provincial-specific experience of HIV on mortality. In provinces where the CARe 3.2 model produced the logit fits that did not fit the data appropriately, the Brass General Standard with beta set to 1 was used (namely, Mpumalanga, Gauteng, Free State and Eastern Cape).

The mean age at child-bearing for the women was obtained from the average of the mean age at their childbearing from the 2001 and 2011 censuses for each of the provinces. For the men, the mean age of the father at conception of the child was obtained by adding seven years, the default difference between fathers and mothers, to the mean age of mothers as suggested in the United Nations (2002) report.

After obtaining the input data for the orphanhood method, the orphanhood workbook (Timæus 2013a) was used to estimate the level of mortality at provincial level.

4 RESULTS

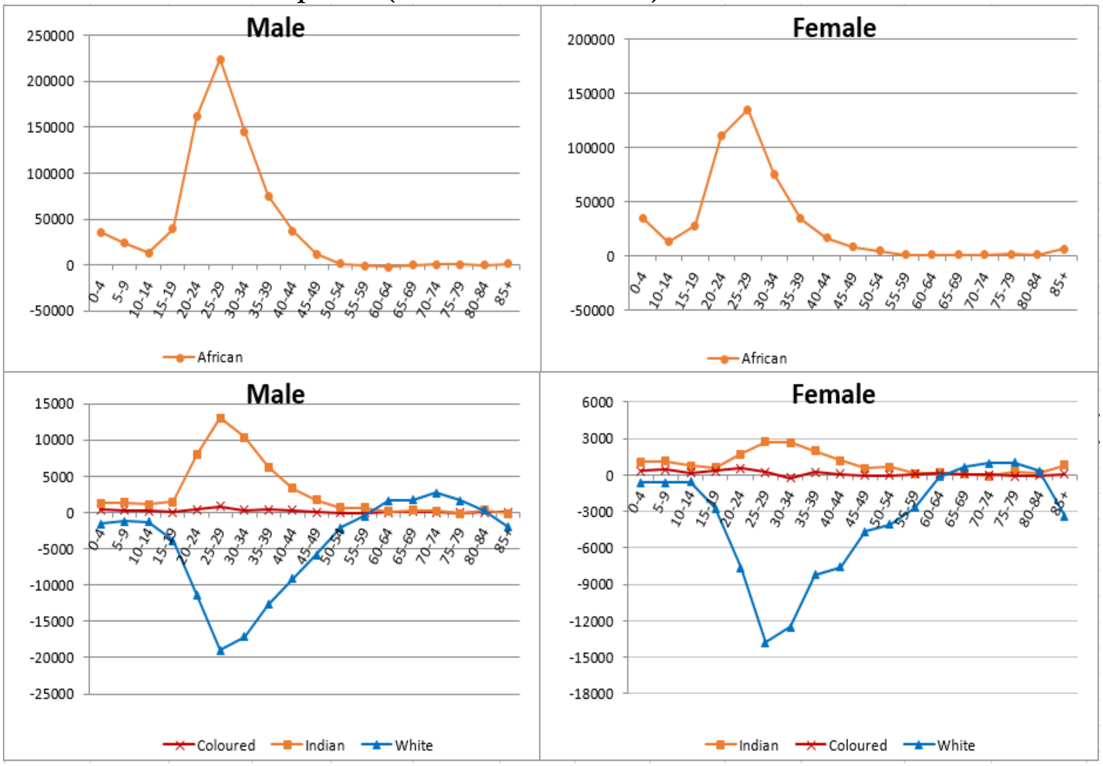
This chapter consists of 3 sections which aim to present the estimates of adult mortality derived at district level in South Africa in the year prior to the 2011 census. The first section deals with migration estimates. It also concentrates on the estimates of completeness of registered deaths by population group for each sex in the intercensal period of 2001 to 2011 and the estimates of the level of completeness in each consecutive year. The second section concentrates on the age-sex specific adjustment factors for adjusting household deaths and the estimates of adult mortality at district level. The third section assesses the reasonableness of the estimates of the mortality, where the estimates are compared to estimates of adult mortality at district level in 2001 derived by Chinogurei (2017) and childhood mortality estimates derived by Zewdie (2014).

4.1 Migration estimates

The level of net immigration of the African population was much higher than the other population groups as indicated in Figure 4.1. The distribution of migrants by age for the Africans and Indians is similar though the level of migration is different. For these two population groups, peaks are observed at economically active ages (that is, between ages 15 and 49) for both males and females. However, it can be noted that the number of African male migrants aged between 20 and 34 is approximately twice the number of female migrants and of interest, the difference between the number of Indian male and female migrants at ages 20 to 34, where the number of male migrants is approximately three times higher than the female migrants.

Net migration loss is observed for the African male migrants aged between 55 and 69 and this could possibly be explained by foreign-born migrants returning home due to retirement or ailments. An uptick for the female population is observed for both the Indian and African population group at ages 85+ which might be due to age exaggeration.

Figure 4.1 Net migrants estimates by age, sex and population group during the intercensal period (2001 and 2011 census)



Different migration patterns are observed for the Coloureds and Whites. The number of migrants for the Coloured population are so low as to be trivial, thus it was assumed that migration for that population group was insignificant. Net migration loss is observed for the White population particularly at economically active ages as well as at ages 85 and above, with the highest number of emigrants at ages between 20-29, and this could be due to out migration related to initial employment, education or violent crime in South Africa and possibly for those aged 85 and above it could be due to parents joining their children overseas.

4.2 Estimation of the level of completeness

Application of the GGB method to each of the population groups using the age ranges shown in Table 4.1 indicates points in the diagnostic plots that lie close to the straight line with the exception of Coloured and Indian males. The estimates derived from the GGB method were then used mainly as a check on the reasonableness of the estimates obtained from the SEG method.

The results produced by the SEG methods are as shown in Figure 4.2-Figure 4.5 and Table 4.1. For all population groups, the level of completeness was estimated on the age range 25-64, as the estimates were relatively uniform over that age range and a decline in completeness at younger ages was observed whereas at older ages the estimates were quite volatile, thus, it was decided to use the 25-64 age range in all cases.

Table 4.1 Estimates of the level of completeness by population groups, GGB and SEG method

	Males			Females		
	SEG method		GGB method		SEG method	GGB method
	Age	Level	Level	Age	Level	Level
African	5-84	63%	61%	5-84	64%	64%
Coloured	5-79	68%	68%	5-84	64%	65%
Indian	5-84	85%	84%	5-84	85%	87%
White	5-84	71%	70%	5-79	78%	79%
National	5-84	91%	94%	5-84	94%	94%

The SEG diagnostic plots indicate fluctuations at older ages, particularly for the African population which could be due to age misstatement in the census population. Also, for the African population there was a drop in the level of completeness at ages around 60-64, and this could perhaps be due to men retiring to their rural homes, leading to a reduction in the number of deaths reported.

Figure 4.2 SEG diagnostic plots for completeness of registered death for the African population group.

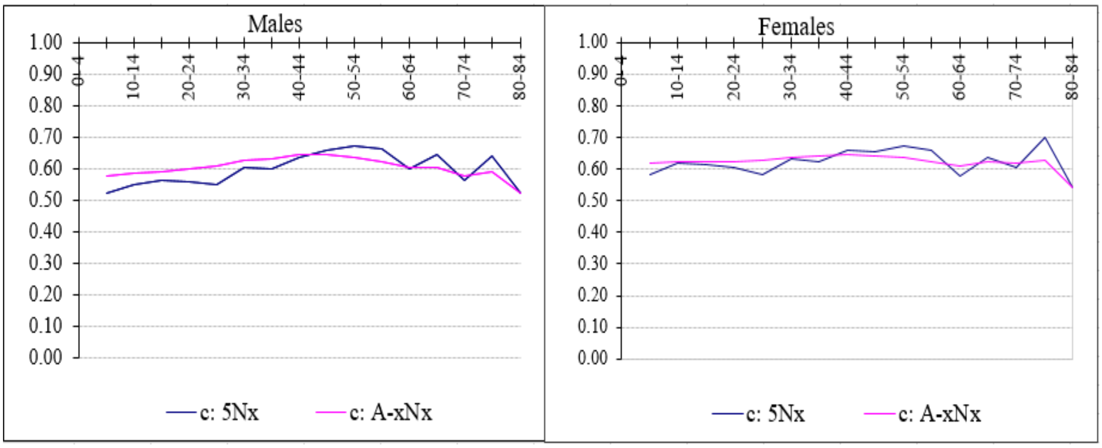


Figure 4.3 SEG diagnostic plots for completeness of registered death for the Coloured population group

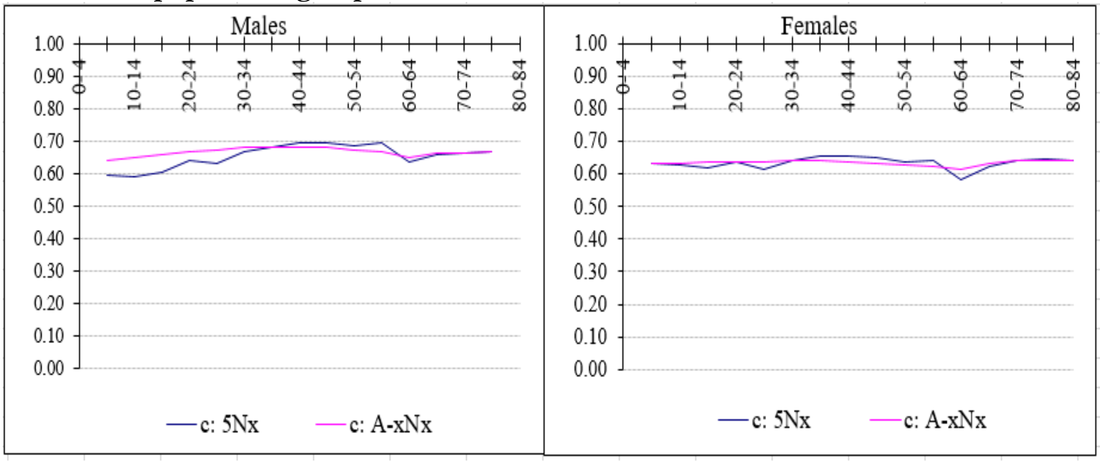


Figure 4.4 SEG diagnostic plots for completeness of registered death for the White population group

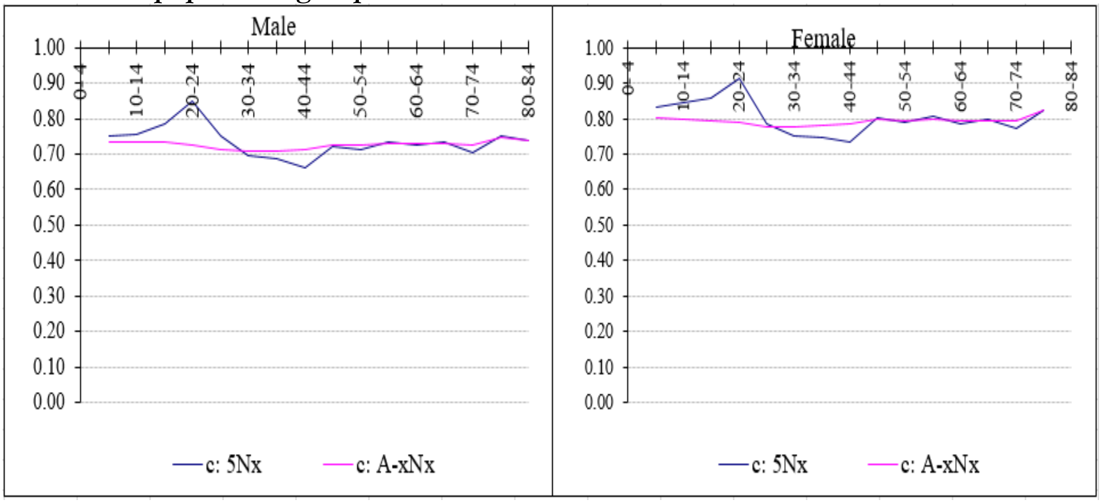
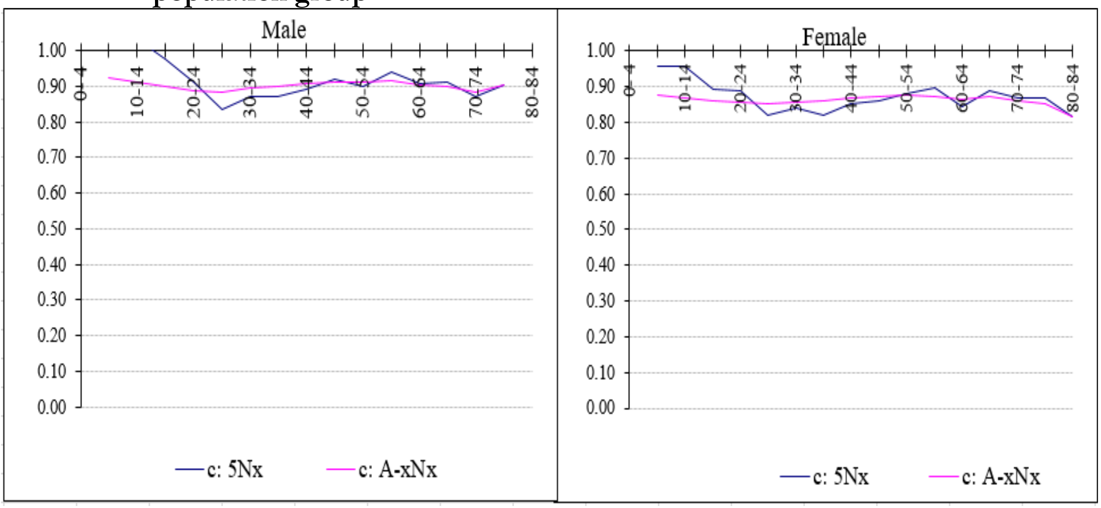


Figure 4.5 SEG diagnostic plots for completeness of registered death for the Indian population group



A drop in completeness in the 5-19 age range is observed for all population groups except for the Indian male population which possibly indicates that there is more likely to be a problem with the population estimate (at least for the Coloured, Indian (females) and the White populations) or possibly it could be explained by underreporting of deaths in this age range (for the African population). Underreporting of deaths might be due to the fact that there are generally less incentives to report these deaths, unlike an adult death which may be important to process things such as wealth re-distribution and insurance among others.

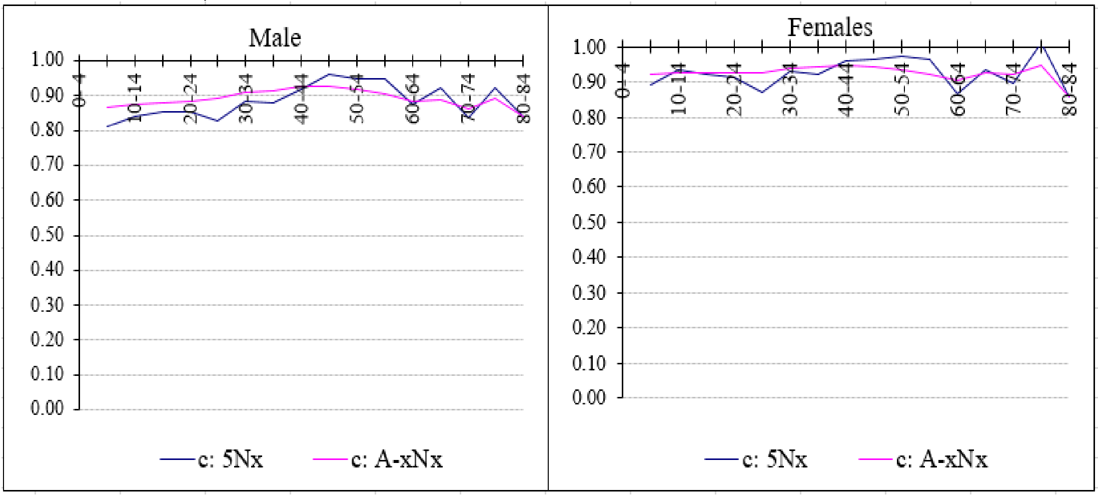
The level of completeness for males and females are close to each other for all population groups (except for the White population where estimates of female are about 7 per cent higher than the male estimates) which validates the estimates (as highlighted in past studies that the results would be sufficiently close to validate the estimates if the male and female estimates are similar).

In addition, estimates obtained from both methods, that is, the GGB method and the SEG+delta method, are consistent for both males and females which also, supports the results.

In general, though fluctuations were noted in the level of completeness across all age groups, these seem to be relatively small enough to suggest that completeness is constant by age (for adult ages), consistent with this observation are the GGB diagnostic plots as they indicate a straight line when fitted to all age groups.

Fluctuations were noted from the diagnostic plots of both males and female as indicated in Figure 4.6. At younger ages, the level of completeness appeared to be decreasing up to ages around 25. Thus, to avoid biases ages 25 to 64 were used in estimating the level of completeness for the population as a whole. Estimates produced are given in Table 4.1.

Figure 4.6 Application of SEG method to the national population as a whole (2001-2011)



As explained in section **Error! Reference source not found.**, it was decided to investigate if urban and non-urban African population differ sufficiently to treat separately. The results produced by investigating whether the assumption that completeness is level by age for the African population as a whole is valid were as shown in Figure 4.7, Figure 4.8 and Table 4.2

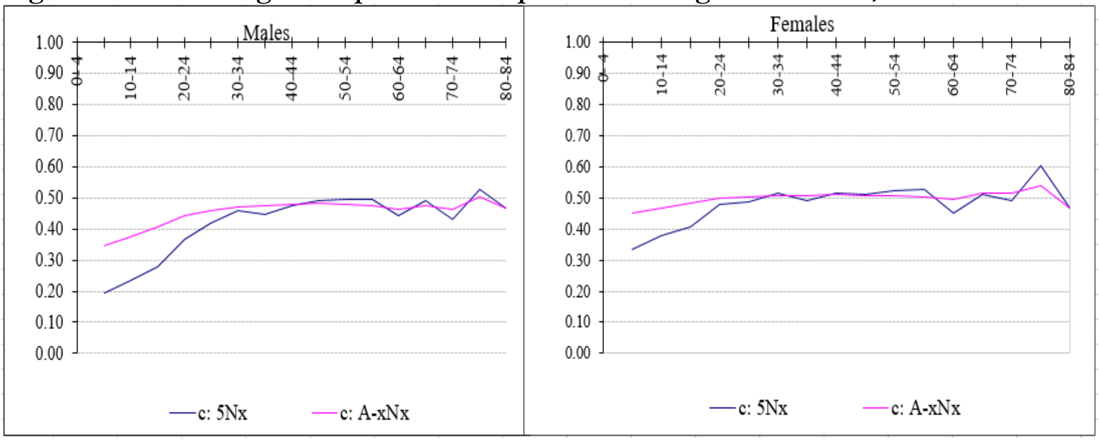
Table 4.2 Estimates of the level of completeness and age ranges for the African population GGB and SEG method[11]

	Males			Females		
		SEG	GGB		SEG	GGB
	Age	Level	Level	Age	Level	Level
Urban	25-84	82%	82%	25-84	83%	88%
Non- urban	25-84	46%	45%	25-84	51%	51%

The SEG diagnostic plots indicate fluctuations at older ages, for both urban and rural areas and this might be due to age misstatement in the census population. A drop in the level of completeness at ages below 25 is observed for the rural population which is in contrast with the urban population where the levels of the estimate of completeness are higher for those aged below 25. This could be due to inaccurate data on migration at these ages. In order to avoid distortions, it was decided to use ages above 25 to determine the level of completeness in all instances as indicated in Table 4.2.

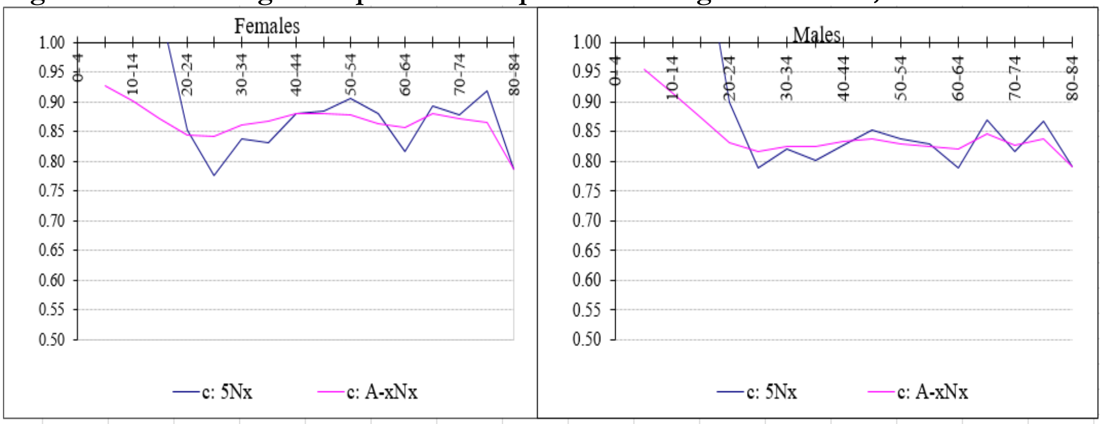
[11] To investigate if urban or non-urban rural differ sufficiently to be kept separate, it was decided to treat African urban and African non-urban as two separate population groups.

Figure 4.7 SEG diagnostic plots for completeness of registered death, non-urban



The level of completeness for males and females are close to each other, as indicated in Table 4.2 except for non-urban males and females with male level of completeness lying below 50 per cent.

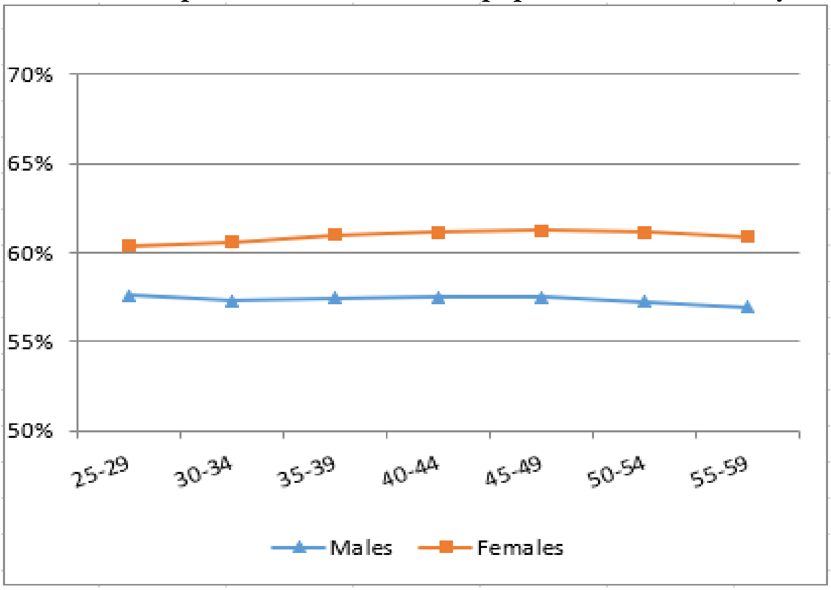
Figure 4.8 SEG diagnostic plots for completeness of registered death, urban



Of interest is the huge difference in the level of completeness between ages 20 to 64, with at least 25 per cent difference between the urban population and the rural population, indicating that estimation of the level of completeness for the African population as a whole might violate the underlying assumption of DDMs of constant underreporting by age.

However, two sets of deaths, that is, the vital registration deaths and the true number of aggregated urban and non-urban deaths estimated using the level of completeness obtained for the urban and non-urban African population by sex were compared, as shown in Figure 4.9. It can be observed that the ratio of these two is quite level by age, thus, although there were some variations as indicated in the results of DDMs, they probably would not make much difference to the estimates of adult mortality of the African population as a whole, and thus it was decided to prefer the estimates derived using the level of completeness for the African population as a whole rather than using the two separate groups (urban and non-urban).

Figure 4.9 True aggregated rural and urban deaths as a ratio of vital registration deaths expected from the African population as a whole by sex.



4.3 Estimating annual completeness of registration of deaths in the year prior to the 2011 census

The parameters of the logistic curves found by minimising the absolute percentage error (for the Indian and White population) and the estimates obtained as well as the fitted logistic curves were as shown in Figure 4.10, Figure 4.11 and Table 4.3. No parameters of the logistic curves were derived for the Coloured (as it was assumed that completeness was constant (at an average of 68 per cent) throughout the intercensal period) and for the African population annual estimates were derived from the estimates of completeness for the White, Indian and Coloured population as mentioned in section **Error! Reference source not found.**.

Table 4.3 Fitted logistic curve parameters

	Indian	Whites
L	0.92	0.77
k	1.37	0.65
A	0.366	1.83
x0	1999	1998

Figure 4.10 Estimates of completeness, Whites

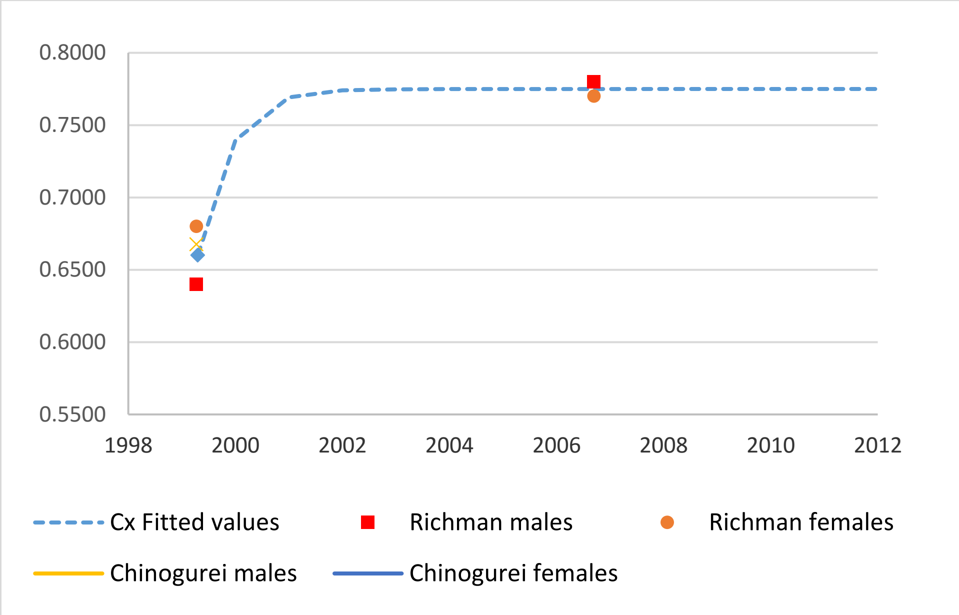
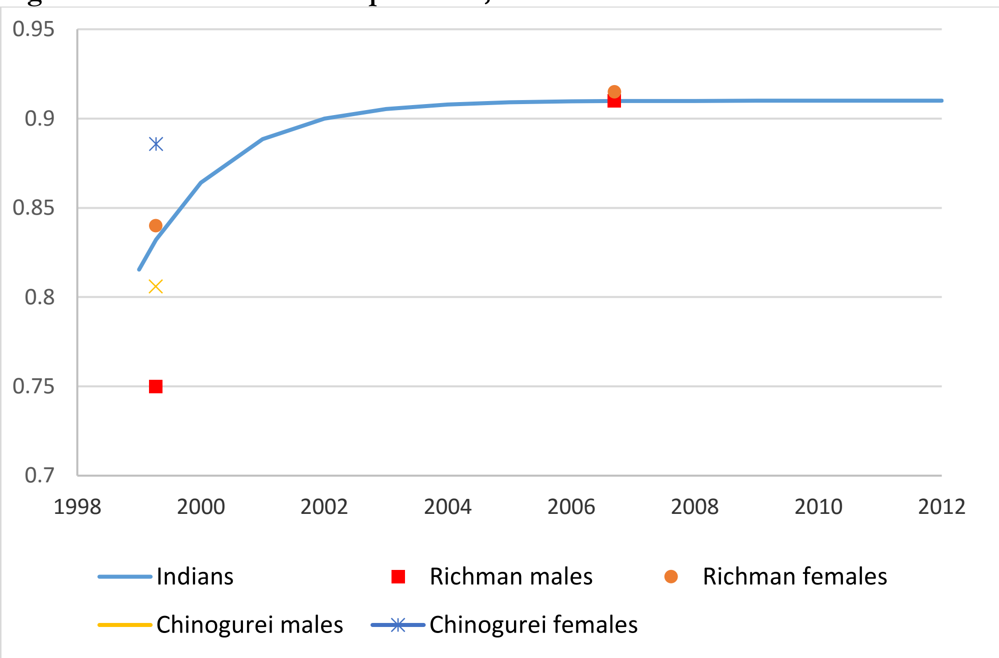


Figure 4.11 Estimates of completeness, Indian



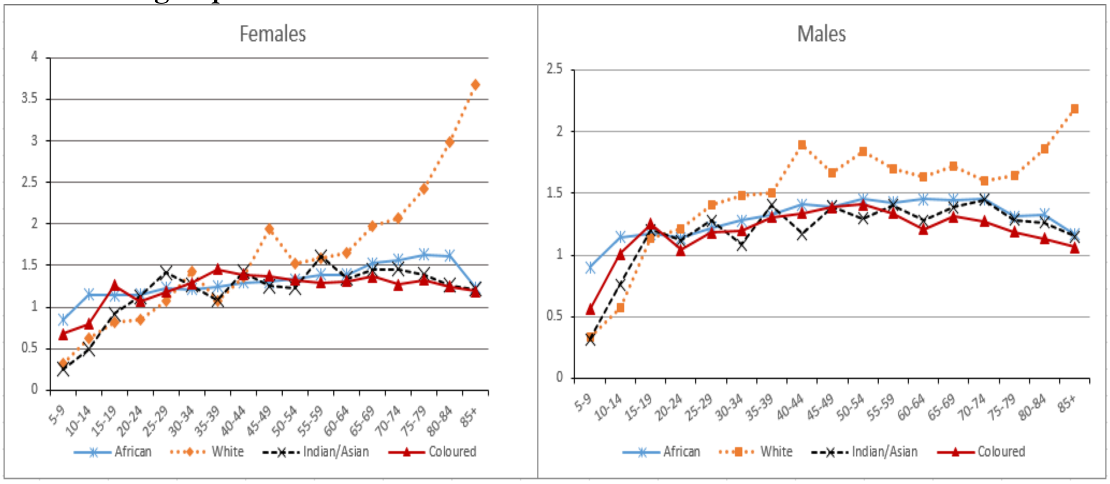
The annual estimates of completeness obtained for the Coloured, Indian and White population were adjusted to allow for the noticeable decrease in the number of deaths missing population group from the year 2010 onwards (thus, indicating that the level of reporting of population group increased in 2010) and a drop in the level of completeness of reporting of about 1.5% for the years 2011-2013 (personal communication, Prof. R. E. Dorrington). To obtain the level of adjustment, the number of deaths that would have been expected to be registered in 2011 was projected based on the assumption that the number of deaths that were registered as occurring in each year increased at a constant rate per annum, and that rate was obtained by fitting an exponential function to the number of deaths in 2001 to 2009. Estimates of completeness in the year prior to the 2011 census were then adjusted by the percentage difference in the number of vital registration deaths compared to the expected number of deaths. The percentage change in deaths reported with population group is 7 %, -5 % and 4% for the Whites, Indians and Coloured respectively and the adjusted estimate of completeness in the year 2011 (see Appendix C). These estimates appear to be reasonable and are consistent with the estimates derived by Richman (2017)

The annual estimates of completeness of registration of the African population were as shown in Appendix C. The estimate of completeness for the year 2007 (63 percent) is consistent with the average estimate of males and females produced using DDMs, that is 64 per cent in 2007.

4.4 Age-sex specific factors for correcting deaths reported by household at district level.

Figure 4.12 shows the age-sex specific adjustment factors. The adjustment factors indicate that deaths for all ages are under-reported with the exception of deaths of young adults, which is perhaps not surprising because young adults are regarded as living in extended households thus leading to multiple reporting of deaths. Second, the adjustment factors for the White population tend to be much higher, particularly at the oldest ages which could be explained by the fact that many deaths occur within retirement homes and as a result they are not reported by households in the census. Third, for the Coloured and African population groups, the ratios tend to be slightly closer to 1 at ages greater than 15, rising with age for the African females. Though, in general, for the African population the ratios appear to be consistent they show that there is under-reporting of household deaths. Indian adjustment factors for both men and women indicate an erratic pattern, which could be due to the small sample size of the Indian population is small, leading to fluctuations.

Figure 4.12 Ratio of adjustment factors in the 2011 census by age for each population group and sex



4.5 Adult mortality at district level in the year prior to the 2011 census

Table 4.4 shows the mean absolute percentage error of the crude and the graduated district mortality estimates compared to those of the province derived from orphanhood data and the conclusion on whether to use crude or graduated estimates in each province (and assumed to apply for each district estimate as well). In general, it can be noted that the deviation of the graduated rates was lower than that of ungraduated rates and that the estimates of ${}_{45}q_{15}$ for women deviated more from the orphanhood estimates (Table E1, Appendix E) than those for men.

Table 4.4 Mean absolute percentage error of ungraduated and graduated district estimates of ${}_{45}q_{15}$ to those of the province derived from orphanhood data

	Male			*Female*		
Province	*Ungraduated*	*Graduated*	*Decision*	*Ungraduated*	*Graduated*	*Decision*
WC	17.2%	16.9%	Graduated	19.0%	22.5%	Ungraduated
NC	17.1%	15.2%	Graduated	23.9%	23.5%	Graduated
EC	23.0%	20.4%	Graduated	21.6%	21.0%	Graduated
FS	10.6%	9.9%	Un-graduated	16.9%	16.3%	Graduated
KZN	18.5%	19.0%	Graduated	17.0%	16.0%	Graduate
MP	18.5%	18.0%	Graduated	16.8%	18.4%	Ungraduated
LP	10.1%	8.2%	Graduated	18.8%	17.2%	Graduated
NW	16.9%	15.1%	Graduated	24.2%	22.5%	Graduated
GT	30.1%	28.5%	Un-graduated	19.1%	19.4%	Graduated

The logit fits to the district data indicate that the CARe model standards that work for the province does not work for some of its districts (see Appendix D) particularly where there are districts that are non-urban that lie in provinces that are predominantly urban. For instance, inspection of the logit fit to a district such as Central Karoo indicate that the standard does not fit well and this could be explained by the fact that this district lies in a predominantly urban province, that is, the Western Cape.

Further inspection of logit fits in the Western Cape indicate that the standard fitted all the other districts with the exception of West Coast (for females) and Eden, which could be explained by the fact that these districts have small population leading to random fluctuations.

In Limpopo, it can be noted that the standard did not fit properly in Greater Sekhukhune and Waterberg, but this might be due to the fact that these districts have a different mortality shape compared to the standard model. An HIV hump is observed at younger ages, (though, however, the standard fits well at older ages) which implies that the standard exhibits high mortality rates at younger ages whereas in these districts there are low prevalences of HIV/AIDS.

A similar pattern is observed in two districts in the North West province that is, Ngaka Modiri Molema and Dr Ruth Segomotsi district and also for two districts in the Free State (Thabo Mofutsanyane and Mangaung). This might be due to the fact that either the standard exhibits a different shape of mortality to that of the districts or the racial composition differs between districts that lie within the same province (as the bulk of deaths in the White population occur at the ages over 60, which have limited impact on ${}_{45}q_{15}$). Inspection of the female logit fits of these districts indicate worse fits compared to male fits.

In addition, it could be noted that most of the districts in KZN (with the exception of eThekwini), Eastern Cape (Cacadu, Amathole, Chris Hani) and Northern Cape (Namakwa and Pixley KaSeme) the standard model fits at older ages and not at younger ages, as the logit fit exhibit an HIV-AIDS hump particularly at young adult ages. This might be explained by the fact that there are higher mortality rates at young adult ages than the projected standard mortality estimates at these ages.

On the other hand, logit fits of districts like Siyanda (the Northern Cape) and O.R Tambo (Eastern Cape), the standard fits well at younger ages and not at older ages indicating that there might be a problem with the data at older ages in these districts.

In general, the assumption that the provincial standard could be used at district level does not hold for some districts and as a result, this might distort the estimates produced for the districts where the standard does not fit well. Thus, the results need to be interpreted with care.

The ranking of districts in 2011 for male and female was compared to the ranking of districts by Chinogurei (2017) in 2001 and by and large the rankings were consistent. Most of the districts in 2001 that were amongst the 15 districts with lower mortality were also observed to be amongst the 15 lowest mortality districts in 2011 for both males and females.

However, a different case was observed for districts with highest mortality as only a few districts were found to be amongst the 15 highest mortality districts in 2011. This could possibly be due to the fact that rolling out of anti-retroviral drugs in some provinces was slower than to other provinces (Johnson 2012). The percentage differences between the 2001 and 2011 estimates of adult mortality are shown Table E2, Appendix E. These indicate that there appears to have been an improvement in the mortality particularly in KZN province, where for iLembe, eThekwini, uGu, uMkhanyakude, uMgungundlovu mortality decreased by 13 percent or more for males from 2001 to 2011.

In general, it could be observed that mortality in most districts improved for males than for females. Of interest, are the percentage differences in Eastern Cape with some of the districts indicating a drop in mortality and other districts indicating elevated mortality levels from 2001 to 2011 (suggesting that rolling out of ARVs in this province was not uniform). For instance, Alfred Nzo recorded the most negative percentage difference for both males and females, whereas Buffalo Metropolitan (both males and females) and OR Tambo (males only) indicated a significant positive percentage change in mortality (Table E2, Appendix E) between the 2001 and 2011. The correlation coefficients of male and female rates were 80 and 77 per cent respectively, indicating a degree of consistency between the 2001 and 2011 adult mortality estimates. Thus, consistency in both the level and ranking of districts between estimates derived in this study and those of Chinogurei (2017) gives some confidence in our estimates.

Childhood mortality estimates (for both sexes) at municipality level derived by Zewdie (2014) using the 2011 census were weighted by estimates of the population aged between 0 and 5 years to obtain mortality estimates at district level as shown in Table 4.6. These estimates were compared to adult mortality estimates for both sexes combined. Aggregated males and females true number of household deaths (corrected using the age-sex correction factors) were divided by the sum of exposure to obtain adult mortality estimates at district level. The ranking of adult and child mortality estimates are as shown in Table 4.5. It can be observed that all of the districts that lie in

Western Cape, Gauteng and Limpopo province recorded lowest level of adult mortality (males and females combined) and this was consistent with the ranking of childhood-mortality estimates. Also, districts with the highest levels of mortality ranked consistently for adult and child mortality estimates, as most districts with high levels of mortality observed to be in KZN (with the exception of eThekwini), Free State and Eastern Cape, with a few disparities). In general childhood mortality was consistently higher in most districts than adult mortality which is to be expected since children are more prone to HIV-Aids (and HIV infection often progresses quickly to AIDS in children than adults). Thus, to some extent, this shows that estimates derived in this study are in line with estimates from other studies.

Table 4.5 Comparison of 20 lowest and highest-ranking of districts for adult mortality against child mortality in South Africa in 2011

	Lowest mortality				*Highest mortality*			
	45q15		*U5MR*		*45q15*		*U5MR*	
Rank	*District*	*Province*	*District*	*Province*	*District*	*Province*	*District*	*Province*
1	Overberg	WC	CPT	WC	uThungulu	KZN	OR Tambo	EC
2	Cape Winelands	WC	Eden	WC	Pixley kaSeme	NC	uMgungundlovu	KZN
3	West Coast	WC	Overberg	WC	Mangaung	FS	Dr Kenneth Kaunda	NW
4	Tshwane	GT	Cape Winelands	WC	Xhariep	FS	iLembe	KZN
5	CPT	WC	Tshwane	GT	Cacadu	EC	Joe Gqabi	EC
6	Eden	WC	JHB	GT	John Taolo	NC	Lejweleputswa	FS
7	JHB	GT	Nelson Mandela	EC	Ngaka Modiri	NW	Ugu	KZN
8	Sedibeng	GT	Vhembe	LP	Lejweleputswa	FS	Ngaka Modiri	NW
9	Vhembe	LP	West Coast	WC	OR Tambo	EC	Sisonke	KZN
10	Ekurhuleni	GT	Central Karoo	WC	UMzinyathi	KZN	Xhariep	FS
11	West Rand	GT	Buffalo Metro.	EC	Dr Ruth Seg	NW	uThungulu	KZN
12	Waterberg	LP	Capricorn	LP	Sisonke	KZN	Alfred Nzo	EC
13	Central Karoo	WC	Mopani	LP	uThukela	KZN	uMkhanyakude	KZN
14	eThekwini	KZN	Sekhukhune	LP	Amathole	EC	Thabo Mofutsanyana	FS
15	Nelson Mandela	EC	Namakwa	NC	Zululand	KZN	Gert Sibande	MP
16	Mopani	LP	Ekurhuleni	GT	Amajuba	KZN	Amajuba	KZN
17	Bojanala P	NW	West Rand	GT	Thabo Mofutsanyana	FS	Dr Ruth Seg	NW
18	Capricorn	LP	Sedibeng	GT	Chris Hani	EC	uThukela	KZN
19	Sekhukhune	LP	Waterberg	LP	Joe Gqabi	EC	Zululand	KZN
20	Siyanda	NC	Francis Baard	NC	Alfred Nzou	EC	uMzinyathini	KZN

The correlation coefficient of adult and childhood mortality estimates is 80 per cent, indicating a reasonable level of consistency between the two sets of estimates.

It can be observed that men have higher mortality than women as would be expected. Looking at Table 4.6 it can be noted that for men, the level of mortality is lower in districts in Western Cape, and Gauteng, the same holds for women, though, for women districts in Limpopo also have lower mortality. Most of the districts in KZN and the Eastern Cape and to a lesser extent in the Free State had higher mortality for both men and women. Of particular interest, is the rate of mortality in uMkhanyakude district (falling within KwaZulu-Natal province), as it was ranked among one of the lowest district, which was somewhat surprising considering the fact that the level of HIV prevalence in KwaZulu-Natal is high. However, comparison of mortality estimates derived in this study for uMkhanyakude district to estimates derived by African HDSS[12] (since uMkhanyakude is the district into which the African Centre Health and Demographic Surveillance System falls) indicates relative consistency for both males and females, thus lending confidence to estimates obtained for this district.

It can also be seen from Table 4.6 that there is inconsistency between the levels of $_{45}q_{15}$ among the districts of the same province with the clear exception of Western Cape, Gauteng and KZN. Great disparities in adult mortality are observed in districts in Northern Cape, with Namakwa having lower level of mortality compared to other districts. Similarly, districts in the North West and KZN (despite having registered the highest mortality rate), that is Bojanala, uMgungundlovu and eThekwini, respectively, have lower mortality levels compared to other districts in these provinces. Spatial disparities between districts are also observed in the Eastern Cape, with most districts indicating high mortality and few reflecting lower mortality.

[12] Adult mortality eestimates (*45q15*) obtained by African HDSS for uMkhanyakude district were 56.1 and 37.3 for males and females respectively. These estimates were obtained from INDEPTHStats (http://www.indepth-ishare.org/indepthstats/index.php/download-desktop-version)

Table 4.6 District level estimates of adult and under-five mortality rates

				Both Sexes Combined	
Province	**District**	**Males**	**Females**	**U5MR**	$_{45}q_{15}$
WC	WestCoast	30,5	20,6	34,6	26,5
WC	CapeWinelands	30,0	21,1	30,7	25,5
WC	Overberg	26,4	18,2	30,4	22,7
WC	Eden	30,4	22,1	29,3	28,1
WC	CentralKaroo	41,3	35,9	37,3	39,0
WC	CPT	32,7	21,5	24,0	26,9
NC	Namakwa	40,3	26,8	42,9	46,7
NC	PixleykaSeme	54,2	48,7	48,7	52,3
NC	Siyanda	45,8	39,2	50,3	43,6
NC	FrancisBaard	54,1	41,3	47,3	48,8
NC	JohnTaoloGaetsewe	59,6	45,9	59,5	53,9
EC	Cacadu	44,6	38,4	52,6	53,6
EC	Amathole	71,9	53,3	62,2	62,6
EC	ChrisHani	73,2	58,1	62,9	65,4
EC	JoeGqabi	70,2	57,4	70,4	65,5
EC	ORTambo	65,4	47,3	63,6	57,9
EC	AlfredNzo	80,2	58,7	76,2	70,8
EC	BuffaloMetropolitan	51,6	37,8	38,6	44,2
EC	NelsonMandelaBayMetropolitan	45,8	46,9	32,3	40,2
FS	Xhariep	66,5	53,4	74,8	53,5
FS	Lejweleputswa	58,9	52,7	72,8	56,3
FS	ThaboMofutsanyana	72,2	57,3	78,7	64,9
FS	FezileDabi	54,0	44,8	60,2	49,2
FS	MangaungMetropolitan	60,3	44,7	57,8	52,7
KZN	Ugu	64,7	45,3	73,0	50,1
KZN	uMgungundlovu	59,1	41,0	64,0	50,2
KZN	uThukela	71,5	49,9	89,9	61,2
KZN	uMzinyathi	73,2	42,2	98,6	58,4
KZN	Amajuba	72,8	54,8	84,4	63,1
KZN	Zululand	75,7	51,2	97,7	62,9
KZN	uMkhanyakude	62,7	32,1	77,8	44,6
KZN	uThungulu	66,3	41,5	76,1	51,5
KZN	Sisonke	72,8	49,5	74,4	59,9
KZN	iLembe	58,8	42,6	69,1	51,2
KZN	eThekwini	46,6	34,7	50,8	39,8
MP	GertSibande	56,3	49,0	79,4	47,7
MP	Nkangala	49,0	42,2	54,2	45,9
MP	Ehlanzeni	63,0	39,6	55,3	46,0
LP	Mopani	50,5	33,4	42,2	41,4
LP	Vhembe	45,8	25,3	33,9	33,5
LP	Capricorn	53,0	35,6	40,5	43,0
LP	Waterberg	50,9	33,9	46,6	38,7
LP	Sekhukhune	65,7	36,9	42,7	43,3
NW	BojanalaPlatinum	43,3	39,9	48,2	41,5
NW	NgakaModiriMolema	59,6	48,5	73,7	54,1
NW	DrRuthSegomotsiMompati	64,2	51,1	84,9	58,5
NW	DrKennethKaunda	52,2	42,7	68,0	48,0
GT	Sedibeng	45,7	39,8	45,6	31,4
GT	WestRand	37,1	33,5	44,7	35,4
GT	Ekurhuleni	37,2	31,1	43,7	34,2
GT	JHB	33,2	25,7	32,2	29,4
GT	Tshwane Metropolitan	30,6	23,0	31,3	26,8

5 DISCUSSION OF RESULTS

5.1 Introduction

This chapter reflects on whether the method proposed by Dorrington, Moultrie and Timæus (2004) produced reasonable estimates at district level through discussing estimates of adult mortality at district level obtained from this study. It then highlights drawbacks that must be taken into account when interpreting the results in the research and thereafter suggests possible areas of further research. Finally, overall conclusions on whether the method suggested by Dorrington and Timæus (2015) produces plausible estimates at sub-provincial level are drawn.

5.2 Discussion of results

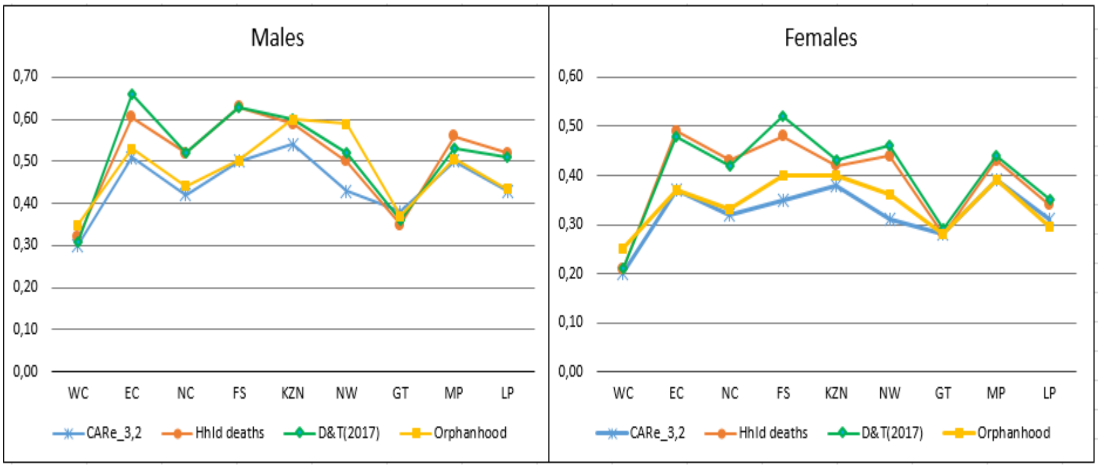
The overall objective of this research was to assess if the method proposed by Dorrington and Timæus (2015) produces reasonable estimates at district level. Thus, the first important step is to see how the estimates derived in this study compare with those presented elsewhere. In this regard, national level estimates are compared to estimates derived by and the orphanhood estimates. Provincial level estimates are compared to estimates derived by Dorrington and Timæus (2017), the CARe_3.2 model and the orphanhood method by sex (see Table 5.1), as well as estimates from SANBD report (Msemburi, Pillay-van Wyk, Dorrington *et al.* 2014) for both sexes combined. District estimates of adult mortality are compared to estimates derived by Chinogurei (2017) as well as estimates of childhood mortality derived by Zewdie (2014).

Table 5.1 Estimates of adult mortality ($_{45}q_{15}$) at national level from different sources

	Males	Females
Dorrington and Timaeus (2017)	0.49	0.39
SEG method (this project)	0.49	0.37
Orphanhood method	0.52	0.39
CARe_3.2	0.43	0.32

Comparison of estimates derived at the national level from different sources indicate that the male mortality estimates from orphanhood method are higher than the other estimates at national level. This could be due more to deficiencies in the data on orphanhood rather than a problem with the method. This could also be because orphanhood estimates apply to the year in the middle of the intercensal period, that is, 2006.7 and not the 12 months prior to the 2011 census.

Figure 5.1 Comparison of adult mortality estimates (${}_{45}q_{15}$) from different sources against estimates derived in this study



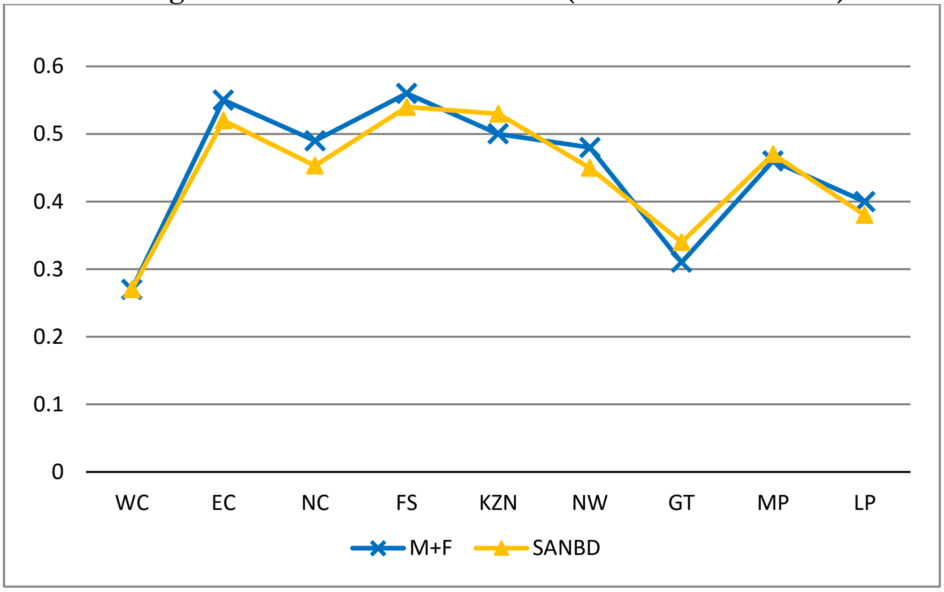
The first thing to note is that provincial estimates derived by Dorrington and Timæus (2017) are relatively consistent (both in level and ranking) to estimates produced in this study. This is to be expected since this research uses a similar approach with the one used by Dorrington and Timæus (2017).

However, estimates from the CARe_3.2 model as well as the orphanhood method are too low particularly for females, with the exception for a few provinces, namely, Western Cape, Gauteng, KZN and North West. This could be explained by the fact that either the background mortality assumption or modelling of HIV/AIDS in the CARe_3.2 model might be wrong whereas for the orphanhood method it could be due to the fact that the orphanhood method estimates refer to the time in the middle of the intercensal period or it could be due to deficiencies in the data for the orphanhood.

Interestingly, the orphanhood estimates match those from the CARe model in all but two provinces for male and three for females.

In most cases, the ranking of mortality by province is consistent with the ranking of mortality from the results of the CARe_3.2 model as shown in Figure 5.1. The CARe_3.2 model indicates that Mpumalanga, Free State and Eastern Cape have the same level of mortality for males, which is unlikely as Mpumalanga has higher HIV prevalence than other provinces.

Figure 5.2 Comparison of adult mortality estimates (${}_{45}q_{15}$) derived in this study against estimates from SANBD[13] (both sexes combined).



The estimates from SANBD (Msemburi, Pillay-van Wyk, Dorrington *et al.* 2014) are broadly similar to estimates obtained in this study though their estimates are slightly lower in most provinces, with percentage differences ranging from 0 to 3 per cent, which gives some support in the estimates derived in this study.

Wide disparities are observed in adult mortality among municipal districts within the same province and in some cases, these disparities appear to be surprisingly high compared to neighbouring districts.

We expect low mortality in districts that either have low HIV-prevalence or that are wealthy (urban) and high mortality in districts that either have high HIV prevalence or in poor (rural) districts. Metropolitan districts have better services since they are amongst the wealthy districts, however, quite surprisingly, Mangaung Metropolitan and Nelson Mandela Bay are amongst the districts with high mortality.

On the other hand, Central Karoo (Western Cape), Mopani and Vhembe (Limpopo) indicated low mortality whilst they are observed to be amongst disadvantaged districts, which might be attributed to that they have low HIV prevalence (Figure E3, Appendix E) whereas most rural districts (Ugu, OR Tambo, Alfred Nzo, Joe Gqabi) were observed to have high mortality. This is in line with our expectations as well as Chinogurei (2017) findings, indicating consistency between these two studies.

Comparison of estimates derived by Chinogurei with the estimates derived in this study indicate, if the estimates are accurate, that mortality improved in most district that

[13] SANBD refers to South African National Burden of Disease.

lie in Gauteng, KwaZulu-Natal and the Western Cape, which is to be expected since there was high provision of ARVs in these provinces (Nattrass 2006). Particularly, in KZN, provision of ARVs was quite extensive (Johnson 2012) and mortality improved tremendously in almost all districts (in particular for males, with 6 districts showing improvement above 10 per cent) with the exception of female mortality in the Zululand district where it increased, but by only 1 per cent.

Also, Gauteng and Western Cape provinces (with the exception of Sedibeng and Central Karoo respectively), indicate improvement in some of its districts, with the City of Tshwane Metropole being the district with highest mortality improvement of all the districts for males. This could also be explained by the fact that HIV prevalence in these provinces is low and also most districts that lie in these provinces are urban districts hence access to health care facilities is not a challenge. However, adult mortality estimates in these districts should be interpreted with care considering the fact that these provinces are predominantly urban and thus they are more likely to be associated with out migration of the sick or dying people going back to their provinces of origin to be taken care of.

Wide disparities between male and female mortality are observed in the districts of the Eastern Cape province. The results indicate that female mortality increased in all districts except two, that is, Amatole and the Buffalo Metropole, whereas for males it increased in only two districts, that is, Chris Hani and Alfred Nzo (recording the highest percentage increase in mortality of all districts). This could be due to the fact that since the province is dominantly rural, women might have challenges in accessing healthcare services leading to high mortality (since they are at a high risk of HIV acquisition) or perhaps due to poverty as it has been associated with high levels of HIV prevalence (Muula 2008). Also this could be attributed to low provision of ARVs as well as increase in HIV prevalence at young adult ages in this province (NDOH 2012).

To a larger extent, comparison of estimates derived in this study with estimates derived by Chinogurei (2017) indicate that in most districts, male mortality improved which is consistent with expectations, whereas for females, mortality improvement was not quite as convincing. However, one should bear in mind that estimates derived by Chinogurei (2017) also have their own limitations which can lead to biased conclusions.

One might expect high child mortality in districts with high adult mortality and low child mortality in districts with low child mortality. Comparison of child mortality estimates derived by Zewdie (2014) and adult mortality estimates derived in this study as well as ranking of districts indicates quite a high consistency (Figure E2, Appendix E) between these two studies, with all of the districts in the Western Cape, Gauteng and Limpopo being among the 20 districts with lowest mortality and most districts in KwaZulu-Natal being among the 20 districts with highest mortality (with only a few exceptions).

5.3 Limitations of the study

The estimates of adult mortality in this study are derived from census data, which often suffer from their own inherent problems, such as age misreporting, age exaggeration and differential reporting of deaths as mentioned in section 3.2. Thus, the adult mortality estimates produced might be biased due to these problems.

The numbers of deaths reported by households in the census were corrected for completeness, but 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- (or over-) count in the census, this is useful for broad adjustments (for instance, at national level) as the estimates come from a small sample, however, these estimates may not identify under- (or over-) enumeration accurately at the district level.

The orphanhood method was used to assess the estimates obtained at provincial level instead of death distribution methods since estimates of completeness produced by death distribution methods at the sub-national level are unreliable due to the difficulty of accounting for inter-regional migration. The major challenge with the orphanhood estimates is that they are not the gold standard since, apart from problems with the accuracy of responses, the data may not be entirely representative of the province of the respondent since not all parents live in the province of the respondent. However, we do not expect the bias to be very big in most cases.

To smooth mortality rates, the standard life table derived from the mortality rates from the CARe 3.2 demographic projection model for each province for 2011 were used. It was assumed that the provincial standard life table applies at district level, which does not appear to be the case for some of the districts. However, in most cases the distortions are confined to certain age ranges (see Appendix D) and we do not expect it to have impact on estimating indicators such as ${}_{45}q_{15}$.

In general, the above-mentioned limitations are not expected to be significant since this research used age-sex specific correction factors for each population group to adjust death estimates from census data.

5.4 Areas of further study

In this study completeness was estimated at national level for each population group. However, there is a possibility that the level of completeness might also differ by geographical area (rural/ urban) hence it would be useful to investigate if estimating the level of completeness by whether urban or non-urban with and without population group makes any difference to the estimates produced in this study.

The standard life tables for the provinces (from the CARe_3.2 model) were used to graduate adult mortality estimates at district level on the assumption that the provincial standard life table applies to the districts that lie within the same province. Further investigation is needed to explore if a different standard should be used for some of the districts.

To assess the improvement of mortality from 2001 to 2011, we compared estimates derived by Chinogurei (2017) to estimates derived in this study. To reach a valid conclusion there is a need to investigate the extent to which differences in mortality estimates by district are due to the impact of ARV treatment, as the earlier study is for the pre-ARV era and the latter study is after the introduction of ARV's. Also, there is a need to investigate the extent to which the differences between the estimates produced in this study and those from CARe_3.2 model are due to the non-HIV mortality.

In this study the correlation between antenatal prevalence of HIV and adult mortality was used to explain mortality differences in districts. Further investigation is needed to explore other factors like proximity to health facilities, inequality within districts (among others) and also allowing for the level of provision of ARVs to help in understanding the differences in adult mortality by geographical area.

5.5 Conclusion

This research set out to investigate if the method proposed by Dorrington, Moultrie and Timæus (2004) produces plausible estimates at district level. Weighted estimates produced at the provincial level were compared against estimates produced by the orphanhood method at district level. However, there was insufficient evidence to reach a definitive conclusion on the reasonableness of the method since estimates produced by the orphanhood method refer to the time approximately four years before the year prior to the 2011 census. Also, the estimates produced might not entirely represent the province as a whole since not all parents live in the province of the respondent. In general, the orphanhood method produced estimates that are lower than the estimates produced in this study, however, the ranking of districts appeared to be consistent with the exception of a few provinces (Mpumalanga and North West), which to some extent support the use of the method.

In addition, the method produces weighted provincial estimates that are consistent (in both the level (percentage differences of under 9 per cent) and ranking) with estimates produced by Dorrington and Timæus (2017), which lends further credence to our estimates.

Also, combined male and female weighted mortality estimates are quite consistent, in both the level and ranking of provinces, with the estimates from Msemburi, Pillay-van Wyk, Dorrington *et al.* (2014). This gives enough evidence to conclude that although the method may not produce accurate estimates for all districts, it works well at sub-national level.

Comparison of estimates of child mortality derived by Zewdie (2014) with adult mortality estimates lends confidence to the estimates produced as these indicate that there is high child mortality in districts with high adult mortality with the exception of a few districts.

In general, we can conclude that the method produces plausible estimates at district level after adjusting household reported deaths with age specific adjustment factors. Findings of this research strengthens confidence in the use of the method.

REFERENCES

Bah, Sulaiman. 2016. "HIV/AIDS in South Africa in the light of death registration data," in Zuberi, T, A Sibanda and E.O Udjo (eds). *The Demography of South Africa*. Routledge, pp. 120-129.

Bah, Sulaiman M. 1998a. "A critical analytical and historical review of South African published life tables", *PSC Discussion Papers Series* **12**(2):1-8.

Bah, Sulaiman M. 1998b. "The making, unmaking and remaking of a national but stratified vital statistics system in the Republic of South Africa", *PSC Discussion Papers Series* **12**(3):1-8.

Bangha, Martin W and Sandile Simelane. 2013. "Spatial differentials in childhood mortality in South Africa: evidence from the 2001 census", *African Population Studies* **22**(2):4-19.

Bennett, Neil G and Shiro Horiuchi. 1981. "Estimating the completeness of death registration in a closed population", *Population index* **47**(2):202-221.

Bennett, Neil G and 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.

Bogue, Donald J. 1965. "The Pros and Cons of "Self-Enumeration"", *Demography* **2**(1):600-626.

Bradshaw, Debbie and Ian M Timæus. 2006. "Levels and trends of adult mortality," in Jamison, D.T, R.G Feachem, M.W Makgoba, R.B Eduard, K Florence, J.H Karen and O.R Khama (eds). *Disease and Mortality in Sub-Saharan Africa*. The International Bank for Reconstruction and Development, pp. 31-42.

Brass, William. 1975. *Methods for estimating fertility and mortality from limited and defective data*. U.S.A: University of North Carolina.

Brass, William and Kenneth H Hill. 1973. "Estimating adult mortality from orphanhood", *International Union for the scientific study of population* **35**(2):111-123.

Chinogurei, Chido. 2017. "Evaluating the adequacy of the method of using vital registration and census data in estimating adult mortality when applied sub-provincially." Unpublished MPhil thesis, University of Cape Town.

Dorrington, R. 2013a. "Estimation of migration from census data," in Moultrie, T., Dorrington R., Hill K., Hill A., Timæus I. and Zaba B. (eds). *Tools for Demographic estimation*. Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/estimation-migration-census-data, Accessed: 15/04/17.

Dorrington, R. 2013b. "The Generalized Growth Balance Method," in Moultrie, T., Dorrington R., Hill K., Hill A., Timæus I. and Zaba B. (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/generalized-growth-balance-method, Accessed: 10/02/2017.

Dorrington, R. 2013c. "The Synthetic Extinct Generations Method," in Moultrie, T., Dorrington R., Hill K., Hill A., Timæus I. and Zaba B. (eds). *Tools for Demographic estimation*. Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/synthetic-extinct-generations-methods, Accessed: 22/02/17.

Dorrington, R, I Timaeus and S Gregson. 2006. "Adult mortality in Southern Africa using deaths reported by households: some methodological issues and results.

Extended abstract," Paper presented at Population Association of America 2006, Los Angeles California, 30 March-1 April 2006. .
Dorrington, R. E, D. Bourne, D. Bradshaw, R. Laubscher *et al.* 2001. *The impact of HIV/AIDS on adult mortality in South Africa*. Burden of disease Unit, Medical Research Council Cape Town. www.mrc.ac.za/bod/complete.pdf, Accessed: 22/03/17.
Dorrington, R. E., T.A Moultrie and I. M. Timæus. 2004. *Estimation of mortality using the South African Census 2001 data*. Monograph. Centre for Actuarial Research: University of Cape Town. http://unstats.un.org/unsd/censuskb20/KnowledgebaseArticle10546.aspx, Accessed: 20/02/17.
Dorrington, R. E. and I. 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 African Population Conference, Johannesburg, South Africa, 30 Nov 2017-4 Dec 2017.
Dorrington, R. E., I. M Timæus, T.A Moultrie and N. Nannan. 2004. "Estimates of provincial fertility and mortality in South Africa, 1985—1996", *Southern African Journal of Demography* **9**(2):25-57.
Dorrington, R. E. and Ian M. Timæus. 2017. "Estimation of sub-national adult mortality rates by region using incomplete civil registration to correct census data on recent deaths," Paper presented at International Population Conference, Cape Town South Africa, 29 Oct 2017-4 Nov 2017.
Dorrington, R.E, D. Bradshaw, R. Laubscher and N. Nannan. 2016. *Rapid mortality surveillance report 2015*. Cape Town, South Africa: South African Medical Research Council. http://www.mrc.ac.za/bod/reports.htm, Accessed: 11/02/17.
Dorrington, R.E, I. M Timæus and T. A Moultrie. 2008. "Death distribution methods for estimating adult mortality: sensitivity analysis with simulated data errors, revisited," Paper presented at PAA Conference, New Orleans, United States, 17-19 April 2008.
Dorrington, Rob E, Debbie Bradshaw and 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", *Cape Town: South African Medical Research Council*
Gakidou, Emmanuela, Margaret Hogan and Alan D Lopez. 2004. "Adult mortality: time for a reappraisal", *International Journal of Epidemiology* **33**(4):710-717.
Henry, Louis. 1960. "Mesure indirecte de la mortalité des adultes", *Population (french edition)* **12**(1):457-466.
Hill K. 2017. *Analytic Methods to Evaluate the Completeness and Quality of Deaths Registration* United Nations Population Division Technical Paper. Accessed:
Hill, Kenneth and Y Choi. 2004. "Performance of GGB and SEG given various simulated data errors," Paper presented at Marconi Conference Center Marin County, California.
Hill, Kenneth, Yoonjoung Choi and Ian M Timæus. 2005. "Unconventional approaches to mortality estimation", *Demographic Research* **13**(12):281-299.
Hill, Kenneth and James Trussell. 1977. "Further developments in indirect mortality estimation", *Population Studies* **31**(2):313-334.
Hill, Kenneth, Danzhen You and Yoonjoung Choi. 2009. "Death distribution methods for estimating adult mortality: sensitivity analysis with simulated data errors", *Demographic Research* **21**:235-254.

Johnson, Leigh F. 2012. "Access to antiretroviral treatment in South Africa, 2004-2011", *Southern African Journal of HIV Medicine* **13**(1):22-25.
Joubert, Jane, Chalapati Rao, Debbie Bradshaw, Rob E Dorrington *et al.* 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. 2015. "3 An Assessment of Quality of Death Registration Data for Mortality Estimation in South Africa," in Odimegwu, C.O. and J. Kekovole (eds). *Social Demography of South Africa: Advances and Emerging Issues*. New York: Routledge, pp. 34.
Kok, PC. 2003. *Post-apartheid patterns of internal migration in South Africa.* Pretoria: Human Science Research Council Publishers.
Luy, Marc. 2009. *Estimating mortality differentials in developed populations from survey information on maternal and paternal orphanhood*. Monograph, Vienna Inst. of Demography. Accessed: 21/03/17.
Mathers, Colin and Ties Boerma. 2010. "Mortality measurement matters: improving data collection and estimation methods for child and adult mortality", *PLoS Medicen* **7**(4):e1000265.
Mathers, Colin D, Doris Ma Fat, Mie Inoue, Chalapati Rao *et al.* 2005. "Counting the dead and what they died from: an assessment of the global status of cause of death data", *Bulletin of the World Health Organization.* **83**(3):171-177c.
Msemburi, W, V Pillay-van Wyk, R Dorrington, I Neethling *et al.* 2014. *Second national burden of disease study for South Africa: Cause-of-death profile for South Africa, 1997–2010*. http://www.mrc.ac.za/bod/SouthAfrica2010.pdf, Accessed: 04/04/17.
Murray, Christopher JL, Brodie D Ferguson, Alan D Lopez, Michel Guillot *et al.* 2003. "Modified logit life table system: principles, empirical validation, and application", *Population Studies* **57**(2):165-182.
Murray, CJL, JK Rajaratnam, J Marcus, T Laakso *et al.* 2010. "What Can We Conclude from Death Registration? Improved methods for evaluating completeness.", **7**(4):1549-1676.
Muula, Adamson S. 2008. "HIV infection and AIDS among young women in South Africa", *Croatian Medical Journal.* **49**(3):423.
Nattrass, Nicoli. 2006. "South Africa's "rollout" of highly active antiretroviral therapy: a critical assessment", *JAIDS Journal of Acquired Immune Deficiency Syndromes* **43**(5):618-623.
NDOH. 2012. *The 2011 National Antenatal Sentinel HIV & Syphilis Prevalence Survey in South Africa*. Pretoria: Cititas Building: National Department of Health. https://www.health-e.org.za/wp-content/uploads/2013/05/f0980fb5107a7ce543a8bd5730e52333.pdf, Accessed: 05/01/18.
Phillips, David E, Rafael Lozano, Mohsen Naghavi, Charles Atkinson *et al.* 2014. "A composite metric for assessing data on mortality and causes of death: the vital statistics performance index", *Population Health Metrics.* **12**(1):14.
Pillay-van Wyk, V, R. E Dorrington and D Bradshaw. 2017. "Guest Editorial: Rapidly changing mortality profiles in South Africa in its nine provinces", *South African Medical Journal* **107**(3):168-169.
Pillay-van Wyk, Victoria, Ria Laubscher, William Msemburi, Rob E Dorrington *et al.* 2014. *Second South African National Burden of Disease Study: Data Cleaning, Validation and SA NBD List*. http://www.mrc.ac.za/bod/SANBDReport.pdf, Accessed: 26/03/17.

Preston, Samuel, Ansley J Coale, James Trussell and Maxine Weinstein. 1980. "Estimating the completeness of reporting of adult deaths in populations that are approximately stable", *Population index* **46**(2):179-202.

Richman, Ronald David. 2017. "Old age mortality in South Africa, 1985-2011." Unpublished Masters thesis, University of Cape Town.

Schmertmann, Carl and Marcos Roberto Gonzaga. 2018. "Bayesian estimation of age-specific mortality and life expectancy for small areas with defective vital records",

Setel, Philip W, Sarah B Macfarlane, Simon Szreter, Lene Mikkelsen *et al.* 2007. "A scandal of invisibility: making everyone count by counting everyone", *The Lancet* **370**(9598):1569-1577.

StataCorp. 2015. Stata Statistical Software: Release 14. College Station. TX, StataCorp LP.

Stats SA. 2000. *South African Life Tables 1985-1994 and 1996*. Pretoria: Statistics South Africa: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/255/.../3941, Accessed: 23/02/17.

Stats SA. 2005. *Mortality and causes of death in South Africa, 1997 - 2003: Findings from death notification*. Pretoria South Africa: Stats SA. https://www.statssa.gov.za/publications/P03093/P03093.pdf Accessed: 25/03/17.

Stats SA. 2012. *Mortality and causes of death in South Africa, 2011: Findings from death notification*. http://www.statssa.gov.za/publications/P03093/P030932011.pdf, Accessed: 31/03/17.

Stats SA. 2014. *Mortality and causes of death in South Africa, 2013: Findings from death notification*. Pretoria South Africa: Stats SA. http://www.statssa.gov.za/publications/P03093/P030932013.pdf, Accessed: 28/05/17.

Stats SA. 2016. *Community Survey Report* Pretoria South Africa: Stats SA. http://cs2016.statssa.gov.za, Accessed: 30/05/17.

Timæus, Ian M. 1991. "Measurement of adult mortality in less developed countries: a comparative review", *Population index* **57**(4):552-569.

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

Timæus, Ian M. 2013a. "Indirect Estimation of adult mortality from orphanhood," in Moultrie, Tom, R Dorrington, Kenneth Hill, Allan Hill, I Timaeus and Zaba Basia (eds). *Tools for Demographic Estimation.* . Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/indirect-estimation-orphanhood-multiple-inquiries, Accessed: 30/04/17.

Timæus, Ian M. 2013b. "Introduction to adult mortality analysis," in Moultrie, Tom, R Dorrington, Kenneth Hill, Allan Hill, I Timaeus and Zaba Basia (eds). *Tools for Demographic Estimation*. Paris International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/introduction-adult-mortality-analysis, Accessed: 03/05/17.

Timæus, Ian M and Momodou Jasseh. 2004. "Adult mortality in sub-Saharan Africa: evidence from Demographic and Health Surveys", *Demography* **41**(4):757-772.

Timæus, Ian M and Andrew J Nunn. 1997. "Measurement of adult mortality in populations affected by AIDS: an assessment of the orphanhood method", *Health Transition Review* **7**(2):23-43.

Udjo, Eric O. 1998. *Additional evidence regarding fertility and mortality trends in South Africa and implications for population projections*. Pretoria South Africa: Statistics SA. https://apps.statssa.gov.za/census01/Census96/HTML/Metadata/Docs/fertility_udjo.htm, Accessed: 26/03/17.

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 Zuberi, T, A Sibanda and E.O Udjo (eds). *The Demography of South Africa*. Routledge, pp. 120-129.

Udjo, Eric O and Pinky Lalthapersad-Pillay. 2014. "Estimating maternal mortality and causes in South Africa: National and provincial levels", *Midwifery* **30**(5):512-518.

UN Population Division. 2009. *World Population Prospects: The 2009 Revision*. New York: United Nations Department of Economic and Social Affairs. Accessed: 27/03/17.

United Nations. 2002. *Methods of estimating adult mortality*. United Nations Population Division. www.un.org/esa/population/publications/adultmort/Complete.pdf, Accessed: 03/08/17.

United Nations. 2014. *Principles and recommendations for a vital statistics system, Third Revision*. New York: United Nations Department of Economic and Social Affairs. Accessed: 14/03/17.

Van Eeden, IJ and JL Van Tonder (1975) Report No. S-34. Pretoria: Human Sciences Research Council.

Vincent, Paul. 1951. "La mortalité des vieillards", *Population (french edition)* **6**(2):181-204.

Zewdie, Samuel Abera. 2014. "Spatial analysis of child mortality in South Africa in relation to poverty and inequality: evidences from the 2011 census." Unpublished MPhil thesis, University of Cape Town.

Appendix A: Assessment of data quality

Figure A. 1 Sex ratios for the year 2001 and 2011

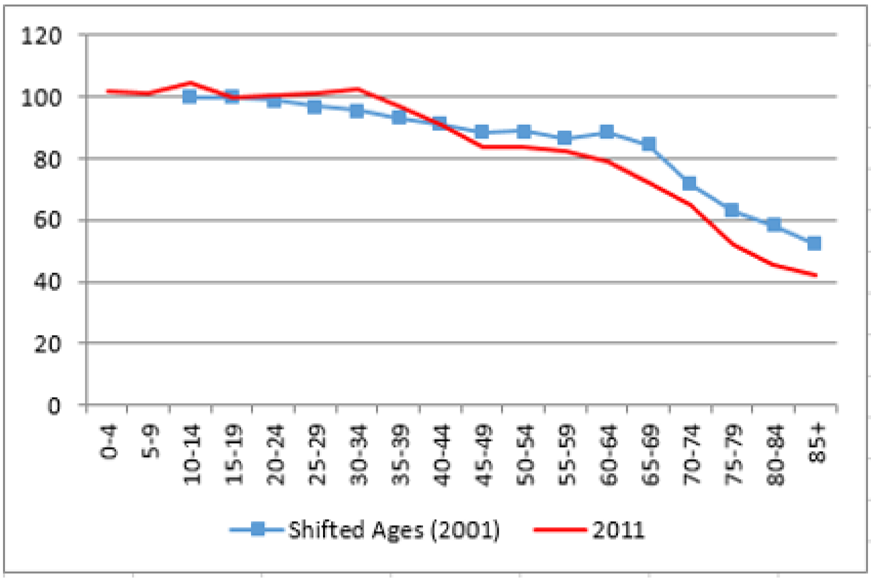
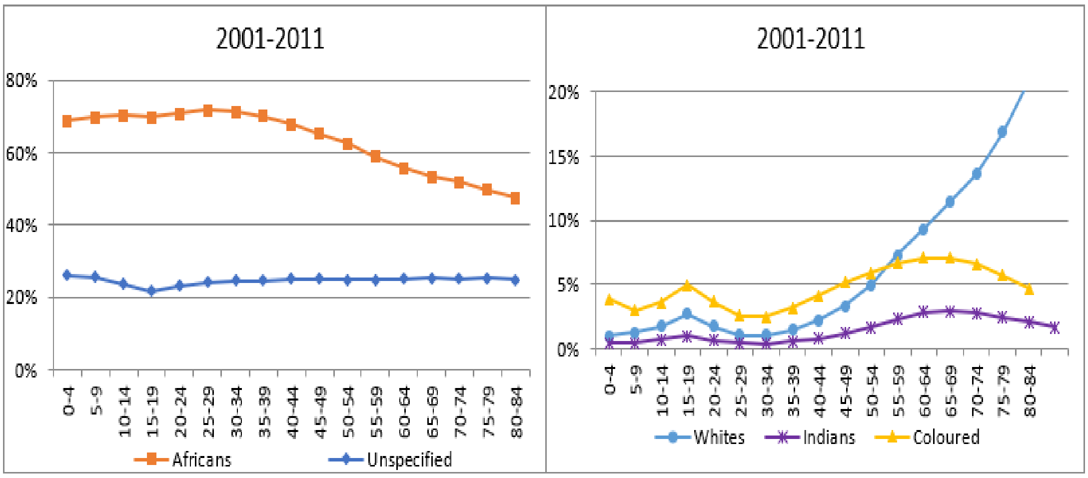


Figure A. 2 Proportion of deaths by population group and age



Appendix B: Emigration inflows and outflows over the intercensal period

Table B. 1 Emigrants inflows and outflows over the intercensal period, Women[14]

Australia	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	3 365	3 548	2863	3 562	2 888	2 403	2 708	3 412	5 518	5 499
Outflows	17	0	0	362	386	399	431	410	354	422
Difference	3 348	3 548	2 863	3 200	2 502	2 004	2 277	3 002	5 164	5 077

New Zealand	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	2 364	1 464	930	686	747	922	1 051	1 519	881	636
Outflows	110	122	120	187	184	211	209	258	333	339
Difference	2 254	1 342	810	499	563	711	842	1 261	548	297

Canada	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	1 073	823	626	599	479	562	604	555	612	596
Outflows	429.2	329.2	250.4	239.6	191.6	224.8	241.6	222	244.8	238.4
Difference	643.8	493.8	375.6	359.4	287.4	337.2	362.4	333	367.2	357.6

United States	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	1 921	1 962	1 120	1 670	2 267	1 660	1 492	1 327	1 563	1 416
Outflows	384.2	392.4	224	334	453.4	332	298.4	265.4	312.6	283.2
Difference	1 537	1 569	896	1 336	1 814	1 328	1 193	1061	1 250	1 133

United Kingdom	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	5 230	8 400	8 000	12000	9 600	6 400	5 200	5 600	3 200	2 000
Outflows	2 287	2 250	4 050	3600	4 500	5 850	4 050	2 700	1 800	1 800
	2 942	6 150	3 950	8 400	5 100	550	1 150	2 900	1 400	200

Germany	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	674	656	524	479	430	421	404	437	377	344
Outflows	269.6	262.4	209.6	191.6	172	168.4	161.6	174.8	150.8	137.6
	404.4	393.6	314.4	287.4	258	252.6	242.4	262.2	226.2	206.4

Austria	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	42	68	48	60	36	52	40	41	44	42
Outflows	18.36	37	29	25	28.91	24.33	16.52	22.95	27.99	20.19
	23.64	30.82	19.08	35.21	7.083	27.67	23.47	18.05	16.1	21.80

France	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	67.47	66.55	63.34	50.49	55.99	79	102	69	96	94
Outflows	0	0	0	0	0	0	0	0	0	0
	67.47	66.55	63.34	50.49	55.99	79	102	69	96	94

Netherlands	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	410.4	316.3	270.4	234.9	217.8	232.2	270	392.85	445	421
Outflows	321	198	210	198	201	210	197	361	410	410
	89.4	118.35	60.45	36.9	16.8	22.2	73	31.85	35	11

Norway	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	18	26	23	23	21	28	51	34	31	24
Outflows	3	3	17	19	25	24	10	23	23	24
	15	23	6	4	-4	4	41	11	8	0

[14] Assuming that these flows are for the year starting 1 July of the year in question to 1 July in the preceding year (that is year 2001 in the Table represents year 1 July 2001 to 1 July 2002)

Spain	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	32	32	24	44	46	50	59	49	37	40
Outflows	0	0	2	2	2	13	22	29	19	17
	32	32	22	42	44	37	37	20	18	23

Sweden	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	70	67	61	45	39	38	30	48	42	38
Outflows	7	15	23	27	22	59	30	18	15	15
	63	52	38	18	17	-21	0	30	27	23

Korea	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	153	223	140	108	105	178	313	433	534	454
Outflows	0	0	51	134	151	83	160	229	329	432
	153	223	89	-26	-46	95	153	204	205	22

Italy	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	0	41	39	85	45	24	29	57	42	43
Outflows	0	11	7	37	7	3	5	0	0	6
	0	30	32	48	38	21	24	57	42	37

Chile	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	13.5	12.15	9.9	11.7	13.05	22.05	13.5	31.5	35.55	26
Outflows	0	0	0	0	0	0	0	0	0	0
	13.5	12.15	9.9	11.7	13.05	22.05	13.5	31.5	35.55	26

Denmark	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Inflows	29.25	29.25	14.4	18	10.35	6.75	11.25	17.55	10.35	13.95
Outflows	17.1	27.45		18	13.05	20.25	10	21	12	17
Net Diff.	12.15	1.8	14.4	0	-2.7	-13.5	1.25	-3.45	-1.65	-3.05

Table B. 2 Emigrants inflows and outflows over the intercensal period, Males

Australia	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	3 448	3 644	3 019	3 570	2 816	2 385	2 670	3 524	5 753	5 606
Outflows	13	0	0	391	384	437	438	414	379	473
Difference	3 435	3 644	3 019	3 179	2 432	1 948	2 232	3 110	5 374	5 133
New Zealand	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	2 377	1 450	880	699	783	895	1 020	1 568	861	589
Outflows	220	134	132	170	191	230	220	235	341	397
Difference	2375	1 316	748	529	592	665	800	1 333	520	192
Canada	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	1 017	808	619	576	509	549	596	568	576	642
Outflows	254.25	202	154.75	144	127.25	137.25	149	142	144	160.5
Difference	762.75	606	464.25	432	381.75	411.75	447	426	432	481.5
United States	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	2 169	1 899	1 090	1 700	2 269	1 541	1 496	1 396	1 608	1 342
Outflows	433.8	379.8	218	340	453.8	308.2	299.2	279.2	321.6	268.4
Difference	1 735	1 519	872	1 360	1 815	1 232	1 197	1 117	1 286	1 074
United Kingdom	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	7 845	12 600	12 000	18 000	14 400	9 600	7 800	8 400	4 800	3 100
Outflows	2 795	2 750	4 950	4 400	5 500	7 150	4 950	3 300	2 200	2 200
Difference	5 049	9 850	7 050	13 600	8 900	2 450	2 850	5 100	2 600	900
Germany	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	499	570	565	491	412	476	440	449	322	354
Outflows	99.8	114	113	98.2	82.4	95.2	88	89.8	64.4	70.8
Difference	399.2	456	452	392.8	329.6	380.8	352	359.2	257.6	283.2
Austria	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	30	56	24	36	35	42	27	24	31	33
Outflows	0	0	0	0	0	0	0	0	0	0
	30	56	24	36	35	42	27	24	31	33
France	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	79.527	78.445	74.658	59.51	66.002	83	82	99	97	80
Outflows	0	0	0	0	0	0	0	0	0	0
Netherlands	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	501.6	386.65	330.55	287.1	266.2	283.8	330	480.15	368.5	322.3
Outflows	107	302	291	280	260	288	238	200	166	205
	67.473	66.555	63.342	50.49	55.998	79	102	69	96	94
Norway	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	34	53	24	25	21	51	92	40	44	35
Outflows	19	30	18	21	25	47	51	29	36	35
	15	23	6	4	-4	4	41	11	8	0

Spain	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	37	40	39	65	82	77	86	83	46	64
Outflows	0	1	4	0	5	9	29	39	34	42
	37	39	35	65	77	68	57	44	12	22

Sweden	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	101	98	78	61	54	53	44	52	54	40
Outflows	6	27	37	26	31	50	35	29	24	22

Korea	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	167	207	109	85	151	159	269	355	375	311
Outflows	0	0	279	133	140	102	137	207	255	296
	167	207	-170	-48	11	57	132	148	120	15

Italy	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	0	32	27	62	33	31	22	42	35	26
Outflows	0	3	13	47	10	8	6	4	6	1
Net Diff.	0	29	14	15	23	23	16	38	29	25

Chile	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	16.5	14.85	12.1	14.3	15.95	26.95	16.5	38.5	43.45	29.15
Outflows	0	0	0	0	0	0	0	0	0	0
Net Diff.	16.5	14.85	12.1	14.3	15.95	26.95	16.5	38.5	43.45	29.15

Denmark	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Inflows	35.75	35.75	17.6	22	12.65	8.25	13.75	21.45	12.65	17.05
Outflows	20.9	33.55	20.9	22	15.95	24.75	26.4	25.85	29.15	26.4
Net Diff.	14.85	2.2	-3.3	0	-3.3	-16.5	-12.65	-4.4	-16.5	-9.35

Appendix C: Annual adjustments to the estimates of completeness of reporting of vital registration deaths

Table C. 1 Annual adjustments to the estimates of completeness of reporting of vital registration deaths

	African	**Indian**	**White**	**Coloured**
2001	0.61	0.888	0.769	0.67
2002	0.61	0.900	0.774	0.67
2003	0.62	0.905	0.775	0.67
2004	0.61	0.908	0.775	0.67
2005	0.62	0.909	0.775	0.67
2006	0.63	0.910	0.775	0.67
2007	0.63	0.910	0.775	0.67
2008	0.63	0.910	0.775	0.67
2009	0.62	0.910	0.775	0.67
2010	0.64	0.910	0.775	0.67
2011	0.70	0.910	0.775	0.71

Appendix D: Logit fits

Figure D. 1 Male logit fits- Western Cape

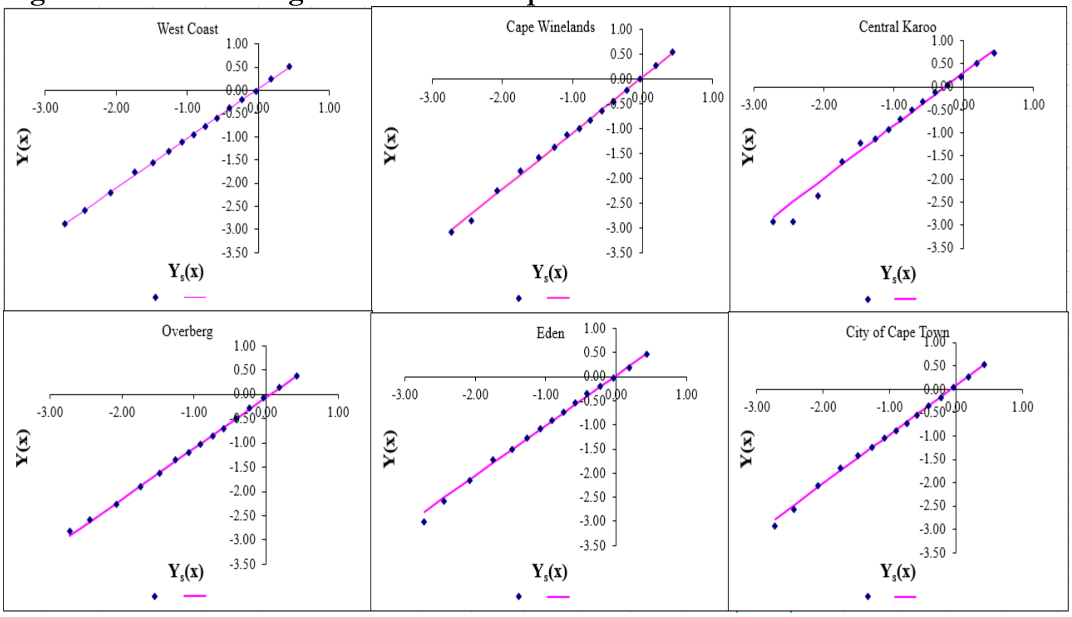


Figure D. 2 Female logit fits- Western Cape

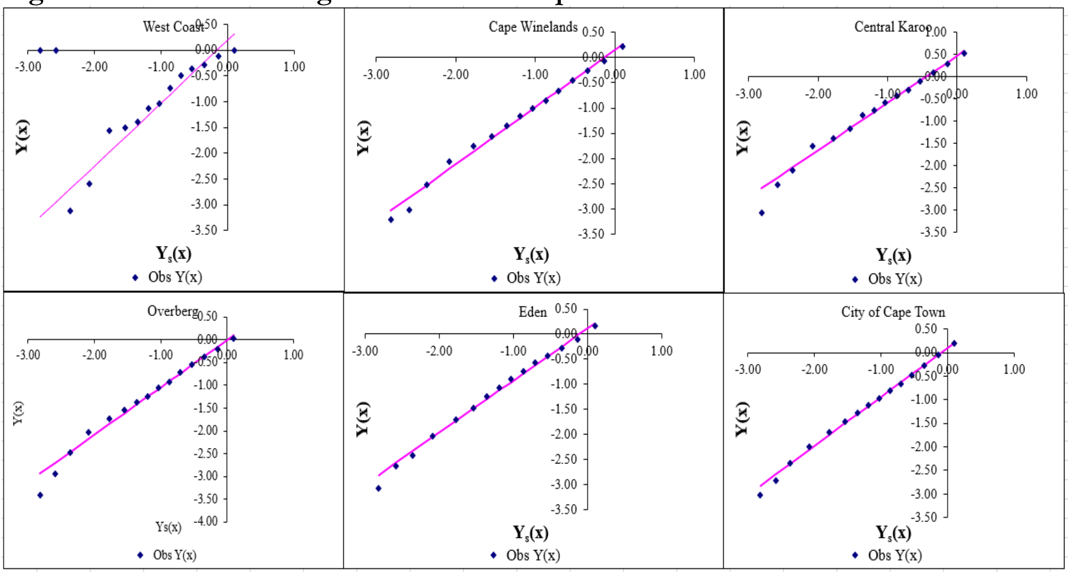


Figure D. 3 Males logit fits -Northern Cape

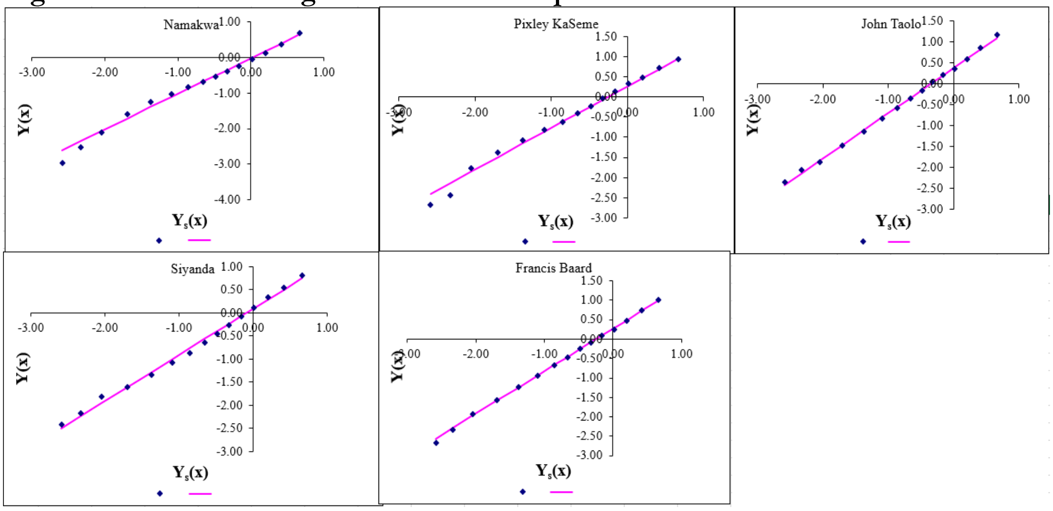


Figure D. 4 Female logit fit Northern Cape

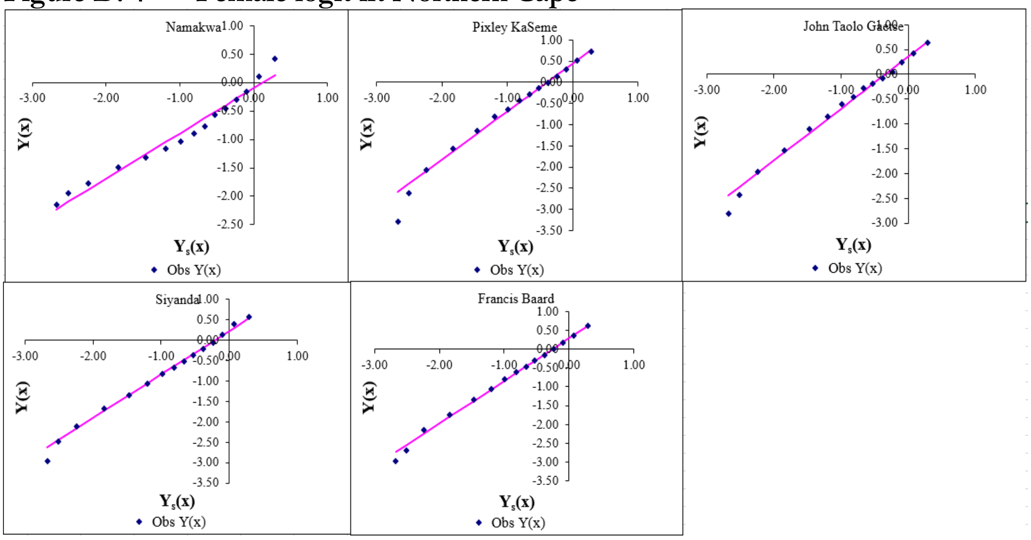


Figure D. 5 Male Logit fit Eastern Cape

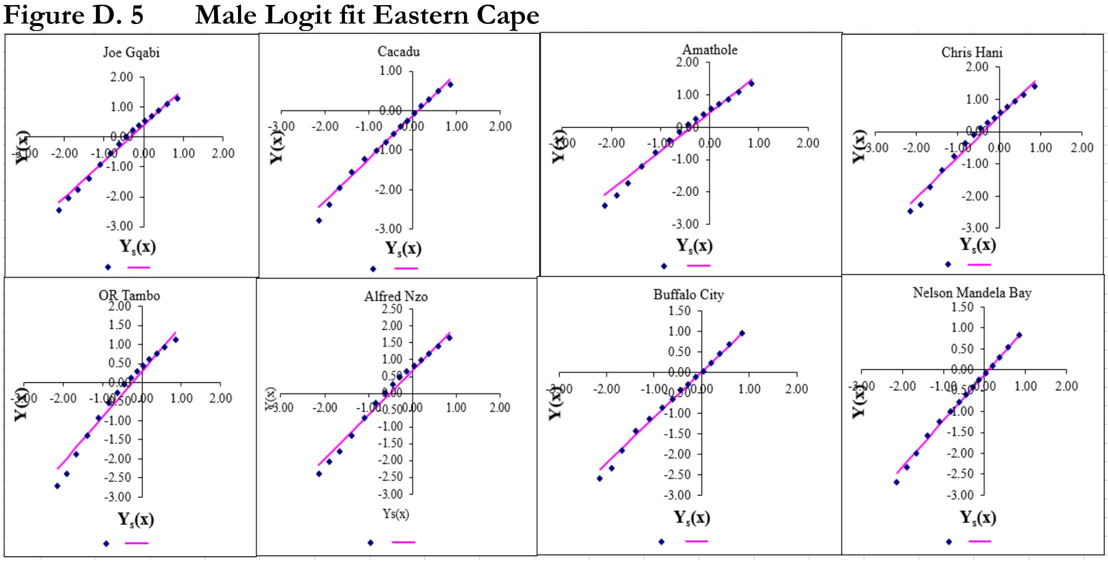


Figure D. 6 Female logit fit- Eastern Cape

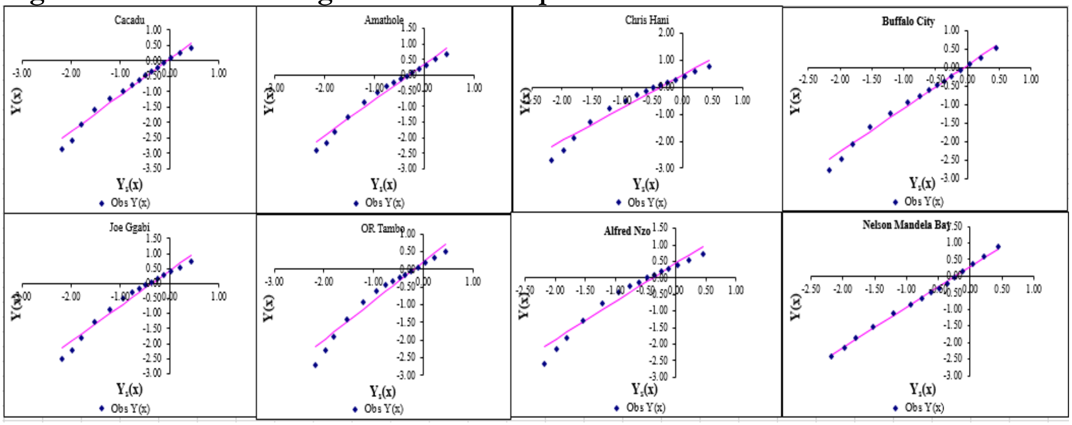


Figure D. 7 Male logit fit- Eastern Cape

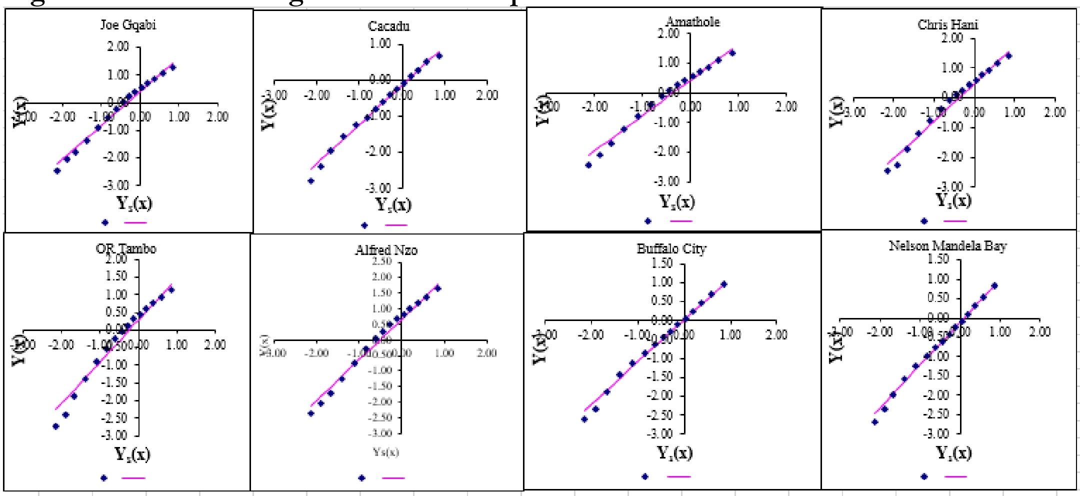


Figure D. 8 Male logit fit- Free State

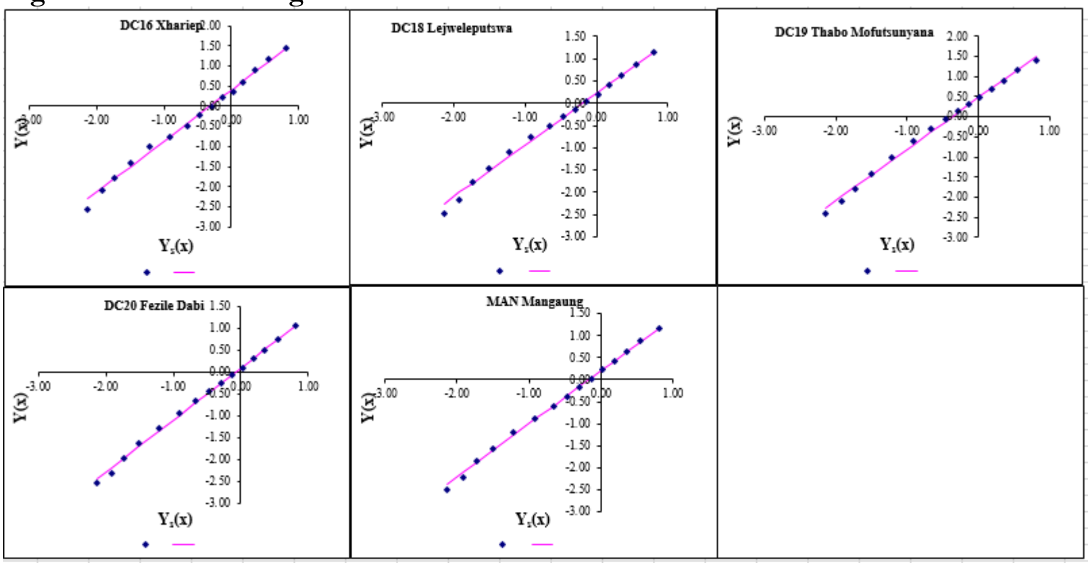


Figure D. 9 Female logit fit- Free State

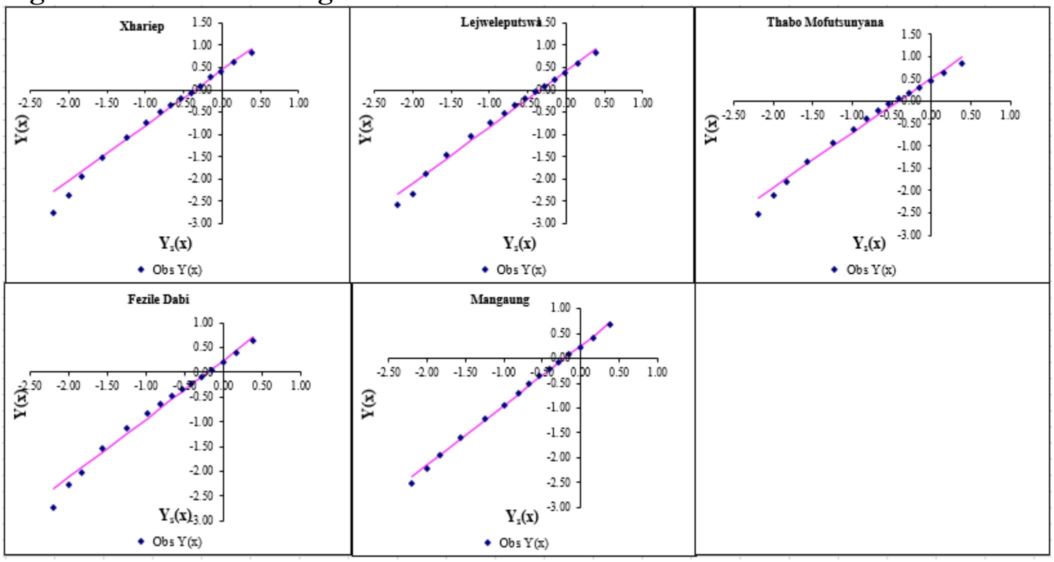


Figure D. 10 Female logit fit- KZN

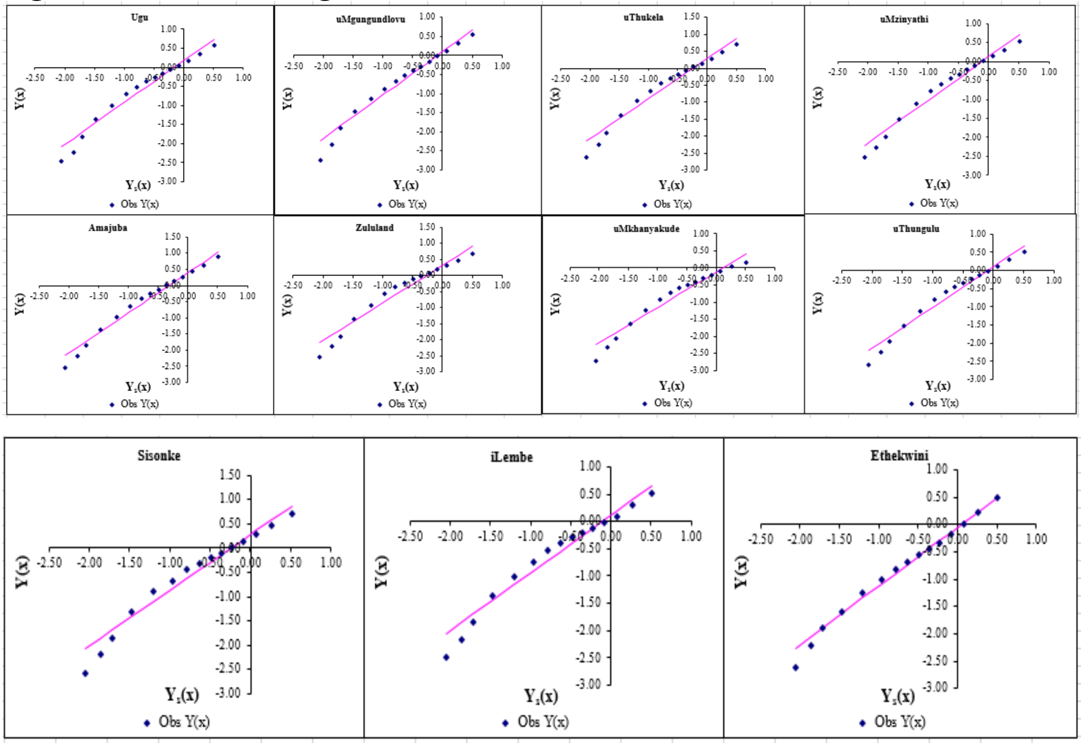


Figure D. 11 Male logit fit- KZN

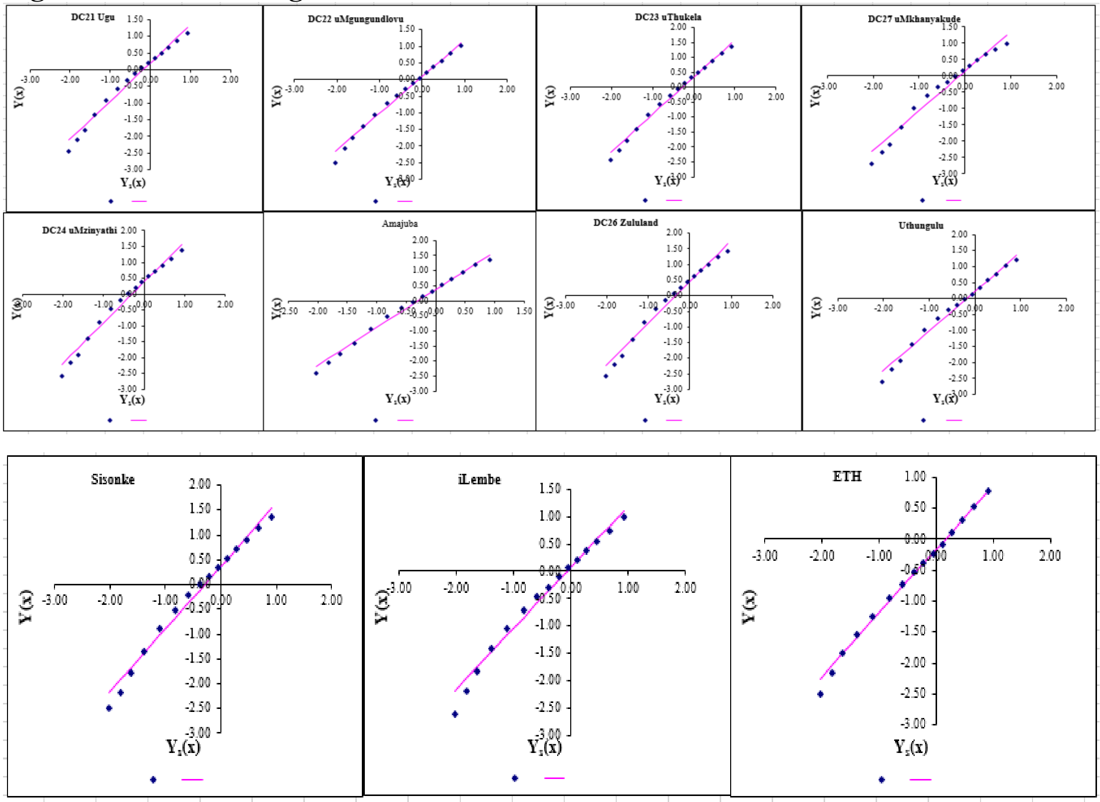


Figure D. 12 Male logit fit Mpumalanga

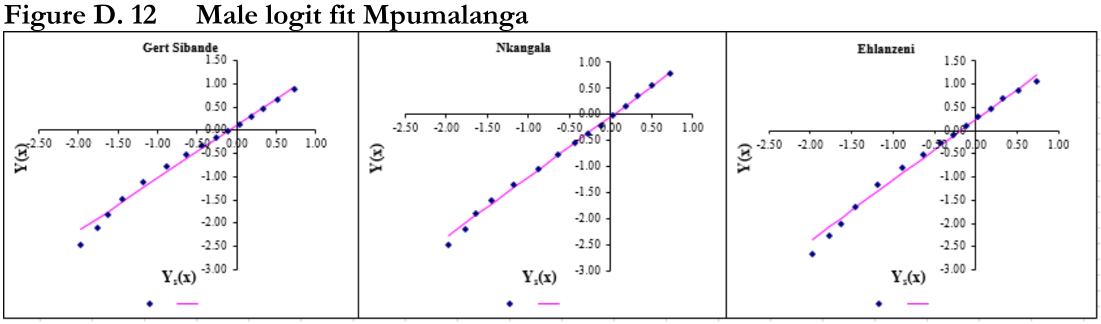


Figure D. 13 Female logit fit- Mpumalanga

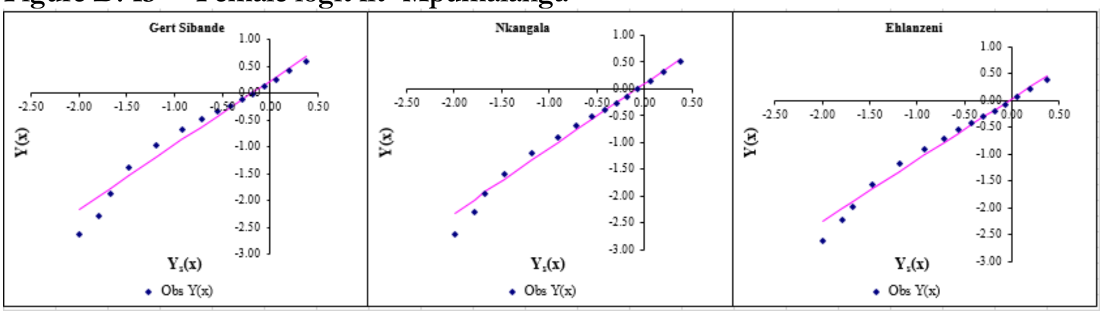


Figure D. 14 Male logit fit Limpopo

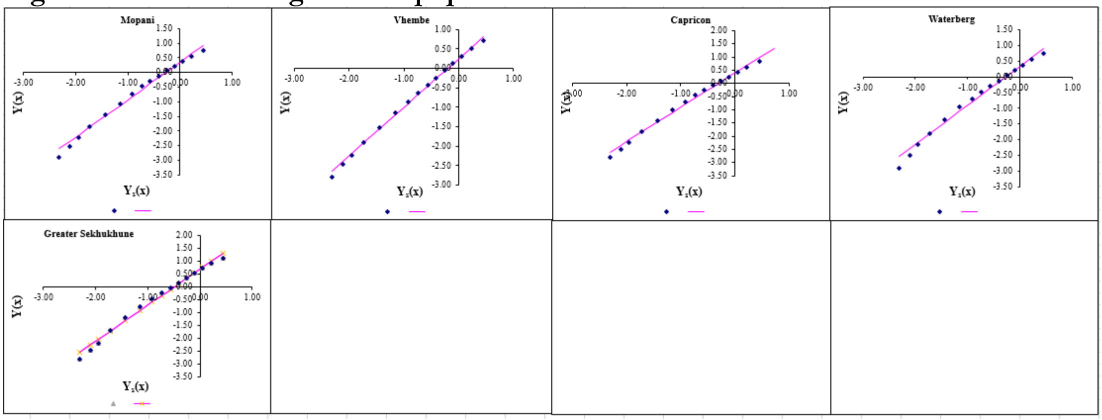


Figure D. 15 Female logit fit Limpopo

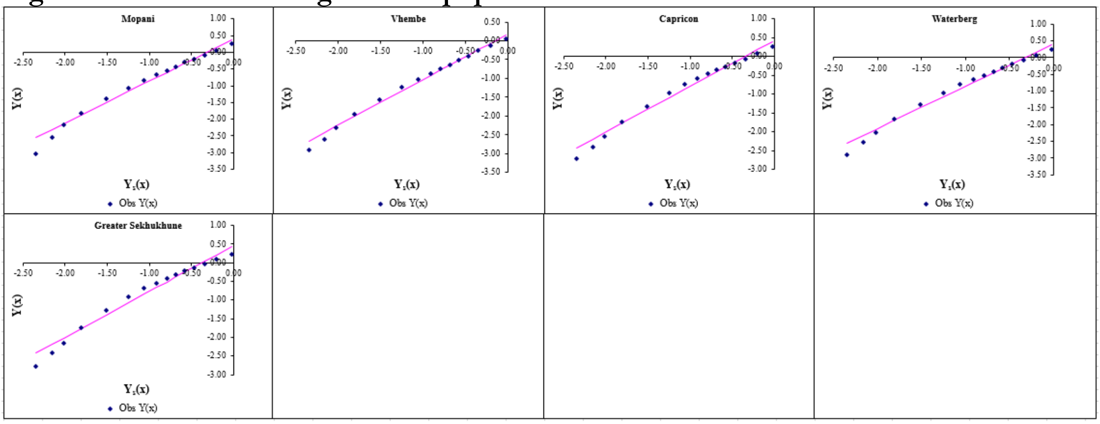


Figure D. 16 Male logit fit North West

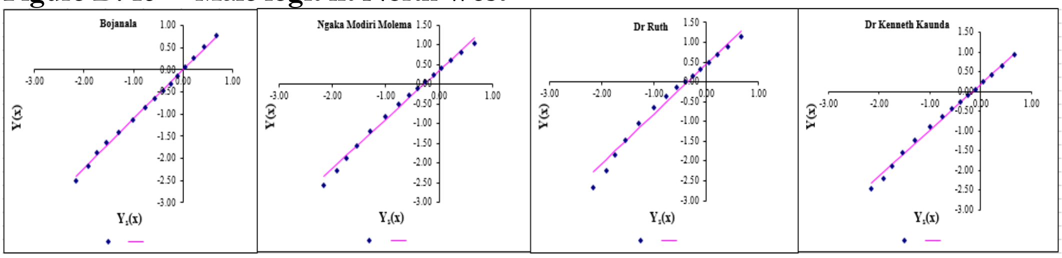


Figure D. 17 Female logit fit North West

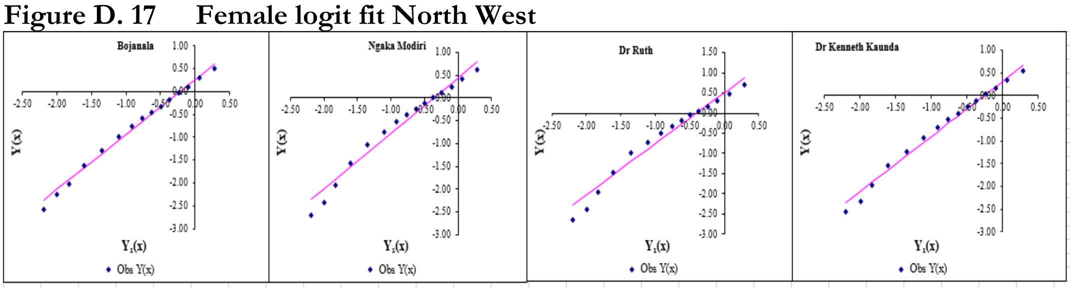


Figure D. 18 Male logit Fit Gauteng

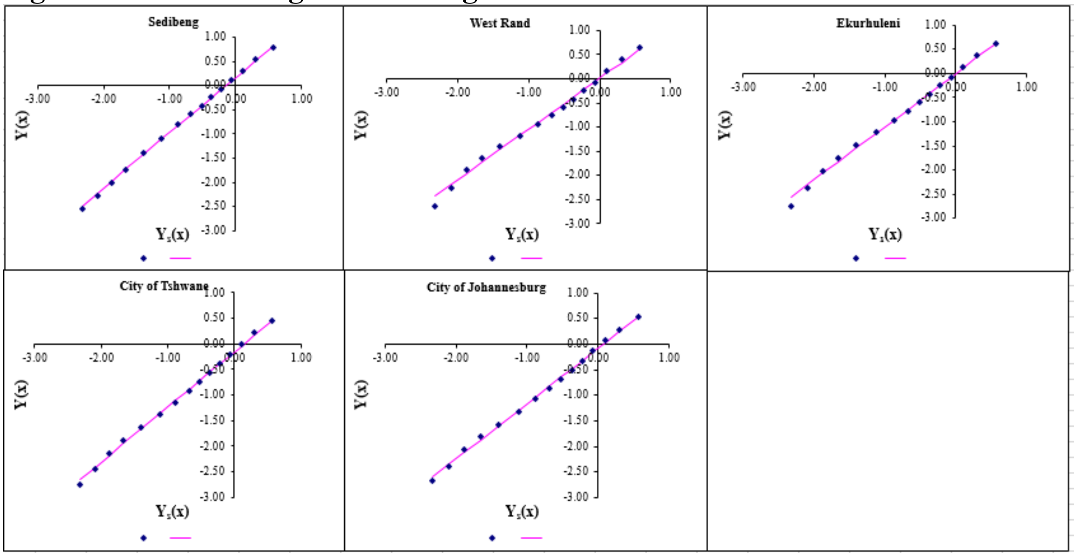
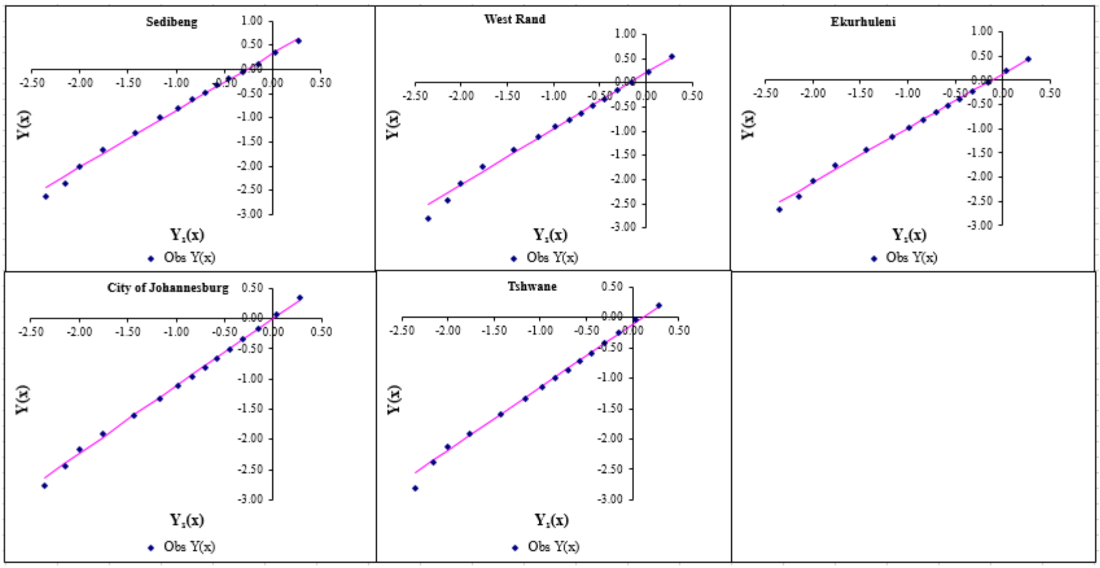


Figure D. 19 Female logit Fit Gauteng



Appendix E: Results

Table E 1 Opharnhood estimates at provincial level

Province	*Male*	*Female*
WC	0.35	0.25
NC	0.44	0.33
EC	0.51	0.37
FS	0.50	0.40
KZN	0.60	0.40
MP	0.53	0.39
LP	0.43	0.28
NW	0.59	0.36
GT	0.37	0.29

Figure E 1 Graduated and ungraduated adult mortality estimates at district level

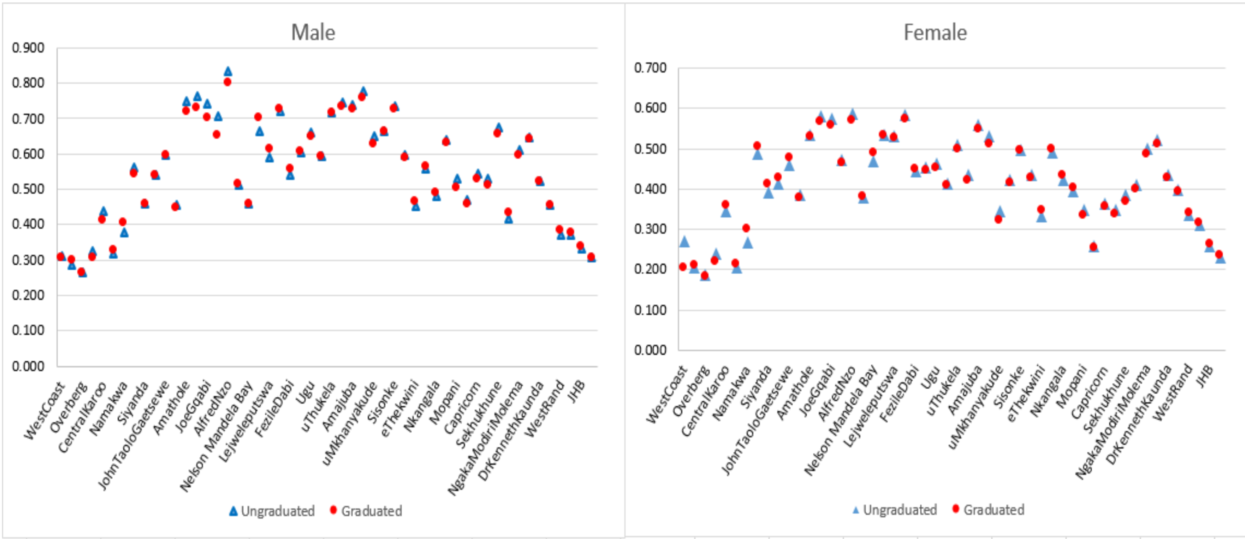


Figure E 2 Adult and child mortality estimates at a district level

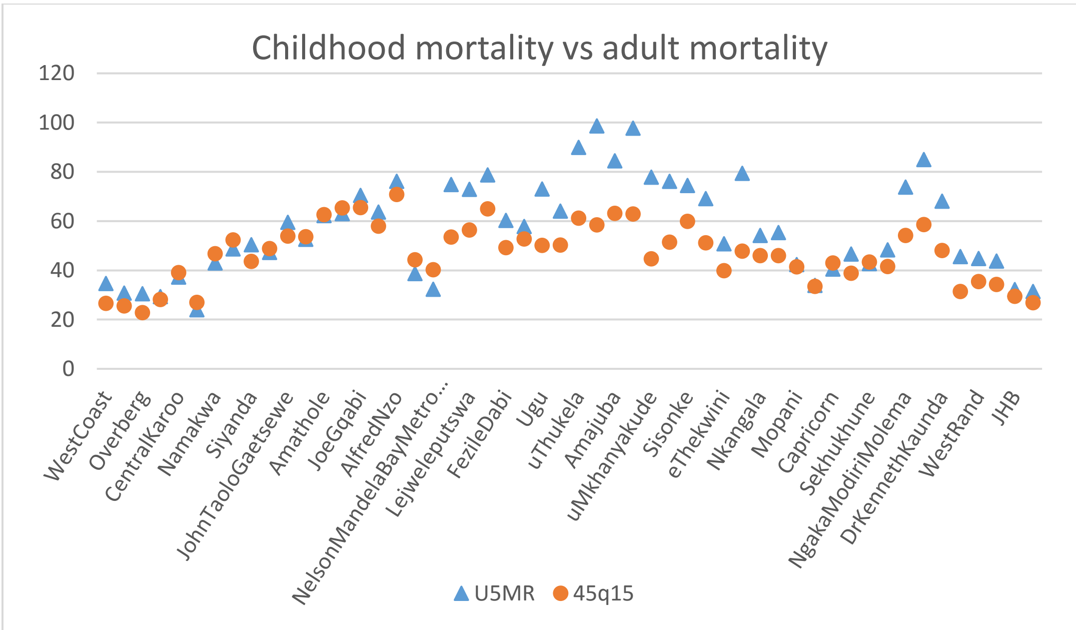


Table E 2 Percentage differences between adult mortality estimates derived in this study and estimates derived by Chinogurei (2017)

Districts	Male	Female	Districts	Male	Female
WestCoast	11%	0%	Ehlanzeni	-6%	-3%
CapeWinelands	6%	3%	Mopani	3%	-10%
Overberg	6%	4%	Vhembe	0%	-6%
Eden	8%	2%	Capricorn	3%	-8%
CentralKaroo	-3%	-18%	Waterberg	-7%	-6%
CPT	5%	3%	Sekhukhune	-13%	-11%
Namakwa	-2%	-12%	BojanalaPlatinum	3%	-2%
PixleykaSeme	6%	-11%	NgakaModiriMolema	0%	-4%
ZFMgcawu	8%	0%	DrRuthSegomotsiMompati	0%	-6%
FrancisBaard	5%	7%	DrKennethKaunda	5%	7%
JohnTaoloGaetsewe	-5%	-8%	Sedibeng	-3%	-11%
Cacadu	10%	-3%	WestRand	1%	-8%
Amathole	1%	3%	JHB	9%	2%
ChrisHani	-3%	-12%	CityofTshwaneMetropolitan	23%	9%
JoeGqabi	10%	-4%	Ekurhuleni	8%	6%
ORTambo	21%	-8%	uMgungundlovu	13%	11%
AlfredNzo	-27%	-23%	uThukela	9%	2%
BuffaloMetropolitan	19%	9%	uMzinyathi	7%	2%
NelsonMandelaBayMetropolitan	9%	-11%	Amajuba	3%	6%
MangaungMetropolitan	-7%	1%	Zululand	8%	-1%
Xhariep	2%	3%	uMkhanyakude	15%	8%
Lejweleputswa	-12%	-10%	uThungulu	6%	4%
ThaboMofutsanyana	18%	-7%	Sisonke	14%	4%
FezileDabi	-6%	-5%	iLembe	19%	8%
Ugu	15%	4%	eThekwini	15%	10%
GertSibande	8%	0%			
Nkangala	3%	-8%			

Figure E 3 Estimates of ${}_{45}q_{15}$ for women against the HIV prevalence from antenatal