FACULTY OF HEALTH SCIENCES

SCHOOL OF PUBLIC HEALTH

Department of Epidemiology and Biostatistics

Socio-Demographic Predictors of Cardiovascular Mortality in South Africa: Cross-Sectional Analysis of Stats-SA Data from 2006 – 2011

By:

Emery Ladi NGAMASANA

Student Number: 875503

A Research Report submitted to the Faculty of Health Sciences, University of the Witwatersrand of Johannesburg, in partial fulfilment of the requirements for the Msc degree in the field of Epidemiology and Biostatistics

November 2015

ABSTRACT:

Background: Worldwide, more people die annually from cardiovascular diseases than from any other cause. Evidence seems to support that these conditions account for nearly 30% of all deaths worldwide, with more than 10 million of deaths occurring among people aged less than 70 years old. Low and middle income countries are the most affected by these conditions, with proportions of more than 80% of cardiovascular diseases deaths occurring in these countries. Although reducing major risk factors such as tobacco use, obesity, hypertension and lifestyle are key steps toward an effective prevention of these man-made diseases, it is also highly recommended to document demographic and socioeconomic factors associated with these deaths. However, considerable uncertainties regarding the epidemiology of CVDs persist in the African continent.

Aims: This study aims to investigate the demographic and socioeconomic factors associated with cardiovascular diseases mortality in South Africa from 2006 – 2011. Secondly, the study wanted to document geographic variations of CVDs deaths across South Africa.

Methods: Using repeated cross-sectional data from the vital registration system in South Africa, we investigated demographic and socioeconomic factors associated with CVDs deaths in South Africa, using simple logistic regression model, Cox proportional hazard model and Cox frailty, accounting for a random effect at district level. The hazard ratios of the Cox shared frailty model were mapped to show geographical variation across the nine provinces.

Results: 3,454,868 deaths were registered at the Department of Home Affairs from 2006 – 2011. Of these deaths, about 500,633 were attributable to CVDs. The final sample contained 598,013 deaths for which relevant data were available. Gender, age, race, education, tobacco use were associated with CVDs death. Males had 6% reduced risk for dying from CVDs (HR: 0.94; 95% C.I: 0.92 – 0.95). The risk for CVDs mortality increased with education level, confirming lifestyle effect among those with tertiary education (HR: 2.19; 95% C.I: 2.11 – 2.28). Smokers had 31% increased risk (HR: 1.31, 95% C.I: 1.29 – 1.33) for dying from CVDs. Most important significant geographic variations were displayed regarding CVDs mortality, with people who lived in Free State and Kwazulu Natal having the highest risk for dying from CVDs (HR: 1.17; 95% C.I: 1.05 – 1.30 and 1.19; 95% C.I: 1.07 – 1.31 respectively) compared to those who lived in Northern Cape.

Conclusion: The results of this study highlight important socioeconomic and demographic differentials regarding CVDs mortality and provide supporting arguments for significant geographical variations of CVDs mortality at provincial level.

DECLARATION

I, Emery Ladi NGAMASANA, hereby declare that this research report is my own unassisted work except as mentioned in the reference list and acknowledgements. It is submitted in partial fulfilment of the requirements for the Degree Master of Sciences in Epidemiology in the field of Epidemiology and Biostatistics at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any other degree or examination at this or any other university.

Acknowledgement

"Success is the result of perfection, hard work, learning from failure, loyalty and persistence", Colin Powell

I would like to take this opportunity to acknowledge that this work would not come to a successful end without the commitment and generosity of many participants who accommodated me in their timetables and freely agreed to share their precious knowledge with me.

In the first position I would like to acknowledge God, the Everlasting Father, for his graces and protection, all praise and thanksgiving be to Him.

I would like to convey my gratitude to Professors Tobias CHIRWA and Ngianga Bakwin KANDALA who provided meaningful guidance and assistance throughout this study.

Distinct thanks goes to my wife Charlotte Yankaya NGAMASANA and daughter Olivia NGAMASANA, who have carried the heaviest burden of this adventure. May your sacrifice be rewarded well beyond your expectations. Without being exhaustive, I would like to express special thanks to my entire family for their support and care: my parents Georges NGAMASANA and Marie KANDALA, my sisters Lydie, Irene and Christelle and my brothers Fiston and Georges, my uncles Mons, Zola, Emery, Bienvenu and Theophile.

I would like to convey my gratitude to Statistics South Africa for providing the data we have analysed.

I really hope this work, not only open my paths toward new great adventures, but also provide great insights to inform public health policy accordingly.

LIST OF ABBREVIATIONS

ACME	: Automated Classification of Medical Entities
BRICS	: Brazil Russia Indian China and South Africa
CHD	: Coronary Heart Diseases
CVDs	: Cardiovascular Diseases
GDP	: Gross Domestic Product
HR	: Hazard Ratio
NCDs	: Non Communicable Diseases
NCHS	: United States Centre for Health and Statistics
OR	: Odds Ratio
SADHS1998	: South Africa Demographic and Health Surveys edition 1998
WHO MONICA	: WHO Multinational Monitoring of Trends and Determinants of Cardiovascular diseases
WHO	: World Health Organization

LIST OF TABLES

LIST OF FIGURES

LIST OF MAPS

Table of Contents

CHAPTER 1: INTRODUCTION

1.1. Background

There is increasing evidence of variations across the world of the pace of epidemiologic transition, with non-communicable diseases (NCDs) increasing and gradually displacing infectious diseases as the leading causes of morbidity and mortality. In developed countries, communicable diseases represent a marginal proportion of deaths (5.67% in the European region, 10.37% in the American region). The Public health agendas of these two regions, mainly, focus on reducing risk-factors associated with chronic diseases. In developing countries, infectious diseases are still the leading causes of morbidity and mortality (61.21% in the African region). The public health agendas of the latter are globally dominated by the fight against infectious diseases. However, this general trend hides a great menace of NCDs, whose magnitude is increasing steadily in the latter (WHO, 2013).

Indeed, more than three quarters of the world's population (82.5%) is living in low and middle-income countries and these countries account for more than three quarters (78%) of all deaths worldwide (Bureau, 2013). Although, for these countries, public health priorities include addressing issues of HIV pandemic and comorbidities, infant mortality, maternal mortality, lack of access to clean water, inadequate sanitation facilities and occupational diseases, it is also important to pay particular attention to chronic diseases whose burden is rising gradually.

For instance, the World Health Organization (WHO) reports that annually Cardiovascular Diseases (CVDs) are responsible for over 17.3 million deaths which represent 31% of all deaths worldwide (Shanti Mendis, 2011).

When looking at the African region, it is noticed that "populations are undergoing a rapid health transition with high burden of infectious diseases and a growing burden of non-communicable diseases" as society's demographic, economic and social structures are changing (Kandala et al., 2013b). According to the (WHO, 2013), in the African region non-communicable diseases represented about 30% of all deaths with cardiovascular diseases accounting for nearly 36% of deaths of NCDs. However, these estimates may not reflect the true situation since we know that administrative, logistic and technical loopholes surrounding the collection, storage and dissemination of these data, particularly in this region, may lead to underreporting. Nevertheless, research on the epidemiology of stroke and high blood pressure in Africa (Mensah, 2008) reveals that despite this lack of comprehensive data on cardiovascular conditions in Africa, the available data show that the age-standardized mortality, case fatality, and prevalence of disabling stroke in Africa are similar or higher than in most high-income regions.

South Africa is one of the African countries where substantial efforts have been made to collect health-related data and conduct supporting research. For instance, a study reports that in South Africa, approximately 195 people die daily due to cardiovascular conditions (Maredza et al., 2011). Data from (WHO, 2009) providing estimates of the burden of diseases, indicate that NCDs accounted for 28% of all deaths in South Africa in 2004, quite close to the average for the African region. Then, the proportion of deaths due to CVDs represented 20% of non-communicable diseases deaths and about 5% of all deaths. Besides, current trends on CVDs contrast with a widespread misconception that cardiovascular diseases are conditions inherent to very old people living developed countries (Mensah, 2008).

Admittedly, evidence suggests that the dynamic of epidemiologic transition requires significant attention to the surveillance of non-communicable diseases in Africa. South Africa, for instance, has a high burden of infectious diseases coupled with an increasing burden of chronic diseases. Researchers argue that demographic transition through migration, urbanisation and globalisation (tobacco smoking, physical inactivity and unhealthy diets) contribute to a large extent to the increase in the burden of NCDs in this setting (Maredza et al., 2011, Jha and Peto, 2014, Blot and Tarone, 2015).

Therefore, it is important from a public health perspective to investigate the factors associated with CVDs mortality in South Africa in order to reduce its burden as it is affecting not only older people but also middle age groups, torpedoing productivity, depleting human resources, reducing human capital, imperilling economic perspectives of the country and boosting healthcare outlays.

1.2. Research question

What are the socioeconomic and demographic factors associated with cardiovascular diseases mortality in South Africa from 2006 – 2011?

1.3. Justification of the study

The General Assembly of United Nations on prevention and management of non-communicable conditions highlighted the necessity of designing and implementing surveillance programs for these conditions in each country and at international level in order to reduce its burden (WHO, 2011). Such policy encompasses three key elements: surveillance, prevention and treatments. This research should be circumscribed in the surveillance and prevention arms of the strategy aiming to analyse not only the trends, but also the demographic and socioeconomic determinants of these conditions at national, provincial and district levels in South Africa.

Actually, there is still considerable uncertainty in the epidemiology of NCDs in Africa, especially that of cardiovascular diseases. For instance, less is known about the demographic and

socioeconomic predictors for cardiovascular mortality in South Africa and the district specific risks for cardiovascular diseases mortality, while many related research projects only focus on the trends in mortality and morbidity over time.

Thus, we argue that currently, what is known about cardiovascular morbidity and mortality is purely on trends over time. Actions and different related policies and the management of this emerging epidemic merely rely on these trends. However, as these conditions are considered to be "man-made diseases", it is important to further investigate the demographic and socioeconomic risk factors associated with the occurrence of these conditions and risk for death across different groups, as underlined in the WHO MONICA project (Evans et al., 2001). Furthermore, it is of a great relevance to investigate district-specific risk factors for cardiovascular diseases mortality in order to inform district level policies. The relevance of such studies is unquestionable as highlighted by (Corsi et al., 2014), as it may also have implications on resource allocation.

1.4. Hypothesis

- Deaths from cardiovascular diseases are not equally distributed across South Africa
- Age, gender, race, place of residence, education level, occupation and smoking status are associated with cardiovascular diseases mortality in South Africa during the period under study (2006 – 2011).

1.5. Aim and Specific Objectives

The study aims to determine the demographic and socioeconomic factors associated with cardiovascular diseases mortality in South Africa from 2006 – 2011.
Specific Objectives:

- To determine the cause specific deaths rate for cardiovascular diseases in South Africa from 2006 – 2011.
- To identify demographic and socioeconomic factors associated with cardiovascular diseases mortality in South Africa from 2006 – 2011.
- To determine geographic variations of cardiovascular disease mortality in South Africa from 2005 – 2011.

1.6. Literature Review

1.6.1. General Context of cardiovascular mortality worldwide

Several behavioural changes in the lifestyle are due to economic transition through urbanisation, industrialisation and globalisation. Notably the following risks factors: tobacco use, physical inactivity and unhealthy diet, alcohol abuse and uncontrolled stress, have increased the burden of CVDs because of ineffective prevention programmes and lack of treatment in

developing countries. Unfortunately, these conditions are affecting even middle age groups which bring in socio-economic implications (labour force depletion, low human productivity, decline in saving and investments, etc).

In 2008 more than 3 million of the 17 million deaths from cardiovascular diseases were premature (before 60 years) (Shanti Mendis, 2011). Comparative figures of cardiovascular mortality between developed and less developed countries project 60% of deaths due to cardiovascular to occur in developing world by 2020 (Kandala et al., 2013a).

Although such statistics increase the awareness on the magnitude of the problem, the pace of responses to the epidemic, largely differs across regions and countries. WHO's global atlas on CVDs indicates that in most developed countries, population-wide intervention and individual health-care have contributed to reduce the trends of cardiovascular mortality (Shanti Mendis, 2011). Undoubtedly, these programmes were mounted in the aftermath of WHO MONICA project designed in 1979 with the main objective of "monitoring trends and determinants of cardiovascular conditions at multinational level on a longitudinal basis to overcome limitations of cross-sectional studies" (Evans et al., 2001). As a matter of fact, the socio-economic and cultural context of developing countries imperils the implementation of effective programs. These structural impediments rather contribute to increase mortality rates among these countries.

1.6.2. Epidemiology of cardiovascular diseases in the African continent

Africa is characterised by low standard of living, low level of labour productivity, high level of socioeconomic inequalities, absolute poverty, adverse geography (desserts), larger rural population undergoing a rapid migration to cities and lower level of human capital (Todaro, 2012).

In such a context, the collection, storage and dissemination of health related data for research purposes is challenging. Thus, the epidemiology of cardiovascular diseases is less documented in Africa.

The WHO provides general information concerning the morbidity and mortality aspects of these conditions in the continent. Evidence indicates that cardiovascular diseases have nearly reached the threshold of an epidemic in the African continent with proportions of deaths due to cardiovascular diseases increasing from 8.5% in 2000 to 11% in 2011 for all deaths (WHO, 2013). The work of (Kadiri, 2005, Mensah, 2008) indicates that hypertension, stroke, cardiomyopathies and rheumatic heart diseases are the most prevalent chronic conditions in Africa.

As part of its global programs, the WHO is currently conducting programmes in countries and sub-regional communities in order to promote healthcare, prevent and control risks factors. In Africa, countries like Benin, Cameroon, Mali, Zambia, Zimbabwe, and Algeria are involved in a

programme for prevention and control of Rheumatic fever and Rheumatic heart disease (WHO, 2014a). Under this programme, countries are encouraged to reduce tobacco use by developing anti-tobacco legislation. South Africa is one of the few countries in Africa that has managed to get this progressive tobacco legislation with “the Tobacco Products Control Act of 1993” (Maredza et al., 2011).

1.6.3. Epidemiology of cardiovascular diseases in South Africa

Overall in Africa, less is known about the determinants of cardiovascular mortality compared to developed countries. Lack and dysfunction of vital registration systems in many countries renders it difficult for researchers to undertake comprehensive investigation in these settings. South Africa is one of the few countries in Africa with reliable vital statistics but still, less is known about the growing epidemic of chronic diseases. (Maredza et al., 2011) argue that the burden of cardiovascular diseases in South Africa constitutes a “hidden menace” for the economy as these conditions are “taking its toll even amongst the younger age groups” with a cause-specific mortality due to cardiovascular disease projected to increase by over 40% in the 34-65 years age group by 2030.
This constitutes an alarming trend in many regards. These conditions affect not only the population, but also undermine the country’s economic prospect. For instance, (Maredza et al., 2011) reveal that the cost related only to stroke treatment was reported to represent 2 to 3% of the South African GDP in 1991.

Although this figure might be outdated, it underlines a challenge as there is little evidence that this trend has decreased over the two decades. Whether this work of (Maredza et al., 2011) had the merit of highlighting the cost of inadequate policies to address the burden of these conditions, it lacked statistical evidence to support effective policies that would entail a national program involving provincial, district and individual at different levels. For instance, in this paper the “population-based intervention” aimed to address socio-economic determinants of these cardiovascular conditions without any statistical analysis of the selected risk factors.

Studies by (Kandala et al., 2013a, Kandala et al., 2013b) investigated the geographical distribution of cardiovascular comorbidity and of hypertension particularly in South Africa. The studies found that hypertension, coronary heart disease, stroke and hypercholesterolemia were the most prevalent cardiovascular diseases in South Africa. The first study described how shared behaviour (lifestyle variables) influences the geographic distribution of these four major cardiovascular conditions using data from SADHS 1998.

Hypercholesterolemia prevalence were evenly distributed across the country, whereas hypertension and coronary heart disease were highly prevalent (36% in Northern Cape) in the south western districts and the north-eastern districts recorded the lowest prevalence for hypertension and CHD.

Adjusting their model for other exposures such as socio-demographic covariates, the study found that gender (male), was associated with a reduced risk of hypertension while smoking and alcohol were not associated with adverse effects on cardiovascular diseases. Although investigating factors associated with chronic conditions was an important step beyond trends, these studies missed an important predictor which is socio-economic status.

Actually, one would argue that the socioeconomic status of people might explain significantly their adherence to a prevention program. For instance, work by (Helakorpi et al., 2008) found that the impact of an anti-tobacco legislation in Finland differed across socioeconomic groups with those in lower socioeconomic class being less likely to adhere to the legislation.

On the other hand, stroke and hypercholesterolemia were reported as rare conditions in the setting (Kandala&al., 2013). However, such finding is somehow restrictive as emphasized by the authors as far as data quality is concerned. This is supported by (Steyn, 2007) in a compiled analysed of cardiovascular conditions in South Africa in which he asserts that no accurate data exist on the prevalence and treatment status of common heart conditions in South Africa.

Results from both studies by (Kandala et al., 2013a, Kandala et al., 2013b) suggest that effective policies to address the burden of cardiovascular conditions must take into account these geographic differences and address all the other covariates accordingly.

As underlined by (Maredza et al., 2011) the focus on cardiovascular conditions is outbalanced by the overwhelming concerns about HIV/AIDS and tuberculosis. In fact, these two conditions merely draw the focus of policy-makers and funders thereby underestimating the impact of epidemiological transition in South Africa, whereas its effectiveness is obvious.

Therefore, it is obvious that there is a gap between what is known and what is still to be investigated about the socioeconomic and demographic determinants of cardiovascular diseases in South Africa and their distributions.

1.6.4. Classification of Cardiovascular diseases

According to (Hurst et al., 1999), cardiovascular diseases are classified based on their:

- **Aetiology** (causes or origin of the condition: atherosclerotic coronary heart disease; other types of coronary disease; hypertensive heart disease; rheumatic heart disease; congenital heart disease; aortic stenosis of the elderly; mitral valve prolapsed due to myxomatous degeneration; dilated cardiomyopathy (including idiopathic, alcoholic, post myocarditis, drug induced such as from Adriamycin); restrictive cardiomyopathy (give cause if known such as amyloid); hypertrophic cardiomyopathy (specify the subaortic, apical, or mid-ventricular type); infective endocarditis; chronic cor pulmonale (give causes of lung disease), etc.

- **Anatomy**: relates to the heart size and other anatomic abnormalities: for instance any cardiac or vascular abnormality that a pathologist can see without a microscope.
- **Physiology**: angina pectoris, heart failure (transient and chronic), rhythm disturbances (arrhythmias or conduction disturbances), syncope, and cardiogenic shock.

1.6.5. Risks factors for cardiovascular diseases

A work by (KLAG, 2013) suggests a relevant classification of the risk factors for cardiovascular conditions as follows:

- Modifiable risk factors (e.g. diabetes, smoking, obesity, physical inactivity, hypertension, alcohol, diet, socioeconomic status, inflammatory markers, infection, homocysteine, psychological factors) and
- Non modifiable (age, gender, ethnicity, family history, genetic polymorphism, coronary artery calcification.

Due to lack of variables in the dataset, we only looked at the following non modifiable risks: age, gender, ethnicity and some modifiable risks such as education, occupation and smoking status.

CHAPTER 2: METHODOLOGY

2.1. Study design and data source

The present investigation is a secondary data analysis of a series of cross sectional surveys conducted by Statistics South Africa (Statistics – SA) from 2006 – 2011.

Data on mortality and causes of deaths are collected routinely as per terms of the Births and Deaths Registration Act. No. 51 of 1992 as amended, and is administered by the Department of Home Affairs using forms BI-1663 and DHA-1663: notification/register of death/stillbirth(Statistics-SA, 2011). After collecting these forms, Stats-SA conducts data management. This includes sorting (by year, month, date and surname of the deceased person to avoid duplicates), pasting, precoding (of all demographic and socioeconomic variables because of the complexity of coding of diseases), coding of causes of deaths using ICD-10 guidelines and capturing of information on the deceased people. After all processes have been undertaken, the data are captured and duplicates are verified to ensure that all the death notification forms removed as duplicates during sorting were indeed duplicates and to identify any that were mistakenly taken out as duplicates. The data are then edited and inconsistent cases manually verified by more experienced coders.

The underlying cause of death is then derived automatically using Automated Classification of Medical Entities (ACME); a software developed by the United States Center for Health and Statistics (NCHS).

2.2. Study Population and Sample

The study was nationally-representative because the target population includes all registered deaths that occurred in South Africa from 2006 to 2011. Although administrative data of such nature is more likely to underestimate the true number of deaths because of underreporting and late registration, the Statistics – SA data on causes of deaths assumes 94% level of completeness of death registrations.

However, since we could not assume arbitrary observations on the covariates of interest, we disregarded missing values on the main outcome and the selected covariates. Thus, our final sample was made of death notifications for which observations on the main outcome and the selected covariates were available.

According to the latest census 2011, South Africa's population stood at 51,770,560 inhabitants distributed across nine (9) provinces: Northern Cape, Free State, Western Cape, North West, Eastern Cape, Gauteng, Kwazulu Natal, Mpumalanga and Limpopo. It is reported that one third (31%) of the population is younger than 15 years old and 7.7% is 60 years or older. In terms of gender, females account for 51.3% of the population (Government, 2014).

2.3. Setting

Data are collected by Statistics – SA at a national level by cross-sectional surveys and will cover the period from 2006 to 2011.

2.4. Data Analysis

Stata13 was used for data management (cleaning, dealing with missing values and recoding of variables that form common scale, or generating new variables and labelling variables) and for analysis. Besides, we used Microsoft Office Paint for mapping the hazard ratios of the robust model (Cox shared frailty) at provincial level.

Outcome of interest: death due to cardiovascular diseases according to the "International Classification of Diseases 10th Revision codes used by the Department of Home Affairs" (Statistics-SA, 2012). These include: acute rheumatic fever, chronic rheumatic fever, hypertensive disease, ischaemic heart disease, pulmonary heart disease and diseases of pulmonary circulation, other forms of heart disease, cerebrovascular disease, and diseases of arteries, arterioles and capillaries, diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified, other unspecified disorders of the circulatory system.

An outcome (event) of interest was coded "Yes" when one died from a cardiovascular condition and "No" if otherwise. The following covariates were considered: gender, race, marital status, occupation and education, place of residence (province), smoking status of next of kin and smoking status of the deceased. For the third objective, the survival time covered the period from birth to the occurrence of the event of interest (death).

The first objective was to determine the cause-specific mortality rate for cardiovascular diseases in South Africa. To achieve this objective, we used cross-tabulation of deaths due to cardiovascular diseases with the mid-year population of 2011. Frequency tables were used to describe the characteristics of the study population with respect to the selected covariates.

The second objective was addressed in three steps:

- In a descriptive analysis using frequency tables for categorical variables (level of education, age group, race, district of residence, marital status, etc.) we looked at the proportion of CVDs in each category. Bar graphs were used to show pictographically the distribution of sex, race, education and cardiovascular mortality. The only continuous variables (age) was described using mean and standard deviation and bar charts portrayed its distribution between the two categories of the outcome. The central limit theorem was applied to the age distribution of the deceased since we had a very large sample.
- Since our outcome of interest was a binary variable, we fitted a logistic regression model to relate the response variable with aforementioned covariates. This allowed us

to find the best fitting, parsimonious and epidemiologically plausible model to quantify the relationship. A bivariate analysis using logistic regression modelling allowed investigating significant predictors for cardiovascular diseases mortality. Only significant covariates (at 20% level of significance in the bivariate analysis) were included in the multiple logistic regression models.

- We then carried out a multiple logistic regression model to adjust for the effect of other covariates. Our final model was obtained by forward elimination, using the likelihood ratio test. The final model was refitted and a Hosmer-Lemeshow test was carried out to test for lack of fit. Marginal probabilities from this final model were used to aid interpretation.

The third objective looked at the pattern of time to death from cardiovascular diseases in South Africa while accounting for latent dependency at district level. To achieve this objective, we fitted a Cox proportional hazard model and Cox survival model (Cox shared frailty) to investigate a district random effect on CVDs deaths. This helped in identifying latent district frailty.

For the Cox survival regression model, proportionality assumptions were tested using *Schoenfeld residuals* and *Grambsch & Therneau* tests and equality of the survival functions was investigated using log-rank tests for each covariate. The overall goodness of fit for the survival model was tested using Cox-Snell residuals and martingale residuals. We plotted Nelson-Aelen cumulative hazard vs Cox-Snell residuals (figure 12).

2.4.1. Defining the logistic regression model

Under this section, we defined the stochastic and systematic structures of the model to ease the understanding of the regression model and the interpretation.

A. Stochastic structure of the data

The choice of logit model under this study is justified by the fact that our response variable "death due to cardiovascular disease" only takes two possible values representing success or failure. Failure meant "death due to cardiovascular event" and success meant death due to other conditions.

Since our data are collected at individuals level, for practical convenient, we used Bernoulli distribution for individual 1/0 data. The result of our model was a Generalized Linear Model, "GLM" with binomial response and link logit.

The stochastic structure of the model was discussed in terms of Bernoulli distribution.

Assuming that our attribute of interest y_i was a binary variable coded 1 or 0, we could note:

$$y_i = \begin{cases} 1\ if\ the\ deceased\ person\ died\ from\ a\ CVD \\ 0\ if\ otherwise\ (died\ of\ other\ causes) \end{cases} \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Equation (1) implied that y_i could be defined as a random variable Y_i that took the value 1 or 0 with a probability $\pi_i\ and\ 1 - \pi_i$. In other words,

For a given value of the outcome of interest y_i we had:

$$y_i = 1 \ldots\ldots\ldots\ldots . Y_i \approx \pi_i$$

And for $y_i = 0 \ldots\ldots\ldots\ldots . Y_i \approx 1 - \pi_i$

Admittedly, the expected value and the variance of Y_i could be defined respectively as:

$$E(Y_i) = \mu_i = \pi_i \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$Var(Y_i) = \sigma_i = \pi_i(1 - \pi_i) \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Given these two parameters, we argued that any factor affecting the probability of the occurrence of the event of interest would alter both the mean and variance.

B. Logit transformation

The next step in defining our model concerned the systematic structure. The idea was to have situation where the probability of π_i (probability of dying from a cardiovascular event) depends on a vector of the observed covariates x_i (age, gender, race, etc.).

To achieve such a situation, we let π_i be a linear function of the covariates so that:

$$\pi_i = X'\beta \ldots\ldots\ldots . . (4)$$

Since equation 4, known as linear probability model poses a problem with the right-hand-side probability which can be out of range, because it can take any value, we transformed the probability to remove the range restrictions so that we could model the transformation as a linear function of the covariates.

First step: We moved the probability π_i to the odds:

$$Odds_i = \frac{\pi_i}{1 - \pi_i} \ldots\ldots\ldots\ldots\ldots\ldots\ldots . (5),$$

Equation 5 can be defined as the ratio of favourable to unfavourable cases. Thus the odds can take any positive value and therefore have no ceiling restriction.

Second step: We took the log-odds (logit)

Let η_i be the logarithm of the odds:

$$\eta_i = logit\,[\frac{\pi_i}{1-\pi_i}] \ldots\ldots\ldots\ldots\ldots\ldots\ldots (5),$$

Equation 5 removes the floor restriction, thereby mapping the probabilities from the range (0, 1) to the entire real line.

From equation 5, we could define the logistic regression model, assuming that the logit of the probability π_i follows a linear model.

C. The logistic regression model

Given that we have k independent observations of the outcome of interest $y_1 \ldots y_k$, and that the i-th observation can be treated as a random variable Y_i, we then assumed that Y_i has a Binomial distribution with two parameters:

$$Y_i \sim B(n_i, \pi_i)$$

This defines the stochastic structure of the model. Then, the logit of the underlying probability π_i is a linear function of the predictors:

$$logit(\pi_i) = X_i^{'}\beta \ldots\ldots\ldots\ldots (6)$$

Where X_i is a vector of covariates and β a vector of regression coefficients. Equation (6) defines the systematic structure of the model. The model defined in equations 5 and 6 is a generalised linear model with binomial response and link logit.

Hence, in terms of interpretation, β_i represents the change in the logit of the probability associated with a unit change in the j-th predictor holding all other predictor constant.

2.4.2. Defining the spatial survival model

Our model sought to explain how in South Africa, the risk (hazard) of dying from a cardiovascular disease at a given time is affected by the selected demographic and socioeconomic variables. In this regards, the hazard is the instantaneous risk of an individual dying from cardiovascular disease at a given time assuming that it has survived up to that time (Darmofal, 2009).

Although cardiovascular diseases survival rates are improving in South Africa as more effective prevention measures and therapies are introduced, considerable patient-heterogeneity remains conditional on treatment and known prognostic factors. Under this study we investigated if spatial effects may account for this heterogeneity (district variation).

To achieve this objective, we started with a standard survival analysis, ignoring any spatial variation across district. Data consisted of survival time t, death due to cardiovascular event (censoring indicators, δ) and a set of covariates.

Thus we postulated:

- t_{ij}, as the time to death for subject j in district i, with $j = 1, \dots, n_i$ and $i = 1, \dots m_i$.
- $x_{ij,}$ as a vector of individual-specific covariates.

With these two parameters, we defined a semi-parametric proportional hazard model with the following stochastic structure (Henderson et al., 2002, Banerjee et al., 2003):

$$h(t_{ij}, x_{ij}) = h_0(t_{ij}) \exp(\beta^T x_{ij}) \dots\dots (10)$$

Where:

- h_0 is the baseline hazard (risk in the absence of any covariate)
- β^T a vector of regression coefficients
- Equation 10 is a semi-parametric Cox model, which assumes no parametric distribution for the baseline hazard.

However, equation (10) assumes that X_{ij} represents a vector of all covariates affecting the hazard (risk) of dying from a cardiovascular condition. Admittedly, such assumption is very restrictive since it ignores any spatial dependency. To overcome the effect of such dependency, we then introduced shared frailty terms because we believed that deceased persons are clustered such that subjects within the same cluster or stratum (districts) share the same frailty while frailties are independent across strata. Thus, according to (Darmofal, 2009, Banerjee et al., 2003) equation (10) becomes:

$$h(t_{ij}, x_{ij}) = h_0(t_{ij})\, \omega_i \exp(\beta^T x_{ij}) \dots\dots (11)$$

$$h(t_{ij}, x_{ij}) = h_0(t_{ij}) \exp(\beta^T x_{ij} + \ W_j) \dots\dots (12)$$

Where

- Individual i is now nested in stratum j and
- $W_j \equiv \log \omega_i$ is the stratum-specific frailty term.

Typically, as mentioned by (Banerjee et al., 2003) equation 12 assumes that this random effect term W_i follows a normal distribution with two parameters mean and variance $W_i \sim N(0, \theta^2)$.

If the variance $\theta^2 = 0$, then the frailty term does not affect the model, meaning there is no variation across district and equation (11) can be reduced into equation (10). However, if $\theta^2 \neq 0$, it suggests spatial variability, which required to factor in the frailty term (Henderson et al., 2002, Banerjee et al., 2003). Under this study, we postulated variation across provinces and districts as far as deaths due to cardiovascular diseases were concerned.

2.5. Limitations

The present investigation is expected to have some limitations. Although data from Statistics – SA are up-to-date, they have quality issues. Constraints include underreporting of cases mainly in rural areas, and lack of ability to ascertain deaths. Many other risk factors cannot be assessed as they are not recorded. For example, socioeconomic status as a composite indicator comprised many variables such as income, occupation, household prestige, etc. is missing and was not included in this analysis. Nevertheless, some of its components such as occupation and education level were used as surrogates.

Another limitation was related to the third objective. An ideal situation would consist of investigating the survival time running from CVD diagnosis to death. However, the dataset does not contain such a variable. Hence, use of survival time running from birth to death may not capture many variations occurring during the lifetime of the deceased people. However, we think that this survival time (from birth to death) might be used as a surrogate of the time from diagnosis to death.

This study would have been enhanced if risks factors of psychological spectrum were available in the dataset.

Another major limitation of this study is related to the nature of the geographical variations analysis. Although the study wanted to document the existence of latent geographical variations of CVDs across the provinces in South Africa, such analysis would be achieved much better with appropriate spatial model using sophisticated Geographic Information System (GIS) techniques of spatial interpolation such kriging, Splines, etc. However, a full discussion of spatial modelling of CVDs mortality is beyond the scope of this dissertation. Therefore, further studies in this area need to be done.

2.6. Ethical issues

Data from Stats-SA are vital statistics from the entire country held by the department of Home Affairs. Even though these data are available in the public domain, an ethical clearance certificate (No. M140912) was obtained from Wits Human Research Ethics Committee and an authorization to use the data was obtained from the gatekeeper. Confidentiality of the study participants was insured my anonymising the data and removing any personal identifier.

2.7 Dissemination of results

After completion and submission of this research to the Faculty of Health Sciences for degree purpose, results will be published in a peer reviewed journal or a conference presentation to inform public health policies in South Africa, at national, provincial and district level.

CHAPTER 3: RESULTS

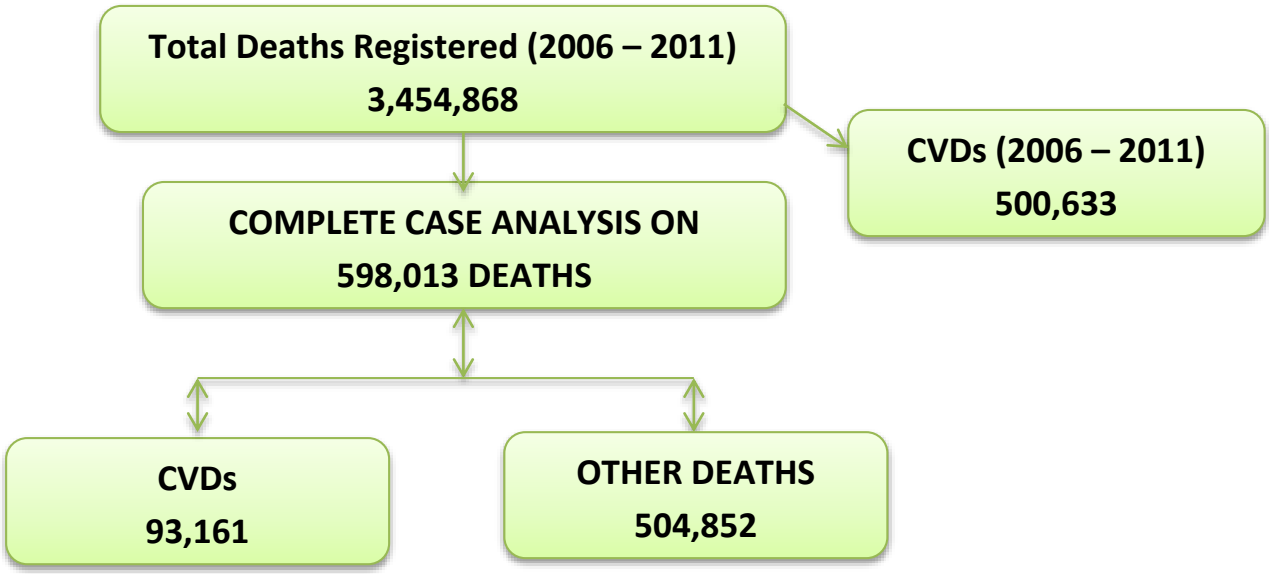
3.1. Introduction

This chapter presents the results of the descriptive and inferential analyses of the study. It will be followed by a chapter on discussion and conclusion.

In the first stage of analysis, we described the socioeconomic and demographic characteristics of the study population (Table 1 & Table 2). Thereafter, we described cross sectional trends of these baseline characteristics by cardiovascular deaths status from 2006 to 2011 (Table 3) and we described the main types of cardiovascular diseases (Table 4). The chapter ends by presenting results of the logistic regression model, Cox proportional hazard and Cox shared frailty models (Table 5). Much of the tables were followed by figures for graphical illustration.

Data Flow Chart

Data Flow Chart 1



3.2. Overall mortality in South Africa from 2006 – 2011

The Department of Home Affairs registered about 3,454,868 deaths from 2006 to 2011. Of the 3,454,868 deaths, 500,633 deaths were attributable to cardiovascular diseases. Thus, for 2011 for instance, there were 82,988 CVDs deaths which gives an estimate for a CVD death rate of $\frac{82,988}{50,586,757} * 100,000 = 164\ CVDs\ deaths\ per\ 100,000\ deaths$ in 2011, using the mid-year population of 2011.

However, our final sample included 598,013 deceased people for whom data on the underlying cause of death and other covariates were available. Of the 598,013 deaths, there were about

305,994 males representing a proportion of 51.17% against 292,019 females representing 48.83% of the study sample (Fig.1.). The median age of the deceased was 46 years with inter quartile range of 31 (IQ1: 34 – IQ3: 65). The mean age was 50 (95% C.I.: 49.97 – 50) with a standard deviation of 24.58 years.

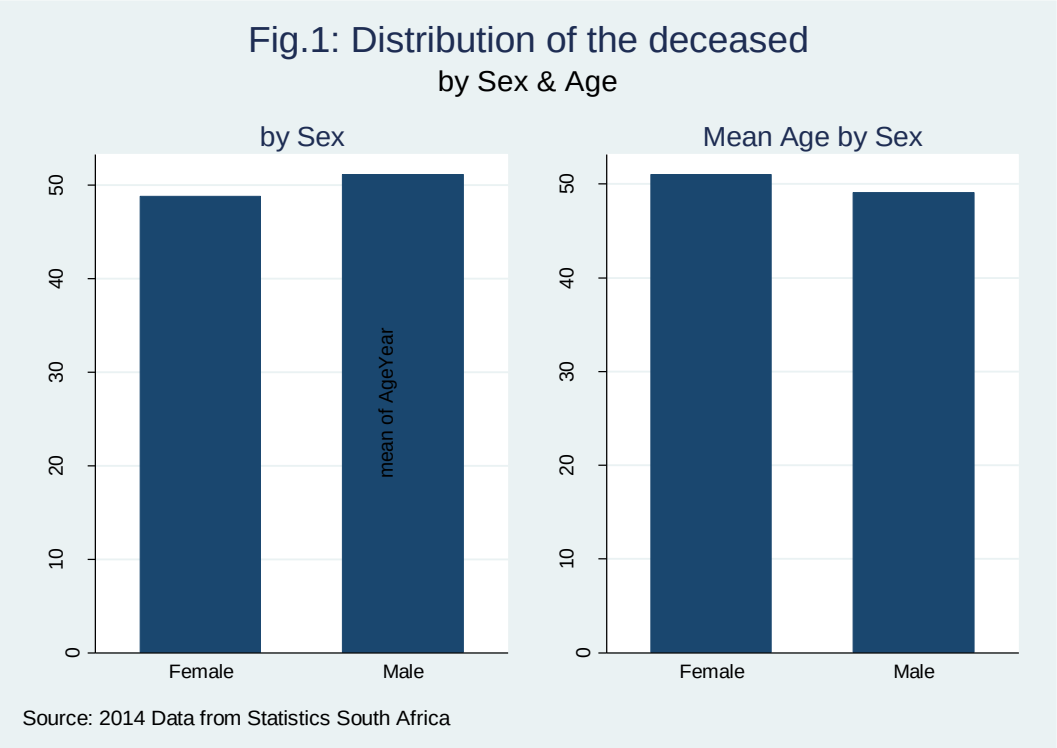


Figure 1: Distribution of the deceased by Sex and Age

Table 1: Distribution of deaths by socioeconomic and demographic characteristics (Statistics- SA, 2014)

Variable	N=598,013
	n (%)
Gender	
Female	292,019 (48.83)
Male	305,994 (51.17)
Age category	
< 25	41,828 (6.99)
25 – 34	119,923 (20.05)
35 – 44	117,341 (19.62)
45 – 54	91,276 (15.26)
55 – 64	74,620 (12.48)
65+	153,025 (25.59)
Education level	
No Schooling	106,959 (17.89)
Primary	145,264 (24.29)
Secondary	326,455 (54.59)

Tertiary	19,335 (3.23)
Marital Status	
Single	331,248 (55.39)
In Union	206,481 (34.53)
Separated	60,284 (10.08)
Race	
Black	513,166 (85.81)
White	42,639 (7.13)
Indian	15,660 (2.62)
Coloured	26,548 (4.44)
Occupational exposure	
Extreme temperature	45,377 (7.59)
Chemical exposure	58,259 (9.74)
Sedentary jobs	494,377 (82.67)
Smoking	
Non Smokers	414,869 (69.37)
Smokers	183,144 (30.63)
Province of residence	
Western Cape	21,142 (3.54)
Eastern Cape	84.483 (14.13)
Northern Cape	17,355 (2.90)
Free State	74,474 (12.45)
Kwazulu Natal	162,376 (27.15)
North West	35,898 (6.00)
Gauteng	63,906 (10.69)
Mpumalanga	77,966 (13.04)
Limpopo	60,413 (10.10)

Of those who died from 2006 - 2011, people aged 65 years and older represented approximately 26% and those aged less than 25 years represented only 7% (Fig. 2).

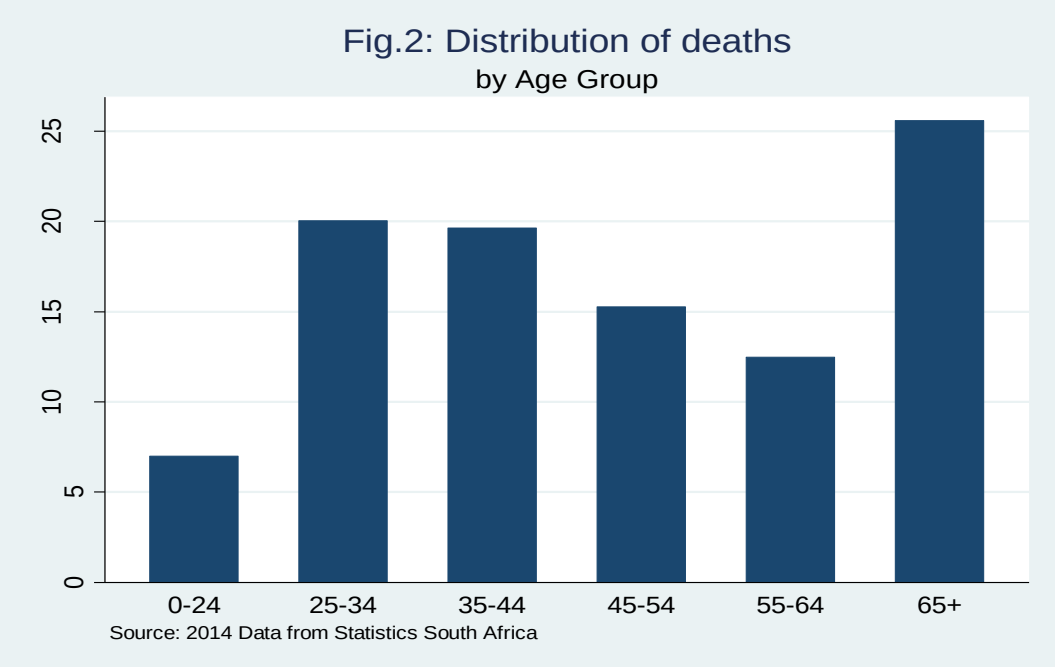


Figure 2: Distribution of deaths by Age Group

In terms of education, about 55% had a secondary education level against only 3.23% who had a tertiary education level.

In terms of race, the largest majority of those who died were black Africans (86%) compared to 7% White, 4% Coloured and 3% Indians (Fig.3). Concerning the age of the deceased by population group, Blacks had the lowest mean age (48 years), and White had the highest mean age (69 years).

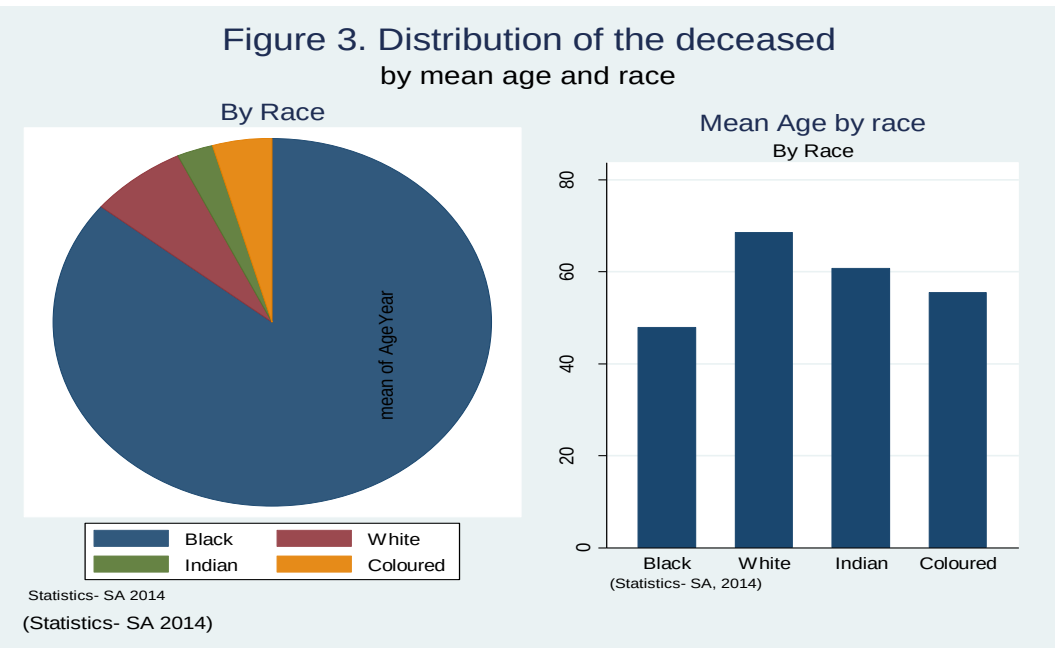


Figure 3: Distribution of the deceased by mean age and race

Regarding the type of occupation, about 82% of the deceased worked in a sedentary sector (including professional and managerial position), compared to 8% who were exposed to extreme temperatures (heat or cold) and 10% who were exposed to chemical substances during their work life.

Of those who died, about 183,144 smoked cigarettes at least five years before they died accounting for 31% of the deaths. Of these 183,144 former smokers; 42,179 (23%) resided in Kwazulu Natal against 9,157 (5%) who lived in Western Cape. Besides, of the 292,019 females who died between 2006 and 2011, approximately 36,560 (12.52%) were former smokers. Of the total number of female smokers, 25,471 (70%) were less than 65 years old.

Concerning the area of usual residence of the deceased, Kwazulu Natal hosted about 27% of the deceased, Eastern Cape (14%) and the smallest proportion of the deceased lived in Northern Cape (3%).

3.3. Cardiovascular diseases mortality in South Africa from 2006 - 2011

Table 2 shows the overall baseline characteristics of the deceased by cardiovascular diseases mortality and table 3 shows the trends of cardiovascular diseases mortality throughout the 6 years under study.

On overall, for the period under study 93,161 deaths were attributed to cardiovascular diseases as defined by the ICD-10 codes, representing 15.58% of the overall deaths. Of these CVD related deaths, 52,235 occurred among females (representing 17.89% of all female deaths) and 40,926 affected males (13.37% of all male deaths).

CVDs related deaths were more prevalent among people aged 65 years and older (34.29% of the deaths in this age group). In addition, cardiovascular related deaths were more prevalent among people with no education and tertiary education (25.05%; 16.78% respectively) and less prevalent among people with secondary education (12.27%).

Concerning race, cardiovascular deaths occurred in a great proportion of Indians and whites (31.27% and 30.93% respectively) whereas only 13.53% of deaths among blacks were attributable to cardiovascular diseases. Exposure to smoking accounted for 13.53% of cardiovascular related deaths while among non-smokers, cardiovascular diseases accounted for 16.48% of deaths.

Table 2: Baseline Characteristics by cardiovascular diseases mortality (Statistics-SA, 2014)

Characteristics	CVD deaths (n=93,161)	Non CVDs deaths (n=504,852)	p-value
Sex			
Female	52,235 (17.89)	239,804 (82.12)	0.001
Male	40,926 (13.37)	265,088 (86.63)	
Age Group			
<25	1,307 (3.12)	40,521 (96.88)	0.001
25 – 34	3,769 (3.14)	116,154 (96.86)	
35 – 44	6,683 (5.70)	110,658 (94.30)	
45 – 54	11,984 (13.13)	79,292 (86.87)	
55 – 64	16,953 (22.72)	57,667 (77.28)	
65+	52,465 (34.29)	100,560 (65.71)	
Age (mean ± SD)	65.63 (19.19)	47.16 (24.39)	0.001
Education Level			
No schooling	26,795 (25.05)	80,164 (74.95)	0.001
Primary	23,081 (15.89)	122,183 (84.11)	
Secondary	40,041 (12.27)	286,414 (85.73)	
Tertiary	3,244 (16.78)	16,091 (83.22)	
Marital Status			
Single	30,071 (9.08)	301,177 (90.92)	0.001
Married/in relationship	45,264 (21.92)	161,217 (78.08)	
Separated	17,826 (29.57)	42,458 (70.43)	
Race			
Black	69,429 (13.53)	443,737 (86.47)	0.001
White	13,188 (30.93)	29,451 (69.07)	
Indian	4,897 (31.27)	10,763 (68.73)	
Coloured	5,647 (21.27)	20,901 (78.73)	
Type of occupation			
Exposure to extreme temperature	7,347 (16.19)	38,030 (83.81)	0.001
Exposure to chemical substances	9,701 (16.65)	48,558 (83.35)	
Sedentary work	76,113 (15.40)	418,264 (84.60)	
Smoking Status			
Smoker	24,7772 (13.53)	158,372 (86.47)	0.001
Non Smoker	68,389 (16.48)	346,480 (83.52)	
Next of kin smoking Status			
Smoker	19,355 (15.37)	106,602 (84.63)	0.019
Non Smoker	73,806 (15.64)	398,250 (84.36)	
Province of usual residence of the deceased			
Western Cape	5,312 (25.13)	15,830 (74.87)	0.001

Eastern Cape	12,925 (15.30)	71,558 (84.70)
Northen Cape	2,896 (16.69)	14,459 (83.31)
Free State	11,005 (14.78)	63,469 (85.22)
Kwazulu-Natal	23,188 (14.28)	139,188 (85.72)
North West	6,023 (16.78)	29,875 (83.22)
Gauteng	11,610 (18.17)	52,296 (81.83)
Mpumalanga	11,283 (14.47)	66,683 (85.53)
Limpopo	8,919 (14.76)	51,494 (85.24)

We also were interested in comparing occupation exposure during the lifetime of the deceased. They were then categorized into three main groups, depending on the occupational hazard: exposure to extreme temperature (heat or cold), exposure to chemical substances, and sedentary jobs (including managerial position, office work, sellers, etc.). Table 2 below, suggests that these occupational hazard are associated with CVDs.

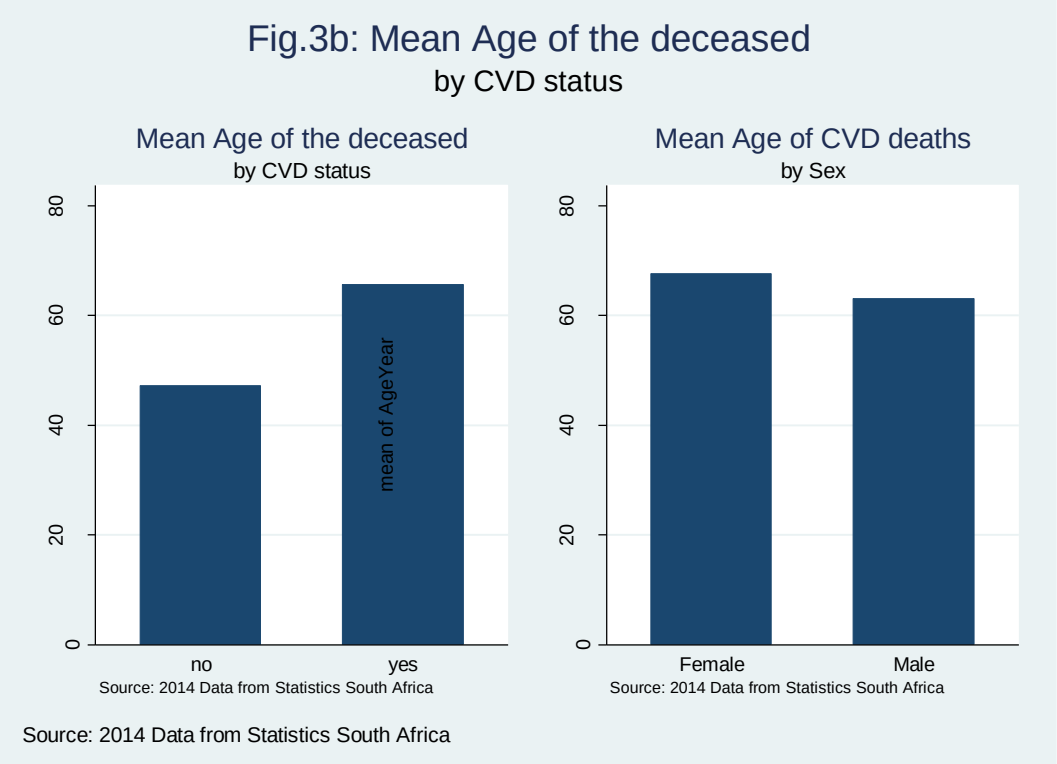


Figure 4(3b): Mean Age of the deceased by CVDs Status

Of those who died between 2006 and 2011 from cardiovascular diseases, on average CVDs deaths occurred around 65 years whereas other deaths occurred around 47 years of age. Of those who died from CVD, female died on average around 67 years whereas male died on average around 63 years of age (Fig.3b).

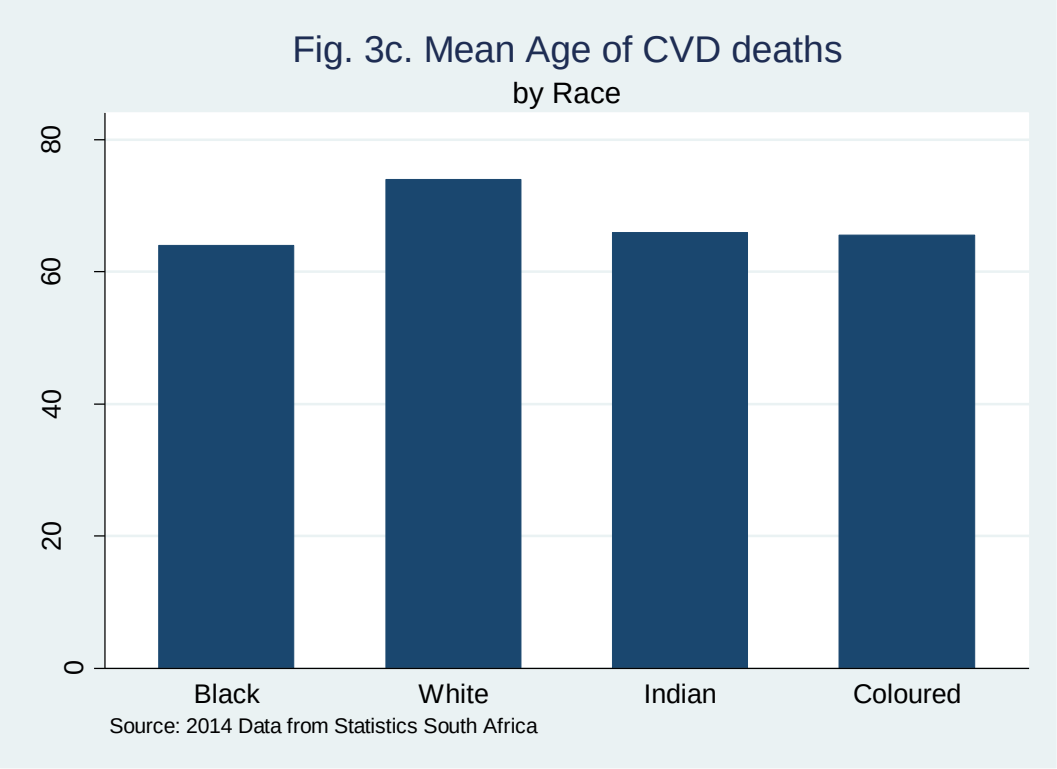


Figure 5(3c): Mean Age of CVD deaths by Race

3.4. Trends of cardiovascular diseases mortality in South Africa from 2006 - 2011

We looked at the trends of cardiovascular diseases deaths over the 6 years under study. Table 3 shows the trends by socioeconomic and demographic covariates of interest.

Table 3: Trends of proportion of cardiovascular diseases mortality by socioeconomic and demographic characteristics in South Africa 2006 – 2011 (Statistics-SA, 2014)

Variable	2006 (N=17,053)	2007 (N=17,005)	2008 (N=17,137)	2009 (N=17,741)	2010 (N=15,058)	2011 (n=9,167)
Sex						
Female	16.85	17.05	17.31	18.02	18.86	21.54
Male	12.97	12.87	12.93	13.27	13.69	16.15
Age category						
< 24	3.06	3.43	2.78	3.35	3.13	2.68
25 – 34	3.10	3.06	3.00	3.12	3.40	3.48
35 – 44	5.61	5.53	5.45	5.62	6.20	6.26
45 – 54	13.76	13.30	12.38	12.56	12.92	14.76
55 – 64	23.30	22.29	22.17	22.60	22.70	23.82
65+	34.16	34.18	34.76	33.65	33.77	35.91
Education level						
No Schooling	24.19	24.74	25.14	25.13	25.17	27.14
Primary	14.88	14.92	15.20	16.09	17.00	19.78
Secondary	11.63	11.61	11.68	12.18	13.06	15.29

Tertiary	16.75	14.94	17.24	16.41	16.87	21.06
Marital Status						
Single	8.63	8.79	8.67	8.89	9.61	11.29
In Union	19.89	20.16	20.39	21.07	24.63	27.24
Separated	30.13	28.98	30.13	30.50	21.37	24.26
Race						
Black	12.79	13.01	13.14	13.62	14.16	16.13
White	30.74	29.71	30.64	31.20	30.81	33.51
Indian	30.98	29.79	30.37	31.59	31.92	35.08
Coloured	21.61	21.16	21.63	19.86	20.18	24.18
Occupational exposure						
Extreme temperature	16.49	16.05	16.32	17.36	16.80	15.38
Chemical exposure	14.86	14.16	14.56	14.79	16.04	24.83
Sedentary jobs	14.81	14.92	15.05	15.53	16.15	18.24
Smoking						
Smokers	13.17	12.87	13.14	13.38	13.91	16.18
Province of residence						
Western Cape	25.92	24.64	26.76	24.82	23.36	25.36
Eastern Cape	14.37	15.06	14.86	14.77	16.25	18.13
Northern Cape	16.15	17.66	15.96	15.19	17.23	21.50
Free State	13.34	13.71	14.74	15.03	15.64	17.75
Kwazulu Natal	13.28	14.26	13.53	14.40	15.10	18.27
North West	17.42	16.42	16.21	16.50	17.40	16.82
Gauteng	17.92	15.88	17.62	18.61	18.79	23.84
Mpumalanga	13.53	13.87	14.04	14.82	14.39	17.63
Limpopo	15.15	13.74	14.49	14.80	15.41	15.80

The proportions of deaths due to cardiovascular diseases increased steadily throughout the period under study. For instance, among females CVD related deaths increased by 1.19% between 2006 and 2007, whereas it increased by 4.7 a couple of years later between 2009 and 2010. The overall increase among female was 27.83% between 2006 and 2011 compared to a 24.52% increase among males (Fig. 4).

Of those who died between 2006 and 2011, the proportion of CVD related deaths increased by 32.93% among people with primary education and by 31.47% among people with secondary education. Proportions of these CVD related deaths increased by 25.73% among those with tertiary education between 2006 and 2011.

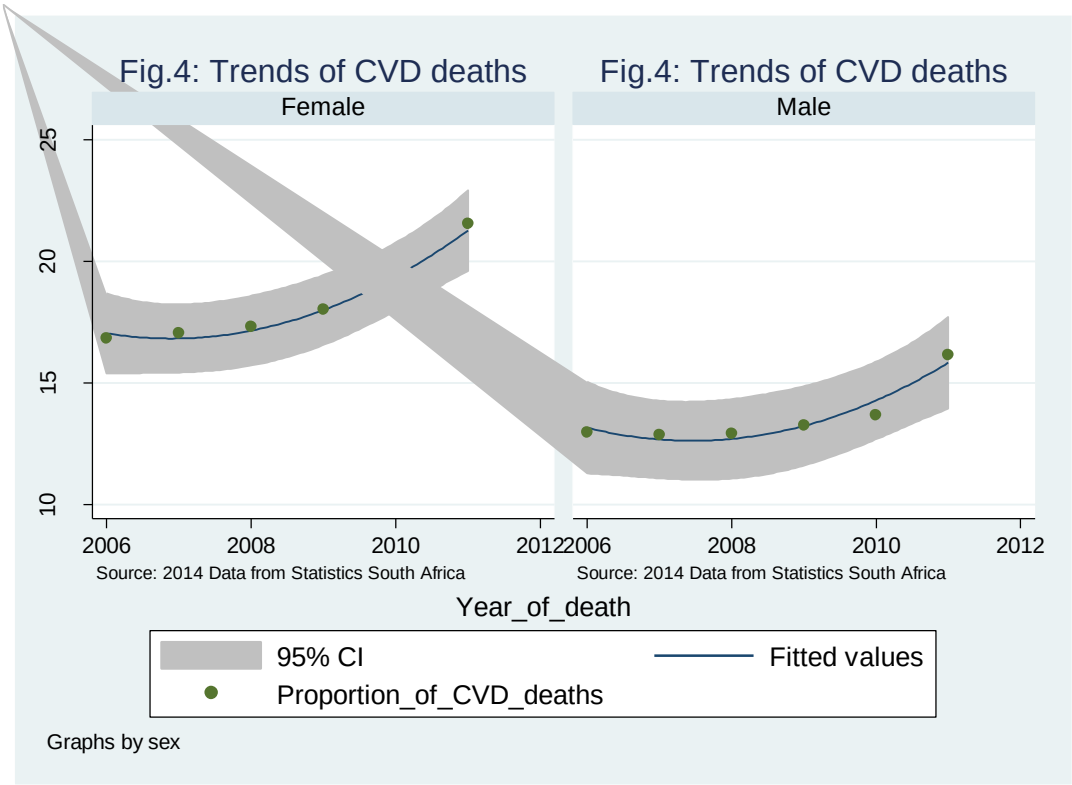


Figure 6(4): Trends of CVD deaths for Males and females

For the period under study, cardiovascular deaths were unevenly distributed with respect to race. The greatest proportions of cardiovascular disease related deaths occurred among Indians for the 6 years under study with an overall increase of 13.23% between 2006 and 2011. The smallest proportions occurred among black Africans. However, the relative increase for the whole period was 26.11% among black Africans, 13% among Indians, 11.89% among Coloured and 9.01% among Whites.

On average, one in four deaths (25.14%) in Western Cape was attributable to cardiovascular events, 18.78% in Gauteng and 14.71% in Mpumalanga. In relative terms, cardiovascular deaths dropped by 2.16% in Western Cape and by 3.44% in North West between 2006 and 2011 whereas it increased by 30.30% in Mpumalanga and 37.58% in Kwazulu Natal.

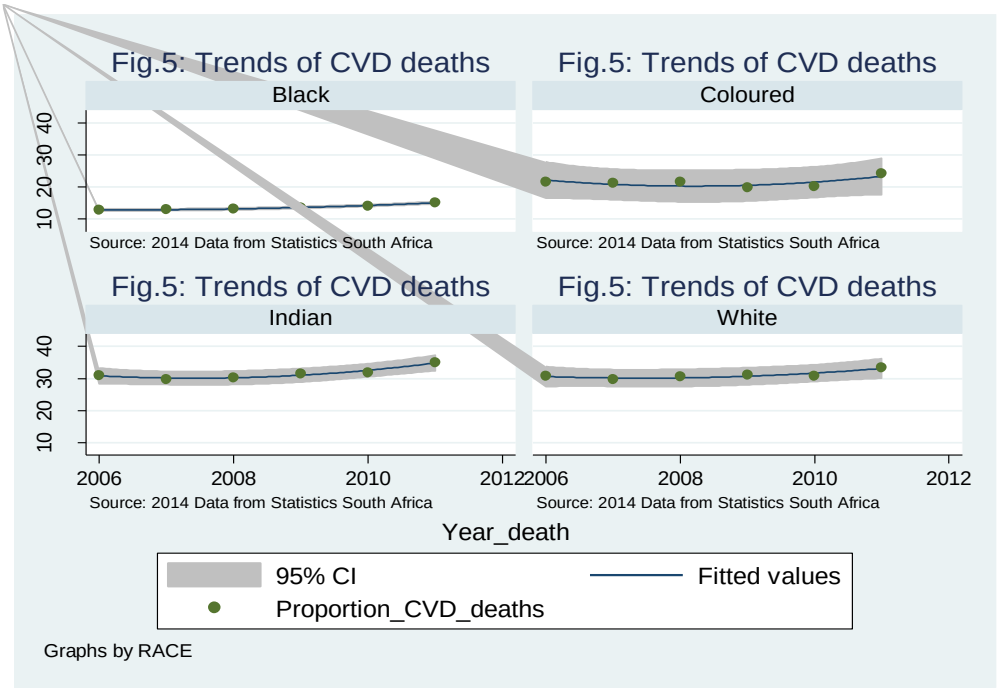


Figure 7(5): Trends of CVD deaths by Race

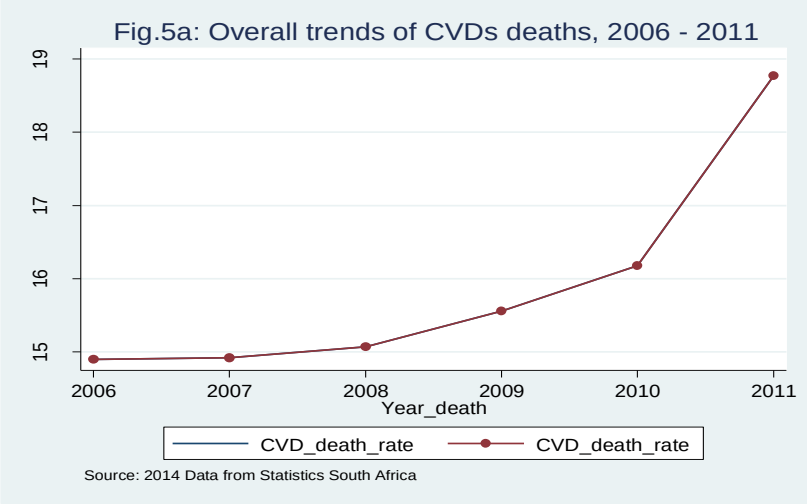


Figure 8(5a): Overall trends of CVDs deaths, 2006 - 2011

3.5. Distribution of cardiovascular diseases deaths in South Africa, by type (2006 – 2011) (Statistics- SA, 2014).

As per the ICD-10 classification, the table below shows the annual trends of cardiovascular diseases mortality over the 6 years under study. It is noticed that cerebrovascular diseases accounted for 31.28% of cardiovascular deaths with stroke (not specified as haemorrhage or infarction) representing 88% of cerebrovascular deaths. Other forms of heart diseases accounted for 29.83% of cardiovascular deaths with heart failure accounting for 64.14% of all deaths attributable to other forms of heart diseases. In addition hypertensive diseases

accounted for 17.18% of cardiovascular deaths with primary hypertension and hypertensive heart diseases accounting for 46.26% and 46.30% of hypertensive deaths respectively.

In relative terms, hypertensive diseases deaths increased by 24.74% whereas cerebrovascular diseases dropped by 1.24% between 2006 and 2011.

Table 4: Distribution of the main CVDs underlying causes, by sociodemographic variables (Statistics-SA, 2014)

Characteristics	I10 – I15 (n=16,008)	I20 – I25 (n=14,753)	I30 – I52 (n=27,788)	I60 – I69 (n=29,142)	Other CVDs (n=5,470)
Sex					
Female	10,038 (19.22)	6,303 (12.07)	15,630 (29.92)	17,461 (33.43)	2,803 (5.37)
Male	5,970 (14.59)	8,450 (20.65)	12,158 (29.71)	11,681 (28.54)	2,667 (6.52)
Age Category					
<24	54 (4.13)	49 (3.75)	684 (52.33)	318 (24.33)	202 (15.46)
25 – 34	262 (6.95)	274 (7.27)	1,862 (49.40)	959 (25.44)	412 (10.93)
35 – 44	872 (13.05)	855 (12.79)	2,572 (38.49)	1,783 (26.68)	601 (8.99)
45 – 54	2,000 (16.69)	2,138 (17.84)	3,437 (28.68)	3,609 (30.12)	800 (6.68)
55 – 64	3,146 (18.56)	3,137 (18.50)	4,566 (26.93)	5,062 (29.98)	1,022 (6.03)
65+	9,674 (18.44)	8,300 (15.82)	14,667 (27.96)	17,391 (33.15)	2,433 (4.64)
Education Level					
No Schooling	5,404 (20.17)	2,353 (8.78)	8,199 (30.60)	9,687 (36.15)	1,152 (4.30)
Primary	4,597 (19.92)	2,476 (10.73)	7,078 (30.67)	7,688 (33.31)	1.242 (5.38)
Secondary	5,673 (14.17)	8,798 (21.97)	11,729 (29.29)	11,024 (27.53)	2,817 (7.04)
Tertiary	334 (10.30)	1,124 (34.71)	782 (24.11)	743 (22.90)	259 (7.98)
Race					
Black	13,549 (19.51)	5,471 (7.88)	22,596 (32.55)	24,085 (34.69)	3,728 (5.37)
White	973 (7.38)	5,485 (41.59)	2,997 (22.73)	2,615 (19.83)	1,118 (8.48)
Indian	551 (11.25)	2,233 (45.60)	1,059 (21.63)	879 (17.95)	175 (3.57)
Coloured	935 (16.56)	1,564 (27.70)	1.136 (20.12)	1,563 (27.68)	449 (7.95)
Smoking Status					
Smokers	3,703 (14.95)	4,871 (19.66)	7,233 (29.20)	7,043 (28.43)	1,922 (7.76)

Table 5: Trends of cardiovascular diseases deaths from 2006 – 2011.

Type of cardiovascular disease	2006	2007	2008	2009	2010	2011	2006 – 2011
	(N=17,053)	(N=17,005)	(N=17,137)	(N=17,741)	(N=15,058)	(N=9,167))	(N=93,161)
Acute rheumatic fever (I00-I02)	9 (0.05)	5 (0.03)	5 (0.03)	3 (0.02)	1 (0.01)	3 (0.03)	26 (0.03)
Chronic rheumatic heart diseases (I05-I09)	81 (0.47)	71 (0.42)	72 (0.42)	67 (0.38)	54 (0.36)	32 (0.35)	377 (0.40)
Hypertensive diseases (I10-I15)	2,619 (15.36)	2,742 (16.12)	2.979 (17.38)	3,256 (18.35)	2,656 (17.64)	1,756 (19.16)	16,008 (17.18)
Ischaemic heart diseases (I20-I25)	2,885 (16.92)	2,720 (16.00)	2,607 (15.21)	2,727 (15.37)	2,325 (15.44)	1,489 (16.24)	14,753 (15.84)
Pulmonary heart disease and diseases of pulmonary circulation (I26-I28)	512 (3.00)	508 (2.99)	558 (3.26)	514 (2.90)	452 (3.00)	302 (3.29)	2,846 (3.05)
Other forms of heart disease (I30-I52)	5,159 (30.25)	5,103 (30.01)	5,206 (30.38)	5,324 (30.01)	4,461 (29.63)	2,535 (27.65)	27,788 (29.83)
Cerebrovascular diseases (I60-I69)	5,377 (31.53)	5,442 (32.00)	5,306 (30.96)	5,412 (30.51)	4,750 (31.54)	2,855 (31.14)	29,142 (31.28)
Diseases of arteries, arterioles and capillaries (I70-I79)	252 (1.48)	264 (1.55)	245 (1.43)	270 (1.52)	206 (1.37)	105 (1.15)	1,342 (1.44)
Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified (I80-I89)	137 (0.80)	120 (0.71)	122 (0.71)	132 (0.74)	128 (0.85)	76 (0.83)	715 (0.77)
Other and unspecified disorders of the circulatory system (I95-I99)	22 (0.13)	30 (0.18)	37 (0.22)	1,001 (0.17)	25 (0.17)	14 (0.15)	164 (0.18)

3.6. Logistic, Cox Proportional Hazard and Cox shared frailty Models

Table 5 below presents the results of the bivariate and multivariate models using logistic regression modelling, Cox proportional hazard and Cox shared frailty models. As highlighted under the analysis plan section, the second Cox regression model accounted for a random effect of cardiovascular diseases deaths at district level while the first Cox model used a semiparametric model.

Table 6: Bivariate and multivariate logistic, Cox survival and Cox shared frailty regression models, Stats-SA (Combined data from 2006 – 2011)

Characteristics	Unadjusted Logistic Model OR & 95% C.I	Adjusted Logistic Model OR & 95% C.I	Adjusted Cox PH Model HR & 95% C.I	Adjusted Cox Frailty Model HR & 95% C.I
Sex				
Female	1.00	1.00	1.00	1.00
Male	0.71 (0.70 – 0.72)	0.69 (0.68 – 0.71)	0.86 (0.84 – 0.87)	0.94 (0.92 – 0.95)
Age Group				
<25	1.00	1.00		
25 – 34	1.01 (0.94 – 1.07)	0.99 (0.93 – 1.06)		
35 – 44	1.87 (1.76 – 1.99)	1.84 (1.73 – 1.95)		
45 – 54	4.69 (4.42 – 4.97)	4.45 (4.19 – 4.72)		
55 – 64	9.11 (8.60 – 9.66)	8.20 (7.73 – 8.71)		
65+	16.18 (15.29 – 17.10)	13.21 (12.46 – 14.02)		
Education Level				
No schooling	1.00	1.00	1.00	1.00
Primary	0.57 (0.55 – 0.58)	0.93 (0.91 – 0.95)	1.50 (1.48 – 1.53)	1.49 (1.46 – 1.51)
Secondary	0.42 (0.41 – 0.43)	0.94 (0.92 – 0.96)	2.22 (2.18 – 2.26)	2.19 (2.15 – 2.31)
Tertiary	0.60 (0.58 – 0.63)	0.92 (0.89 – 0.97)	2.07 (1.99 – 2.16)	2.19 (2.11 – 2.28)
Marital Status				
Single	1.00	1.00	1.00	1.00
Married/in relationship	2.81 (2.77 – 2.86)	1.18 (1.16 – 1.20)	0.73 (0.72 – 0.74)	0.73 (0.72 – 0.74)
Separated	4.21 (4.12 – 4.29)	1.19 (1.17 – 1.23)	0.56 (0.55 – 0.57)	0.55 (0.54 – 0.56)
Race				
Black	1.00	1.00	1.00	1.00
White	2.86 (2.80 – 2.93)	1.00 (0.97 – 1.04)	0.45 (0.44 – 0.46)	0.56 (0.55 – 0.57)
Indian	2.91 (2.81 – 3.01)	1.18 (1.11 – 1.24)	1.00 (0.96 – 1.04)	1.08 (1.04 – 1.12)
Coloured	1.73 (1.67 – 1.78)	1.30 (1.23 – 1.36)	0.99 (0.95 – 1.03)	1.00 (0.97 – 1.04)
Employment Status				
Type of occupation	1.00	1.00		1.00
Exposure to chemical substances	1.03 (1.00 – 1.07)	1.01 (0.97 – 1.04)		
Sedentary work	0.94 (0.92 – 0.97)	0.97 (0.95 – 1.00)		
Smoking Status				
Non Smoker	1.00	1.00	1.00	1.00
Smoker	0.79 (0.78 – 0.80)	0.92 (0.91 – 0.95)	1.34 (1.32 – 1.36)	1.31 (1.29 – 1.33)
Next of kin smoking Status				
Non Smoker	1.00			
Smoker	0.98 (0.96 – 1.00)			
Province				
Northern Cape	1.00	1.00	1.00	1.00
Western Cape	2.01 (1.95 – 2.08)	1.09 (1.03 – 1.15)	0.89 (0.85 – 0.93)	0.95 (0.85 – 1.06)
Eastern Cape	1.08 (1.06 – 1.11)	1.01 (0.96 – 1.0.7)	**0.95 (0.91 – 0.99)**	**1.01 (0.91 – 1.11)**

Kwazulu Natal	1.20 (1.15 – 1.25)	1.14 (1.08 – 1.20)	1.22 (1.17 – 1.27)	1.19 (1.07 – 1.31)
Free State	1.04 (1.02 – 1.07)	1.22 (1.16 – 1.29)	1.25 (1.19 – 1.30)	1.17 (1.05 – 1.30)
North West	1.21 (1.17 – 1.25)	1.28 (1.21 – 1.35)	1.10 (1.05 – 1.15)	1.10 (0.99 – 1.22)
Gauteng	1.33 (1.30 – 1.37)	1.23 (1.16 – 1.29)	1.18 (1.13 – 1.23)	1.10 (0.99 – 1.22)
Mpumalanga	1.01 (0.99 – 1.04)	1.18 (1.12 – 1.24)	1.26 (1.20 – 1.31)	1.15 (1.04 – 1.28)
Limpopo	1.04 (1.01 – 1.07)	1.01 (0.96 – 1.07)	0.96 (0.92 – 1.01)	1.06 (0.95 – 1.18)
Interaction Sex & Race				
Female Black		1.00	1.00	
Female White		1.00	1.00	
Female Indian		1.00	1.00	
Female Coloured		1.00	1.00	
Black Male		1.00	1.00	
White Male		1.65 (1.56 – 1.72)	1.60 (1.54 – 1.67)	
Indian Male		1.40 (1.76 – 2.04)	1.25 (1.18 – 1.32)	
Coloured Male		1.04 (0.97 – 1.10)	1.11 (1.05 – 1.17)	

Note: Age-groups were not considered as covariate in the Cox models because they perfectly predict the survival time (age in complete year).

Under the logistic regression model, between 2006 and 2011, the baseline likelihood of dying from cardiovascular diseases in South Africa was 0.18 (0.18 – 1.9) which represents 82% reduced probability of dying from cardiovascular diseases while ignoring any external risk factor.

Results of the logistic bivariate regression analysis show that the odds of males dying from CVDs was lower by 29% (OR=0.71, 95% CI: 0.70-0.72) compared to their female counterparts. The odds of dying from CVD increased steadily with age below 45 years, thereafter the increase is sharp. For instance, people aged 65 years and older had about 16 times the probability of dying from cardiovascular diseases compared to those aged less than 25 years old. People living separated (widowed, divorced) and those in a relationship (married and in cohabitation) were more likely to die from cardiovascular diseases compared to those living alone (single). Indians and Whites were almost 3 times more likely to die from cardiovascular diseases (OR: 2.91, 95% CI: 2.81 – 3.01 and OR: 2.86, 95% C.I: 2.80 – 2.92) compared to Blacks. As a matter of fact, under the logistic model, smoking had been protective against CVDs deaths (OR: 0.92; 95% C.I: 0.91 – 0.95).

The final logistic regression model reveals that all predictors were statistically significant except for the smoking status of the next of kin. Most important, it reveals consistent geographic variation in the odds of dying from CVDs at provincial level. For instance, those who lived in North West province had 28% increase likelihood of dying from CVDs (OR: 1.28, 95% C.I: 1.21 – 1.35) compared to those living in Northern Cape.

In terms of occupational exposure, the logistic model revealed slight differences across the above occupational categories. For instance, those who were exposed to chemicals in the

environment with regular exposure to either carbon disulfide, nitrate esters, carbon monoxide, solvent, lead, cobalt or arsenic being more likely to die from CVD (OR: 0.1.01, C.I: 0.97 – 1.04) compared to those who were exposed to extreme temperature during their work. Deceased who had sedentary jobs were less likely to die from cardiovascular disease compared to those who were exposed to extreme temperature (OR: 0.97; 95% C.I: 0.95 – 1.00).

Results from the Cox proportional hazard model reveal that the mean survival time was 85 years for female and 84 years for males. The overall mean survival time was 85 years. Considering the time running from birth to death from CVDs, there were 4 and 3 CVDs related deaths per 1,000 deaths respectively among females and males. In terms of races, there were 5 CVDs deaths per 1,000 deaths among Indians and Whites, 4 among coloured and 3 among blacks. (Fig. 6) illustrates the survival functions of the deceased by sex and figure 7 portrays the survival functions by race.

From the survival model, we found a 6% constant reduced risk of dying from cardiovascular diseases among males (H.R:0.94, 95% C.I: 0.93-0.95) compared to females. While accounting for a random effect at district level, the probability of dying at a given time did not vary much with respect to gender. In the Cox proportional hazard model, the risk of dying increased with education level. For example, people with secondary and tertiary education had 8 and 15 increased risk of dying from cardiovascular diseases compared to those with no education (H.R: 7.89; 95% C.I: 7.36 – 8.46 and 14.77 95% C.I: 13.28 – 16.44 respectively).

In terms of race, Whites had about 55% (H.R: 0.45; 95% C.I: 0.44 – 0.46) reduced risk of dying from cardiovascular conditions compared to Blacks and Coloured had only 1% reduced risk compared to Blacks. While accounting for a random effect at district level, Indians had 8% increased risk of dying from CVDs compared to Blacks and Whites had 44% reduced risk compared to Blacks.

Smoking increased the risk of dying from cardiovascular diseases by 34% (1.34 95% C.I.: 1.32 – 1.36). When accounting for a random effect at district level, smoking increased the risk of dying from CVD by 31% (H.R: 1.31; 95% C.I: 1.29 – 1.33).

However, none of the covariate satisfied the proportional hazard assumption. And (Fig. 11) below suggests that our Cox PH survival model fairly fits the data.

In the final survival model, accounting for spatial dependency at district level, we found significant autocorrelation at district level (θ=0.02, p-value: 0.001).

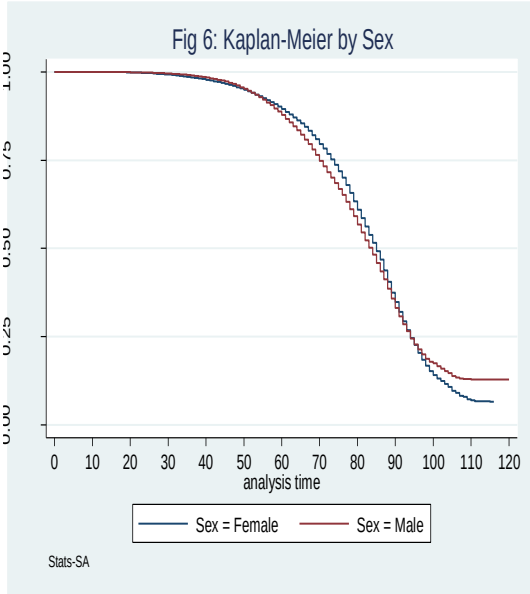


Figure 9(7): Kaplan Meier by Race

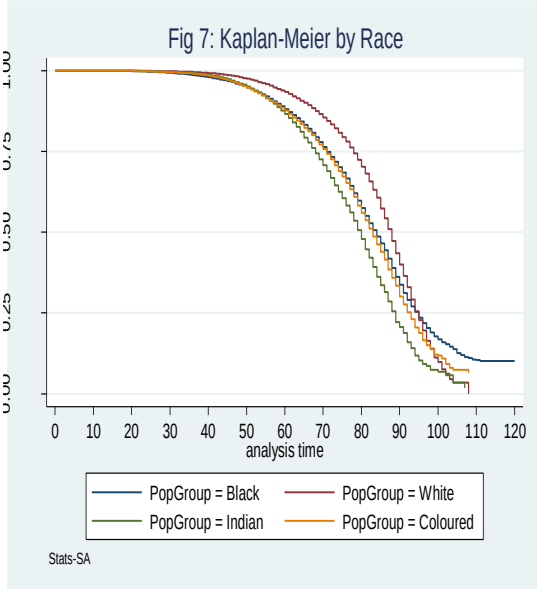


Figure 10(6) : Kaplan-Meier by Sex

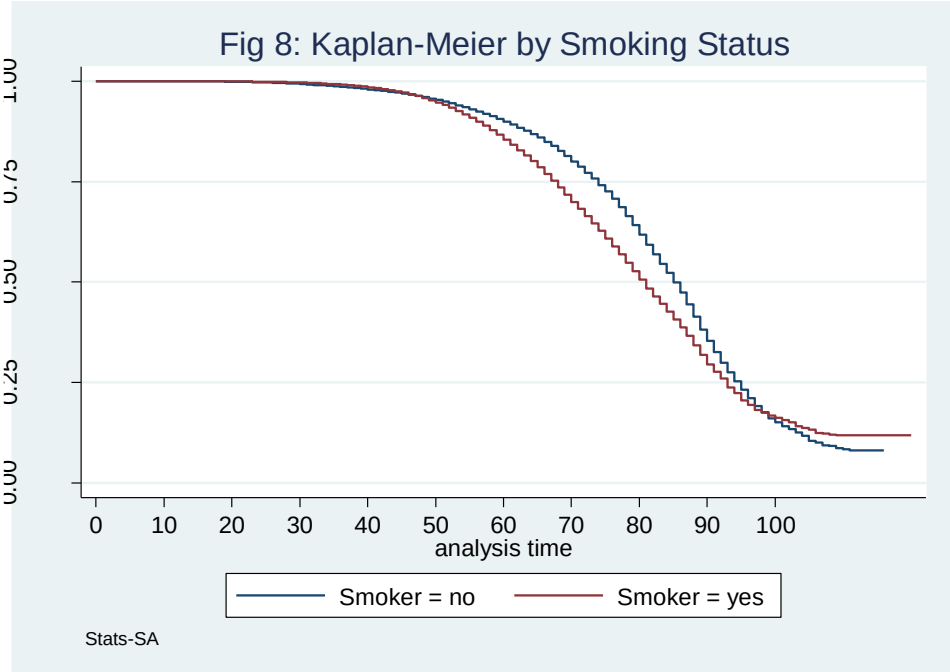


Figure 11(8): Kaplan Meier by Smoking Status

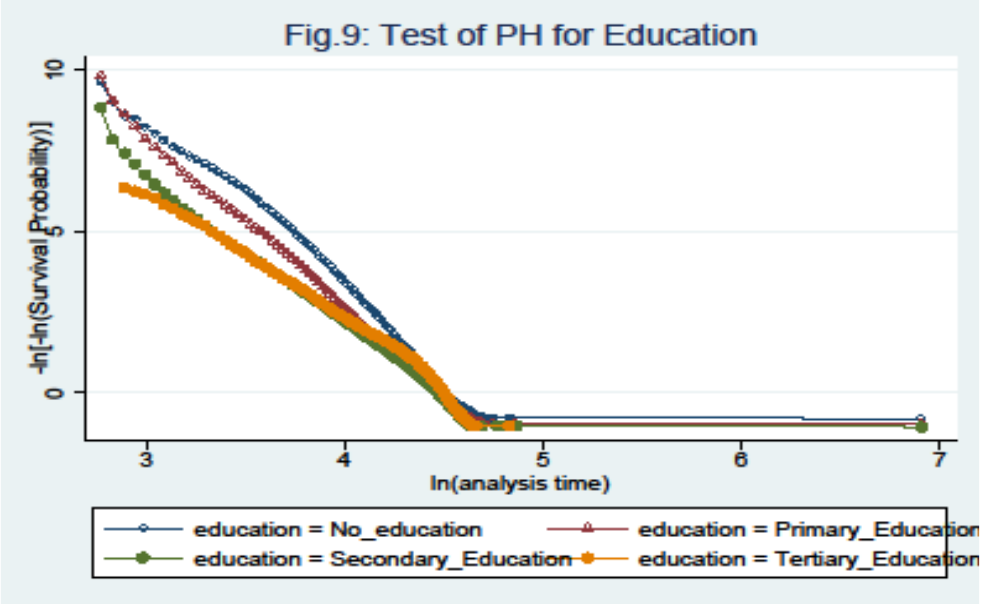


Figure 12(9): Test of PH for Education Level

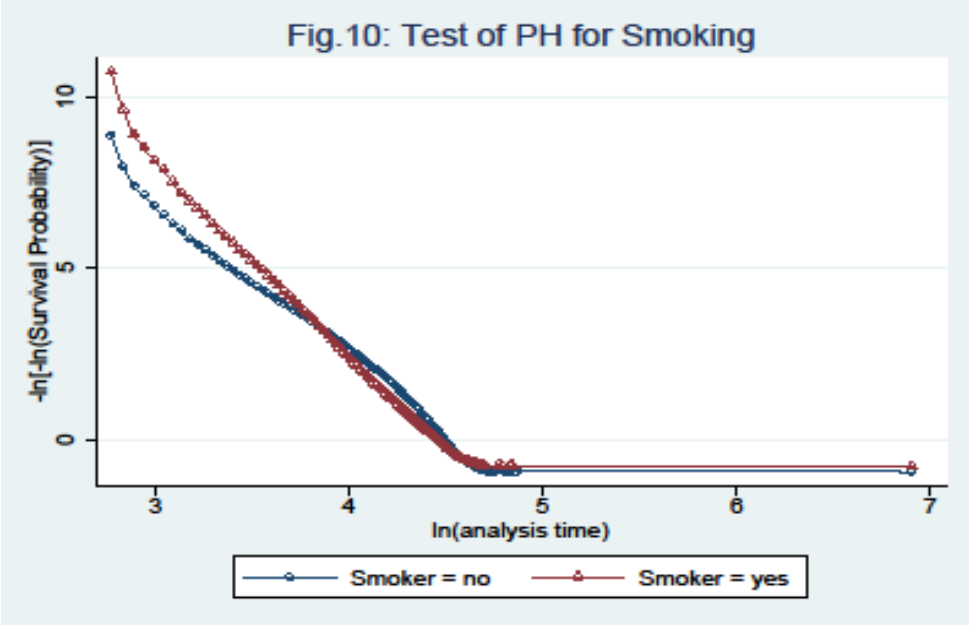


Figure 13(10): Test of PH for Smoking Status

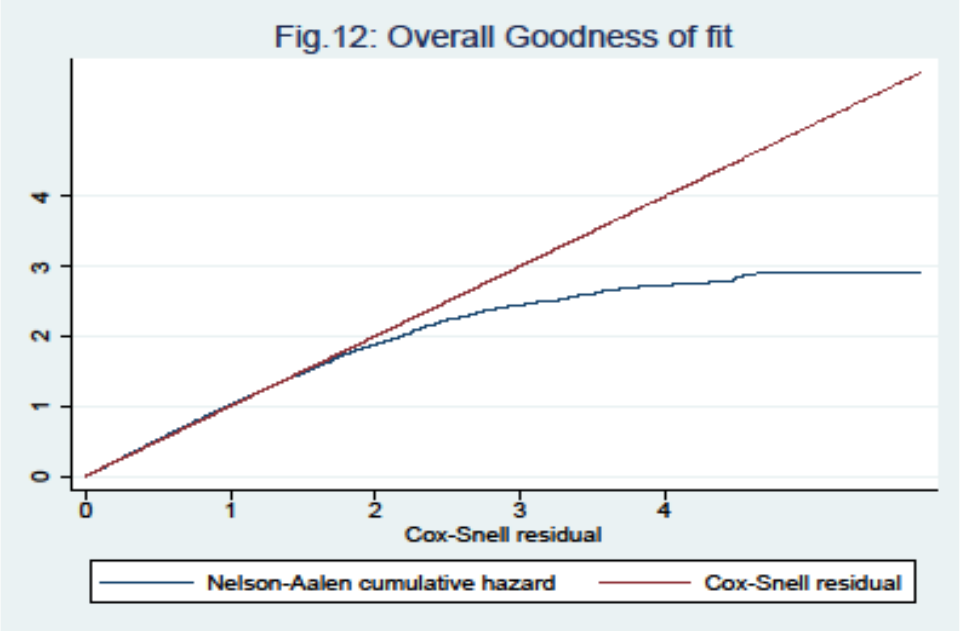


Figure 14(12): Overall Goodness of fit

Although we note a deviation from the norm around, it should be noticed that for practical purpose this is a fairly good model. In fact, from the literature it is known that for models with censored data, at large values of time the Nelson-Aelen cumulative hazard curve may depart from the line of the Cox Snell residuals (UCLA).

CHAPTER 4: DISCUSSION

This chapter discusses the main findings of this study as far as each of the three objectives is concerned. It should also be highlighted that the discussions merely focus on the result of the most robust model (Cox shared frailty model). Efforts were made to assess the consistency of these findings in line with the existing literature. Besides, efforts were made to emphasize on critical findings in order to inform possible public health interventions.

It should be recalled that the present study, mainly, sought to determine the socioeconomic and demographic risk factors for cardiovascular diseases deaths from 2006 to 2011.

4.1. Socioeconomic and demographic characteristics of the deceased

Cardiovascular diseases constitute a major public health concern around the world. Although its magnitude varies across the globe, evidence seems to be that CVDs are becoming a pandemic affecting all continents. Decades ago, CVDs used to be more prevalent among old people living in developed countries; research has shown significant spread of these conditions around the globe, suggesting increasing trends among poor developing countries, especially from Africa.

The main interest of the analysis was to investigate proportions of deaths attributable to CVDs. Due to lack of data on the population alive; the study could not estimate mortality adjusted measures such as: age adjusted death rates, age specific death rates. However, as discussed earlier, analysing proportions of deaths seemed relevant because not only can it inform possible health transition but it also has implications on resource allocations.

The CVDs death rate was estimated using mid-year population of 2011. Given that the South African mid-year population was estimated at 50,586,757 people, considering data on causes of deaths, the study found that about 164 per 100,000 deaths were attributable to CVDs, in 2011, 168 per 100,000 deaths in 2009 and 165 per 100,000 deaths in 2006.

In terms of the age of the deceased for the period under study, the study found that the overall mean age of the deceased for all causes of deaths was 50 years with a narrow 95% confidence interval of 49.97 – 50 years. It can be argued that in the general population in South Africa, the mean age of the deceased lies within this range, which indeed is quite close to the South Africans' life expectancy: 52.8 in 2011 and 56 in 2014 (UNDP, 2014, UNDP, 2011). Admittedly, this life expectancy is low compared to other upper middle income countries, mainly those from the BRICS group (China 75.3, India 66.4, Brazil 73.9 and Russia 68.0 in 2014), because of the interplay of HIV/AIDS and comorbidities (UNDP, 2014).

Of the total number of those who died between 2006 and 2011, the mean age of those who died from CVDs was significantly higher (65 years) than that of those of died from other causes (47 years), which confirms an age effect as far as CVDs are concerned (figure 3b).

As shown in the previous chapter, of the 598,013 deaths analysed in this study, 93,161 were attributable to cardiovascular diseases, which represent roughly 16% of the overall deaths. Whether this proportion is low compared to 29% worldwide (Anthony Mbewu & Jean-Claude Mbanya, 2006), it goes without saying that it falls within the range of 4% - 42% predicted by (Shanti Mendis, 2011). Besides, this proportion of people dying from CVDs (16%) in South Africa is higher compared to an overall estimate of 11.3% at Sub-Saharan African region as found by (Mensah et al., 2015).

Of the 93,161 deaths attributable to cardiovascular diseases, the proportion of females who died from CVDs represented 18% against 13% of males who died from CVDs. This confirms sex difference with respect to cardiovascular diseases deaths, which indeed is in concert with other findings in the Sub-Saharan African region (van der Meer et al., 2014, Ogah et al., 2015, Mensah et al., 2015).

Moreover, the mean age of females who died from CVDs was higher (67 years) than that of their male counterparts (63 years). Again, this finding is in line with studies by (Hu and Group, 2003, Kandala et al., 2013a, Juutilainen et al., 2004) which found that on average females were living longer than males.

One of the concerns raised in the study was that proportions of deaths attributable to CVDs were increasing among people aged less than 70 years, thereby increasing premature deaths. This study found that of the 93,161 who died from CVD, roughly 54.59% of deaths were premature deaths (before 70 years of age), which in fact negatively affects the economic perspective of the country by depleting the active labour force, disturbing family expectations. Such great proportion of premature deaths might be attributable to known risk factors such as tobacco, alcohol abuse and bad blood lipids. As discussed by (Norheim et al., 2015), a person who starts smoking earlier and keeps on smoking, loses about 10 years of healthy life. It can be argued that, the combination of smoking, bad blood lipids and alcohol abuse are possible reasons for high premature death from CVDs in South Africa for the period under study.

Concerning the distribution of all causes of deaths with respect to races, the distribution followed the demographic composition of the South African inhabitants. As released by (Statistics-SA, 2014), Black Africans who represent more than 80% of the South African population represented more than three quarters of the deceased (85.81%). The smallest proportions of deaths were registered among Indians (2.62%).

However, when looking at the distribution of deaths by cardiovascular diseases status, figures sharply contrast with proportions for all causes of deaths. Indians and Whites are the most affected by cardiovascular diseases, because these conditions accounted for about 31% of

deaths among Indians and Whites whereas they represented only 14% of deaths among black Africans. This finding highlights a critical issue of lifestyle hazards among Whites and Indians.

Despite these differences in the proportions of CVDs deaths with respect to race, the study revealed that for all causes of deaths, the mean age of the deceased differed significantly across the four races (Anova, p-value: 0.0001) with Whites having the highest mean (69 years, std: 18 years) against Blacks having the lowest mean (47 years, std: 25 years).

Although there was a significant difference in the age of the deceased with respect to their races, such differences were still significant when accounting for CVDs deaths exclusively. For instance, the mean age of Whites who died from CVDs was 74 years while Blacks died from CVDs on average around 64 years. Indians and Coloured who died from CVDs had a similar mean age around 65 years (Figure 3c).

A small proportion of the deceased had a tertiary education level (3.23%), and more than half of the deceased had a secondary education and about 18% of the deceased had no education.

Following the geographical distribution of the population in South Africa, it was expected that provinces with great urban agglomerations will have high proportions of deaths. Hence, 27% of the deceased lived in Kwazulu Natal, 14.13% in Eastern Cape and only 3% lived in Northern Cape.

However, this distribution of the deceased with respect to their provinces of usual residence, sharply contrasts with the distribution of CVDs deaths. For instance, Kwazulu Natal which hosted 27% of the overall deaths only accounted for 14.28% of the CVDs deaths whereas Western Cape Province which hosted only 3.54% of the overall deaths registered 25% of the CVDs deaths. As one might expect, this contrast could be explained by the predominance of White populations, which are the most affected, in the Western Cape Province.

4.2. Trends of CVD deaths over time

As shown in (Fig. 5a), CVDs deaths displayed an increasing trend over the period under study. This finding is closely in line with arguments raised by (Mensah et al., 2015) in a study on mortality from cardiovascular diseases in Sub-Saharan Africa, from 1990 – 2013.

Indeed, (Fig. 5a) shows that in 2006, 14.90% of all causes of deaths were attributable to CVDs. Then 5 years later in 2011, 18.77% of all deaths were attributable to CVDs. This suggests that, the average annual rate of increase over the time interval from 2006 to 2011 was: $\frac{(18.77\%-14.90\%)}{(2011-2006)} = \frac{3.87\%}{5} = 0.774\ \%\ per\ year.$

Furthermore, between 2006 and 2011, proportions of CVDs deaths rate has increased steadily among females with a relative increase of 28% from 2006 to 2011, compared to 24% among males. This suggests that for the period under study, South Africa has seen no significant

decline in CVDs mortality rates, which is consistent with other findings (Mensah et al., 2015, Norheim et al., 2015).

Instead, these rates are increasing at alarming trends whereas they are falling in most of the high and middle income countries, as noticed by (Mensah et al., 2015). In addition, estimations from the WHO reveal that South Africa has a higher rate of premature deaths from NCD (27.7%); compared his BRICS partners (Brazil: 19.8%, China: 19.5% and India: 26.1%) (WHO, 2014b).

On the other hand, a thorough analysis of the trends in terms of races revealed an alarming increase among Indians and Black Africans, as shown in (Fig. 5). Whether the highest annual average increase of the proportions of CVDs was among Indians (0.82%), it should be highlighted that in relative terms, from 2006 to 2011, the relative change in the proportions of Black Africans who died from CVDs was the most remarkable (26%), compared to 13% among Indians, 11.89% among Coloured and 9% among Whites.

It could then be argued that, at these rates for the next 12 years, in relative terms CVDs deaths will increase by $1 + \big((1 - 0.26) * (1 - 0.26)\big) = 1.5476 \approx 54.76\%$ among Blacks, which in fact is an alarming situation. As discussed by (Norheim et al., 2015), throughout the world the under 70 age-standardized mortality has declined by 19% with low and middle income countries having the highest absolute gains. These achievements are mostly due to the success in curtailing infectious and communicable diseases but also progress in the prevention and treatment of non-communicable diseases, which increase the lifespan of patients affected by NCDs. However, these progresses are not distributed equally around the globe. It is clear that these achievements are setback in countries, such South Africa, where the effect of HIV and social inequalities are predominated. As shown in our findings, social inequalities could be one possible reason for such increase in the burden of CVDs among Blacks. Hence further investment toward the prevention and treatment of these conditions are needed among this population groups to reverse the trends.

4.3. Types of Cardiovascular diseases

For the period under study, we took steps to investigate the main types of CVDs which claimed more lives. The study revealed that cerebrovascular diseases were the most common CVDs underlying causes (31.28% of CVDs deaths). Its distribution displayed significant differences with respect to the main socio-demographic variables. For instance, cerebrovascular diseases (I60 – I69, ICD-10 code) were more prevalent among females (33.43% of CVDs deaths) than among their male counterparts (29.71% of CVDs deaths). Of the 29,142 deaths attributable to cerebrovascular diseases, stroke accounted for 87.89% (25,614) of cerebrovascular deaths.

In addition, the study found that although cerebrovascular diseases were the leading type of CVDs underlying causes, this general pattern hide many disparities in terms of races. For

instance, among Blacks cerebrovascular diseases were the leading underlying cause of CVDs deaths (34.69% of CVDs deaths) whereas among Indians and Whites, Ischaemic Heart Diseases were the most prevalent (45.6% and 41.59% respectively). Cerebrovascular diseases accounted for only 17.95% of CVDs deaths among Indians and 19.83 among Whites, whereas ischaemic heart diseases accounted only for 7.88% of CVDs deaths among Blacks. Among Coloured deaths, the most common CVDs were cerebrovascular diseases (27.68%), Ischaemic heart diseases (27.70) and other forms of heart diseases (20.12%). These findings suggest that there might be different risk profiles associated with each race, as far as CVDs are concerned. Thus, more research is still needed in this regard. As discuss in a the global atlas on cardiovascular diseases by (Shanti Mendis, 2011) in the context of South Africa, we could argue that Blacks are more exposed to *behavioural risk factors* (physical inactivity, tobacco use, unhealthy diet, harmful use of alcohol), *metabolic risk factors* (raised blood pressure, raised sugar and dyslipidaemia) and *poverty and low educational status.* On the other hand, among Whites and Indians the spectrum of risk factors includes ageing, tobacco, psychological factors (stress, depression). However, the gap is still huge and need to document such differences is overwhelming.

Indians and Whites were more prone to Ischaemic heart diseases as underlined above. Of the 14,753 deaths attributable to Ischaemic heart diseases, more than three quarters (11,528 deaths: 78%) were attributable to Acute Myocardial Infarction, a condition affecting almost 65% of Whites above 65 years of age, 51% of Indians, but less than 50% of Blacks and Coloured of the same age group. This finding somehow confirms some of the concerns that were raised in the introduction with respect to premature deaths. As argued by (Norheim et al., 2015), these premature deaths can be preventable by controlling some key risk factors such as tobacco use, alcohol abuse and physical exercise.

Other forms of heart diseases (I30 – I52) including heart failure were among the leading underlying causes of CVDs deaths. Of the 27,788 deaths attributable to other forms of heart diseases, heart failure accounted for 64%.

Although findings from the three models yielded different results depending on the model considered, they all converge in terms of their effect on CVDs deaths. Regardless of the model, the study has shown that all predictors (sex, age, education, marital status, race, smoking status and place of residence) displayed significant differences across groups. It should be emphasized that, given the very large sample size that we analysed in this study, different tests used were powered enough, thereby preventing the study from type II error.

4.4. Discussing the robust model (Survival models)

Under this section, we are discussing only the two survival models because of the nature of the data (time to event data, death from CVDs). Although the main interest was to look at time from birth to death, we paid a scant attention to look at logistic regression model to see whether

despite different outcomes (death due to CVDs for logistic and time to death for survival model), the factors associated are similar or not in the two models. This approach has allowed the study to highlight some limitations while relying on the logistic regression model. Apart from the fact that the final logistic regression model displayed lack of fit, some of the results did not line up with known evidence from the literature. For instance, tobacco use which is a known risk factor for CVDs was protective and the likelihood of dying from CVDs decreased with education level. Owing to these reasons, we mainly discussed the results from the survival models.

The model derived from the Cox proportional hazard model and Cox survival models with shared frailties gave more meaningful results which are consistent with many similar studies.

Firstly, descriptive survival results show that there were 4 CVDs related deaths per 1,000 deaths among females against 3 CVDs deaths per 1,000 deaths among males, for the period under study. In terms of races, there were 5 CVDs deaths per 1,000 deaths among Indians and Whites and 3 CVDs deaths per 1,000 deaths among Blacks. Worthy of note, it was found that the survival curves of the deceased in terms of gender and races were all converging around 85 years of age, thereby suggesting a violation of the proportional hazard assumption for these covariates. Besides, the survival curves of the smoking status of the deceased converged around 100 years of age. An attempt was made to get all the covariates (mainly smoking status, education and Marital Status) interacting with time.

As highlighted under the limitation section, the quality of the survival model was affected by the nature of the time variable considered. Ideally, the best time period would cover the time interval from diagnosis of the underlying cause to death. However, such a variable was lacking from the dataset and could not be estimated anyhow. We then used the time interval from birth to death as a surrogate for the survival time. Despite these limitations, results from the Cox showed that the two models were consistent with many other findings in Africa (Mayosi et al., 2009, Mensah, 2008, Mensah et al., 2015).

For instance, the Cox survival model revealed that males had 14% reduced risk for dying from CVDs compared to their female counterparts. Figure 6 portrays the survival functions of the two sexes displaying two trends: from 25 – 50 years, CVDs deaths are occurring faster among females compared to their male counterparts. Thereafter, males are dying faster from CVDs compared to females until their survival curves become similar again around 85 years of age. One possible reason could be maternal risks associated with the pregnancies.

In terms of education, those with secondary and tertiary education had about 2 fold increased risk for dying from CVDs compared to those with no education, which is in line with other findings showing that CVDs are lifestyle diseases. One possible reason of such increased risk among educated people is that those with tertiary education were more likely to have better jobs, therefore earning more and being at high risk of dying from CVDs due to lifestyle issues.

Married deceased and those who were in a formal relationship had 27% reduced risk for dying from CVDs compared to single.

In terms of races, Whites were dying much slower as found in the descriptive analysis earlier and as portrayed in figure 7. Indians had the lowest survival curve suggesting that they were dying faster from CVDs compared to other races. Compared to Blacks, Indians had 8% increased risk for dying from CVDs. Surprisingly, Whites who were the most affected had almost 44% reduced risk for dying from CVDs compared to Blacks. One possible reason could be inequalities in accessing medical treatment and getting support, to the extent that those accessing better healthcare have great potential for prolonging their lifespan. As one might expect, such differences suggest persisting inequalities in the management of patients affected by these conditions.

As expected, former smokers had 31% increased risk for dying from CVDs compared to non-smokers. This implies that smoking is still the major risk factor, which is in concert with worldwide research. As suggested by (Jha and Peto, 2014), those who smoke earlier in their life and keep on smoking, are more likely to lose 10 years of healthy life, thereby shortening their lifespan.

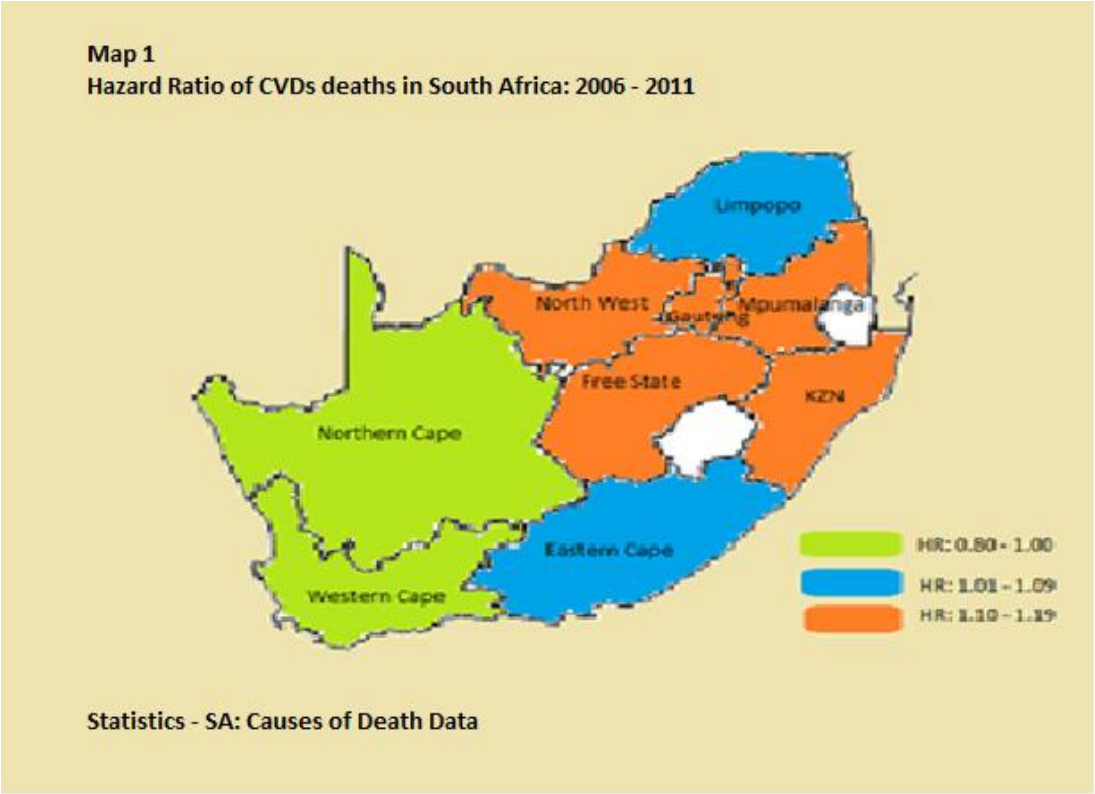
The final combined model, which investigated a district random effect using Cox shared frailty model, displayed significant heterogeneous effect at district level ($\hat{\theta} = 0.023$ and p-value: 0.001), implying that correlation within districts cannot be ignored. This suggests that CVDs frailties (risks) are commonly shared among individuals living in the same district. To portray the geographical variation of these CVDs we mapped the hazard ratio vs the proportions of CVDs deaths. Although a spatial model using sophisticated techniques (such Kriging, Geospatial modelling) would be a gold standard for such exercise, the aim of this study was to show evidence of geographical variations in CVDs mortality at provincial level.

4.5. Geographical variations of proportions and hazard rates of CVDs deaths of CVDs deaths for the combined period (2006 – 2011) at provincial level

The third objective of the study aimed to investigate geographic variation in cardiovascular deaths, while accounting for other risk factors including a random effect at district level. We have used an approach, which consisted in mapping the hazard ratios of the robust model (Cox shared frailty model) to illustrate geographical variations of CVDs deaths across the nine (9) provinces of South Africa for the combined period 2006 – 2011. Results for each year are provided in the appendix.

The purpose of mapping these hazard ratios is to point out the existence of latent geographical variations in the hazard of CVDs death across the nine provinces in South Africa. Besides, the maps also allowed us to highlight the limitations of a solely analysis based on proportions of deaths which does not reflect the nature of geographical exposures. For instance, the maps clearly show that in some provinces with high proportions for CVDs, the hazard of dying from

CVDs is very low. This implies that any public health policy aiming to address these geographical variations should not rely on the proportions of deaths, but rather on the adjusted hazard ratios.



Map 1: Hazard Ratio of CVDs deaths in South Africa: 2006 - 2011

Evidence from the above map clearly shows a cluster of CVDs deaths for the combined period (2006 – 2011) around the central provinces of South Africa with people living in North West, Free State, Gauteng and Kwa Zulu Natal having the highest risk for dying from CVDs. On the other hand, the Southern provinces including Northern Cape and Western Cape seemed to have the low risk during the same period.

For example, a study by (Kandala et al., 2013b) found that hypertension was more prevalent in the northern provinces (North West, Free State and Northern Cape). Our study also found that the hazard of CVDs were higher in the northern provinces of North West, Free State, Gauteng, Kwazulu Natal, and Mpumalanga. However, the hazard of CVDs mortality was lower in the southern provinces of Northern Cape and Western Cape (HR: 0.80 – 1.00). As discussed by (Kandala et al., 2013b), most of the provinces with high hazard of CVDs mortality seem to be regions with high concentration of poor quintiles classes. This argument was further supported by a highly significant association (p-value: 0.001) between races and provinces. For instance, 81% of Indians CVDs deaths and 26% of Black CVDs occurred in Kwazulu Natal; whereas only 4.15% of Indians CVDs deaths and 0.45% Blacks CVDs deaths occurred in Western Cape.

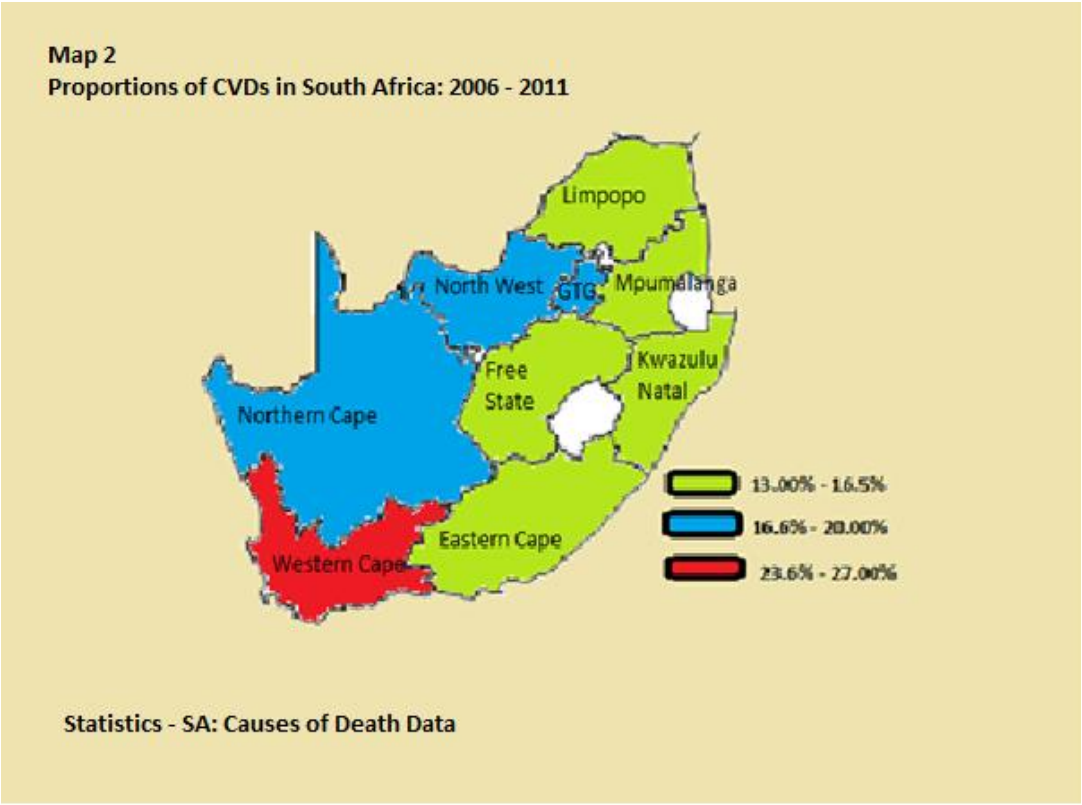
In addition to that, other possible reasons could include lifestyle parameters, lack of access to health facilities and the consequences thereof.

It goes without saying that, a cross sectional analysis of each of the six years confirms this general trend. For instance, for the six years under study, the three provinces of Gauteng, Free State and Kwazulu Natal were having the highest risk profile whereas the southern provinces of Western and Northern Cape had the lowest hazard ratio.

However, these geographical variations have fluctuated over the 6 years under study, with a clear increased trend as shown on the maps in the Appendix (Maps 3-14).

For instance, in 2006, high risk provinces of Free State, North West, Gauteng and Mpumalanga had hazards fluctuating between 1.10 – 1.19, whereas two years later (2008), the hazard ratio of people who lived in Free State province increased significantly and ranged between 1.30 – 1.47, whereas the Southern provinces remained at low risk.

However, the above map sharply contrasts with the maps of proportions of deaths due to CVDs. For instance for the overall period under study the map below, portraying proportions of CVDs by province, provides a different pattern. It is seen Western Cape, which had the highest proportions of CVDs followed by Northern Cape and North West provinces are having the lowest hazard ratio. This clearly shows that analysing CVDs based on the proportions of deaths, is misleading and could result in formulating wrong recommendations.



Map 2: Proportions of CVDs in South Africa: 2006 - 2011

In addition, synoptic maps in the appendix support these variations with a clear increased trend for all the provinces over the six years under study. For instance in 2011, geographic variations of the hazards of CVDs mortality reveal a constant level of the hazards in the Northern Cape Province over the six years. The rest of the 8 provinces experienced a constant increase in the hazards of dying over the six years.

Sharp increases were observed in 2011, compared to the other years in both proportions and hazards of CVDs deaths. Although these results of 2011 may be due to some bias, there is no evidence that the exact situation is different from what has been found. Thus, further analysis should be conducted to document these changes.

CONCLUSION AND RECOMMENDATIONS

The present study aimed to investigate socio-demographic determinants of CVDs deaths in South Africa, from 2006 – 2011, using data from the vital register system gathered by Statistics South Africa. The study postulated significant differences in CVDs mortality across sex, age, marital status, education, smoking status and geographical place of residence of the deceased.

After analysis, results from the study confirm these differences suggesting that all these predictors were statistically significant risk factors for CVDs mortality in South Africa.

Therefore, interventions must be implemented to control the known risk factors and reduce these avoidable massive premature deaths. Such measures should be in line with the WHO goal of reducing mortality before age 70 by 40%. In the context of this study, such a target is achievable since the main conditions affecting people are known and the socio demographic differentials are now well documented. Thus, a subset of the priorities should include continuation of national and international programs inspired by the Millennium Development Goals (1, 2, 3 and 6). In addition, a passive surveillance system for CVDs is needed to insure close monitoring of these conditions. This will provide reliable data for analysis and dissemination, which in turn will inform public health policies. Limitations raised in this study clearly highlight the relevance of such need.

In addition, significant efforts should target preventive measures such as anti-tobacco legislation (plain packaging policy which has shown effective in Australia (Mitchell, 2010), regulation on alcohol consumption (effective in Russia (Zaridze et al., 2009, Leon et al., 2007, Neufeld and Rehm, 2013)), campaign against bad blood lipids and high blood pressure and encouraging people to do physical exercises. Also secondary prevention measures should target patients with early history of vascular diseases, by providing them with long term treatment of statin and aspirin, which have proven to be effective in decreasing the likelihood of vascular events.

Finally, efforts must be done to reduce socioeconomic inequalities to ensure an equal access to healthcare across the four population groups. This will definitely curb the increasing trend of CVDs deaths. Special care must be given to regions with high risk profile (central provinces including Free State, North West, Gauteng and Kwazulu Natal). Significant efforts must be done to implement individual and community-based preventive measures in these regions.

To Statistics South Africa, we recommend an adjustment of the BI-1663 and DHA-1663 to include duration of disease for those who have been diagnosed for a long time. In addition, it is very important to include information on alcohol intake in the forms. To minimize incomplete filling of forms, we suggest that Statistics South Africa should engage more with practitioners filling the forms to improve their acceptability of the reporting and also explain to them the importance of filling correctly the form.

REFERENCE LIST

ANTHONY MBEWU & JEAN-CLAUDE MBANYA 2006. Cardiovascular disease. *In:* JAMISON DT, F. R., MAKGOBA MW, ET. AL (ed.) *Disease and mortality in Sub-Saharan Africa. 2nd edition.* Washington DC: World Bank.

BANERJEE, S., WALL, M. M. & CARLIN, B. P. 2003. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. *Biostatistics,* 4**,** 123-142.

BLOT, W. J. & TARONE, R. E. 2015. Doll and Peto's Quantitative Estimates of Cancer Risks: Holding Generally True for 35 Years. *Journal of the National Cancer Institute,* 107**,** djv044.

BUREAU, P. R. 2013. *2013 World Population Data Sheet* [Online]. World Population Reference Bureau. Available: http://www.prb.org/pdf13/2013-population-data-sheet_eng.pdf [Accessed 23 March 2014.

CORSI, D. J., SUBRAMANIAN, M. A., SMITH, G. D. & SUBRAMANIAN, S. V. 2014. Authors' response to Gupta and Pednekar: Importance of examining cause-specific proportions of deaths as well as mortality rates. *International Journal of Epidemiology,* 43**,** 278-280.

DARMOFAL, D. 2009. Bayesian spatial survival models for political event processes. *American Journal of Political Science,* 53**,** 241-257.

EVANS, A., TOLONEN, H., HENSE, H. W., FERRARIO, M., SANS, S., KUULASMAA, K. & PROJECT, W. M. 2001. Trends in coronary risk factors in the WHO MONICA Project. *International Journal of Epidemiology,* 30**,** S35-S40.

GOVERNMENT, S. 2014. *About SA - South Africa's people* [Online]. Available: http://www.gov.za/aboutsa/people.htm [Accessed 09/05/2014 2014].

HELAKORPI, S., MARTELIN, T., TORPPA, J., VARTIAINEN, E., UUTELA, A. & PATJA, K. 2008. Impact of the 1976 Tobacco Control Act in Finland on the proportion of ever daily smokers by socioeconomic status. *Preventive Medicine,* 46**,** 340-345.

HENDERSON, R., SHIMAKURA, S. & GORST, D. 2002. Modeling spatial variation in leukemia survival data. *Journal of the American Statistical Association,* 97.

HU, G. & GROUP, D. S. 2003. Gender difference in all-cause and cardiovascular mortality related to hyperglycaemia and newly-diagnosed diabetes. *Diabetologia,* 46**,** 608-617.

HURST, J. W., MORRIS, D. C. & ALEXANDER, R. W. 1999. The use of the New York Heart Association's classification of cardiovascular disease as part of the patient's complete problem list. *Clinical Cardiology,* 22**,** 385-390.

JHA, P. & PETO, R. 2014. Global effects of smoking, of quitting, and of taxing tobacco. *New England Journal of Medicine,* 370**,** 60-68.

JUUTILAINEN, A., KORTELAINEN, S., LEHTO, S., RÖNNEMAA, T., PYÖRÄLÄ, K. & LAAKSO, M. 2004. Gender difference in the impact of type 2 diabetes on coronary heart disease risk. *Diabetes care,* 27**,** 2898-2904.

KADIRI, S. 2005. Tackling cardiovascular disease in Africa - Will need much more than just imported measures from more developed countries. *British Medical Journal,* 331**,** 711-712.

KANDALA, N.-B., MANDA, S. O., TIGBE, W. W., MWAMBI, H. & STRANGES, S. 2013a. Geographic distribution of cardiovascular comorbidities in South Africa: a national cross-sectional analysis. *Journal of Applied Statistics***,** 1-14.

KANDALA, N.-B., TIGBE, W., MANDA, S. O. & STRANGES, S. 2013b. Geographic variation of hypertension in Sub-Saharan Africa: a case study of South Africa. *American Journal of Hypertension,* 26**,** 382-391.

KANDALA&AL., N. B. 2013. Geographic distribution of cardiovascular comorbidities in South Africa: a national cross-sectional analysis. *Journal of Applied Statistics,* 41**,** 1203-1216.

KLAG, M. J. 2013. Epidemiology of cardiovascular disease. *???*

LEON, D. A., SABUROVA, L., TOMKINS, S., ANDREEV, E., KIRYANOV, N., MCKEE, M. & SHKOLNIKOV, V. M. 2007. Hazardous alcohol drinking and premature mortality in Russia: a population based case-control study. *Lancet,* 369**,** 2001-2009.

MAREDZA, M., HOFMAN, K. J. & TOLLMAN, T. 2011. A hidden menace: cardiovascular disease in South Africa and the costs of an inadequate policy response: health policy and cardiovascular disease. *SA Heart,* 8**,** 48-57.

MAYOSI, B. M., FLISHER, A. J., LALLOO, U. G., SITAS, F., TOLLMAN, S. M. & BRADSHAW, D. 2009. The burden of non-communicable diseases in South Africa. *The Lancet,* 374**,** 934-947.

MENSAH, G. A. 2008. Epidemiology of stroke and high blood pressure in Africa. *Heart,* 94**,** 697-705.

MENSAH, G. A., ROTH, G. A., SAMPSON, U. K., MORAN, A. E., FEIGIN, V. L., FOROUZANFAR, M. H., NAGHAVI, M. & MURRAY, C. J. 2015. Mortality from cardiovascular diseases in sub-Saharan Africa, 1990-2013: a systematic analysis of data from the Global Burden of Disease Study 2013: cardiovascular topic. *Cardiovascular Journal of Africa: H3Africa Supplement: Supplement 1,* 26**,** S6-S10.

MITCHELL, A. D. 2010. AUSTRALIA'S MOVE TO THE PLAIN PACKAGING OF CIGARETTES AND ITS WTO COMPATIBILITY. *Asian Journal of Wto & International Health Law and Policy,* 5**,** 405-425.

NEUFELD, M. & REHM, J. 2013. Alcohol Consumption and Mortality in Russia since 2000: Are there any Changes Following the Alcohol Policy Changes Starting in 2006? *Alcohol and Alcoholism,* 48**,** 222-230.

NORHEIM, O. F., JHA, P., ADMASU, K., GODAL, T., HUM, R. J., KRUK, M. E., GÓMEZ-DANTÉS, O., MATHERS, C. D., PAN, H. & SEPÚLVEDA, J. 2015. Avoiding 40% of the premature deaths in each country, 2010–30: review of national mortality trends to help quantify the UN Sustainable Development Goal for health. *The Lancet,* 385**,** 239-252.

OGAH, O. S., DAVISON, B. A., SLIWA, K., MAYOSI, B. M., DAMASCENO, A., SANI, M. U., MONDO, C., DZUDIE, A., OJJI, D. B. & KOUAM, C. 2015. Gender differences in clinical characteristics and outcome of acute heart failure in sub-Saharan Africa: results of the THESUS-HF study. *Clinical Research in Cardiology***,** 1-10.

SHANTI MENDIS, B. N. 2011. Global Atlas on Cardiovascular disease prevention and control. Geneva: WHO.

STATISTICS-SA 2011. Mortality and causes of death in South Africa, 2011. *In:* AFFAIRS, D. O. H. (ed.). Pretoria: Stats-SA.

STATISTICS-SA. 2012. *Mortality and causes of death in South Africa* [Online]. Available: http://beta2.statssa.gov.za/publications/P03093/P030932011.pdf.

STATISTICS-SA, S. A. 2014. Mid-year population estimates 2014. *Statistical release P0302.* Pretoria: Statistics-SA.

STEYN, K. 2007. *Heart Disease in South Africa* [Online]. Cape Town: MRC. Available: http://www.mrc.ac.za/chronic/heartandstroke.pdf [Accessed 24 April 2014 2014].

TODARO, M. P. 2012. *Economic Development,* Boston, Addison Wesley Longman.

UCLA. *Background for Survival Analysis* [Online]. Available: http://www.ats.ucla.edu/stat/stata/seminars/stata_survival/default.htm#fit [Accessed 02 March 2015].

UNDP 2011. Sustainability and Equity. *Human Development Report.* New York: United Nations Development Programme.

UNDP 2014. Sustaining Human Progress: Reducing vulnerabilities and building resilience. *Human Development Report 2014.* New York: United Nations Development Programme.

VAN DER MEER, M. G., CRAMER, M. J., VAN DER GRAAF, Y., DOEVENDANS, P. A. & NATHOE, H. M. 2014. Gender difference in long-term prognosis among patients with cardiovascular disease. *European journal of preventive cardiology,* 21**,** 81-89.

WHO 2009. Table 2. Estimated total DALYs ('000), by cause and WHO Member State, 2004 (a, m).

WHO. 2011. *Cadre global mondial de suivi et cibles mondiales volontaires pour la prévention et la lutte contre les maladies non transmissibles* [Online]. Available: http://www.who.int/nmh/events/2011/consultation_dec_2011/2012_Jan_DiscussionPaper_FR.pdf [Accessed 09 May 2014 2014].

WHO. 2013. *Global Health Estimates Summary (GHE): Deaths by age, sex and causes* [Online]. Available: http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/ [Accessed 09/05/2014.

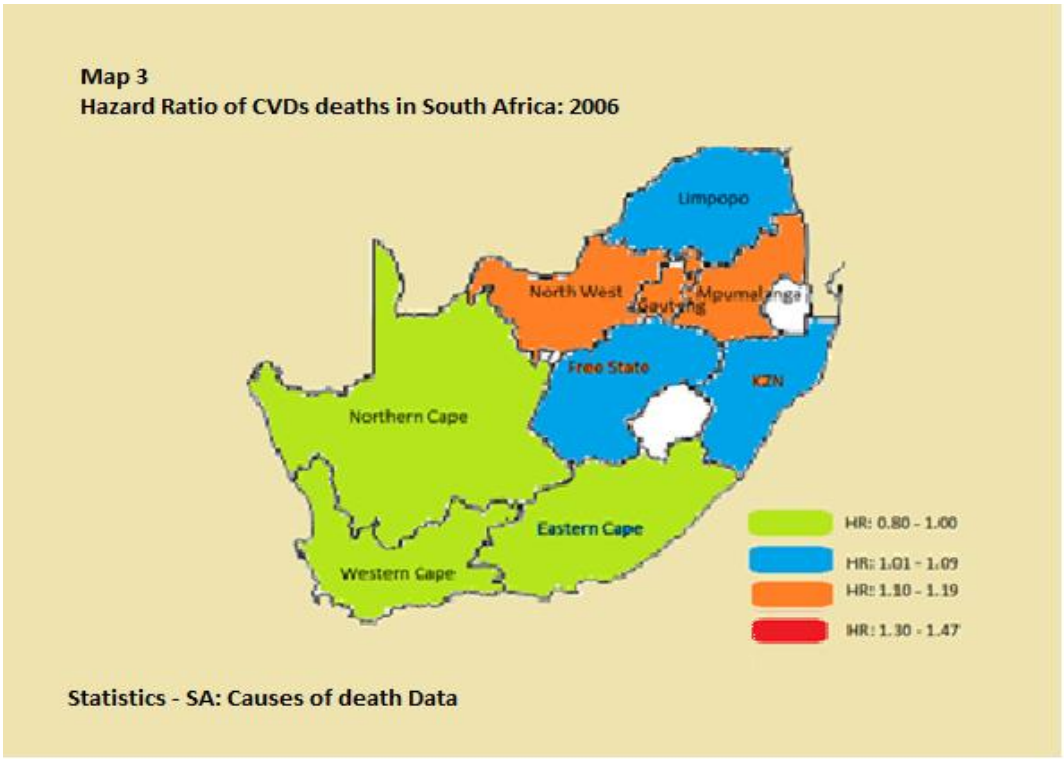
WHO. 2014a. *Cardiovascular disease: Strategic priorities* [Online]. Available: http://www.who.int/cardiovascular_diseases/priorities/en/ [Accessed 14-04-2014 2014].

WHO 2014b. Global Status Report on Noncommunicable diseases 2014. *Global Status Report.* Geneva: WHO.

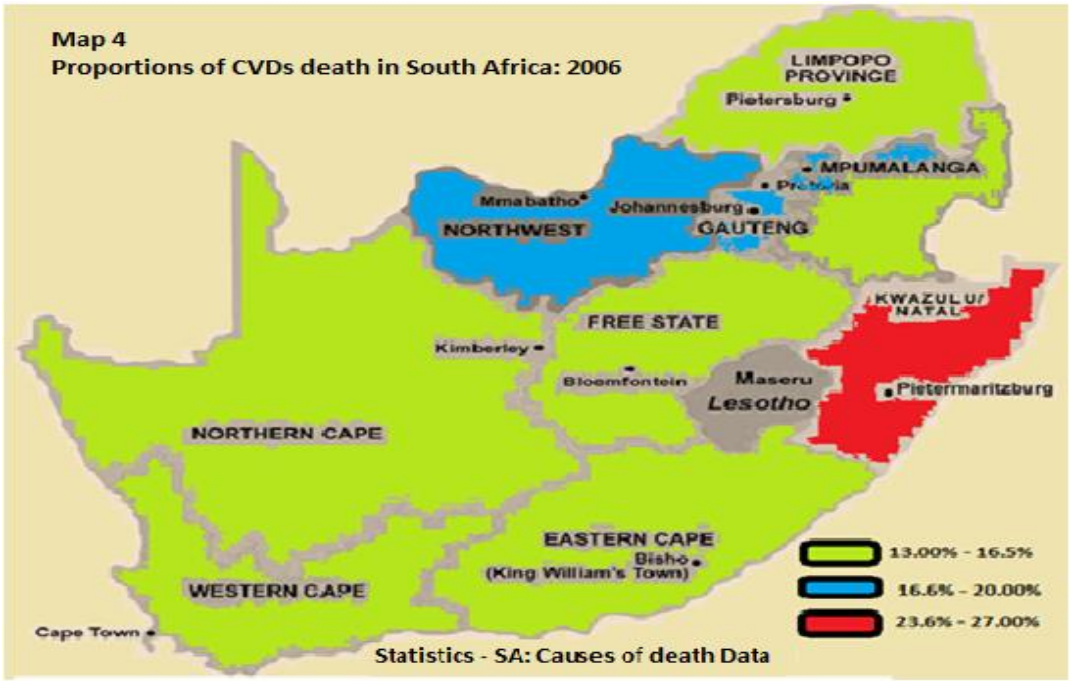
ZARIDZE, D., BRENNAN, P., BOREHAM, J., BORODA, A., KARPOV, R., LAZAREV, A., KONOBEEVSKAYA, I., IGITOV, V., TERECHOVA, T., BOFFETTA, P. & PETO, R. 2009. Alcohol and cause-specific mortality in Russia: a retrospective case-control study of 48 557 adult deaths. *Lancet,* 373**,** 2201-2214.

APPENDIX

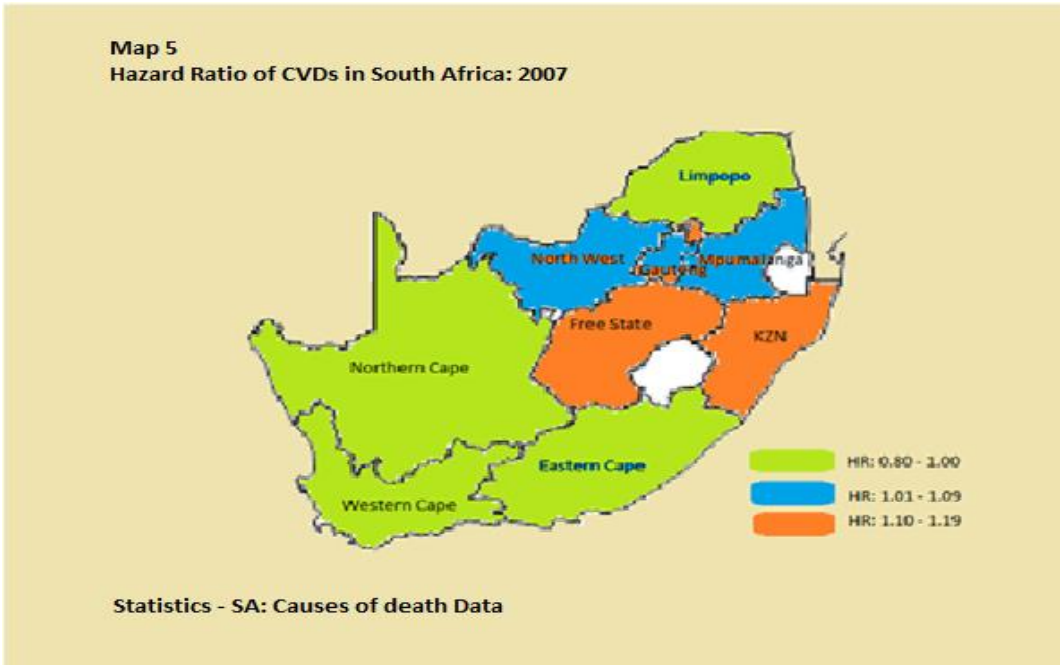
Maps of hazards ratio derived from the Cox frailty model and proportions of deaths attributable to CVDs, by province (Statistics – SA: Causes of death data 2006 – 2011)



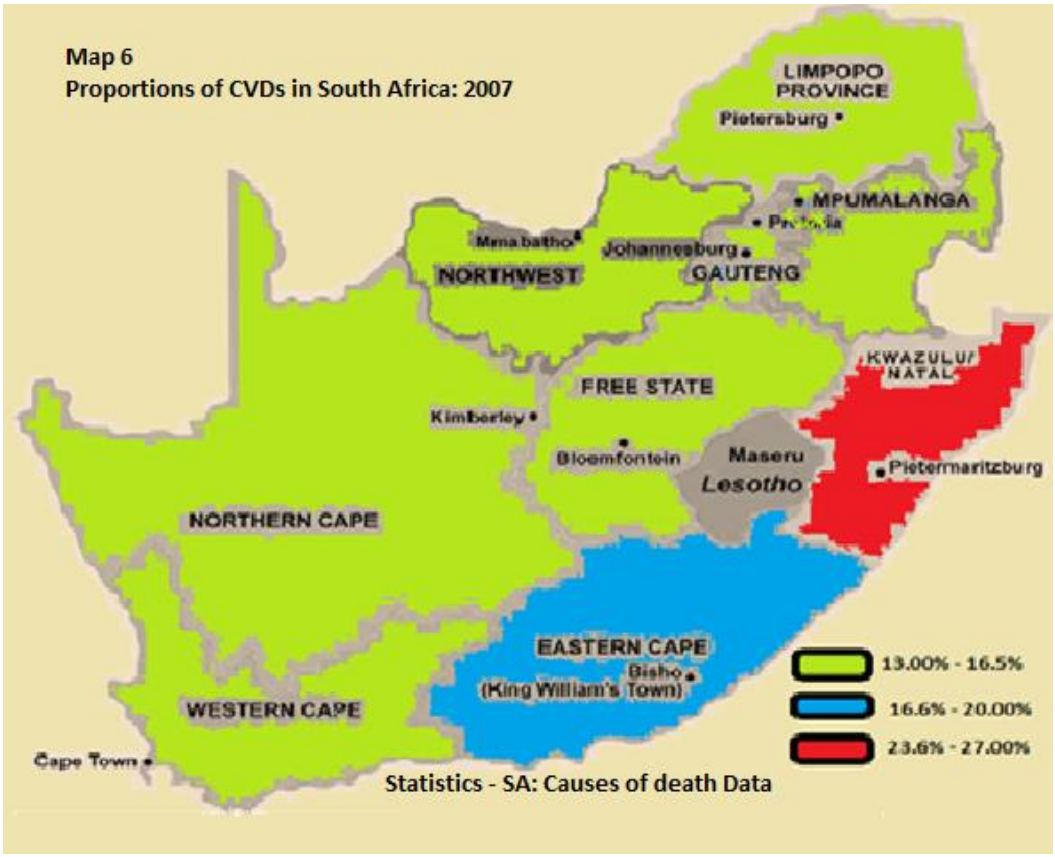
Map 3: Hazard Ratio of CVDs death in South Africe, 2006



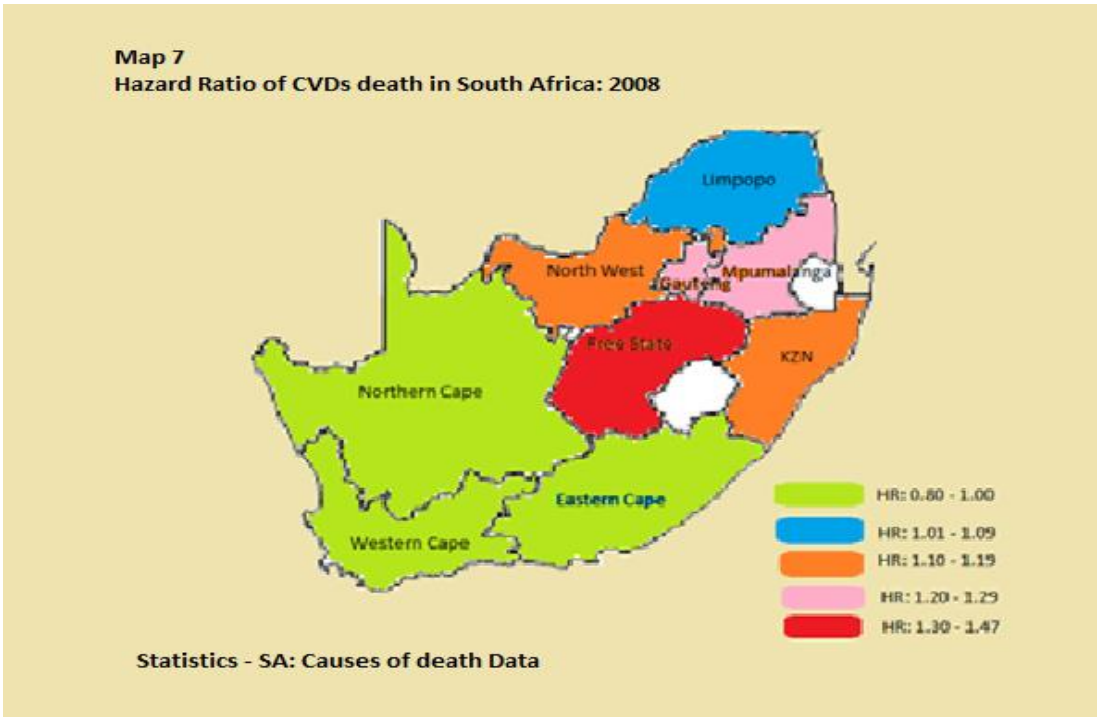
Map 4: Proportions of CVDs deaths in South Africa, 2006



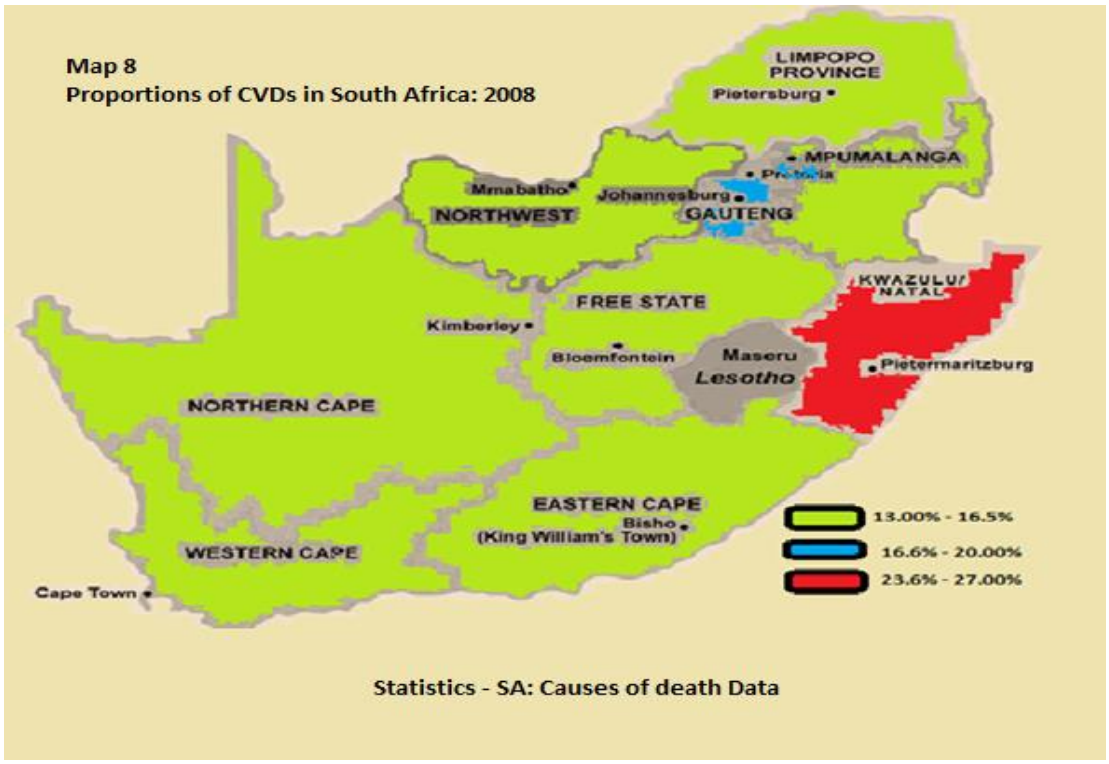
Map 5: Hazard of CVDs deaths in South Africa, 2007



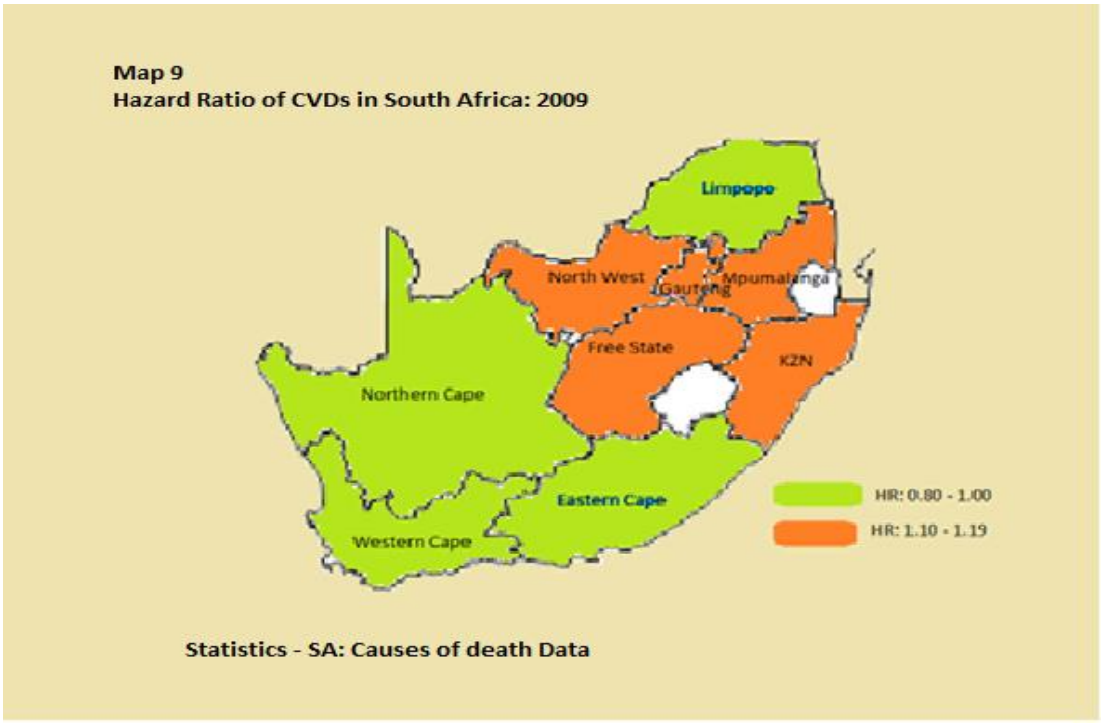
Map 6: Proportions of CVDs deaths in South Africa, 2007



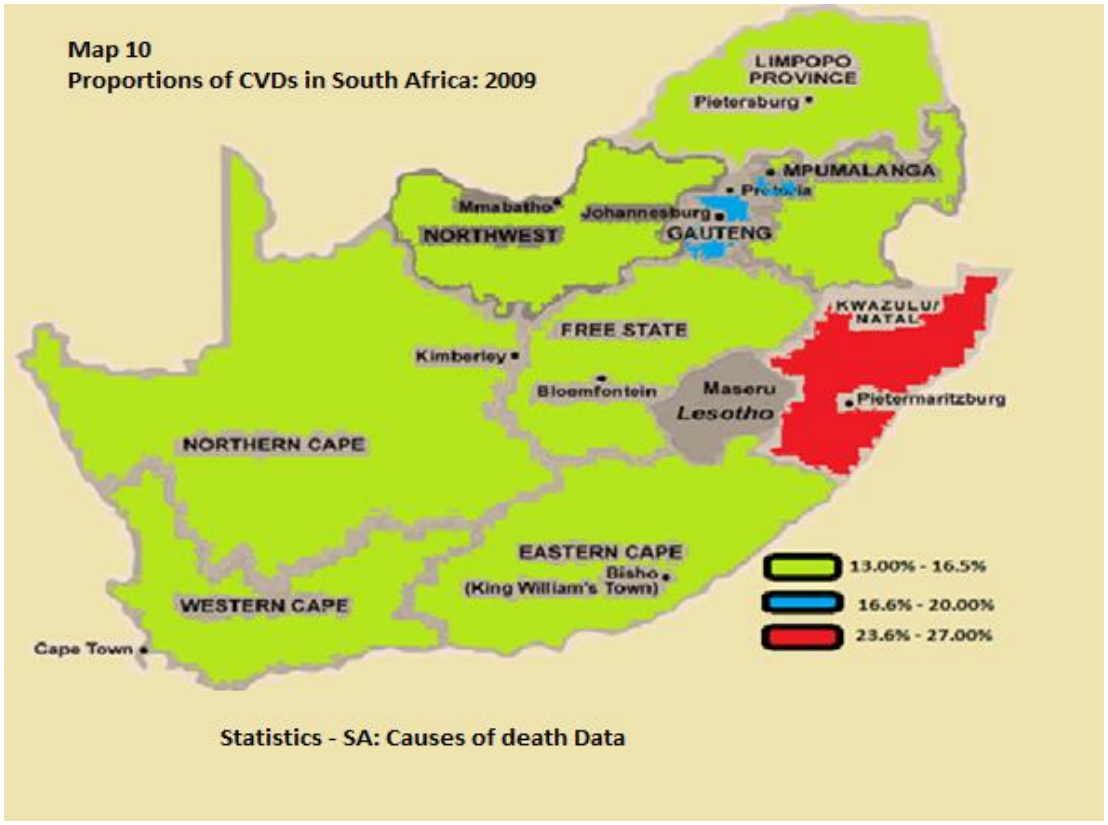
Map 7: Hazard Ratio of CVDs deaths in South Africa, 2008



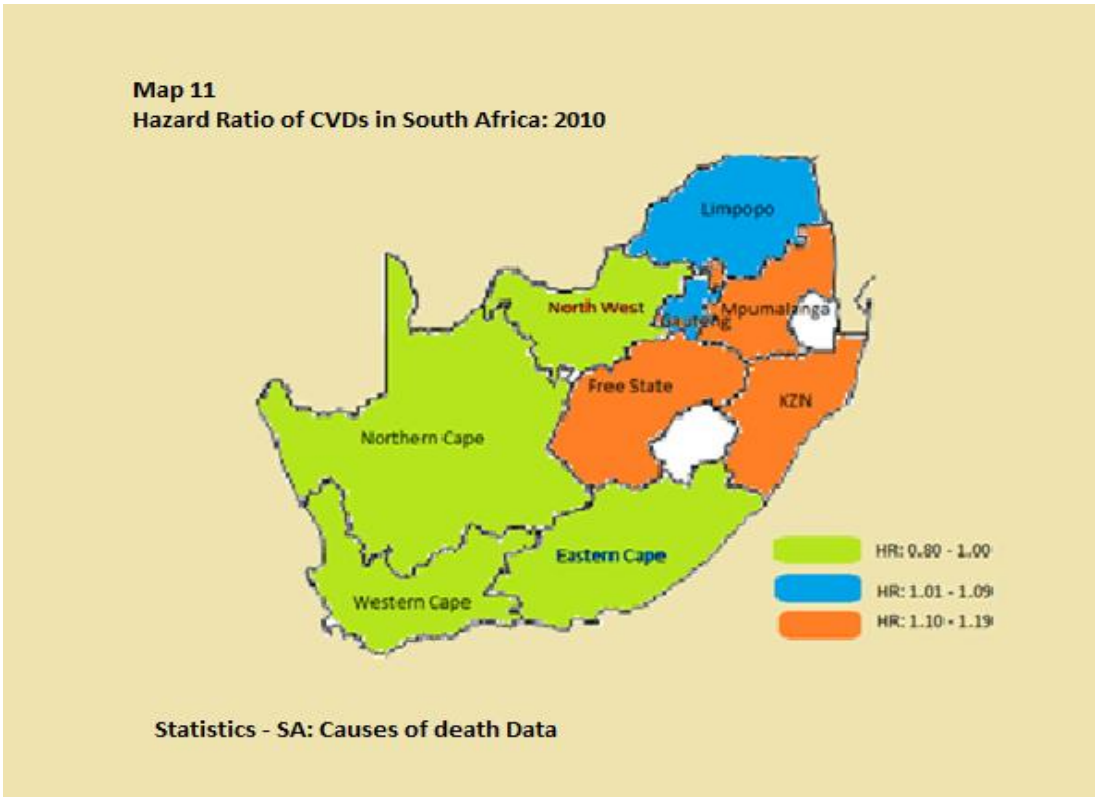
Map 8: Proportions of CVDs deaths in South Africa, 2008



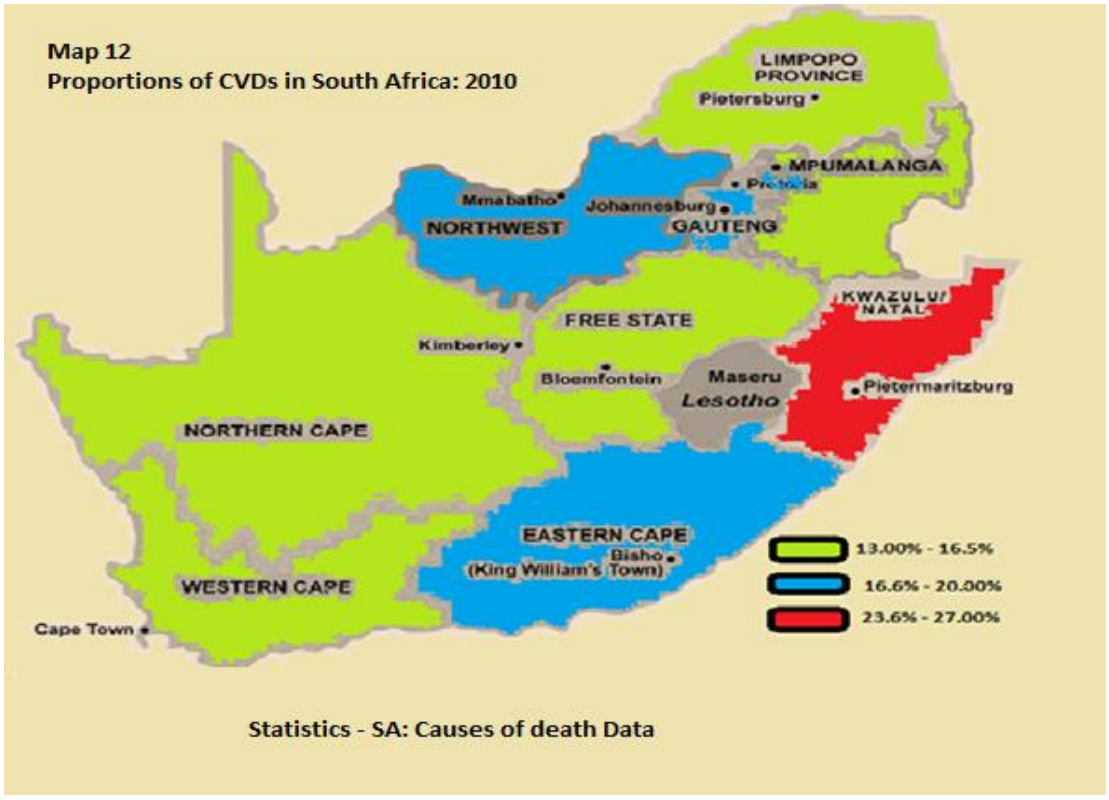
Map 9: Hazard ratio of CVDs deaths in South Africa, 2009



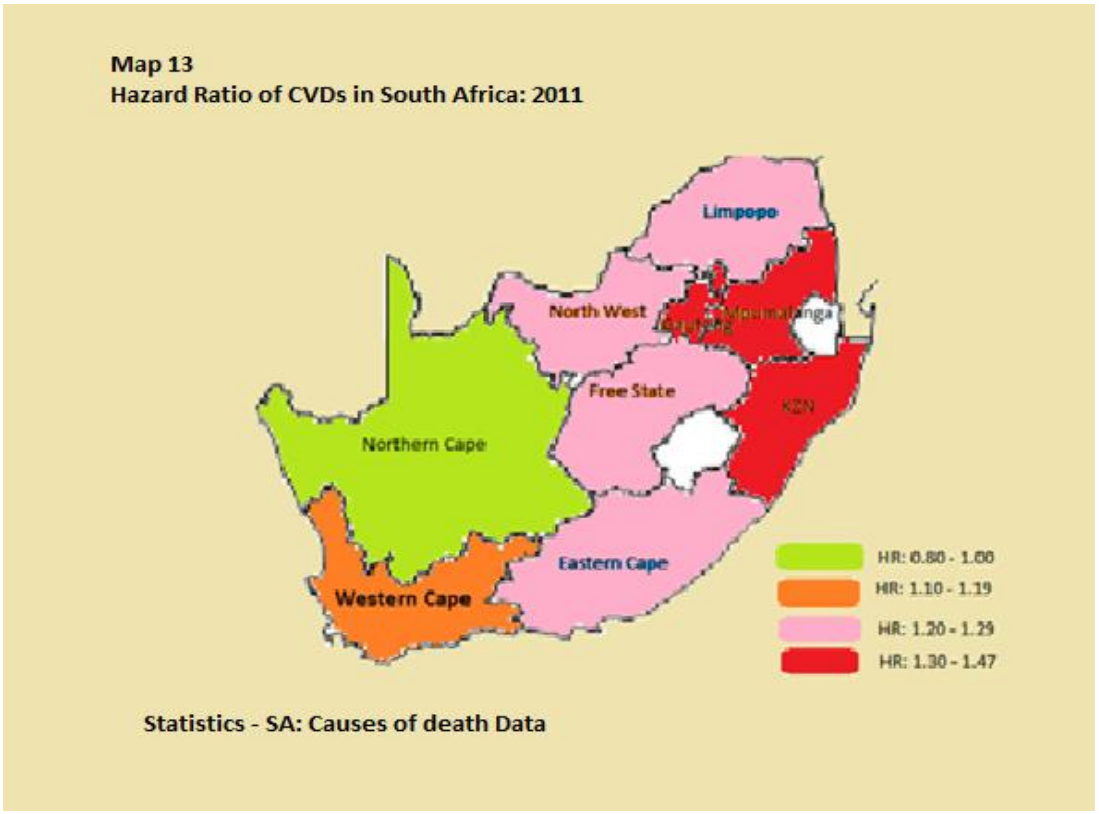
Map 10: Proportions of CVDs deaths in South Africa, 2009



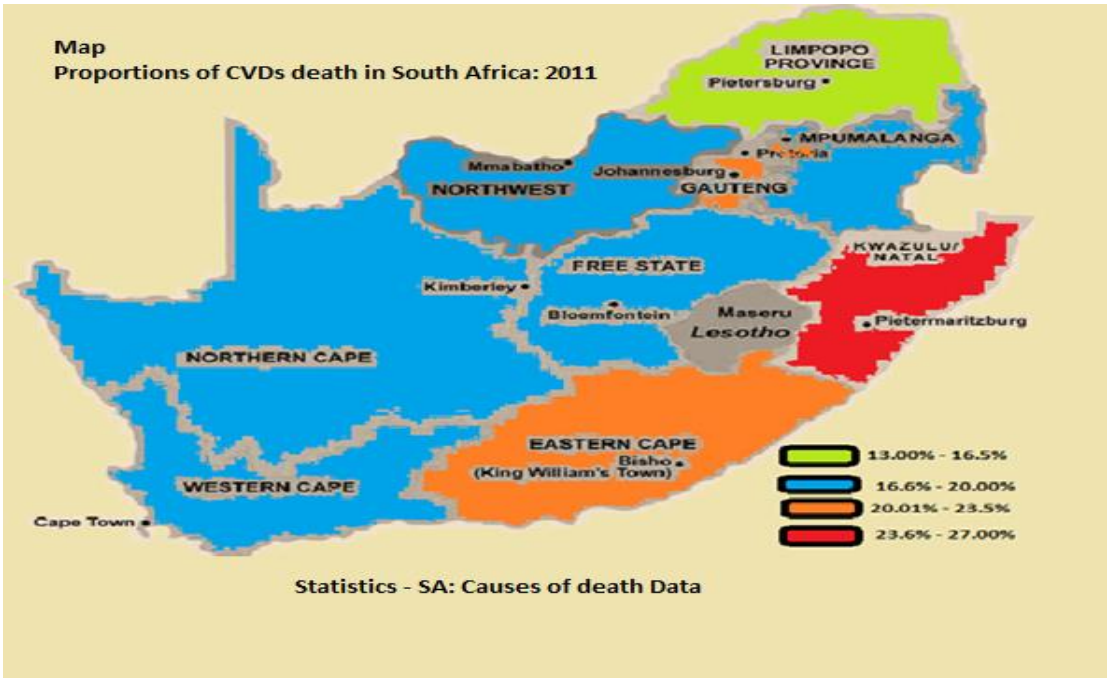
Map 11 : Hazard ratio of CVDs deaths in South Africa, 2010



Map 12: Proportions of CVDs deaths in South Africa, 2010



Map 13: Hazard ratio of CVDs deaths in South Africa, 2011



Map 14: Proportions of CVDs deaths in South Africa, 2011

R14/49 Mr Emery Ladi Ngamasana et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M140912

NAME: (Principal Investigator) Mr Emery Ladi Ngamasana et al

DEPARTMENT: Epidemiology and Biostatistics
School of Public Health

PROJECT TITLE: Socio-Demopgraphic Predictors of Cardiovascular Mortality in South Africa: Cross-Sectional Analysis of Stats-SA Data from 2006-2011

DATE CONSIDERED: 03/10/2014

DECISION: Approved unconditionally

CONDITIONS:

SUPERVISOR: Prof Tobias Chirwa

APPROVED BY:

Professor Cleaton-Jones, Chairperson, HREC (Medical)

DATE OF APPROVAL: 29/10/2014

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

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**.

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES