RISK FACTORS ASSOCIATED WITH DIABETES IN SOUTH AFRICAN ADULTS 15 YEARS AND OLDER FOR THE PERIOD 2011-2012

Masedi Menyatsoe

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

Johannesburg, June 2021

DECLARATION

I, Masedi Menyatsoe declare that this research report is my independent work. It is being submitted in partial fulfilment for the degree of Master of Science in Epidemiology in the field of Biostatistics, Faculty of Health Sciences at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Signature _____  _____

Date: 25 June 2021

ABSTRACT

Introduction: Diabetes is a major public health problem, with the prevalence increasing globally despite being a preventable disease. According to the International Diabetes Federation (IDF), diabetes is a chronic disease that occurs when the body cannot produce enough insulin or cannot use insulin effectively (IDF, 2015). Globally, 442 million people were living with diabetes in 2014 compared to an estimate of 108 million in 1990 (WHO, 2018). In 2019 the IDF estimate was 463 million people. The IDF projected that by 2045 there are expected to be 700 million people having diabetes globally. In 2015, approximately 1.6 million deaths were directly attributable to diabetes (WHO, 2018).

In South Africa, 4.6 million South Africans aged 21-79 years were living with diabetes in 2019. According to the mortality report of Statistics South Africa (StatsSA, 2016), in 2016 tuberculosis and diabetes were the top two causes of mortality in South Africa, with diabetes contributing 5.5% of the deaths. In 2014 diabetes ranked third highest cause of death, contributing 5.1%. In 2016 tuberculosis was the highest cause of death in males, while diabetes was the highest cause of death in females, causing 7.2% of female deaths. StatsSA found that the prevalence of tuberculosis is higher amongst persons with diabetes (StatsSA, 2016).

Objective: The objective of the study was to assess the risk factors associated with diabetes in South Africa among all adults (those aged 15 years and above) from the SANHANES 2011-2012 survey.

Methods: In this study, the SANHANES-1 is used to analyse the risk factors associated with diabetes. SANHANES-1 was a cross-sectional household national survey for the nutrition examination of people living in nine provinces of South Africa collected by the Human Science Research Council (HSRC) (Shisana et al., 2013). The data were collected using a stratified multi-stage disproportionate cluster sample design, with enumerator areas serving as the clusters. The current study was a secondary data analysis from SANHANES-1 survey for the people living in nine provinces of South Africa (Shisana et al., 2013).

Descriptive statistics such as frequency and weighted percentages were used for analysis. The Rao-Scott adjustment to the chi-squared test was used to test for associations between exposures of interest and diabetes. Weighted bivariate survey logistic regression was used and variables statistically significant at $p < 0.2$ were included in the multivariable weighted logistic regression.

Factors associated with diabetes were determined by fitting survey weighted logistic regression models both including and excluding the effect of BMI, since BMI was only measured on a subset of participants.

Results: The sample size for the cross-sectional study was 15 069 with BMI measured for a sample of 6410. The study found the prevalence of diabetes in South Africa was 5.3%. The prevalence of hypertension was 20.5% for the current study. Two provinces with the highest prevalence of diabetes were KwaZulu Natal 7.5% followed by Northern Cape with 6.5%. The prevalence of diabetes was higher in females (6.3%) than in males (4.1%). Prevelence of diabetes by age category showed age 55 to 64 had the highest with 16.8% with age 65+ at 16.2%.

All models found age (aOR 1.04, (P<0.0001 , 95% CI 1.03; 1.05), family history of diabetes (aOR=4.09, P<0.0001, 95% CI 2.84; 5.91) and hypertension (aOR=4.10, P<0.0001, 95% CI 2.69; 6.25) significantly associated with the presence of diabetes. Almost all models found a significant race effect, with Indians at a greater risk of diabetes compared to Africans.

Geotype (type of geographical location) was found significant only in the Generalised linear latent and mixed models (gllamm) model which included data on BMI. Those that were living in urban informal were found to be 32% lower risk of being diabetic compared to those living in urban formal.

Alcohol intake was found significant in some models. There was a significant interaction between alcohol intake and BMI. Controlling for BMI was found to slightly increase the odds of diabetes for those that drink monthly. The effect of BMI (aOR=1.20, P<0.0001 , 95% CI 1.09;1.29) was found significant in some models. Employed status was not found significant in any of the models.

Conclusion: In this study, the risk factors associated with diabetes mellitus include socio-demographic, lifestyle and co-morbidity factors. The findings suggested that age, race, alcohol intake, hypertension and family history of diabetes were statistically significant risk factors associated with diabetes mellitus. Significant interactions were found between alcohol and race, alcohol and hypertension for models (B, C, D and E), race and age, alcohol and BMI.

ACKNOWLEDGEMENTS

I would like to my supervisors Professor Jonathan Levin and Professor Khangelani Zuma for their assistance and patience with me.

My family for their encouragement especially my sisters, thank you so much your support and encouragement meant so much to me.

Also, I would like to thank Sub-Saharan African Consortium for Advanced Biostatistician (SSACAB) for training and funding me.

My colleagues and former colleagues for their encouragement, support and having faith in me.

Many thanks to the Human Science Research Council (HSRC) for allowing me to use their South African National Health and Nutrition Examination Survey (SANHANES-1) data for my analysis

LIST OF ABBREVIATIONS

AIHLD: African Institute for Health and Leadership Development
BMI: Body Mass Index
CHC: Community Health Centre
CVD: Cardiovascular Diseases
DHS: Demographic and Health Survey
DM: Diabetes Mellitus
EA: Enumeration Areas
EC: Eastern Cape
FS: Free State
GDM: Gestational Diabetes Mellitus
Geotype: Type of geographical location
GHS: General Household Survey
GP: Gauteng Province
HREC: Human Research Ethics Committee
HSRC: Human Sciences Research Council
HST: Health System Trust
IDF: International Diabetes Federation
JEMDSA: Journal of Endocrinology, Metabolism and Diabetes of South Africa
KZN: KwaZulu Natal
LP: Limpopo Province
MP: Mpumalanga Province
NC: Northern Cape
NCD: Non Communicable Disease
NW: North West
PHC: Primary Health Care
SA: South Africa
SADHS: South African Demographic and Health Survey
SANHANES: South Africa National Health and Nutritional Examination Survey
SEMDSA: Society of Endocrinology Metabolism and Diabetes of South Africa
StatsSA: Statistics South Africa
T2DM: Type 2 Diabetes Mellitus
TB: Tuberculosis
WC: Western Cape

WHO: World Health Organisation

WITS: University of Witwatersrand

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

DEFINITION OF TERMS

Type 1 diabetes is caused by an autoimmune reaction where the body's immune system attacks the insulin-producing beta cells in the islets of the pancreas gland.

Type 2 diabetes is defined as hyperglycaemia which is the result of inadequate production of insulin and inability of the body to respond fully to insulin.

Gestational diabetes is defined as high blood glucose level that is first detected during pregnancy.

Contents

CHAPTER ONE
INTRODUCTION

This chapter describes the burden of diabetes globally, in Africa and in South Africa, and states the relevance of the study to policy makers in South Africa. The literature on factors that are associated with diabetes in South Africa is reviewed. The chapter concludes by stating the aim and objectives of the study.

1.1 Background

1.1.1 Burden of Diabetes

Diabetes is a major public health problem, with its prevalence increasing globally, despite it being a preventable disease. Diabetes is a chronic disease that occurs when the body cannot produce enough insulin or cannot use insulin effectively (IDF, 2015). Globally, 442 million people were living with diabetes in 2014 compared to an estimate of 108 million in 1990 (WHO, 2018). The IDF projected that by 2045 there are expected to be 700 million people having diabetes globally. In 2015, approximately 1.6 million deaths were directly attributable to diabetes (WHO, 2018).

In 2017 the WHO diabetes key facts showed that diabetes has been increasing more rapidly in middle- and low-income countries compared to the increase in high-income countries.

1.1.2 Epidemiological transition in Africa

In 2019 IDF estimated that more than 19 million people living with diabetes were from the African region. This number is expected to increase to 47 million people by 2045. The four countries with the highest prevalence of diabetes in Africa are South Africa, the Democratic Republic of Congo (DRC), Nigeria and Ethiopia (Amod et al., 2017; IDF, 2019). The highest prevalence is found to be from South Africa (4.6 million with estimated confidence limits between [1.4-5.3 million]), Nigeria 2.7 million 95% CI [0.9 – 9.2 million]) and Democratic Republic of Congo (1.8 million 95% CI [1.5 – 2.2 million]) (IDF, 2019). On average 66.7% of people with diabetes in sub-Saharan Africa are undiagnosed, as some governments may not prioritise screening for diabetes (IDF, 2015; Amod et al., 2017). The Sub-Saharan Africa projections do not take into account the result of urbanization and ageing (Wild et al., 2000). In addition, these numbers do not account for obesity that has been increasing across urban Africa (Wild et al., 2000).

In 2015 more than 321,100 deaths were as a result of diabetes. The highest proportion of these deaths was among people under 60 years of age (IDF, 2015).

The increasing numbers of people with type 2 diabetes worldwide remains a major concern (Idemyor, 2010). This represents an added challenge due to the scarcity of financial resources in Sub-Saharan Africa (IDF, 2015). The increase is due to physical inactivity, obesity and urbanization (Idemyor, 2010). High rates of mortality from diabetes are attributed to late diagnosis and poor care (IDF, 2015). In many communities in sub-Saharan Africa, being overweight or obese is a symbol of wealth or prosperity (Idemyor, 2010).

1.1.3 The burden of diabetes in South Africa

In 2019, IDF estimated 4.6 million South Africans aged 21-79 years were living with diabetes (IDF, 2019). According to Statistics South Africa (StatsSA, 2016), in 2016 tuberculosis and diabetes were the top two causes of mortality in South Africa, with diabetes contributing 5.5% of the deaths. In 2014, diabetes ranked third highest cause of death, contributing 5.1%. In 2016 tuberculosis was the highest cause of death in males, while diabetes was the highest cause in females, causing 7.2% of female deaths (StatsSA, 2016).

WHO found risk factors that are associated with diabetes are obesity, unhealthy diet, high blood pressure, inadequate fruit and vegetable intake, physical inactivity, smoking, alcohol and a family history of diabetes. Prevention measures against diabetes include a simple lifestyle, maintaining healthy body weight and being physically active (WHO, 2018).

One of the factors that have been found to be associated with diabetes is an unhealthy diet. The first South African National Health and Nutrition Examination Survey (SANHANES-1) found many factors that contributed to food consumption patterns of the household, these factors are household income, nutritional knowledge, and educational level (Shisana et al., 2013).

This study aims to assess factors that are associated with diabetes using a list of potential risk factors identified by WHO.

1.2 Problem statement

Obesity has been identified as one of the major risk factors for diabetes, which suggests that diet needs to be improved. SANHANES-1 (Shisana et al., 2013) reported that the prevalence of being overweight and the prevalence of obesity were significantly higher in females than in males in South Africa, with the prevalences being 24.8% and 39.2% in females compared to 20.1% and 10.6% in males.

The South African Demographic and Health Survey stated that overnutrition is an ongoing problem in South Africa (SADHS, 2016). SADHS also showed a high obesity rate among high income-earning women and adolescent girls (SADHS, 2016). Only 4.6% of the adolescent population consumed the amount of fruit and vegetables recommended in the guidelines (Shisana et al., 2013). The patterns of food consumption from the SANHANES-1 could be analysed to examine dietary intake.

Previous studies have shown a rising trend of overweight among youth in all races in South Africa (Peer et al., 2012; WHO, IDF, 2018; Reddy et al.,2012; Shisana et al., 2013), which showed that dietary lifestyle in South African has changed drastically resulting in a high risk of obesity related diseases. The relationship between diabetes as the outcome variable and Body Mass Index (BMI) as the primary exposure variable will be examined, with secondary exposure variables including family history of diabetes, tobacco usage, employment status and race.

1.3 Justification of the study

The Journal of Endocrinology, Metabolism and Diabetes of South Africa (Amod et al., 2017) found that in Sub-Saharan Africa diabetes poses major public health and socio-economic challenges. Disadvantaged communities face greater risk. The prevalence is expected to be the largest in developing regions because of ageing and urbanisation (Whiting et al., 2011; Amod et al., 2017). This rise will result in a higher burden of future diabetes (Amod et al., 2017). Thus, a need exists for effective intervention (Skelly et al., 2008) to determine these challenges (Amod et al., 2017) caused by an increasing prevalence of diabetes. Development of initiatives is essential in order to reduce the incidence and mortality rates of diabetes (Amod et al., 2017).

Africa is faced with major challenges of morbidity and mortality due to diabetic patients, which results in additional socio-economic costs (Mbanya et al., 2010). In addition, lack of access to medication especially insulin leads to avoidable metabolic complications (Mbanya et al., 2010).

Furthermore, there is a need for studies that will highlight the risk associated with lifestyles and socio-demographic factors on the prevalence of diabetes in South Africa. There are major risk factors that are associated with diabetes identified by WHO.

The study will assess the risk factors that are associated with diabetes with BMI as the primary exposure variable and family history of diabetes, tobacco usage, employment status and race as secondary exposures using the SANHANES-I survey.

1.4 Literature Review

1.4.1 South African Health System

The public healthcare system in South Africa provides access to healthcare for the majority of the population (Gray et al., 2013). The South African public healthcare system is divided into primary, secondary and tertiary facilities, which are all managed by provincial departments (Gray et al., 2013).

The Health System Trust in their South African health review for 2012/13 found evidence of a growing burden of Non-Communicable Diseases (Gray et al., 2013). In September 2011 there was a summit to prevent and control Non-Communicable diseases that approved action against different risk factors including behavioural (lifestyle), environmental and structural factors. The strategies proposed include preventing NCDs, promoting health and developing policies for NCD management (Gray et al., 2013).

The general household survey (GHS) carried out by StatsSA in 2011, found that 61.2% of households used public clinics to seek medical help, with 24.3% using private doctors and 9.5% using public hospitals (GHS, 2011). The survey found that 81.3% of black Africans and 63.1% of the coloured population group used public clinics, while 88% of whites and 64.1% of Indian / Asians used private health care facilities. Only 10% of whites used public health facilities, while only 17.2% of black Africans used private health care. The StatsSA general household survey collected data on diabetes and found 4.7% of participants aged 25 years or older had been diagnosed with diabetes by a medical practitioner or nurse. The prevalence was 5.2% among females and 4.1% among males (GHS, 2011).

1.4.1.1 Diabetes Care in South Africa

Whiting et al. (2011) modelled global estimates for diabetes for the period from 2011 to 2030. The analysis showed the prevalence of diabetes increasing as a result of aging and risk factors such as obesity (Whiting et al., 2011). Growing population, aging, urbanization and lifestyle change are likely to increase the number of diabetes cases worldwide by 50.7% by the year 2030 (Whiting et al., 2011).

A study was conducted in KwaZulu-Natal to assess doctors' compliance with the Society for Endocrine Metabolism and Diabetes of South Africa (SEMDSA) diabetes guidelines. The study found that compliance of doctors with the current diabetes guidelines was poor (Igbojiaku et al., 2012).

Diabetes and other NCDs are causing premature death in people aged 60 and younger causing major financial issues in many countries (IDF, 2017; WHO, 2018; SANHANES, 2012; Shisana et al., 2013). IDF, WHO and SANHANES found evidence of the risk factors for NCD and diabetes are well known and preventable (IDF, 2017; WHO, 2018; SANHANES, 2012; Shisana et al., 2013).

1.4.1.2 Classification of Diabetes

"Type 1 diabetes is defined as diabetes which is caused by an autoimmune reaction where the body's immune system attacks the insulin-producing beta cells in the islets of the pancreas gland". This type of diabetes is characterised as an absolute deficiency of insulin (IDF, 2017; ADA, 1997; WHO, 2018). It accounts for 10% of diabetes cases (IDF, 2017; ADA, 1997; WHO, 2018). The risk factors are not fully understood (IDF, 2018). Type 1 diabetes is managed by proper daily treatment, regular monitoring of blood glucose, maintaining a healthy diet (IDF; WHO, 2018).

"Type 2 diabetes is defined as hyperglycaemia which is the result of inadequate production of insulin and inability of the body to respond fully to insulin" (IDF, 2017). This type of diabetes is characterised as insulin resistance. It accounts for 90% of diabetes cases (Hall et al., 2011;

IDF; ADA, 1997; WHO). Risk factors are obesity, poor diet and nutrition, physical inactivity and prediabetes (as a result of family history) (IDF; ADA, 1997; WHO). Type 2 diabetes is managed by a healthy diet, physical activity, and maintenance of healthy body weight (IDF; ADA, 1997; WHO).

Gestational diabetes is defined as “high blood glucose level that is first detected during pregnancy”. It accounts for 75–90% of cases of high blood glucose during pregnancy (IDF, ADA,1997, WHO). Risk factors are older age, overweight or obesity, excessive weight gain during pregnancy, a family history of diabetes and a history of stillbirth (IDF, ADA,1997, WHO).

1.4.2 Lifestyle Factors

1.4.2.1 Nutritional Knowledge

The SANHANES-1 report recommended that health care promotions related to dietary knowledge and nutritional support could provide information that would improve the population’s knowledge and hence could change their diet and related behaviour (Shisana et al., 2013).

SANHANES-1 assessed the nutritional knowledge of participants and produced a knowledge score given by the number of correct answers; a score of 0-3 was considered low, a score of 4-6 as medium and a score of 7-10 as high. The overall mean knowledge score was 5.3; of the participants, 22.6% achieved a high score, 62.9% achieved a medium score and 14.5% achieved a low score (Shisana et al., 2013). SANHANES-1 fat score ranged for low score was 0-7, moderate 8-14 and high 15-20 as a person eating a lot of fat in their foods. The dietary intake of the SANHANES-1 shows the mean fat score was 7.37 for males and 7.15 for females. The mean score differs between rural informal with score from 5.57 to 8.32 in urban formal (Shisana et al., 2013). For race groups of participants, the mean fat score was highest in whites 8.54 compared to other race groups of black Africans 7.09 and Indians 7.09 (Shisana et al., 2013). SANHANES-1 “fat score” was a separate scale to the knowledge score.

One of the factors that have been found to be associated with diabetes is an unhealthy diet. The SANHANES-1 found many factors that contributed to food consumption patterns of the household, these factors are household income, nutritional knowledge, and educational level (Shisana et al., 2013).

A study was carried out in the DRC to assess the knowledge about diabetes among patients with type 2 diabetes, and found that overall the knowledge score was poor (Ntontolo et al., 2017). The respondents scored 39.3 % for risk factors knowledge. Poor knowledge has been identified as one of the barriers to diabetes prevention programmes (Ntontolo et al., 2017). The study suggested health education intervention programmes for diabetic patients (Ntontolo et al., 2017).

1.4.2.2 Smoking status

Smokers have a higher risk of developing type 2 diabetes compared to those that do not smoke (SGR, 2014; Chang, 2012). The risk increases with increasing the number of cigarettes smoked per day (Chang, 2012). Smoking causes damage to cells which may be highly related to increasing risk of diabetes (SGR, 2014). Some studies have confirmed that people exposed to high levels of nicotine are more likely to have type 2 diabetes (SGR, 2014). Smoking is a well known risk factor for various diseases (Chang, 2012). It has been shown, that there is a positive association between smoking and insulin resistance (SGR, 2014; Chang, 2012).

1.4.2.3 Alcohol intake

In their study, (Soo- Jeong et al., 2012), explained how chronic use of alcohol is a potential risk factor for the development of type 2 diabetes. Chronic use of alcohol causes insulin resistance and pancreatic β-cell dysfunction that can result in the onset of type 2 diabetes (Soo- Jeong et al., 2012).

1.4.3 Socio-Demographic Factors

Several studies have found obesity to be a major risk factor for non-communicable diseases (NCD) in South Africa such as diabetes and cardiovascular disease (CVD) (Zemlin, 2016; WHO, IDF, 2018); Peer et al.,2012; Reddy et al.,2012; Ponesai et al., 2011).

Ponesai et al. (2011) conducted a case-control study in Zimbabwe using structured questionnaires on a sample of 68 cases and 68 controls. Cases were any person, 18 years and above with diabetes, hypoglycaemia and nephropathy and controls were any person, 18 years and above with diabetes without hypoglycaemia and nephropathy. Hypoglycaemia and nephropathy are complications of diabetes and the case control study aimed to investigate factors associated with diabetic complications. They found socio-demographic factors associated with diabetic complications were being unmarried, widowed, having primary education, having a diet containing more sugar, living in an urban area and distance from the health facility. According to Ponesai et al., (2011) risk factors for diabetic complications among diabetic patients are largely patient's knowledge and health service factors.

1.4.3.1 Gender

Shisana et al. (2013) found that 62.4% of male participants were physically fit, while 27.9% of males and 45.2% of females were found to be unfit by physical examination. SANHANES-1 self-reported prevalence of high blood sugar was 20.7% and 74.7% of South Africans believed that what they ate has an impact on their weight. In 2019, IDF found more deaths associated with diabetes in women (2.3 million) than in men (1.9 million) (IDF, 2019).

1.4.3.2 Race

The SANHANES-1 study found that amongst the SA population, the prevalence of DM was higher for the Coloured and Indian population groups.

1.4.3.3 Age

Some studies found age as a risk factor for diabetes (Ntontolo et al., 2017; Mbanya et al., 2010). IDF estimated that by 2030 the highest prevalence will be in the oldest age-group (60–79 years). A study carried out in sub-Saharan Africa found that the prevalence of diabetes increases with age (Mbanya et al., 2010). In 2015 IDF estimated that globally 5 million deaths among people aged 20-79 were attributable to diabetes (IDF, 2015; Amod et al., 2017).

1.4.3.4 BMI

The study in the DRC by Ntontolo et al. (2017) showed that obesity was one of the major risk factors for the development of type 2 diabetes. The BMI is calculated by dividing weight by height squared. An individual is classified as underweight if BMI <18 kg/m squared, normal when the BMI is between 18 and 25 kg/m squared, between 25 and 30 kg/m squared is considered overweight and that of 30 or above kg/m squared is obese according to the WHO criteria.

Sartorius et al. (2015) found that the prevalence of obesity in South Africa had increased from 23.5% in 2008 to 27.2% in 2012; they also showed that the prevalence of obesity was higher among females in South Africa. Giday et al., (2014) assessed the risk factors associated with type 2 diabetes in central zone of Tigray, North Ethiopia. They used a sample of 73 cases and 290 control face to face interviews (Giday et al., 2014). They found that the odds of developing type 2 diabetes were 9.6 times higher for individuals who were overweight or obese and were 6 times higher for individuals with hypertension (Giday et al., 2014).

1.4.3.5 Geographical location

Some researchers found that the prevalence of diabetes and its risk factors (mainly obesity) was higher in urban areas than in rural areas (Mbanya et al., 2010; Shisana et al., 2013). SANHANES found the highest prevalence of diabetes in rural informal dwellers (11.9%) and urban formal dwellers (11.3%). There is a rapid urbanization in sub-Saharan Africa which has been shown to be a major determinant in the rise of burden of diabetes and other cardiovascular diseases (Mbanya et al., 2010).

1.4.4 Co-Morbidity Factors

1.4.4.1 Hypertension

In Tanzania, a cross sectional study was carried out to determine prevalence and knowledge about diabetes and hypertension (Mbuya et al., 2014). The authors found that age was a risk factor for both diabetes and hypertension, with an increase in age identified as a risk factor for diabetes by 38.1% respondents while hypertension was identified by 36.7% (Mbuya et al., 2014). The study showed that most of the participants were aware of diabetes and hypertension but knowledge on causes, signs and symptoms, risk factors and associated complications were poor (Mbuya et al., 2014).

A study was carried out in Botswana among 401 patients for three months, which found 61.2% of diabetes patients had hypertension and 56.4% were obese (Mengesha, 2007). The results showed hypertension was associated with age, gender, type of diabetes and BMI (Mengesha, 2007). It further found the percentage of hypertension increased to 70.9% among diabetic population aged older than 50 years (Mengesha, 2007). Several studies found hypertension tended to affect females who are diabetic more than males (overweight and obese diabetic patients) (Mengesha, 2007; Alqudah et al., 2017; Balogun et al., 2011).

Some researchers found that hypertension and diabetes were co-morbidities (Mengesha, 2007; Alqudah et al., 2017; Alanazi et al., 2017). Alqudah et al. (2017) in their study showed that hypertension is associated with the duration of diabetes. Both share common factors such as insulin resistance, aging, obesity (Balogun et al., 2011; Nouh et al., 2017; Alanazi et al., 2017).

1.4.4.2 Family History

A family history of diabetes has been shown to be an independent risk factor for the type 2 diabetes (Mbanya et al., 2010). The Centre for Disease Control (CDC) stated that if a close relative of one individual has type 2 diabetes that individual is likely to be prediabetic (CDC,2018). This prediabetes can be prevented by eating healthy, being physically active, and maintaining a healthy weight.

Peer et al. (2012) conducted a study among urban-dwelling black South Africans to investigate the rising diabetes prevalence. They found women with diabetes had a high rate of physical inactivity compared to women without diabetes. They also found that diabetes may be associated with specific diabetes risk factors (Peer et al., 2012). They further stated that these high prevalence rates may be a result of impaired glucose tolerance (IGT) and obesity. In their study diabetes was significantly associated with family history (Peer et al., 2012).

1.5 Research question

What are the risk factors associated with diabetes in South African adults (those aged 15 years and above) in the SANHANES 2011-2012 survey?

1.6 Research aim and objectives

1.6.1 Research aim

To assess the risk factors associated with diabetes in South African adults (those aged 15 years and above) from SANHANES 2011-2012 survey.

1.6.2 Research objectives

1. To describe the prevalence of diabetes among all adults (those aged 15 years and above) in South Africa 2011-2012 SANHANES survey
2. To determine risk factors associated with diabetes among all adults (those aged 15 years and above) of South Africa SANHANES-1 2011-2012 survey
3. To compare different models used to determine risk factors associated with diabetes for all adults (those aged 15 years and above) of South Africa SANHANES-1 2011-2012 survey

1.7 Thesis overview

The research paper is presented as follows:

Chapter 1: Introduction, justification and literature of the research study
Chapter 2: Study Design and Methodology followed for the research study
Chapter 3: Results
Chapter 4: Discussion
Chapter 5: Conclusion and recommendations for future studies

CHAPTER TWO
STUDY DESIGN AND METHODOLOGY

2. Overview

This chapter describes the methodology of the study. It includes the study design, setting, population, sampling, data management process, data analysis and ethical considerations of the study.

2.1 Study design

2.1.1 Primary Study design

SANHANES-1 was a cross-sectional national household survey with the aim of examining the health and nutritional status of people living in the nine provinces of South Africa (Shisana et al., 2013). The study was conducted by the Human Sciences Research Council (HSRC). The data were collected using a stratified multistage sample design, with unequal selection probabilities, with enumerator areas serving as the primary sampling units or clusters, within which households were selected. A total of 500 enumerator areas were sampled, and within each enumerator area 20 households were sampled. A sample of individuals within the selected households were interviewed, all persons in households were eligible to participate giving a total sample size of n=25,532 individuals with completed interviews (Shisana et al., 2013).

2.1.2 Secondary Analysis

In this study SANHANES-1 was used to determine the risk factors associated with diabetes. The study reports on a secondary data analysis from SANHANES-1 survey for the people living in nine provinces of South Africa (Shisana et al., 2013).

2.1.3 South African demography

The 2018 mid-year population estimates revealed that South Africa has approximately 57.73 million people with approximately 51% being female (StatsSA, 2018). The area of South Africa is about 1,220,813 square kilometres. It’s a multi-racial country with 80.9% of the population being black. The life expectancy at birth was estimated to be 61.1 years for males and 67.3 years for females. The country is divided into nine provinces namely Eastern Cape (EC), Free

State (FS), Gauteng (GP), KwaZulu-Natal (KZN), Limpopo (LP), Mpumalanga (MP), Northern Cape (NC), North West (NW) and Western Cape (WC). The largest province in terms of population size was Gauteng with approximately 25.4% of the population (StatsSA, 2018).

2.2 Study population

People living in all nine provinces of South African households were eligible to participate in the 2011-2012 SANHANES nutrition survey. Those individuals staying in educational institutions, old-age homes, hospitals, homeless people, and uniformed-service barracks were excluded from the primary study. For the secondary analysis all participants in the primary study were included, provided that their diabetes status was known as either yes or no, and provided that their sampling weight was not missing.

2.3 Data Management

2.3.1 Data collection

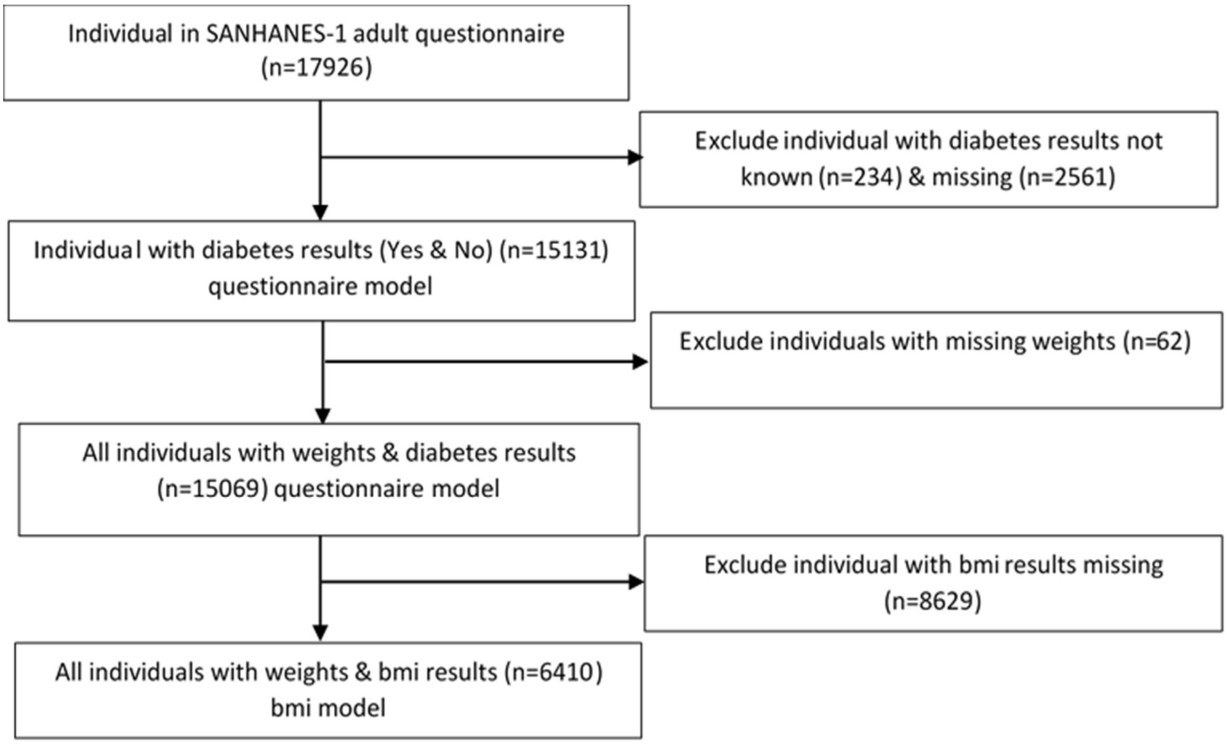
The data was collected by HSRC using HSRC's 2007 Master Sample. The HSRC sampled 500 enumeration areas (EAs) from their 2007 Master Sample, with 20 visiting points randomly selected from each EA, so the sample consisted of 10,000 visiting points. The HSRC's 2007 Master Sample was allocated into EAs according to municipality and province.

2.3.2 Data Analysis

Data for people living in the nine provinces of South Africa, that participated in the 2011-2012 SANHANES-1 and have diabetes results was analysed. The data was analysed using Stata Release 15. The datasets (SANHANES-1) were obtained from HSRC through data curation systems (http://curation.hsrc.ac.za/Datasets-XKAHAA.phtml). Three datasets were obtained (1) adult questionnaire consisting of participants geographic information, biographic details, TB, nutrition and non-communicable disease, (2) those that completed the adult questionnaire underwent physical examination such as height, weight provided and blood specimens (Physical examination data), and lastly SANHANES VP data points consisting of visiting points and EA weights.

Out of the 25,532 individuals that completed interviews 17,926 were all adults (those aged 15 years and above), of whom 2,561 had missing results for diabetes status and an additional 234 respondents did not know their diabetes status. Observations with missing sampling weights were also omitted from the analysis (62). A total sample size of 15,069 adults (those aged 15 years and above) were included in the analysis of whom 6,410 had BMI measurement as shown in Figure 2.1.

Figure 2.1 Study inclusion & exclusion graph



2.3.3 Study variables and codes used

This section contains a summary of the variables, divided into the survey design variables and then the study variables, followed by a description of the codes used for the variables.

2.3.4.1 Study design variables

Table 1: Survey design variables

	Variable	
Strata	Province	
Sample weights (Questionnaire)	INDV_WGT_ANSW_QSTN_BENCH_NEW	Adult weight
Sample weights (physical examination)	INDV_WGT_ANSW_PHYS_BENCH_NEW	Bmi weight
Primary sampling unit (PSU)	Enumeration area	EA
EA weight	EA_wgt	

- INDV_WGT_ANSW_QSTN_BENCH_NEW: analysis weight for the questionnaire/interview (used when analysing variables from the questionnaire).
- INDV_WGT_ANSW_PHYS_BENCH_NEW: analysis weight for the physical examination.
- EA_wgt: analysis weight for the Enumeration area

2.3.4.2 Study variables

Table 2: Variable description and measurement

Variable	Type of variable	Measurement
Diabetes	Outcome	Binary
Age	Confounder	Continuous
Gender	Confounder	Binary
Alcohol intake	Confounder	Categorical
Province	Demographic	Categorical
BMI	Primary exposure	Continuous
Tuberculosis	Secondary exposure	Binary
Hypertension	Secondary exposure	Categorical
Family history	Secondary exposure	Binary
Smoking status	Secondary exposure	Categorical
Employment	Secondary exposure	Categorical
Geotype	Secondary exposure	Categorical
Race	Secondary exposure	Categorical

2.3.4.3 Codes used for the study variables

- Diabetes: The variable was initially coded as 1=Yes, 2= No and 3 = don't know. The variable was changed into binary variable with 2 categories, which were coded as 0=No, 1=Yes. Those that didn't know their diabetes results were dropped from the study.
- Age as category: This variable was already categorised in the dataset into 6 classes of 10 years interval starting with 1=15 to 24, 2= 25 to 34, 3=35 to 44, 4=45 to 54, 5=55 to 64 and 6= 65+.
- Race: The population group was already categorised into 5 categories which were coded as 1=African, 2=White, 3=Coloureds, 4=Indians, and 5=other (Missing race was combined with other).
- Smoking status: This variable already was categorised into 5 categories, which were coded 1=Yes, daily, 2=Yes, less than daily, 3=Yes, but not now, 4=No, not at all and 5=Other (combined Don't know and Missing)
- Gender: This variable was already categorised into 2 categories, which were coded as 1=Male, 2= Female.

- Family history: the variable for family history was those that had relative with diabetes, variable was already categorised into 3 categories, which were coded as 1=Yes, 2= No, 3=Other (combined don't know and Missing).
- Alcohol intake: This variable was already categorised into 4 categories, which were coded as 0= Never, 1= Monthly or less, 2=2-4 times a month, 3= 2 or more times a week.
- Geotype: This variable was already categorised into 4 categories, which were coded 1= Urban formal, 2= Urban informal, 3= Rural informal (tribal), 4= Rural formal (Farms).
- Province: This variable was already categorised into 9 categories, which were coded as 1=Western Cape, 2= Eastern Cape, 3= Northern Cape, 4= Free State, 5= KwaZulu-Natal, 6=North West 7=Gauteng, 8= Mpumalanga, 9=Limpopo.
- Employment: The variable employment status had 14 categories in the database. For this analysis the variable was recoded into 4 categories. Employed (coded as 1) included full time employees, those who were self employed, those who were employed part time, those who did seasonal work and those who worked in the informal sector. Unemployed (coded as 2) included those who were unemployed, those looking for work and those who were housewives, homemaker. Unable to work and other (coded as 3) included those who were sick, disabled and unable to work, those that retired and those that specified their job status as other. Not looking for work (coded as 4) included housewife, homemaker, not looking for work, (unemployed but not looking for work) and student / pupil/ learner
- Hypertension: the variable for those that were told they have/had High blood Pressure, variable already categorised into 3 categories, which were coded as 1=Yes, 2= No, 3=Other (combined don't know and Missing).

2.4 Data Analysis

2.4.1 Descriptive analysis

Objective One: To describe the prevalence of diabetes among all adults (those aged 15 years and above) in South Africa 2011-2012 SANHANES survey.

The prevalence of diabetes was given overall and by gender, province and race. This was given as the unweighted frequency together with the weighted percentage. Categorical explanatory variables were summarised using unweighted frequencies and weighted percentages, together with the diabetes status of participants for each level of the factor. For continuous explanatory variables the mean and 95% confidence limits were given separately by diabetes status.

2.4.2 Objective two

To determine risk factors associated with diabetes among all adults (those aged 15 years and above) of South Africa SANHANES-1 2011-2012 survey.

2.4.2.1 Rao-Scott adjustment to chi-square test

The Rao-Scott adjustment to the chi-square test was used to test the association between individual risk exposures such as smoking status, alcohol intake, race, TB status and hypertension and the outcome, namely diabetes.

2.4.2.2 Weighted Logistic Regression Model

Unadjusted weighted logistic regression models were fitted to examine risk factors associated with diabetes. Only risk factors with p-values <=0.2 in the unadjusted linearized weighted logistic regression were included in the multiple linearized weighted logistic Regression model. The unadjusted odds ratios were computed together with 95% confidence limits and the unadjusted P-values. The "svy" command was used for survey weighted multivariate model, which computes standard errors using the linearized variance estimator. This linearized estimator estimates variance within each stratum, these variance estimators use a first-order Taylor series linear approximation (Wolter 2007). Kackar and Harville (1984) proposed a correction to the variance/covariance matrix of the fixed effects to adjust for the underestimation of the standard errors of fixed effects. This correction approximates bias in the variance/covariance matrix with quantities approximate to the second-order Taylor series approximation to the bias term. The correction fixes the covariance matrix at the end of the estimation process.

The logistic regression model was formulated as:-

$$\text{logit}(y) = In(odds) = In\left(\frac{p}{1-p}\right) = \alpha + \beta x, \quad p = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} \qquad (1)$$

p= $P(Y = 1 / X = x)$

2.4.2.3 Weighted Multiple Logistic regression models

Multivariable weighted logistic regression models were fitted to identify the risk factors associated with diabetes, examine the statistical difference of proportions and to control for confounding variables. These models were fitted in a similar way to the unadjusted weighted logistic regression (2.4.2.2) adjusting for risk factors associated with diabetes.

$$logit(y) = In(odds) = In(\frac{p}{1-p}) = \alpha + \beta_1 x_1 + ... + \beta_k x_k, \; p = \frac{e^{\alpha + \beta_1 x_1 + ... + \beta_k x_k}}{1 + e^{\alpha + \beta_1 x_1 + ... + \beta_k x_k}} \quad (2)$$

Therefore p= $P(Y = 1 / X_1 = x_1, ..X_k = x_K)$,where *p* is the probability of the outcome of interest (diabetes), α is the intercept parameter, β is a regression coefficient, and x is a predictor.

2.4.3 Generalized linear mixed-effects model

Breslow and Clayton (1993) formulated the Generalized Linear Mixed Model (GLMM) which provides the extension of the Linear Mixed Model. The GLMMs are also known as multilevel or hierarchical generalized linear models. Mixed-effects models represent the covariance structure related to the grouping of data by associating the common random intercept to observations that have the same level of a grouping variable in this case enumeration areas. These models are popular for multilevel data in which elementary units (here survey respondents) are nested in clusters (here census EAs).

$$\log(\frac{p_{ij}}{1 - p_{ij}}) = \alpha_j + \beta x_{ij} \quad (3)$$

$$\alpha_j = \alpha + u_j, u_j \sim N(0, \sigma^2)$$

Where i was the level 1 individual and j is the level 2 such as EA. The random intercept u_j measures EA effects which are assumed to be randomly distributed. α and β are fixed effects and u_j are random effects (Ke-Sheng Wang et al., 2017*).*

2.5 Estimation for GLMMs

The marginal Likelihood is the joint probability of all observed responses given the observed covariates. For linear mixed models, the marginal likelihood can be evaluated and maximized relatively easy.

2.5.1 Pseudo Maximum Likelihood Estimation Method (PML)

The PML estimates are obtained by maximizing the weighted log-likelihood. The sampling weights are incorporated into the likelihood for estimation of parameters and their standard errors (Rozi et al., 2017). The weighted pseudo likelihood has to be maximised with respect to the regression parameters. (Rozi et al., 2017). This approach was used in performing Multilevel weighted model using stata command "melogit" which uses reciprocals of selection probabilities. Only the EA level weights were used for the melogit model.

2.6 Sampling weights

Complex sampling results in sampled individuals having varying inclusion probabilities which are often characterized by disproportionate representation (Rozi et al., 2017; Rabe-Hesketh and Skrondal 2006). As a results sampling weights are necessary in the survey to adjust for different sample representation (West et al., 2018). For example let π_i for i=1, N be the probability that unit i in the population will be included in the sample.

The sampling weight for any sample design was defined as the reciprocal if π_i given by

$$w_i = \frac{1}{\pi_i}$$

2.7 Weighted Multivariate Model (WMM)

Six models were fitted, Model A is the weighted logistic regression model which does not control for BMI and uses the adult weights. Model B is the weighted logistic regression model that does control for BMI and uses the physical examination weights. Model A and B were fitted using “svy” stata command for both unadjusted and multivariable weighted logistic regression models.

2.8 Weighted Multilevel Model (WGLMM)

Model C and model D were fitted using the "melogit" command; model C did not control for BMI while model D did control for BMI. The EA level weights were used for both models. Melogit fits mixed-effects logistic models containing both fixed and random effects. The effect of clustering is summarised by the intra cluster correlation (ICC). The random effects are estimated using adaptive Gauss-Hermite quadrature. Melogit default integral method (mvaghermite) which is the mean–variance adaptive Gauss–Hermite quadrature uses log pseudolikelihood to estimate parameters.

Model E and F were fitted using the GLLAMM framework (Generalized Linear Latent And Mixed Models), which was developed by Rabe-Hesketh (2004). Scaling method 1 was used; this method scales the weights so that the sum of the scaled weights over all respondents in a cluster is equivalent to the cluster size (Rabe-Hesketh and Skrondal (2006); Rozi et al., 2017). Both individual level and EA level weights were scaled. For the model not controlling for BMI, both adult weight (model E) and EA weights were used and for the model controlling for BMI physical examination weights (model F) and EA weights were scaled (Rabe-Hesketh and Skrondal (2006)) using gllamm command in stata. The gllamm uses adaptive quadrature that uses the Newton-Raphson (NR) algorithm (Rabe-Hesketh *et al.*, 2005). When fitting this model 12 quadrature points were used.

For complex survey designs studies recommend scaling weights to reduce the variation in the magnitude of the weights and to minimise bias (Rabe-Hesketh and Skrondal 2006; Carle 2009; Pfefferman et al. 1998). Model E and F in which the weights were scaled can be compared to model C and D (melogit) in which weights were not scaled.

2.9 Objective three

To compare different models used to determine risk factors associated with diabetes for all adults (those aged 15 years and above) of South Africa SANHANES-1 2011-2012 survey. The above models were compared (model A-F).

2.9.1 Introduction to survey methods

The survey employs multistage sampling designs where primary sampling unit are clusters and are sampled at the first stage (Rabe-Hesketh and Skrondal (2005)). The multilevel models are often used to analyse complex surveys. This survey estimate using pseudolikelihood produced from incorporating sampling weights.

The generalized linear mixed models use weighted iterative quasilikelihood algorithm which is not expected to perform well since unweighted penalized quasi-likelihood often produces biased estimates. Two-level models study the effect of clustering. This model shows the importance of the group effects.

Gllamm estimates Generalized Linear Latent And Mixed Models. Complex sampling designs commonly incorporate unequal selection probabilities. Failing to incorporate for design may lead to biased parameter estimates. The multilevel model may lead to biased estimates if samples include unequal probability of selection. To rectify this problem, scaling methods has been suggested (Asparouhov, 2008). Rabe-Hesketh and Skrondal (2005)) suggested scaling of the weights if level one weights that are used vary between units in the cluster. Scaling methods normalize weights at each level. There are two scaling methods that are commonly used. Method 1 provides the least biased estimate for between cluster variance and method 2 provides the least biased estimates for estimating slopes, intercepts, and odds ratios Rabe-Hesketh and Skrondal (2005)). The gllamm used scaled EA and level 1 weights. Scaling method 1 determines apparent cluster size equals to effective sample size. Producing unbiased results. Few simulation studies (Carle, 2009; Roze al.,2017; Asparouhov, 2006) recommended scaling method 1 for estimating between cluster variance.

The study compared three different models namely survey logistic regression “svy”, multilevel mixed-effects logistic regression “melogit” and Generalized Linear Latent and Mixed Models “gllamm”. The motivation behind comparing three models was to confirm if scaling weights improves the results.

2.10 Ethical consideration

Approval to conduct the study was obtained from the Human Research ethics committee ((Medical) of the University of Witwatersrand. (Clearance Certificate number: M190323 (Appendix I)). Both datasets (SANHANES-1) were obtained from HSRC through data curation systems (http://curation.hsrc.ac.za/Datasets-XKAHAA.phtml).

CHAPTER 3
RESULTS

3.1 Introduction

This chapter describes the findings of the study. Baseline characteristics of the participants in the study are summarised, including socio-demographic characteristics, lifestyle factors and other factors believed to be related to diabetes. The characteristics of participants who have data with BMI were compared to those of participants who do not have data with BMI. Analyses carried out include the Rao-Scott adjustment to the chi-square test, weighted survey logistic regression model and weighted multilevel models.

3.2 Descriptive Analysis

Objective one: The prevalence of diabetes was given overall as well as by gender, province and race

3.2.1 Description of the study sample by diabetes status

Table 3.1 shows the associations between socio-demographic factors and diabetes, with a formal test carried out using the Rao-Scott adjustment to the Pearson chi-square test. In the table the percentages (prevalences) are weighted using the survey weights.

Table 3.1 (a) Description of the Socio-demographic factors by diabetes status

Factor	Level	No Diabetes n (%)	Diabetes n (%)	Rao-Scott Test Statistic/P-value
Diabetes		14071(94.7%)	998(5.3%)	
Gender	Male	5896(95.9)	331(4.1)	F=10.068 P < 0.0001
	Female	8102(93.7)	662(6.3)	
	Missing	73(94.8)	5(5.2)	
Race	African	9433(95.5)	531(4.5)	F= 15.274 P < 0.0001
	White	663(93.0)	47(7.0)	
	Coloured	2755(93.4)	227(6.6)	
	Indian	1097(83.4)	190(16.6)	
	Other	123(98.3)	3(1.7)	

In this study, participants were N= 15069, with 998 (5.3%) being diagnosed with diabetes and 14071 (94.7%) not being diagnosed with diabetes (Table 3.1). The prevalence of diabetes was estimated as 5.3% (95% confidence interval 5 %– 6%). The prevalence was higher in females (6.3%) than in males (4.1%).

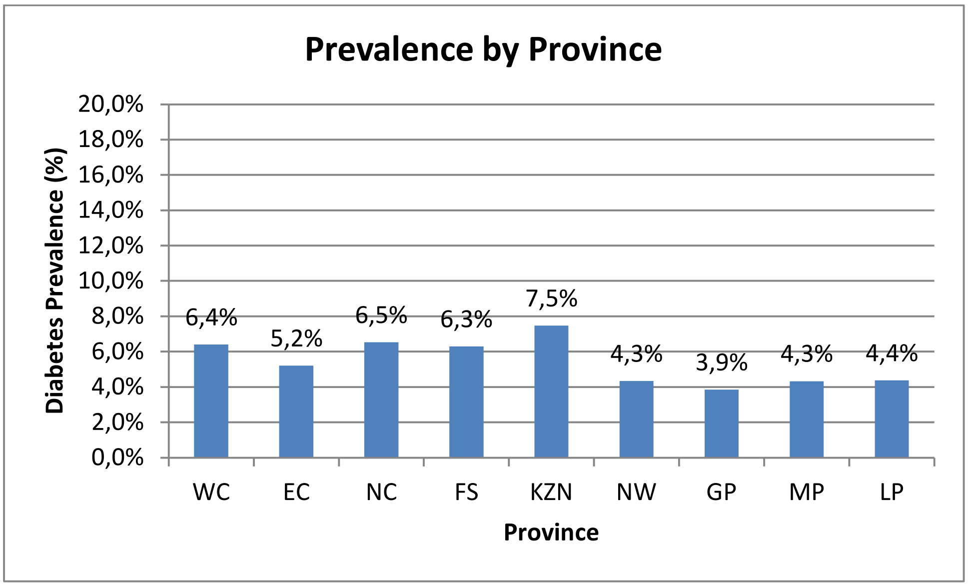
Table 3.1 (b) Description of the Socio-demographic factors by diabetes status

Factor	Level	No Diabetes n (%)	Diabetes n (%)	Rao-Scott Test Statistic/P-value
Province	Eastern Cape	1512(94.8)	103(5.2)	F=3.444
	Free State	776(94.0)	57(6.0)	P = 0.001
	Gauteng	2473(96.2)	119(3.8)	
	KwaZulu Natal	2241(96.2)	272(7.5)	
	Limpopo	1151(95.6)	70(4.4)	
	Mpumalanga	1224(95.7)	69(4.3)	
	North West	1789(95.7)	92(4.3)	
	Northern Cape	921(93.5)	69(6.5)	
	Western Cape	1994(93.6)	147(6.4)	
Employed	Yes	4758(95.7)	272(4.3)	F=13.059
	No	3887(96.3)	140(3.7)	P < 0.0001
	Other	5240(92.4)	575(7.6)	
	Missing	186(95.6)	11(4.4)	
Geotype	Urban Formal	7509(93.6)	682(6.4)	F=10.939
	Urban Informal	1811(97.1)	64(2.9)	P < 0.0001
	Rural Informal	3001(95.6)	174(4.4)	
	Rural Formal	1750(96.4)	78(3.6)	
Age category	15 to 24	4212(99.4)	29(0.6)	F=104.867
	25 to 34	2928(98.7)	37(1.3)	P < 0.0001
	35 to 44	2382(96.8)	104(3.2)	
	45 to 54	2011(88.3)	264(11.7)	
	55 to 64	1408(83.2)	315(16.8)	
	65+	1130(83.8)	249(16.2)	

Among the participants most were from Gauteng followed by KwaZulu Natal and Western Cape (Table 3.1 b). This corresponds with the Statistics population release that shows majority province is Gauteng with 25.3%, followed by KwaZulu Natal with 19.6% then Western Cape and Eastern Cape with 11.5% each. The prevalence of diabetes was found to be highest in

KwaZulu Natal with 7.5%, followed by Northern Cape with 6.5%, then Western Cape with 6.4% (Table 3.1(n); Figure 3.1).

Figure 3.1 Prevalence by province



The above figure shows prevalence of diabetes by province with KZN being the province with the highest prevalence at 7.5% and the lowest prevalence was found in GP with 3.9% as shown on (Table 3.1).

For race the prevalence was highest for Indians with 16.6% followed by white with 7%. For employment status the prevalence was highest for those that reported other (those who were sick disabled and unable to work, retired and those that specified their employment status as other) at 7.6 % followed by those that reported being employed at 4.3%. Prevalence was found to be highest among those living in urban formal areas (6.4%) and lowest among those living in rural informal areas (4.4%). Prevalence by age category showed age 55 to 64 was the highest with 16.8% with age 65+ at 16.2% (Table 3.1 (b)).

The prevalence of diabetes was much higher among people with hypertension (20.5%) than among those without hypertension (2.2%).

Table 3.2 Description of the Co-morbidity by diabetes status

Factor	Level	No Diabetes n (%)	Diabetes n (%)	Rao-Scott Test Statistic/P-value
Family History	Yes	2676(85.6)	596(14.4)	F= 156.1811 P < 0.0001
	No	10419(97.2)	344(2.8)	
	Other	976(95.8)	58(4.2)	
Hypertension	Yes	2290(79.5)	710(20.5)	F=402.29 P < 0.0001
	No	11489(97.8)	282(2.2)	
	Other	292(98.9)	6(1.1)	
TB Status	Yes	847(92.4)	81(7.6)	F=3.481 P = 0.032
	No	12827(94.9)	887(5.1)	
	Missing	397(95.4)	30(4.6)	

The prevalence of diabetes was higher for those with a family history of diabetes, for those with hypertension and those with TB Table 3.2).

Table 3.3 Description of the Lifestyle factors by diabetes status

Factor	Level	No Diabetes n (%)	Diabetes n (%)	Rao-Scott Test Statistic / P-value
Smoking status	No, not at all	10838(95.0)	749(5.0)	F= 7.6288 P = 0.0001
	Yes, but not now	295(84.2)	56(15.8)	
	Yes, daily	2593(95.2)	164(4.8)	
	Yes, less than daily	270(95.3)	16(4.7)	
	Other	75(92.0)	13(8.0)	
Alcohol intake	Never	10288(94.2)	813(5.8)	F= 3.6082 P = 0.009
	Monthly or less	1334(95.3)	69(4.7)	
	2-4 times a month	1222(95.2)	62(4.8)	
	2-3 times a week	683(97.8)	24(2.2)	
	4 or more times a week	249(97.7)	12(2.3)	

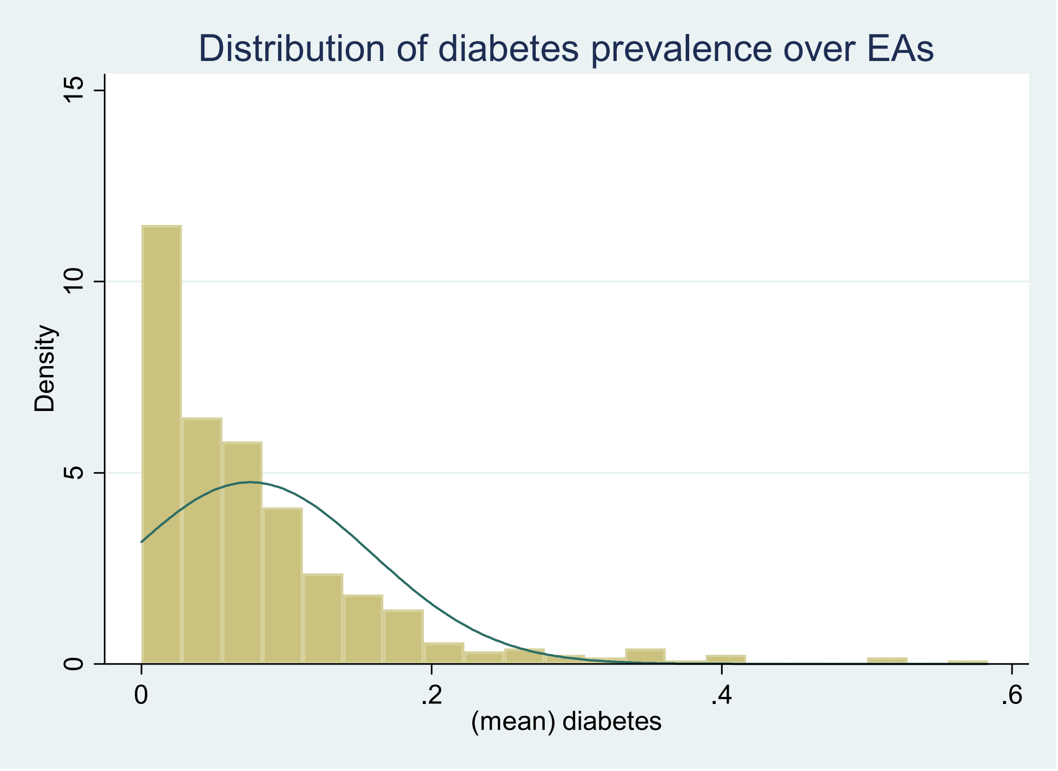
The prevalence of diabetes was highest for ex-smokers, but did not differ between current smokers and non-smokers. Diabetes was associated with alcohol intake, being highest among those who never drank alcohol and decreasing with increasing alcohol consumption (Table 3.3.

The results in tables 3.1 - 3.3 showed statistically significant associations (P<0.05) between the socio-demographic factors (age, race, geotype, gender and employment status) and diabetes. Also there were statistically significant associations (P<0.05) between the co-morbidity factors (family history of diabetes, hypertension and TB status) and diabetes. Lastly there were statistically significant associations between the lifestyle factors (smoking status and alcohol intake) and diabetes. These associations were further investigated by fitting univariable weighted logistic regression models and the factors found to be significant (using p-value <0.2) were included as potential exposure factors.

3.2.2 Variability of diabetes status across the EAs

Figure 3.2 shows distribution of diabetes status across the EAs. There are 458 EAs in the questionnaire data.

Figure 3.2 Distribution of diabetes prevalence over EAs



3.2.3 Status of data with BMI by potential explanatory factors

The purpose of the table below was to compare the characteristics of participants who have data with BMI to those of participants who do not have data with BMI. The formal test was carried out using the Rao-Scott adjustment to the Pearson chi-square test.

Table 3.4 (a) Status of data with BMI by potential explanatory factors

Factor	Level	Data with BMI present n(%)	Full data
Diabetes	Diabetes (Yes)	491(6.7)	998(5.3)
	Diabetes (No)	5906(93.3)	14071(94.7)
Gender	Male	2235(43.6)	6227(45.5)
	Female	4133(56.2)	8764(54.1)
	Missing	31(0.2)	78(0.5)
Race	African	4516(77.5)	9964(77.3)
	White	129(9.1)	710(10.1)
	Coloured	1394(9.9)	2982(9.1)
	Indian	302(2.7)	1287(2.7)
	Other	56(0.8)	126(0.8)
Employed	Yes	1584(30.5)	5030(35.8)
	No	1946(32.9)	4027(27.9)
	Other	2776(36.5)	5815(35.0)
	Missing	91(1.1)	197(1.2)
Geotype	Urban Formal	3101(52.8)	8191(54.4)
	Urban Informal	788(10.3)	1875(9.6)
	Rural Informal	1607(27.9)	3175(27.9)
	Rural Formal	901(9.1)	1828(8.2)

Data with BMI were obtained for 6410 (39.8%) participants and were missing for 8659 (60.2%). Amongst those with diabetes, prevalence of diabetes was found at 6.7% for data with BMI obtained compared to 5.3% amongst those with full data.

Amongst gender, the data with BMI had 56.2% of females compared to 54.1% females amongst those with full data. Amongst those that was employed, 30.5% was employed for data with BMI compared to 35.8% amongst those with full data.

Table 3.4 (b) Status of data with BMI by potential explanatory factors

Factor	Level	Data with BMI present n(%)	Full data
Age category	15 to 24	1156(27.4)	4241(27.6)
	25 to 34	1083(24.6)	2965(24.1)
	35 to 44	975(19.7)	2486(19.2)
	45 to 54	1017(13.3)	2275(13.3)
	55 to 64	848(8.5)	1723(8.6)
	65+	718(6.5)	1379(7.2)
Family History	Yes	1430(21.7)	3272(20.6)
	No	4546(71.4)	10763(72.4)
	Other	421(7.0)	1034(7.0)
Hypertension	Yes	1516(17.7)	3000(16.8)
	No	4734(80.1)	11771(81.3)
	Other	147(2.1)	298(2.0)
TB Status	Yes	524(7.3)	928(6.1)
	No	5716(89.7)	13714(90.9)
	Missing	157(3.0)	427(3.0)
Smoke status	No, not at all	4844(76.4)	11587(78.8)
	Yes, but not now	156(2.8)	351(2.5)
	Yes, daily	1227(18.2)	2757(16.2)
	Yes, less than daily	128(1.9)	286(2.0)
	Other	42(0.7)	88(0.6)
Alcohol intake	2 or more times a week	386(9.3)	967(8.2)
	2-4 times a month	516(9.3)	1284(9.0)
	Monthly or less	565(8.7)	1403(9.8)
	Never	4805(72.7)	11101(72.9)

Amongst those with hypertension, 17.7% of participant in the data with BMI had hypertension compared to 16.8% amongst those with full data. Amongst those with family history of diabetes, 21.7% participant from data with BMI had family history of diabetes compared to 20.6% among those with full data. Amongst those that smoke status 18.2% of participants in the data with BMI smoke daily compared to 16.2% among those with full data. Amongst those who drank alcohol 2 or more times a week 9.3% had data with BMI compared to 8.2% among amongst those with full data.

Table 3.5 Mean BMI and age by diabetes status

Factor	Level	No Diabetes	Diabetes	Adjusted wald Test Statistic / P-value
BMI	Mean (95% C.I.)	26.47(26.06-26.89)	30.82(29.78-31.85)	F= 57.81 p<0.0001
	Coefficient	Ref	4.34(3.22;5.47)	
Age	Mean (95% C.I.)	35.88(34.97-36.75)	51.75(49.05-54.45)	F= 125.04 p<0.0001
	Coefficient	Ref	15.89(13.09;19.68)	

A simple survey linear regression model was performed to compare the mean BMI between participants with and without diabetes. The mean BMI was significantly higher for those with diabetes (30.82 kg/m^2) than for those without (26.47 kg/m^2) (P<0.0001). The coefficient shows the average of BMI was about 4.34 kg/m^2 higher among those with diabetes. This difference was statistically significant.

The mean age was significantly higher for those with diabetes (51.75 years) than for those without diabetes (35.88 years) (P<0.0001). The coefficient shows the average age was about 15.89 years higher among those with diabetes. This difference was statistically significant.

3.3 Factors associated with diabetes

3.3.1 Weighted survey logistic regression model

The association between diabetes and explanatory variables was investigated firstly by using the Rao-Scott adjustment to the chi-square test (see the previous sections) and then by fitting unadjusted and multivariable weighted logistic regression models without controlling for BMI (Model A; Table 3.6). The first model that did not control for BMI was fitted since a large proportion of respondents (57%) do not have data with BMI. All variables were significant hence included in the model and table splitting was for reporting purpose (Model A).

The weighted survey logistic regression model allowed for clustering at EA level and stratification at the provincial level. Participant's individual weights based on the full sample (level 1 weights) were used in the model. Standard errors were estimated using Taylor linearization.

Table 3.6 (a) Results of fitting logistic regression models for the association between diabetes and socio-demographic factors

		Crude OR		Adjusted OR	
Variable	Level	OR(95% CI)	p-value	OR(95% CI)	p-value
Gender	Males	Ref	p=0.001	Ref	p=0.768
	Females	1.55(1.23;1.95)		1.12(0.82,1.53)	
	Missing	1.26(0.42;3.77)		1.11(0.28;4.27)	
Race	African	Ref	p<0.0001	Ref	p<0.0001
	White	1.57(0.96;2.57)		1.28(0.68;2.41)	
	Coloured	1.49(1.20;1.84)		1.06(0.79;1.41)	
	Indian	4.19(2.92;6.02)		2.57(1.75;3.77)	
	Other	0.37(0.10;1.39)		0.43(0.11;1.63)	
Employed	No	Ref	p<0.0001	Ref	p=0.568
	Yes	1.18(0.81;1.71)		0.79(0.52;1.21)	
	Other	2.15(1.52;3.04)		0.74(0.46;1.18)	
	Missing	1.21(0.57;2.56)		1.06(0.42;2.64)	
Geotype	Urban formal	Ref	p<0.0001	Ref	p=0.205
	Urban informal	0.45(0.31;0.63)		0.74(0.50;1.09)	
	Rural informal (tribal)	0.68(0.53;0.88)		0.80(0.61;1.05)	
	Rural formal (farms)	0.54(0.38;0.78)		0.72(0.47;1.11)	
Age		1.06(1.06;1.07)	p<0.0001	1.05(1.04;1.05)	p<0.0001

[1]

In the unadjusted analysis all explanatory factors considered were statistically significantly associated with diabetes. In the adjusted analysis the results showed that age and race were significantly associated with diabetes.

Race

In the adjusted analysis race was significantly associated with diabetes, with Whites, Coloureds and Indians having a higher risk than African. In particular the odds of Indians

[1] All exposure variables are included in the multivariable analysis

having diabetes was 2.57 times higher than the odds of Africans (aOR=2.57, P<0.0001, 95% CI 1.74 ; 3.77).

Age

The odds of having diabetes increased as age increased, after adjusting for other factors. The adjusted OR of 1.05 (P<0.0001 , 95% CI 1.04 ; 1.05) shows that the odds of diabetes increase by 5% for every one year increase in age. To get a better idea of the effect of age, for a 5 year increase in age, the OR is 1.26 (95% CI 1.21 ; 1.30) so the odds of diabetes increase by 26% for a 5 year increase in age.

Table 3.6 (b) Results of fitting logistic regression models for the association between diabetes and co-morbidity factors

		Crude OR		Adjusted OR	
Variable	Level	OR(95% CI)	p-value	OR(95% CI)	p-value
TB status	No	Ref	p=0.025	Ref	p=0.193
	Yes	1.52(1.12;2.06)		1.40(0.96;2.02)	
	Missing	0.88(0.49;1.57)		0.94(0.49;1.79)	
Family History	No	Ref	p<0.0001	Ref	p<0.0001
	Yes	5.93(4.56;7.71)		4.81(3.64;6.36)	
	Other	1.56(1.07;2.27)		1.53(1.00;2.34)	
Hypertension	No	Ref	p<0.0001	Ref	p<0.0001
	Yes	11.20(8.70;14.41)		4.99(3.76;6,62)	
	Other	0.51(0.19;1.34)		0.44(0.16;1.19)	

In the unadjusted analysis, all explanatory factors considered were statistically significantly associated with diabetes. In the adjusted analysis, the results showed that hypertension and family history were significantly associated with diabetes. TB status was not significantly associated with diabetes (Table 3.6 (b)).

Family history of diabetes

In the adjusted analysis, the odds of diabetes among those that had a family history were 4.81 times higher than those who did not have a family history of diabetes (aOR=4.81, P<0.0001, 95% CI 3.64 ; 6.36).

Hypertension

In the adjusted analysis, the odds of diabetes among those that had hypertension were 4.99 times higher than those who did not have hypertension (aOR=4.99, P<0.0001, 95% CI 3.76; 6.62).

Table 3.6 (c) Results of fitting logistic regression models for the association between diabetes and lifestyle factors

		Crude OR		Adjusted OR	
Variable	Level	OR(95% CI)	p-value	OR(95% CI)	p-value
Smoke status	Yes, daily	0.96(0.74;1.25)		0.95(0.66;1.36)	
	Yes, less than daily	0.92(0.47;1.78)		0.89(0.43;1.85)	
	No, not at all	Ref	p=0.013	Ref	p=0.340
	Yes, but	3.55(1.68;7.49)		2.31(0.73;7.27)	
	Other	1.64(0.81;3.30)		0.58(0.17;2.02)	
Alcohol intake	Never	Ref	p=0.001	Ref	p=0.003
	Monthly	0.79(0.49;1.27)		0.94(0.56;1.56)	
	2-4 times a month	0.81(0.55;1.19)		0.97(0.62;1.53)	
	2 -3 times a week	0.37(0.21;0.65)		0.31(0.17;0.56)	
	4 or more times a week	0.38(0.18;0.83)		0.24(0.08;0.70)	

In the unadjusted analysis all explanatory factors considered were statistically significantly associated with diabetes. In the adjusted analysis, the results showed that alcohol intake was significantly associated with diabetes, while smoking status was not significantly associated with diabetes (Table 3.6 (c)).

Alcohol intake

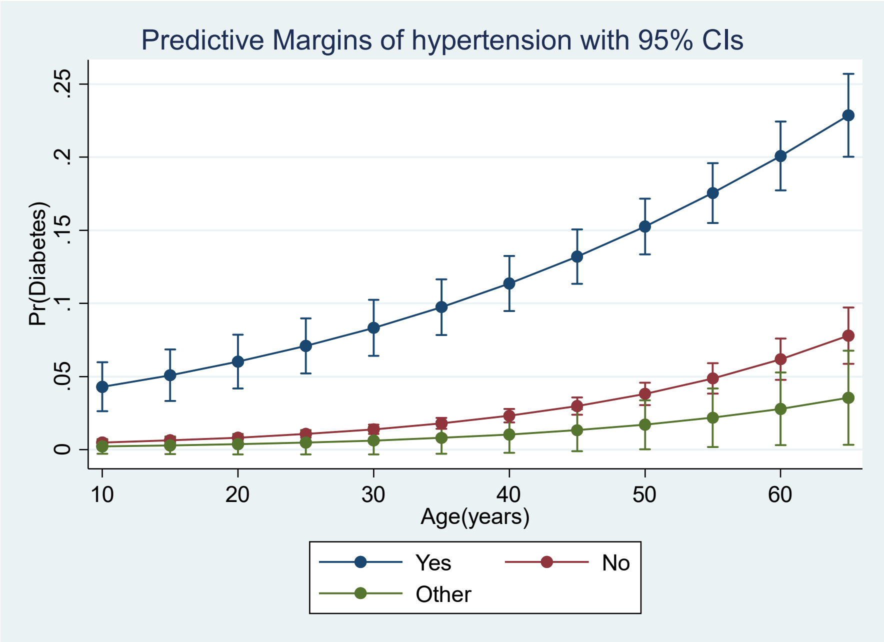
In the adjusted analysis, drinking alcohol was protective against diabetes, with those who drank 2 -3 more times a week having 69% lower odds of having diabetes compared to those who never drank alcohol.

3.3.1.2 Investigation of interactions in the model

All possible two-way interactions involving terms that were significant in the adjusted model were investigated. Each interaction was assessed individually. The interactions were assessed between age and hypertension, hypertension and family history, race and hypertension, age and alcohol intake, race and alcohol intake, race and age, hypertension and alcohol intake and lastly race and family history, alcohol intake and family history, age and family history. Each interaction was assessed individually.

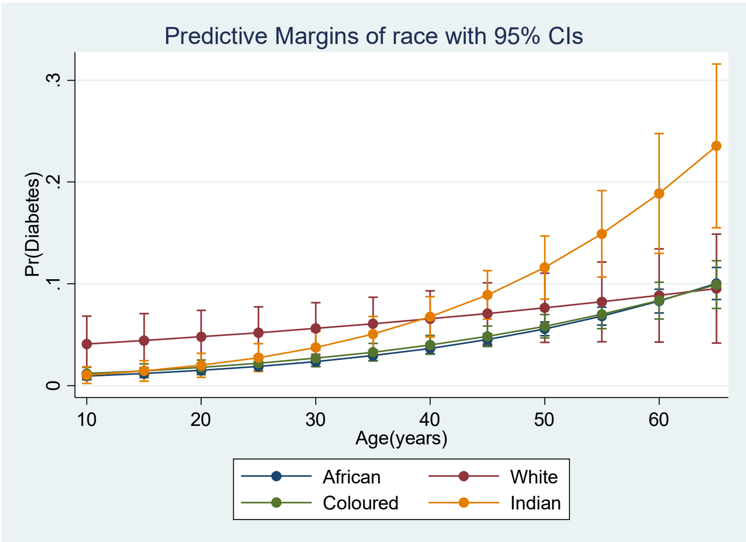
There was a significant interaction between hypertension and age. The effect of age on the odds of diabetes is more marked for those with hypertension than for those without hypertension (Figure 3.3).

Figure 3.3. Margin plot hypertension and age interaction



There was a significant interaction between race and age. The plot of the predictive margins for diabetes by race shows that the probability of diabetes increases with age more rapidly for Indians than it does for the other race groups.

Figure 3.4. Margin plot race and age



The interaction between geotype and alcohol intake was found to be significant. Also, the interaction between geotype and gender was found to be significant.

3.3.2 Weighted survey logistic regression model including BMI

The weighted survey logistic regression model was then fitted including BMI as a potential exposure (Model B). To help interpret the effect of BMI, this was measured as the OR for an increase of 5 kg/m^2 in BMI. The sampling weights for data with BMI were used for this analysis.

Neither TB status nor employment status were found to be significant in the univariable model using the adjusted Wald test, and hence were not included in the multivariable model which included BMI.

Table 3.7 (a) Weighted survey logistic regression model including BMI

		Adjusted OR	
Variable	Level	OR(95% CI)	p-value
BMI		1.05(0.92;1.20)	p=0.469
Gender	Males	Ref	p=0.537
	Females	1.31(0.55;3.13)	
	Missing	1	
Race	African	Ref	p=0.028
	White	2.83(0.59;13.55)	
	Coloured	0.91(0.53;1.56)	
	Indian	2.16(1.19;3.93)	
	Other	0.37(0.06;2.10)	
Geotype	Urban formal	Ref	p=0.084
	Urban informal	0.55(0.29;1.03)	
	Rural informal (tribal)	0.60(0.39;0.93)	
	Rural formal (farms)	0.53(0.29;0.99)	
Age		1.04(1.03;1.05)	P<0.0001

[2]

3.3.2.1 Factors associated with diabetes

In the adjusted analysis the overall Wald test showed age and race were significantly associated with diabetes. Gender, BMI and geotype were not significantly associated with diabetes.

[2] TB status and employment status were not included in the multivariable model which included BMI.

Race

In the adjusted analysis race was significantly associated with diabetes, with Whites, Coloureds and Indians having a higher risk than Africans of having diabetes; in particular the odds of Indians having diabetes was 2.17 times higher than the odds of Africans (aOR = 2.17, p=0.028, 95% CI 1.19;3.93).

Age

The odds of having diabetes increased as age increased, after adjusting for other factors. The adjusted OR of 1.04 (p<0.0001, 95% CI 1.03; 1.05) shows that the odds of diabetes increase by 4% for every one year increase in age. To get a better idea of the effect of age, for a 5 year increase in age, the OR is 1.20 (95% CI 1.14 ; 1.26) so the odds of diabetes increase by 20% for a 5 year increase in age.

Table 3.7 (b) Weighted survey logistic regression model including BMI

		Adjusted OR	
Variable	Level	OR(95% CI)	p-value
Family History	No	Ref	P=0.003
	Yes	3.37(1.58;7.20)	
	Other	1.39(0.65;2.97)	
Hypertension	No	Ref	p<0.0001
	Yes	5.09(3.03;8.53)	
	Other	0.88(0.21;3.74)	

3.3.2.2 Factors associated with diabetes

In the adjusted analysis the overall Wald test showed hypertension and family history were significantly associated with diabetes.

Family history of diabetes

In the adjusted analysis the odds of diabetes among those with family history of diabetes were 3.37 times higher than those who did not have family history of diabetes (aOR = 3.37, p<0.0001, 95% CI 1.58; 7.20).

Hypertension

In the adjusted analysis the odds of diabetes among those that had hypertension were 5.09 times higher than those who did not have hypertension (aOR = 5.09, p<0.0001, 95% CI 3.03, 8.53).

Table 3.7 (c) Weighted survey logistic regression model including BMI

		Adjusted OR	
Variable	Level	OR(95% CI)	p-value
BMI		1.05(0.92;1.20)	p=0.469
Smoke status	Yes, daily	1.04(0.58;1.89)	
	Yes, less than daily	1.49(0.66;3.36)	
	No, not at all	Ref	p=0.107
	Yes, but	1.92(0.89;4.14)	
	Other	0.12(0.02;0.90)	
Alcohol intake	Never	Ref	p=0.004
	Monthly	1.19(0.72;1.99)	
	2-4 times a month	0.92(0.41;2.10)	
	2- 3 more times a week	0.17(0.06;0.47)	
	4 or more times a week	0.14(0.03;0.71)	

3.3.2.3 Factors associated with diabetes

In the adjusted analysis the overall Wald test showed alcohol intake was significantly associated with diabetes. Smoking status was not significantly associated with diabetes.

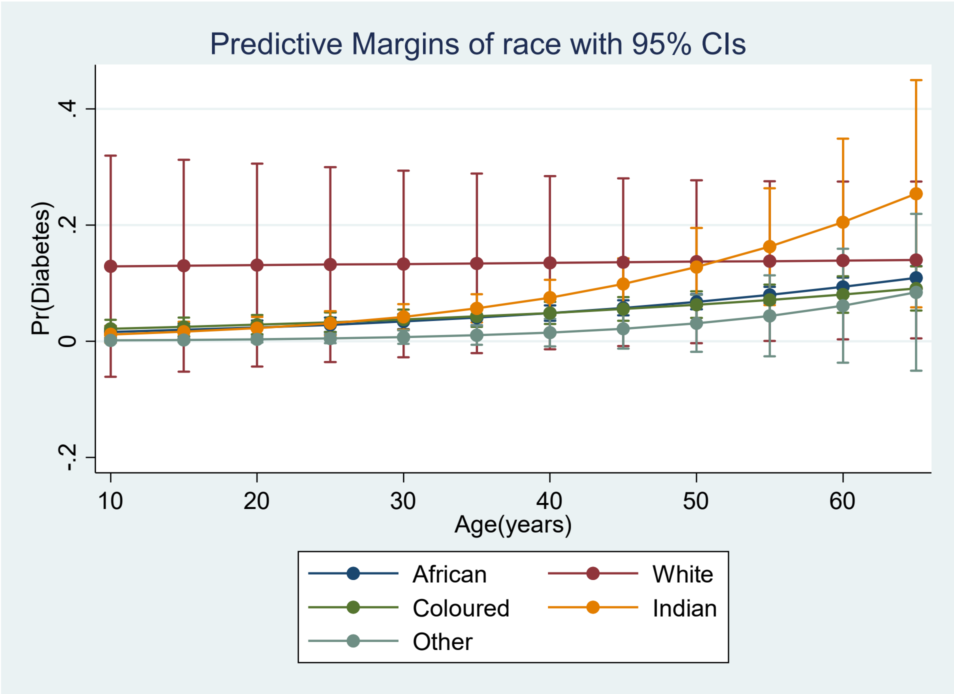
3.3.2.4 Investigation of interactions in the model

All possible two-way interactions involving terms that were significant in the adjusted model were investigated. Each interaction was assessed individually. The interaction was assessed between age and hypertension, hypertension and family history, race and hypertension, age and alcohol intake, race and alcohol intake, race and age, hypertension and alcohol intake and lastly race and family history. Furthermore, interaction between BMI and all significant risk factors were assessed.

The interaction between geotype and alcohol intake was found to be significant. Also, the interaction between geotype and smoke status was found to be significant.

The effect of age was more marked for Indians, while for whites the overall prevalence of diabetes is high, but it does not seem to depend on age (Figure 3.5).

Figure 3.5 Margin plot for race and age interaction



There was a significant interaction between hypertension and the amount of alcohol drunk. For those without hypertension there is little effect of alcohol and the prevalence is generally very low, while for those with hypertension the prevalence is high among those who do not drink alcohol and those who drink up to 4 times a month, while it is low for those who drink twice a week or more as shown in table below.

		Alcohol				
HBP		Never	Monthly or less	2-4 times a month	2-3 times a week	4 or more times a week
	Yes	0.14	0.20	0.28	0.06	0.05
	No	0.0	0.04	0.01	0.001	0.001

There was a significant interaction between race and alcohol intake. For Indians (unlike the other races) the prevalence is highest among those who drink alcohol 4 or more times a week, while Whites prevalence is highest among those who never drank as shown in table below.

		Alcohol				
Race		Never	Monthly or less	2-4 times a month	2-3 times a week	4 or more times a week
	African	0.06	0.06	0.10	0.03	0.02
	White	0.17	0.15	-	-	0.02
	Coloured	0.05	0.10	0.04	0.01	0.10
	Indian	0.11	0.07	-	0.15	0.16

There was a significant interaction between alcohol consumption and BMI. Figure 3.6 shows marginal probability plots of the probability of diabetes versus BMI by alcohol consumption category. Whereas the probability of diabetes increases with increasing BMI among participants who do not consume alcohol, the relationship between diabetes and increasing BMI is much less clear, with the curves being flat in the case of those who drink 2 – 3 times a week.

Figure 3.6 Margin plot for alcohol intake and BMI interaction

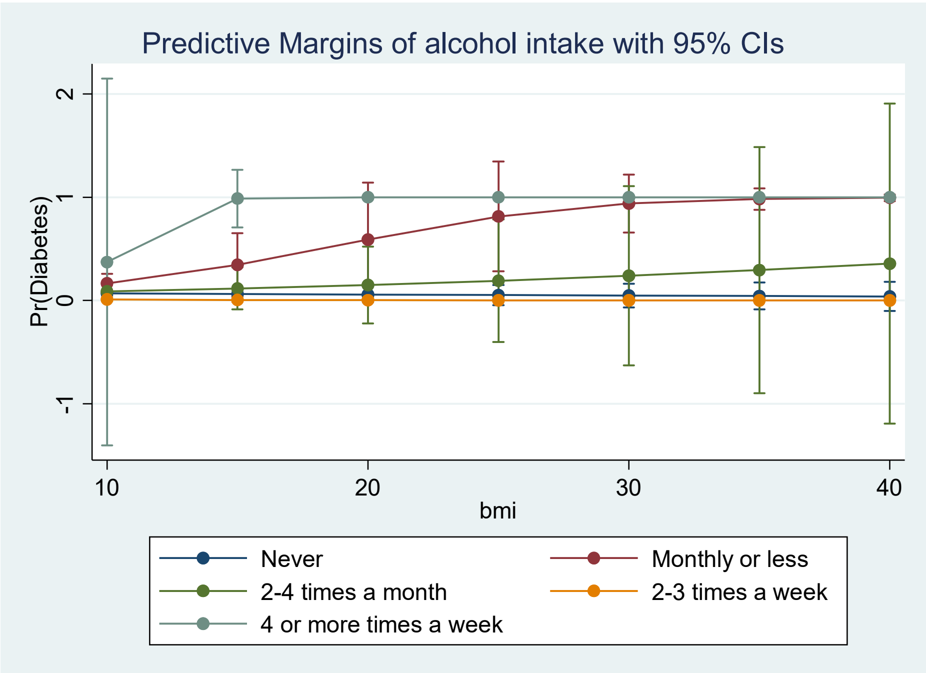
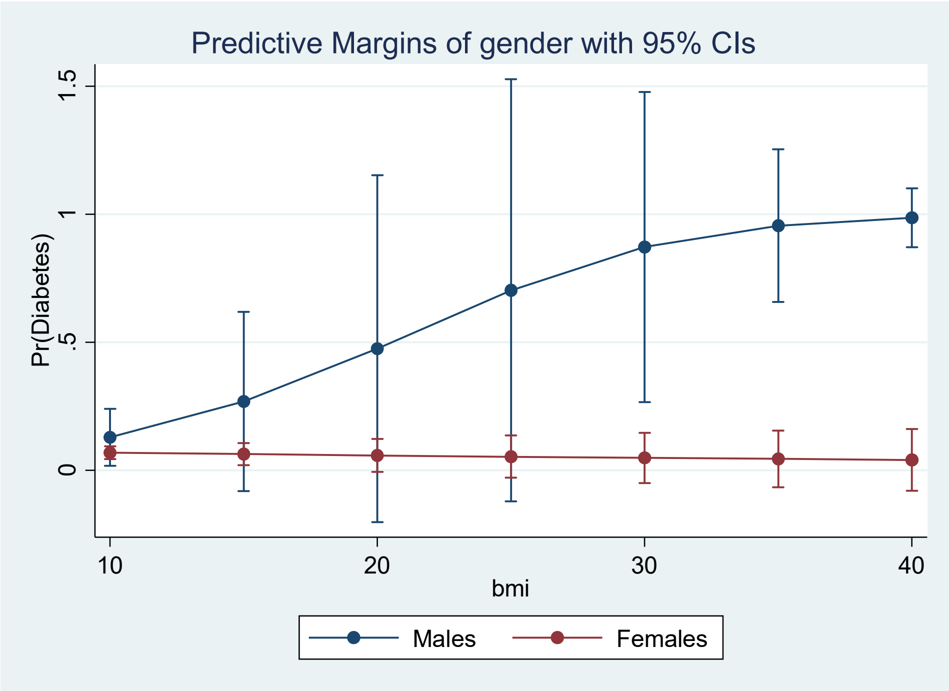


Figure 3.6 shows those drink 4 or more times a week have higher risk as their BMI increases.

The interaction between gender and BMI was found to be significant. For males, the risk of diabetes increases with increasing BMI while for females the risk of diabetes seems to be constant (Figure 3.7).

Figure 3.7: Margin plot by BMI and gender interaction



3.3.2.5 Comparison of two Multivariate models

The weighted logistic regression model that excluded BMI was compared to the weighted logistic regression model that controlled for BMI. In models for the subset of participants who had BMI measured, employment status and TB status were found not to be significant in unadjusted analyses, hence they were not included in the multivariable model.

In the model that included all participants, including those for whom BMI measurements were not available, the odds ratio for diabetes for those with a family history of diabetes was 4.66, while for the analysis including only those for whom BMI was measured, and adjusting for BMI the odds ratio decreased to 3.27 (Table 3.6(b) and Table 3.7(b)).

In the model without BMI the odds ratio for diabetes among those who took alcohol 2-3 times a week compared to those who did not take any alcohol was 0.31, while for the model including BMI it becomes 0.17 (Table 3.6 and Table 3.7). Although BMI itself was not statistically significant, it is an important confounder that impacts on the estimates for a number of variables, so should be included in the final model.

Adding the effect of BMI reduced odds ratio of having diabetes for race (Indian), family history and alcohol intake (2-3 times a week). The risk for diabetes as age increases stayed the same for both models. Gender was found not to be significant for both models.

The interaction between hypertension and age was found to be significant in the model without BMI and not significant in the model with BMI. The interaction between hypertension and race was found to be significant for both model without and the model with BMI.

The interaction between geotype and alcohol intake was found to be significant in both models. Also, the interaction between geotype and smoke status was found to be significant in the model with BMI and not significant in the model without BMI. The interaction between geotype and gender was found be significant in the model without BMI and not significant in the model with BMI.

3.4 Weighted multilevel model

The adjusted multilevel models were fitted using the "melogit" command in Stata. Separate models were fitted to all participants and then those participants with data with BMI, and these models were then compared. These multilevel models were fitted using the EA level weights (level 2 weights) with these models labelled as model C and D (Table 3.10(a)).

Table 3.8 Multilevel structure of the 2010-11 SANHANES-1 data

Level 2	Questionnaire	BMI
	EAs	
	458	413
Level 1	Individual	
	14673	6233

Table above (3.8) shows the multilevel/ Hierarchal structure. At level one there were the individual adults, who were nested within the level 2 enumeration areas.

Table 3.9 Percentage of missing multilevel model

Variable	% Missing Values BMI model	% Missing Values questionnaire model
Gender	0.5%	0.5%
Alcohol intake	2.0%	2.1%
BMI		58.0%
Tuberculosis	2.4%	2.8%
Employed	1.4%	1.3%

3.4.1 Weighted multilevel model adult questionnaire (model C)

Table 3.10 (a) Multilevel analysis for the association between diabetes and socio-demographic factors

		Univariate		Multivariate	
Variable	Level	OR(95% CI)	pvalue	OR(95% CI)	pvalue
Fixed effect					
Gender	Males	Ref	p=0.0002	Ref	p=0.778
	Females	0.56(0.26;1.93)		0.97(0.75;1.26)	
Race	African	Ref	p<0.0001	Ref	p<0.0001
	White	1.19(0.78;1.81)		0.83(0.47;1.47)	
	Coloured	1.40(1.08;1.83)		1.00(0.75;1.35)	
	Indian	3.37(2.53;4.48)		2.04(1.45;2.87)	
	Other	0.25(0.06;1.18)		0.36(0.08;1.63)	
Employed	No	Ref	p<0.0001	Ref	p=0.855
	Yes	1.61(1.14;2.28)		1.08(0.76;1.54)	
	Other	2.70(1.92;3.78)		1.04(0.74;1.48)	
	Missing	1.98(0.92;4.26)		1.48(0.70;3.77)	
Geotype	Urban formal	Ref	p<0.0001	Ref	p=0.112
	Urban informal	0.48(0.34;0.67)		0.76(0.54;1.09)	
	Rural informal (tribal)	0.70(0.54;0.89)		0.74(0.56;1.00)	
	Rural formal (farms)	0.56(0.41;0.76)		0.72(0.51;1.02)	
Age		1.06(1.05;1.07)	p<0.0001	1.04(1.03;1.05)	p<0.0001
Random effects					
EA				0.10(0.01;1.20)	0.13

[3]

Fixed effect

The results of fitting multilevel models are summarized in Table 3.10 (a). In adjusted analysis age, family history of diabetes and race were found to be significantly associated with diabetes. Geotype, gender and employment status were not found to be significantly associated with diabetes

Race

In the adjusted analyses race was significantly associated with diabetes (P<0.0001). In particular, those of Indian race were at a higher risk of diabetes compared to Africans, with an adjusted OR of 2.04 (95% CI 1.04 ; 2.87).

[3] All exposure variables are included in the multivariable analysis

Age

The odds of having diabetes increased as age increased, after adjusting for other factors. The adjusted OR of 1.04 (P<0.0001 , 95% CI 1.04; 1.05) shows that the odds of diabetes increase by 4% for every one year increase in age.

To get a better idea of the effect of age, for a 5 year increase in age, the odds of diabetes increase by 22% for a 5 year increase in age OR is 1.22 (95% CI 1.18 ; 1.26).

Table 3.10 (b) Multilevel analysis for the association between diabetes and co-morbidity factors

		Univariate		Multivariate	
Variable	Level	OR(95% CI)	pvalue	OR(95% CI)	Pvalue
Fixed effect					
TB status	No	Ref	p=0.072	Ref	p=0.703
	Yes	1.48(1.04;2.10)		1.36(0.89;2.06)	
	Missing	0.82(0.48;1.41)		0.78(0.41;1.47)	
Family History	No	Ref	p<0.0001	Ref	p<0.0001
	Yes	6.05(4.75;7.71)		5.00(3.81;6.57)	
	Other	1.62(1.08;2.45)		1.74(1.09;2.78)	
Hypertension	No	Ref	p<0.0001	Ref	p<0.0001
	Yes	12.98(10.17;16.55)		5.59(4.12;7.59)	
	Other	0.59(0.19;1.85)		0.44(0.13;1.47)	
Random effects					
EA	0.07(0.01;0.38)	0.06		0.10(0.01;1.20)	0.13

In the adjusted analysis, the results showed, family history and hypertension were significantly associated with diabetes (Table 3.10 (b)). TB status was not significantly associated with diabetes.

Family history of diabetes

In both analyses, there was a significant association between family history and diabetes. The odds of diabetes among those that had a family history 5.00 higher compared to those who did not have a family history of diabetes (aOR=5.00, P<0.0001, 95% CI 3.81; 6.57).

Hypertension

In both analyses there was a significant association between hypertension and diabetes. The odds of diabetes among those that had hypertension were 5.59 times higher than those who did not have hypertension (aOR=5.59, P<0.0001, 95% CI 4.12; 7.59).

Table 3.10 (c) Multilevel analysis for the association between diabetes and lifestyle factors

		Univariate		Multivariate	
Variable	Level	OR(95% CI)	pvalue	OR(95% CI)	pvalue
Fixed effect					
Smoke status	Yes, daily	0.81(0.64;1.04)		0.84(0.60;1.18)	
	Yes, less than	0.80(0.33;1.91)		0.81(0.35;1.86)	
	No, not at all	Ref	p=0.0003	Ref	P=0.594
	Yes, but	2.94(1.77;4.89)		1.53(0.76;3.09)	
	Other	2.13(1.02;4.47)		0.92(0.30;2.86)	
Alcohol intake	Never	Ref	p=0.040	Ref	P=0.091
	Monthly	0.69(0.47;1.02)		0.93(0.63;1.37)	
	2-4 times a month	0.72(0.45;1.15)		0.95(0.57;1.57)	
	2-3 times a week	0.43(0.22;0.85)		0.53(0.19;1.47)	
	4 or more times a week	0.45(0.18;1.11)		0.36(0.12;1.10)	
Random effects					
EA	0.07(0.01;0.38)	0.06		0.10(0.01;1.20)	0.13

In the adjusted analysis, the results showed, smoking status and alcohol intake, were found to not be significantly associated with diabetes (Table 3.10 (c)).

3.4.1.1 Investigation of interactions in the model

For the adult questionnaire model (Model C)

There was evidence that being hypertensive and drinking alcohol increase the odds of being diabetic. Among hypertensive, those who drank 4 or more times a week were at a reduced risk, while for those who were not hypertensive there was very little relationship between alcohol consumption and the prevalence of diabetes.

		Alcohol				
HBP		Never	Monthly or less	2-4 times a month	2-3 times a week	4 or more times a week
	Yes	0.11	0.12	0.13	0.09	0.06
	No	0.03	0.02	0.02	0.002	0.02

The predicted prevalence of diabetes for those that have hypertension and drink alcohol monthly or less is 12%, whilst among those who drink alcohol 2-4 times a month it is 13%.

Among Indians, those who drank 4 or more times a week were at a higher risk, while for those who were Africans there was a little relationship between alcohol consumption and the prevalence of diabetes from table below.

		Alcohol				
Race		Never	Monthly or less	2-4 times a month	2-3 times a week	4 times or more a week
	African	0.06	0.05	0.05	0.03	0.02
	White	0.05	0.05	0.09	0.002	0.01
	Coloured	0.05	0.06	0.06	0.03	0.06
	Indian	0.09	0.03	0.09	0.10	0.14

3.4.2 Weighted multilevel model including BMI (model D)

Table 3.11 (a) Multilevel analysis for the association between diabetes and socio-demographic factors

		Multivariate	
Variable	Level	OR(95% CI)	pvalue
Fixed effect BMI		1.19(1.09;1.29)	P<0.0001
Gender	Males	Ref	p=0.012
	Females	0.71(0.49,1.04)	
Race	African	Ref	p=0.053
	White	0.75(0.33;1.68)	
	Coloured	0.82(0.54;1.26)	
	Indian	1.62(0.99;3.19)	
	Other	0.52(0.02;1.51)	
Employed	No	Ref	p=0.343
	Yes	1.16(0.72;1.88)	
	Other	1.26(0.78;2.04)	
	Missing	2.68(0.91;7.92)	
Geotype	Urban formal	Ref	p=0.127
	Urban informal	0.70(0.43;1.13)	
	Rural informal (tribal)	0.66(0.44;0.99)	
	Rural formal (farms)	0.69(0.47;1.02)	
Age		1.03(1.02;1.04)	p<0.0001
Random effects			
EA		0.10(0.01;1.20)	0.13

[4]

Fixed effect

The results of fitting multilevel models are summarized in Table 3.11 (a). In adjusted analysis age, family history of diabetes, BMI, and gender were found to be significantly associated with diabetes. Geotype, race and employment status were not found to be significantly associated with diabetes.

BMI

In the adjusted analysis the odds of having diabetes increased as BMI increased, after adjusting for other factors. The adjusted OR of 1.19 (P<0.0001 , 95% CI 1.09; 1.29) shows that the odds of diabetes increase by 19% per 5kg/m^2 increase in BMI.

[4] All exposure variables are included in the analysis

Age

The odds of having diabetes increased as age increased, after adjusting for other factors. The adjusted OR of 1.03 (P<0.0001 , 95% CI 1.02; 1.02) shows that the odds of diabetes increase by 3% for every one year increase in age.

To get a better idea of the effect of age, for a 5 year increase in age, the OR is 1.18 (95% CI 1.12 ; 1.24) so the odds of diabetes increase by 18% for a 5 year increase in age.

Table 3.11 (b) Multilevel analysis for the association between diabetes and co-morbidity factors

		Multivariate	
Variable	Level	OR(95% CI)	Pvalue
Fixed effect			
TB status	No	Ref	p=0.703
	Yes	1.20(0.70;2.06)	
	Missing	0.78(0.30;2.05)	
Family History	No	Ref	p<0.0001
	Yes	4.73(3.30;6.78)	
	Other	2.12(1.18;3.78)	
Hypertension	No	Ref	p<0.0001
	Yes	4.68(3.11;7.04)	
	Other	0.45(0.10;2.09)	
Random effects			
EA		0.10(0.01;1.20)	0.13

In the adjusted analysis, the results showed, family history and hypertension were significantly associated with diabetes for both models (Table 3.11 (b)). TB status was not significantly associated with diabetes for either models.

Family history of diabetes

In both analyses, there was a significant association between family history and diabetes. The odds of diabetes among those that had a family history were 4.73 times higher than those who did not have a family history of diabetes (aOR=4.73, P<0.0001, 95% CI 3.30; 6.78).

Hypertension

In both analyses there was a significant association between hypertension and diabetes. The odds of diabetes among those that had hypertension were 5.59 times higher than those who did not have hypertension (aOR=5.59, P<0.0001, 95% CI 4.12; 7.59) for the questionnaire model and 4.68 higher than those who did not have hypertension (aOR=4.68, P<0.0001, 95% CI 3.11; 7.04). BMI was found to confound the effect of hypertension since in the model including BMI the odds ratio was lower than in the model without BMI.

Table 3.11 (c) Multilevel analysis for the association between diabetes and lifestyle factors

		Multivariate	
Variable	Level	OR(95% CI)	pvalue
Fixed effect			
Smoke status	Yes, daily	0.83(0.51;1.37)	
	Yes, less than	1.27(0.54;2.96)	
	No, not at all	Ref	p=0.544
	Yes, but	1.62(0.59;4.43)	
	Other	0.52(0.12;2.29)	
Alcohol intake	Never	Ref	p=0.403
	Monthly	1.31(0.84;2.06)	
	2-4 times a month	0.97(0.48;1.93)	
	2-3 times a week	0.53(0.19;1.47)	
	4 or more times a week	0.43(0.12;1.56)	
Random effects			
EA		0.10(0.01;1.20)	0.13

In the adjusted analysis, the results showed, smoking status and alcohol intake, were found to not be significantly associated with diabetes for either models (Table 3.11 (c)).

3.4.2.2 Investigation of interactions in the model

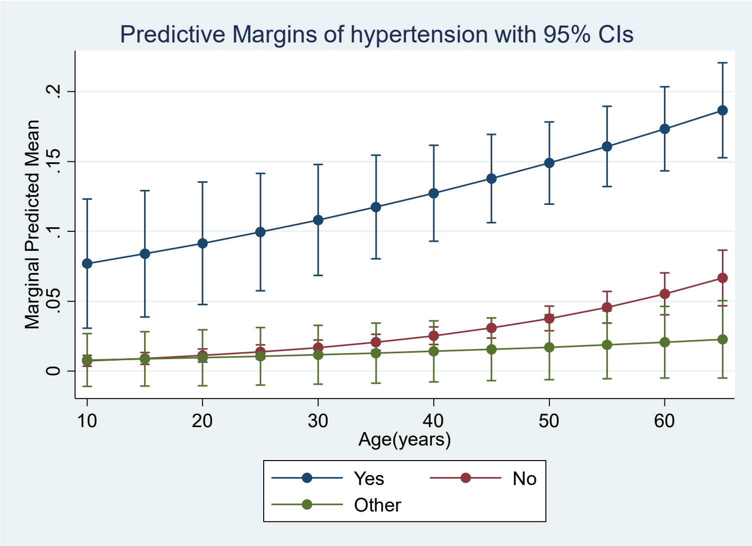
For model including BMI (model D)

The interaction between hypertension and alcohol was significant. From the table below for hypertensive those who drink more than 4 times a week are at a much lower risk while for those without hypertension the effect of alcohol is much less pronounced.

	Alcohol					
HBP		Never	Monthly or less	2-4 times a month	2-3 times a week	4 or more times a week
	Yes	0.12	0.18	0.16	0.14	0.07
	No	0.04	0.04	0.01	0.002	0.01

There was a significant Interaction between hypertension and age. Figure 3.8 below shows that the effect of hypertension on the risk of diabetes increases with increasing age.

Figure 3.8: Margin plot age and hypertension interaction



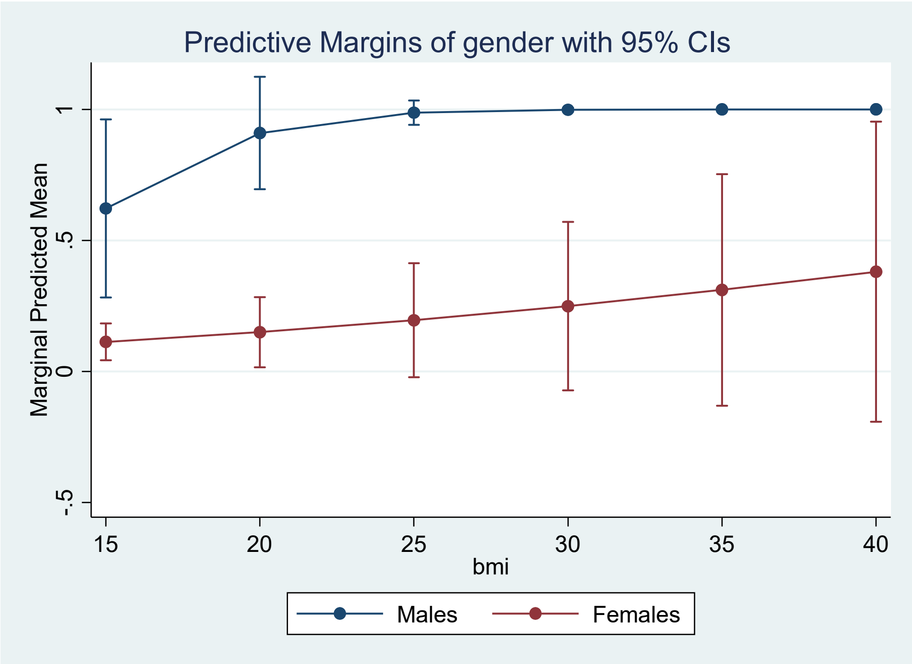
The interaction between alcohol and race was found to be significant. Margin probabilities are shown in table below.

		Alcohol				
Race		Never	Monthly or less	2-4 times a month	2-3 times a week	4 times or more a week
	African	0.07	0.08	0.07	0.05	0.03
	White	0.06	0.16	-	-	0.01
	Coloured	0.06	0.09	0.06	0.01	0.12
	Indian	0.10	0.05	-	0.32	0.27

Africans drinking had a small effect on the prevalence of diabetes, for whites those who drank monthly or less were at highest risk, for coloureds those who drank 4 times or more per week were at the highest risk, while for Indians those who drank 2 or more times a week were at a much higher risk.

For Males the probability of diabetes increases rapidly up to a BMI of 25 per 5kg/m^2, and then levels off, while for females the probability shows a slow and steady increase.

Figure 3.9: Margin plot by BMI and gender interaction



The interaction between geotype and alcohol intake was found to be significant. Also, the interaction between geotype and gender was found to be significant.

The interaction between geotype and hypertension was found to be significant. For those without hypertension there is little effect of geotype and the prevalence is generally very low, while for those with hypertension the prevalence is high among those living in urban informal and urban formal.

		geotype			
HBP		Urban formal	Urban informal	Rural informal (tribal)	Rural formal (Farms)
	Yes	0.14	0.15	0.10	0.09
	No	0.04	0.01	0.03	0.05

3.4.6 Comparison between multilevel models (melogit)

The effect of hypertension was larger in the model from the adult questionnaire than in the model including BMI, with ORs of 5.59 and 4.68 respectively. The BMI plays a confounding role in the effect of hypertension since adding BMI to the model changes the magnitude of the hypertension effect. In addition the odds ratio for the effect of family history of diabetes was higher in the model using adult questionnaire data.

3.4.7 Comparison of interactions between the two models

The interaction between hypertension and age was significant in BMI model and not significant in the questionnaire model.

The interaction between race and alcohol was significant in both models. The BMI was found to have increased the marginal probabilities for both interactions. The models are similar in detecting this interaction.

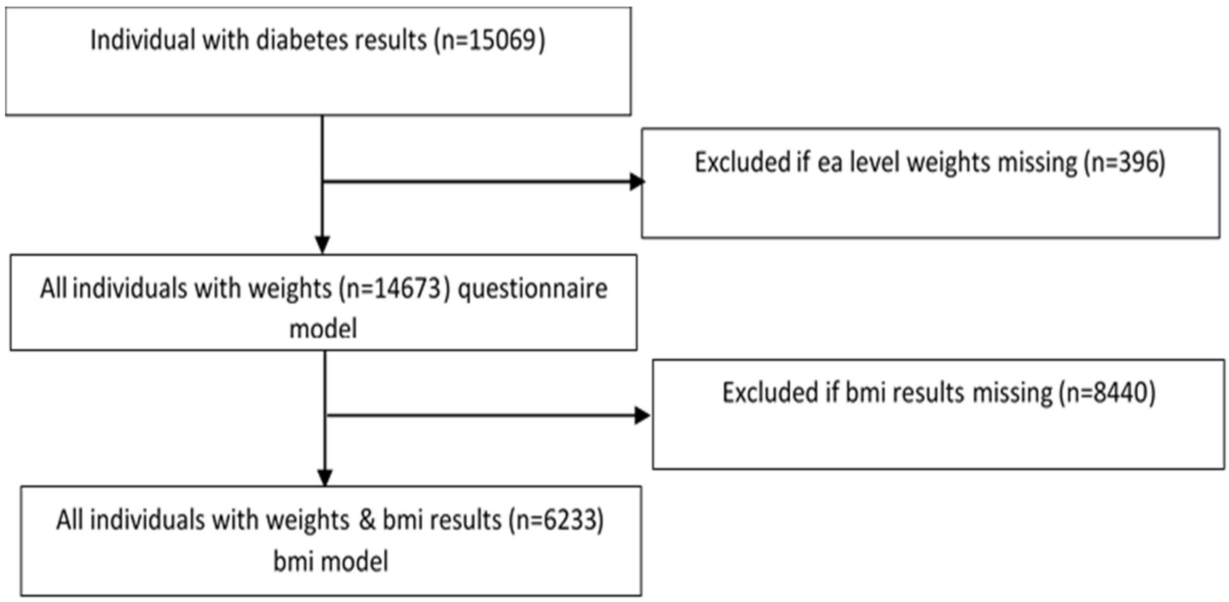
Interaction between race and alcohol indicates that the effect of alcohol is different for different races. In addition BMI was found to have increased the marginal probabilities for both interactions.

The interaction between hypertension and alcohol intake was significant in both models. The two models detected similar interactions.

3.5 Weighted multilevel model

The Generalized Linear Latent and Mixed Models (GLLAMM) of Rabe-Hesketh and Skrondal as applied to multistage surveys were fitted (2006). Rabe-Hesketh and Skrondal suggest alternative strategies for weighting and here we used scaling method 1 for weights. In this method the weights are standardized by summing the scaled weights. The two models were fitted with 413 enumeration areas and 6233 individuals for the model including BMI and 458 enumeration areas and 14,673 individuals for the adult questionnaire model. These models are referred to as Model E and F. In the GLLAMM models both EA level weights and unit level weights are scaled. Multilevel models fitted previously (model C and model D) only allow EA level weights, while GLLAMM (which is used in model E and model F) allows both EA level and individual level weights.

Figure 3.10 Multilevel inclusion & exclusion graph



3.5.1 Weighted multilevel model

Table 3.12 (a) Multilevel analysis for the association between diabetes and socio-demographic factors (model E)

		Univariate		Multivariate	
Variable	Level	OR(95% CI)	pvalue	OR(95% CI)	pvalue
Fixed effect					
Gender	Males	Ref	p=0.0003	Ref	p=0.810
	Females	1.53(1.24;1.89)		1.02(0.78;1.32)	
	Missing	1.15(0.22;5.88)		0.59(0.11;3.34)	
Race	African	Ref	P<0.0001	Ref	P=0.001
	White	1.18(0.77;1.82)		0.94(0.54;1.63)	
	Coloured	1.58(1.22;2.04)		1.09(0.79;1.48)	
	Indian	3.49(2.61;4.66)		1.94(1.40;2.68)	
	Other	0.41(0.08;2.18)		0.38(0.08;1.78)	
Employed	No	Ref	P<0.0001	Ref	p=0.899
	Yes	1.41(1.01;1.98)		0.99(0.98;1.00)	
	Other	2.44(1.78;3.34)		1.02(1.00;1.03)	
	Missing	1.62(0.74;3.52)		1.02(0.99;1.05)	
Geotype	Urban formal	Ref	P<0.0001	Ref	p=0.077
	Urban informal	0.50(0.34;0.70)		0.54(0.55;1.15)	
	Rural informal (tribal)	0.69(0.53;0.88)		0.79(0.54;1.00)	
	Rural formal (farms)	0.55(0.40;0.77)		0.74(0.45;0.95)	
Age		1.06(1.06;1.07)	P<0.0001	1.05(1.04;1.05)	p<0.0001

[5]

Fixed effect

The results of fitting the models are shown in Table 3.12 (a). Age and race were found to be significantly associated with diabetes. Gender, employment and geotype were found to be not significantly associated with diabetes.

Race

In the adjusted analysis race was significantly associated with diabetes. With Whites, Coloureds and Indians having a higher risk than Africans of having diabetes. In particular the odds of Indians having diabetes was 1.94 times higher than the odds of Africans (aOR=1.94 , P<0.0001, 95% CI 1.40 ; 2.68) for the questionnaire model.

[5] All exposure variables are included in the analysis

Age

The odds of having diabetes increased as age increased, after adjusting for other factors. The adjusted OR of 1.05 (P<0.0001 , 95% CI 1.04; 1.05) shows that the odds of diabetes increase by 5% for every one year increase in age. To get a better idea of the effect of age, for a 5 year increase in age, the OR is 1.21 (95% CI 1.14 ; 1.28) so the odds of diabetes increase by 26% for a 5 year increase in age.

Table 3.12 (b) Multilevel analysis for the association between diabetes and co-morbidity factors (Model E)

		Univariate		Multivariate	
Variable	Level	OR(95% CI)	pvalue	OR(95% CI)	Pvalue
Fixed effect					
TB status	No	Ref	P=0.084	Ref	p=0.629
	Yes	1.37(0.96;1.94)		1.15(0.74;1.81)	
	Missing	0.66(0.37;1.18)		0.69(0.36;1.31)	
Family History	No	Ref	p<0.0001	Ref	p<0.0001
	Yes	5.99(4.73;7.59)		4.58(3.54;5.92)	
	Other	1.55(1.01;2.37)		1.51(0.94;1.05)	
Hypertension	No	Ref	Ref	Ref	p<0.0001
	Yes	4.84(3.63;6.45)		4.84(3.63;6.45)	
	Other	0.64(0.36;1.31)		0.64(0.36;1.31)	

In the adjusted analysis, the results showed, family history and hypertension were significantly associated with diabetes (Table 3.12 (b)). TB status was not significantly associated with diabetes.

Family history of diabetes

In both analyses, there was a significant association between family history and diabetes. The odds of diabetes among those that had a family history were 4.58 higher compared to those who did not have a family history of diabetes (aOR=4.58, P<0.0001, 95% CI 3.54;5.92).

Hypertension

In both analyses there was a significant association between hypertension and diabetes. The odds of diabetes among those that had hypertension were 4.84 higher than those who did not have hypertension (aOR=4.84, P<0.0001, 95% CI 3.63; 6.45).

Table 3.12 (c) Multilevel analysis for the association between diabetes and lifestyle factors

		Univariate		Multivariate	
Variable	Level	OR(95% CI)	pvalue	OR(95% CI)	pvalue
Fixed effect					
Smoke status	Yes, daily	0.94(0.74;1.21)		1.03(0.73;1.43)	
	Yes, less than	0.97(0.40;2.36)		1.08(0.46;2.52)	
	No, not at all	Ref	P=0.0001	Ref	P=0.526
	Yes, but	3.21(1.96;5.28)		1.59(0.82;3.06)	
	Other	2.14(1.01;4.51)		1.84(0.68;5.01)	
Alcohol intake	Never	Ref	P=0.0005	Ref	p=0.004
	Monthly	0.68(0.46;1.02)		0.89(0.58;1.33)	
	2-4 times a month	0.74(0.47;1.17)		0.94(0.57;1.01)	
	2-3 times a week	0.37(0.19;0.88)		0.36(0.20;0.98)	
	4 or more times a week	0.73(0.38;1.39)		0.96(0.08;0.98)	
	Missing	0.89(0.39;1.02)		0.89(0.39;1.02)	

In the adjusted analysis, the results showed alcohol intake was found significantly associated with diabetes (Table 3.12 (c)). Smoking status was found not significant.

Alcohol intake

In the adjusted analysis, drinking alcohol was protective against diabetes, with those who drank 2 -3 more times a week having 69% lower odds of having diabetes compared to those who never drank alcohol. Those who drank 2-3 times a week were at reduced risk, while in the model with BMI it was those who drank 4 or more times a week who were at reduced risk.

3.5.2 Weighted multilevel model including BMI

Table 3.13 (a) Multilevel analysis for the association between diabetes and socio-demographic factors (model F)

		Multivariate	
Variable	Level	OR(95% CI)	pvalue
Fixed effect			
Gender	Males	Ref	p=0.073
	Females	0.71(0.46,1.10)	
	Missing	0.10(0.01;1.11)	
Race	African	Ref	p=0.061
	White	1.57(0.75;3.31)	
	Coloured	0.91(0.55;1.50)	
	Indian	1.65(0.98;2.76)	
	Other	0.96(0.92;0.99)	
Employed	No	Ref	p=0.371
	Yes	1.11(0.66;1.87)	
	Other	1.38(0.82;2.32)	
	Missing	2.35(0.80;6.94)	
Geotype	Urban formal	Ref	p=0.006
	Urban informal	0.68(0.44;1.07)	
	Rural informal (tribal)	0.60(0.41;1.07)	
	Rural formal (farms)	0.53(0.36;0.90)	
Age		1.04(1.03;1.05)	p<0.0001

[6]

Fixed effect

The results of fitting the models are shown in Table 3.13 (a). Gender, race and employment were found not to be significantly associated with diabetes. Age and geotype were found to be significantly associated with diabetes.

Age

The odds of having diabetes increased as age increased, after adjusting for other factors. The adjusted OR of 1.04 (P<0.0001 , 95% CI 1.03; 1.05) shows that the odds of diabetes increase by 4% for every one year increase in age. To get a better idea of the effect of age, for a 5 year increase in age, the OR is 1.21 (95% CI 1.14 ; 1.28) so the odds of diabetes increase by 21% for a 5 year increase in age.

[6] All exposure variables are included in the analysis

Gender

In the adjusted analysis gender was significantly associated with diabetes. With females having lower risk than males of having diabetes. In particular the odds of females having diabetes was 29% lower than the odds of males (aOR=0.71 , P=0.030, 95% CI 0.46 ; 1.10).

Geotype

In the adjusted analysis geotype was significantly associated with diabetes. With those living in urban informal having 32% lower risk than those living in urban formal of having diabetes. In particular the odds of having diabetes was 40% lower for those living in rural informal than the odds of living in urban formal (aOR=0.60 , P=0.006, 95% CI 0.41 ; 1.07).

Table 3.13 (b) Multilevel analysis for the association between diabetes and co-morbidity factors (Model F)

		Multivariate	
Variable	Level	OR(95% CI)	pvalue
Fixed effect			
TB status	No	Ref	p=0.629
	Yes	1.17(0.60;2.26)	
	Missing	0.69(0.27;1.76)	
Family History	No	Ref	p<0.0001
	Yes	4.09(2.84;5.91)	
	Other	1.87(1.04;3.37)	
Hypertension	No	Ref	p<0.0001
	Yes	4.10(2.69;6.25)	
	Other	0.82(0.18;3.68)	

In the adjusted analysis, the results showed, family history and hypertension were significantly associated with diabetes (Table 3.13 (b)). TB status was not significantly associated with diabetes for either models.

Family history of diabetes

In adjusted analyses, there was a significant association between family history and diabetes. The odds of diabetes among those that had a family history were 4.09 higher than those who did not have a family history of diabetes (aOR=4.09, P<0.0001, 95% CI 2.84;5.91).

Hypertension

In adjusted analyses there was a significant association between hypertension and diabetes. The odds of diabetes among those that had hypertension were 4.10 higher than those who did not have hypertension (aOR=4.10, P<0.0001, 95% CI 2.69; 6.25).

Table 3.13 (c) Multilevel analysis for the association between diabetes and lifestyle factors

		Mutlivariate	
Variable	Level	OR(95% CI)	pvalue
Fixed effect			
BMI	-	1.20(1.09;1.29)	p<0.0001
Smoke status	Yes, daily	1.09(0.60;1.96)	p=0.623
	Yes, less than	1.52(0.57;4.04)	
	No, not at all	Ref	p<0.0001
	Yes, but	1.33(0.72;4.79)	
	Other	1.14(0.29;6.20)	
Alcohol intake	Never	Ref	p=0.001
	Monthly	1.14(0.64;2.04)	
	2-4 times a month	1.02(0.50;2.08)	
	2-3 times a week	1.02(0.07;0.50)	
	4 or more times a week	0.19(0.07;0.53)	
	Missing	1.03(0.30;3.48)	

In the adjusted analysis, the results showed alcohol intake and BMI was found significantly associated with diabetes (Table 3.13 (c)). Smoking status was found not significant.

BMI

In the adjusted analysis the odds of having diabetes increased as BMI increased, after adjusting for other factors. The adjusted OR of 1.20 (P<0.0001 , 95% CI 1.09;1.29) shows that the odds of diabetes increase by 20% per 5kg/m^2 increase in BMI.

Alcohol intake

In the adjusted analysis, drinking alcohol was protective against diabetes, with those who drank 2 -3 more times a week having 69% lower odds of having diabetes compared to those who never drank alcohol.

3.5 Comparison between all models

Comparison of questionnaire models (A, C and E)

Being employed was found only significant on the model E. Alcohol intake was found significant for model A and E and not significant for model C. For those with family history, odds of being diabetic was found higher for the melogit model C, also those with hypertension odds of being diabetic were found higher for the melogit model C. Smoking status was not found to be statistically significant in any model.

Comparison of BMI models (B, D and F)

Geotype was only found significant in the gllamm model (model F). Being employed was found significant with that model F. For those that were hypertensive, odds of diabetes were found higher in the svy model (model B) and those with family history of diabetes higher odds of being diabetic was found higher in the melogit model (D). Alcohol intake was found significant for mode B and F and not significant in the model D.

Comparison between all models

Overall the interactions between hypertension and age was found significant for model A and D. The interactions between hypertension and alcohol was found significant for model B and C. The interactions between geotype and alcohol was found significant for model A, B and D. The interactions between race and alcohol was found significant for model B, C and D. The interactions between BMI and gender was found significant for model B and D.

3.7 Conclusion

A number of different logistic regression models were fitted. All models found age, family history and hypertension significantly associated with the presence of diabetes. The odds of having diabetes increased with age and the effect was similar for all models.

All models except model F found a significant race effect, with Indians at a greater risk of diabetes compared to Africans. The odds ratios found varied between models. The risk was found higher in the multivariate adult questionnaire model (model A) (2.57) compared to other models.

All models found that family history of diabetes was overwhelmingly associated with diabetes. The odds ratios found varied between models. The melogit model without BMI (model C) showed a larger effect of family histopry of diabetes than model A and model E.

The study found that BMI was an important confounder when assessing the effect of family history, hypertension, gender and alcohol intake on the odds of having diabetes (Table 3.6 and Table 3.7).

In models using only the questionnaire data there was not a significant association between gender and having diabetes. However when BMI was added to the model a significant interaction between BMI and gender was found. Both multilevel models found significant association between gender and having diabetes for effect of BMI.

The study found BMI was an important confounder when assessing the effect of family history, hypertension, gender and alcohol intake on the odds of having diabetes (Table 3.6 and Table 3.7).

All models found a statistically significant association between hypertension and diabetes, although the odds ratios differed between models. The effect of BMI was found to confound the association with diabetes in the melogit model. Also the melogit model showing higher risk of diabetes compared to other models (5.59) without the effect of BMI for those with hypertension.

Geotype was found significant only in the gllamm model with the effect of BMI (Table 3.14a). those that were living in urban informal were found at 32% lower risk of having diabetes compared to those living in urban formal.

Alcohol intake was found significant in some models. The odds were not too different between these models. The effect of BMI was found to slightly increase the odds of diabetes for those that drink monthly for all significant models. The odds for those that drink monthly were found higher in the melogit model (model D) with 31% higher risk of having diabetes with the effect of BMI compared to multivariate model (model B) with 19% higher risk of having diabetes with the effect of BMI.

The effect of BMI was found significant in some models. The odds varied between these models.

CHAPTER 4
DISCUSSION

4. Introduction

In this chapter, results of the study are discussed. The strength and limitations of the study are also discussed. The study assessed the risk factors associated with diabetes in all South African adults (those aged 15 years and above) from SANHANES-1 survey. Participants who reported having diabetes and those who reported not having diabetes were investigated. The risk factors associated with diabetes including socio-demographic, lifestyle and co-morbidity factors of participants were assessed.

4.1 Prevalence of diabetes

The current study found the prevalence of diabetes to be 5.3%. The current study found that the provinces with the highest prevalences were KwaZulu Natal (7.4%), Northern Cape (6.5%), Western Cape (6.4%) and Free State (6.3%). In the SADHS 2016 Eastern Cape was the highest with 18% with Gauteng as the lowest province for both studies. The SADHS used glycated haemoglobin (HbA1c) test to measure diabetes that used an adjusted HbA1c level of 6.5% or above to indicate diabetes. A study in Zambia and Western Cape showed the prevalence of diabetes at 2.9% in Zambia and Western Cape 9.4%, compared to the current study in which the Western Cape had prevalence of 6.4% (Bailey et al., 2015). An earlier study found the prevalence in the Western Cape to be 9.4% (Bailey et al., 2015).

IDF Diabetes Atlas estimated 1 in 2 (232 million) people globally are living with undiagnosed diabetes. With this increasing trend in the prevalence of diabetes, a study in Korea stated efforts should be made to detect diabetes (Lee et al., 2018). In addition, this increase was attributable to change in risk factors (Lee et al., 2018). As such, there’s a need for early detection (Lee et al., 2018). Screening for early detection is essential (Lee et al., 2018). Kao et al., 2018 suggested close follow up for diabetes patients 15-55 years to control and monitor for diabetes.

A study that analysed lifestyle socioeconomic inequalities in South Africa found diabetes prevalence was 11% of which 38% was undiagnosed with 63% of the undiagnosed being

Africans (Mutyambizi et al., 2019). A study of 1099 participants in Urban-Dwelling Black South Africans found prevalence of diabetes to be 13.1% (Peer et al., 2012).

Researchers found the prevalence of diabetes to be increasing globally (Whiting et al., 2018; Hall et al., 2011). The lowest health care costs were estimated to be from the African region.

A study that was conducted to assess the knowledge of T2DM found mean knowledge for diabetes score was poor (3.2 out of a total of 10) with 72.3% having poor general knowledge about diabetes (Ntontolo et al., 2017). It further found education level and gender were associated with poor knowledge about diabetes (Ntontolo et al., 2017) showing that there is a need for diabetes education.

Other studies have found an association between health education and diabetic complications (Ponesai et al., 2015), showing that health education is necessary among people that are at risk.

4.2 Socio-demographic factors

The socio-demographic factors considered in this study included age, gender, race, geotype and employment status.

4.2.1 Age

Age-groups with the highest diabetes prevalence were 55 to 64 at 16.8% and 65+ at 16.2%, showing diabetes increased with age (SADHS, 2016; Ntontolo et al., 2017; Bradshaw et al., 2000; IDF, 2015; Amod et al., 2017).

South African Demographic Health Survey 2016 found 15.3% of the elderly (65 years and older) were diabetic. As such, this study was consistent with other studies that found prevalence of diabetes increases with age (Ntontolo et al., 2017; SADHS, 2016; Mbanya et al., 2010; Bradshaw et al., 2000; IDF, 2015; Amod et al., 2017). Diabetes was found to be the second highest cause of mortality among the elderly (StatsSA, 2016).

In the adjusted analyses age was significantly associated with diabetes for all models. The effect of age was similar in all models.

4.2.2 Gender

In this study the prevalence of diabetes was higher among females (6.3%) than among males (4.1%). Similarly, in the SADHS 2016, prevalence of diabetes by gender was 13% of females and 8% of males. Furthermore, SADHS showed 16.4% males and 14.6% females among elderly (SADHS, 2016). In addition the General Household Survey (GHS) 2016, found the prevalence of diabetes among the elderly to be 16.6%, with a prevalence of 17.6% among females and 15.0% among males (GHS, 2016).

In model B there was a significant interaction between gender and BMI. For females the odds of diabetes increased by 9% for each 5kg/m-sq increase in BMI, while for males the increase was steeper at 15% for each 5 kg/m-sq increase in BMI.

Other researchers such as Kruger et al., 2012, found that females who were underweight and living in rural areas, farms and the informal sector were at a reduced risk. Their study also showed that females were at a greater risk than males of being overweight or obese. Their study further showed increased risk of overweight/obesity was associated with females (Kruger et al, 2012). Sartorius et al., 2015 confirmed white females and African females as a key determinant of higher obesity risk in South Africa.

4.2.3 Race

The current study found that the prevalence of diabetes was highest among Indians. The current study showed prevalence of diabetes was highest among Indians with 16.6% followed by white with 7% then African with 4.5%. Similarly, the GHS 2016 and SADHS 2016 found that diabetes prevalence was highest for Indians, with higher prevalence among the elderly compared to this study.

In the adjusted analyses, race was found to be significantly associated with diabetes in all analyses. The odds ratio for diabetes was found to be higher for Indians compared to Africans. There was a significant interaction between race and age.

4.2.4 Geotype

The current study found diabetes prevalence higher in those living in urban formal at 6.4%. Similarly, diabetes mellitus was higher among elderly women residing in urban areas (16.2%) compared to those who resided in non-urban areas (12.5%) in the SADHS 2016.

A study carried out by Hill at al. (2020) in South Africa found that the prevalence of Type 2 Diabetes Mellitus was higher among people living in urban areas than in rural areas. Another study in Tanzania and Uganda showed high prevalence of diabetes in rural areas results from lack of awareness due to concentration target in urban areas while rural remains unaware of risk factors of diabetes and prevention (Chiwanga et al, 2016).

In Africa people are shifting towards westernized diet leading to increased obesity levels (Bos & Agyemang, 2013). Although the IDF has recognized urbanization as a risk factor for T2DM, this study found that geotype was not statistically significantly associated with diabetes in any of the adjusted models. StatsSA defines formal being structured and organized and informal as squatter camps. The difference with other studies could be because other does not categorise by urban and rural. Geotype was only found to be significant in model F (gllamm with data on BMI).

4.2.5 Employment status

The current study found the highest prevalence of diabetes in those whose employment status was categorised as "other" (which includes those that are sick disabled and unable to work, those that retired and those that specified their job status as other) with a prevalence of 7.6% compared to those that are employed with a prevalence 4.3%.

Employment status was found not to be statistically significantly associated with diabetes.

4.2.6 Summary

All models found age significantly associated with of diabetes. Geotype, and gender were significantly associated with diabetes in some models.

This study showed that diabetes prevalence differs with age and gender, being higher in females than males and being lower in the age groups aged 34 or below. Other studies found similar results showing that diabetes prevalence varies with age and gender (SADHS, 2008; SANHANES, 2013; StatsSA, 2016).

The obesity epidemic is increasing in South Africa, and a strong link has been found between this and type 2 diabetes (Chen et al., 2018). Hence diabetes increased proportionally with age and BMI (Chen et al., 2018). For a 10 year increase in age, the odds of having diabetes increased by 88% (Chen et al., 2018). Furthermore, age was found to be an effect modifier for the effect of BMI on the incidence of diabetes (Chen et al., 2018).

Several studies in informal urban settlements in South Africa suggested that higher socio-economic status was key factors for higher obesity rates amongst females (Case et al., 2009; Sartorius et al., 2015; Puoane et al., 2005). A number of studies also found that black South African women prefer a larger body size (Case et al., 2009); Sartorius et al., 2015; Puoane et al., 2005).

4.3 Co-morbidity factors

The co-morbidity factors for this study included hypertension, family history of diabetes and TB status.

4.3.1 Hypertension

South Africa was found to have the highest prevalence of hypertension in sub-Saharan Africa (SADHS 2016). A study in Cape Town found prevalence of Hypertension at 30% (Hill et al., 2020). In the current study, 20.5% of participants that had hypertension were diabetic, and this was higher among females (20.8%) than males (11.8%).

The current study showed the prevalence of diabetes among those who are hypertensive to be highest in age 55 to 64 (28%) followed by 45 to 54 (27.8) and 65+ years at 27.2%.

In another study hypertension was found to be a common co-morbidity among diabetic patients (Berkowitz et al., 2018). For the current study, in the adjusted regression analysis hypertension was associated with diabetes for all models.

A significant interaction was found between hypertension and age, so the effect of hypertension on the prevalence of diabetes depends on age. The increased risk of diabetes among hypertensives was found to increase with increasing age and the risk of diabetes among those aged 65 or more was much higher for those with hypertension than those without hypertension.

Similar results were found in earlier studies that found hypertension co-existed with diabetes (Giday et al., 2015). The current study showed those that are hypertensive, have family history of diabetes, increasing with age and race were significantly associated with diabetes. The risk of diabetes among those with hypertension was lowest for those who drank heavily (4 or more times a week).

4.3.2 Family history of diabetes

The current study found 14.4% of participants that had family history of diabetes were diabetic. Another study also found that participants with a family history of diabetes were at a higher risk of being diabetic (Chen et al., 2018).

Family history of diabetes was found to be statistically significantly associated with diabetes in all models. Other studies found family history of diabetes were significantly associated with diabetes in men (Peer et al., 2012).

4.3.3 TB status

The prevalence of TB among study participants was 6.1%. The prevalence of diabetes among participants who had TB was 7.6% compared to a prevalence of diabetes of 5.1% among participants who did not have TB. None of the models found a statistically significant association between TB status and diabetes.

4.3.5 Summary

Family history of diabetes and hypertension were both found to be statistically significantly associated with diabetes in all models. On the other hand no evidence of an association was found between TB status and diabetes.

4.4 Lifestyle factors

Lifestyle factors for this study included alcohol intake, smoking and BMI.

4.4.1 Smoking status

The prevalence of diabetes was found to be high (15.8%) in those that used to smoke but no longer smoked, while for those who never smoked the prevalence was 5.8%.

The current study found no association between smoking status and diabetes. Other studies found an association between smoking and diabetes; a study by Aynalem et al., 2015, found that individuals who smoked were about 27 times more likely to be DM positive compared to participants who never smoked in their lifetime. A study by Giday et al., 2015 found that smoking was a significant risk factor for developing type 2 diabetes mellitus

A study carried out in Zambia and the Western Cape found association between smoking and diabetes was not significant in Zambia, but in the Western Cape the association was significant, even though the number of participants who smoked in the Western Cape was low (Bailey et al., 2016).

4.4.2 Alcohol intake

Alcohol intake was categorised as monthly, 2-4 times a month, 2-3 times a week and 4 or more times a week for the current study. Prevalence of diabetes in each alcohol intake category was found highest in those that had alcohol monthly at 10%, followed by 2-4 times a month at 9%. As alcohol consumption increases, the prevalence of diabetes decreases.

Researchers such as Soo- Jeong et al., 2012 found chronic use of alcohol result in the onset of type 2 diabetes, Mutyambizi et al., 2019 found alcohol consumption to be significantly associated with self-reported diabetes, however in this study, alcohol was found to reduce the risk of diabetes.

Alcohol intake was found to be associated with diabetes in the weighted multivariate models (A and B) also the multilevel models (E and F). Furthermore, for model controlling for BMI (model B) Indians (unlike other races) the prevalence is highest among those who drink alcohol 4 or more times a week (16%), while for model C at 14%, highest risk for Indians was found in model D at 27%. Also those that are hypertensive and drink alcohol had higher odds of diabetes.

4.4.3 BMI

The current study found that women had a higher mean BMI (28.91 kg/m^2) than men (23.56 kg/m^2). The mean BMI for those that are hypertensive is higher compared to those that are not with 30.58 kg/m^2 and 25.85 kg/m^2.

The mean BMI for those with a family history of diabetes (28.61 kg/m-sq) was higher than the mean BMI for those without a family history of diabetes (26.29 kg/m-sq).
The mean BMI for those that never drank alcohol were higher compared to those who had alcohol with BMI of 27.30 kg/m^2. Those who had no TB had higher BMI of 27.14 kg/m^2 compared to those who had TB 23.76 kg/m^2. Those that were living in urban formal had higher BMI compared to those that were not with BMI of 27.58 kg/m^2. Age category showed that age groups 45 to 54 and 55 to 64 had higher BMI compared to other groups (29.33 and 29.18 kg/m^2) followed by 65+ with BMI of 28.14 kg/m^2.

The mean BMI was higher for those with diabetes (30.82 kg/m^2) than for those without (26.47 kg/m^2). When BMI was categorized, the prevalence of diabetes was found to be 2.99% in those that were underweight, 2.77% in those with a normal weight, 7.57% among those who were overweight and 11.96% among those who were found to be obese.

In the 2016 SADHS the prevalence of diabetes was found to increase with increasing BMI, rising to 23% among obese women and 24% among obese men.

BMI was found to be associated with diabetes in the weighted multilevel models (D and F) adjusting for cluster effects. The BMI was found to increase the risk of diabetes (19% model D, 20% model F). BMI was found to be associated with diabetes.

A number of studies found that the proportion of participants with diabetes increased with increasing BMI (Matsha et al., 2012; Gray et al., 2015; Tavşanli et al., 2018). Although BMI was not found significant in the weighted multivariate logistic (model B) and weighted multilevel (model D) it was found significant in the scaled weighted multilevel model (model F). A number of studies in SA found obesity to be a major risk factor for NCDS including diabetes (Zemlin, 2016; WHO, IDF, 2018; Peer et al., 2012; Reddy et al., 2012; Ponesai et al., 2011).

A study was carried out in North Ethiopia that assessed risk factors associated with Type 2 Diabetes Mellitus (T2DM), which found individuals that were overweight or obese, were 9.61 times more likely to develop DM (Giday et al., 2015). Sartorius et al., 2015 suggested increasing obesity rates resulting in a rising prevalence of T2DM especially among females in South Africa (Kruger et al., 2012; Micklesfield et al.,2013). Researchers such as Katchunga et al., 2016 confirmed an association between diabetes mellitus and abdominal obesity but not with general obesity.

Case & Menendez (2009) suggested understanding the differences that men and women face in their risks of obesity is a necessary and important first step for effective policy intervention. They stated that it is better to educate women on the risks of obesity. Also, this can change their perceptions of body image. They also stressed the importance of understanding early life and the impact of socio-economic status on obesity (Case & Menendez, 2009).

A study in Korea reported an increase in awareness rate of diabetes from 23.2% to 68% from the previous Korean National Health and Nutrition Examination survey (Lee et al., 2018). As a result, increasing diabetes prevalence was found in those not attending regular health screening and had inverse relationship with BMI (Lee et al., 2018).

In models including BMI, significant interactions were found between geotype on the one hand and race, age and family history of diabetes. For the effect of BMI, geotype was found to significantly interact with race, increasing age and family history of diabetes. As thus, those living in rural area was found to interact with those with family history of diabetes.

The odds of having diabetes increasing at 5% for the multivariate model and at 19% highest increasing odds for the multilevel model per 5 kg/m^2 increase. On average those that never had alcohol had higher BMI compared to those that had alcohol. The effect of BMI showed odds of diabetes were 13% higher in whites and Indians 86% higher being compared to Africans for the multilevel model. On average whites had higher BMI compared to other races.

4.4.4 Summary

The adjusted analyses in model A, B, E and F showed that alcohol intake was significantly associated with diabetes, while there was no evidence of an association between smoking status and diabetes.

A study carried out in South African using SANHANES-1 to examined contribution of lifestyle factors to diabetes inequalities (Mutyambizi et al., 2019). The study showed that lifestyle factors contributed 22% and 35% to socioeconomic inequalities (Mutyambizi et al., 2019).

The adoption of a Western lifestyle in Africa has led to an increase in obesity, diabetes and hypertension (Matsha et al., 2012).

4.5 Strengths of the study

- The results provided diabetes prevalence in South African adults 15 years and older.
- The results will assist policy makers and Health Professionals to develop policies to prevent the development of diabetes
- The findings help to broaden our knowledge of the risk factors for diabetes and hence allow the implementation of targeted strategies for the prevention of diabetes The study had a large sample size

4.6 Limitations of the study

- More than half adult were missing weight/ height (for BMI)
- Missing or invalid probability weights, missing BMI and missing diabetes results may have introduced selection bias and may affected both the internal and external validity
- As this was a secondary data analysis, certain determinants (e.g. physical activity) were not obtained for the current study
- Since this was a cross-sectional study the temporal relationship between exposure and disease can be unclear i.e. the disease might actually come before the exposure
- Cross-sectional studies tend to identify prevalent cases of long duration, since people who die quickly or recover quickly or who are no longer employed in a particular occupation are less likely to be identified.
- In this study we could not differentiate between type 1 and type 2 diabetes

CHAPTER 5
CONCLUSION AND RECOMMENDATIONS

In this chapter conclusions and recommendations are drawn, based on the associations found between socio-demographic and lifestyle factors and co-morbidities; these recommendations could assist policy makers.

5.1 INTRODUCTION

According to StatsSA, diabetes is the second most common cause of death in South Africa and the leading cause of death in females. Diabetes results in life-threatening complications such as disability, stroke and blindness. In this study socio-demographic, lifestyle and co-morbidity risk factors for diabetes mellitus were found. The findings suggested that, race, alcohol intake, family history of diabetes, hypertension and BMI were statistically significant risk factors associated with diabetes mellitus.

SA in 2011 agreed to the Global Action Plan to prevent and control non-communicable diseases by prevention, promoting health and wellness and improving control of NCDs through health system strengthening, monitoring and their risk factors (Department of Health, 2013; SADHS, 2016). Goals were set as part of an international agenda for the prevention and control of NCDs (SADHS, 2016). Anderson et al, 2013 suggested that with increasing prevalence of diabetes impoverished communities are at highest risk.

Beliefs, behaviours and practices contribute to increasing prevalence, particularly lifestyle and rising prevalence of obesity among children (Anderson et al., 2013).

Studies have shown physical activity was low among children (Anderson et al., 2013). Additionally, low socio-economic status, lack of resources, beliefs about weight and non-participation in physical activity contributes to ever-increasing T2DM (Anderson et al, 2013).

Some studies found health-promoting behaviour change was needed in South Africa, targeting diabetes (Anderson et al., 2013). These changes are not only at the individual level, but at a higher level as well such as political (Anderson et al., 2013). These changes are aimed at building networks especially for young people that reinforce positive behaviour change.

Anderson et al. (2013) suggest that since South Africa has scarce resources, prevention strategies are vital. Such strategies should focus on behaviour change, being more physically active, healthy lifestyle and lowering risk of developing T2DM (Anderson et al, 2013). Bos &

Agyemang (2013) found that the high prevalence of obesity was due to lifestyle changes and rising level of urbanization as the study showed that the highest prevalence was found in the urban formal setting, with a prevalence of 6.4%. In Africa people are shifting towards a westernized diet leading to increased obesity levels Bos & Agyemang (2013).

5.2 Conclusions

The prevalence of diabetes has reached high levels in sub-Saharan Africa. Of particular concern, obesity is one of the risk factors associated with the development of diabetes especially in females with StatsSA showing diabetes as the most common cause of deaths in females (StatsSA, 2016). Studies have shown females are at higher risk than men (Giday et al., 2015; Sartorius et al., 2015; Kruger et al., 2012; Micklesfield et al.,2013; Katchunga et al., 2016). Additionally, their body perception needs to change (Case & Menendez, 2009; Kruger et al., 2012). Furthermore, educating women about the risk associated with obesity will change their minds (Case & Menendez, 2009).

The current study showed that diabetes was significantly associated with hypertension, race, age and family history of diabetes. BMI, alcohol intake and geotype were found significant in some models.

5.3 Recommendations

This study has made a number of findings which can help policy makers and health professionals develop ways of preventing the development of diabetes

- Many studies found women had higher odds of diabetes compared to men, also these women were more likely to be overweight/obese showing that there is a need to educate women about risk factors for obesity and it’s association with diabetes and other NCDs. This should be done at community level (local clinics) ensuring that those women know their weight status and the associated risks. Pamphlets about obesity should be available at local clinics.
- Support groups should be started for those women that are overweight and obese on what they can do to reduce their weight and risk factors associated with obesity, supported by the Department of Health using local clinics
- The targeted youth should be subject to regular check-ups and encouraged to adopt a healthy life-style to reduce the risk of diabetes. Other studies suggested targeting young adolescent girls (Sartorius et al., 2015). Educating young girls about obesity and

associated risk factors from young age at school level should be a priority for future generation to avoid future incidence.

- With South African aiming to implement National Health Insurance (NHI) by 2025, the health care sector should give priority to diabetes hence diagnosing the undiagnosed is critical and and this could reduce the annual cost of diabetes (Erzse et al., 2019). As such, testing systems should be strengthened on a regular basis. Rather than just focusing on the costs of treating and managing diabetes other costs should go into prevention of diabetes which will result in cost saving on the treatment and management of diabetes.
- There is a need to educate people on why the tax on sugar-sweetened beverages is necessary, since studies have shown that daily energy intake projects a rising BMI, which in turn influences the incidence of T2DM (Manyema et al., 2015).
- Pamphlets outling the risks of diabetes and ways of preventing diabetes should be made available at public hospitals, schools, libraries, community halls and clinics
- Awareness should be raised in the community on the link between co-morbidities and diabetes
- Advice should be given to hypertensive patients at clinics and hospitals on the need for a healthy lifestyle and the need to reduce their risk of other NCDs. Hypertensive patients, as well as those with a family history of diabetes should be subject to regular blood glucose monitoring, as well as having their BMI monitored
- Target Indian communities for awareness and how they are at higher risk than other race groups
- BMI monitoring mobile application to be developed to ease calculations
- Healthy lifestyle promotions which include physical activity and healthy eating should be made available at schools, libraries and community halls
- Periodic monitoring of fasting blood glucose levels using community health workers.
- Multilevel model determine how different clusters interact and influence a response variable of interest in this case BMI, both multilevel models BMI was significant whereas in the “svy” model it was not. It is best to use the model that accounts for both scaled level 1 and level 2 weights in this case model E and F. With model E not including primary exposure, model F was recommended which in addition found that the effect of geotype was significant.

5.4 Further opportunities for research

- Since hypertension is a co-morbidity factor, examining factors associated with hypertension is neccesary
- Further analysis comparing the food consumption patterns of those who are diabetic with those that are not diabetic.
- Compare interventions of department of health looking at healthy eating and diabetes for those living in rural areas compared to those living in urban areas. The research should check whether people living in rural areas are getting the same interventions and support as people in urban areas, especially in cases where there is a family history of diabetes
- Compare the eating habits of females with those of males, as well as comparing being overweight or obese between males and females, taking into account biological factors such as a family history of being overweight or obese. Compare the eating habits and level of physical activity between Indians and other race groups.
- Further studies are necessary especially in those racial groups with high prevalence to better understand the local factors that contribute to high risk in order to identify community specific interventions that are necessary for each racial group.
- Beliefs and perceptions interventions are necessary at African rural level especially those with high BMI to change the knowledge and attitude towards wealth and body image
- A cohort study can be designed by recruiting youth with family history of diabetes, followed up with regular screening.
- A cohort study can also be designed by recruiting adults that are overweight and obese, monitor their development of diabetes and how this relates to any changes from baseline in eating habits and level of physical activity

REFERENCES

- Alanazi, T.O. Alenezi, Y.M. Al-Alawi, M. Alghamdi, E.A. Alsulami, A.N. Alzahrani, A.J. Albarqi, B.A. Alotaibi, S. AlanaziAhmedOwaid, Z. AlharthiSalehFurayhan, O. Alghanmi, A.H. & Asiri, S.H. (2017). Coexistence of Hypertension and Diabetes Mellitus in Elderly Population of Arar City, Northern Saudi Arabia. *The Egyptian Journal of Hospital Medicine, 69*, 3154-3159.
- Alqudah, B.M.,et al. 2017. Prevalence of hypertension among diabetic type 2 patients attending medical clinic at Prince Hashem bin Abdullaii hospital in Aqaba. Indian Journal of Medical Research and Pharmaceutical Sciences. June 2017;4(6) ISSN: ISSN: 2349-5340. DOI: 10.5281/zenodo.815699 Impact Factor: 3.052
- Amod, A. Ascott-Evans, B.H. and Berg G.I. (2017). The 2017 SEMDSA guideline for the management of type 2 diabetes. JEMDSA, 22(1), pp.S1-S196.Thematic report: Selected health indicators among the elderly: Findings from the SADHS 2016, GHS 2016 and MACoD 2016
- Anderson, C. Jansen van Vuuren, A. & Learmonth, D. (2013). Prevention of Type 2 Diabetes (T2DM) in South Africa - a Focus on Facilitating Physical Activity Behaviour Change in Peri-Urban Youth. *Online Readings in Psychology and Culture, 10*(1). https://doi.org/10.9707/2307-0919.1122
- Aounallah-Skhiri, H. Traissac, P. El Ati, J. Eymard-Duvernay, S. Landais, E. Achour, N. Delpeuch, F. Ben Romdhane, H. & Maire, B. (2011). Nutrition transition among adolescents of a south-Mediterranean country: dietary patterns, association with socio-economic factors, overweight and blood pressure. A cross-sectional study in Tunisia. *Nutrition journal*, *10*, 38. https://doi.org/10.1186/1475-2891-10-38
- Asparouhov, T. (2006) General Multi-Level Modeling with Sampling Weights. Communications in Statistics—Theory and Methods , 35, 439-460.
- Aynalem, S. B. & Zeleke, A. J. (2018). Prevalence of Diabetes Mellitus and Its Risk Factors among Individuals Aged 15 Years and Above in Mizan-Aman Town, Southwest Ethiopia, 2016: A Cross Sectional Study. *International journal of endocrinology*, *2018*, 9317987. https://doi.org/10.1155/2018/9317987
- Bailey, S. L. Ayles, H. Beyers, N. Godfrey-Faussett, P. Muyoyeta, M. du Toit, E. Yudkin, J. S. & Floyd, S. (2016). Diabetes mellitus in Zambia and the Western Cape province of South Africa: Prevalence, risk factors, diagnosis and management. *Diabetes research and clinical practice*, *118*, 1–11. https://doi.org/10.1016/j.diabres.2016.05.001

- Balogun, W.O. and Salako, B.L. et al. 2011, Co-occurrence of diabetes and hypertension: pattern and factors associated with order of diagnosis among Nigerians. *Annals of Ibadan Postgraduate Medicine*. Vol. 9 No. 2 December, 2011
- Bos, M. & Agyemang, C. (2013). Prevalence and complications of diabetes mellitus in Northern Africa, a systematic review. *BMC public health*, *13*, 387. https://doi.org/10.1186/1471-2458-13-387
- Bradshaw, D. Norman, R. Pieterse, D. Levitt, N.S. (2007). South African Comparative Risk Assessment Collaborating Group. Estimating the burden of disease attributable to diabetes in South Africa in 2000. *S Afr Med J*. 2007;97(8 Pt 2):700-706.
- Breslow, N.E, Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421):9_25; 1993.
- Carle, A.C. (2009) Fitting Multilevel Models in Complex Survey Data with Design Weights: Recommendations. *BMC Medical Research Methodology* , 9, 49.
- Case, A. Menendez, A (2009) Sex differences in obesity rates in poor countries: Evidence from South Africa. *Economics & Human Biology* 7: 271–282.
- Cassela, G. & Berger, R.(2001). statistical Inference, 2nd Edition,Wadsworth,Pacific Groves,California.
- Center for Disease Control (CDC). (2018). Family Health History and Diabetes. Available from: https://www.cdc.gov/genomics/famhistory/famhist_diabetes.htm [accessed 13 Jan 2019].
- Centre for Diabetes and Endocrinology (CDE). Diabetes and TB; 2019. Available from: https://www.cdediabetes.co.za/uploads/images/files/TB%20and%20Diabetes_HamishvWyk_MB%20Edit.pdf [accessed 16 Jan 2019].
- Chang, S.A. (2012). Smoking and type 2 diabetes mellitus. *Diabetes Metab J*. 2012 Dec;36(6):399-403. doi: 10.4093/dmj.2012.36.6.399. Epub 2012 Dec 12. PMID: 23275932; PMCID: PMC3530709.
- Chen, Y. Zhang, X-P. Yuan, J. *et al*. 2018. Association of body mass index and age with incident diabetes in Chinese adults: a population-based cohort study. *BMJ Open* 2018;8:e021768. doi:10.1136/ bmjopen-2018-021768
- Cois, A. & Day, C. (2015). Obesity trends and risk factors in the South African adult population. *BMC obesity*, *2*, 42. https://doi.org/10.1186/s40608-015-0072-2
- Diggle, P.J, Moyeed, R.A. & Tawn, J.A. (1998) Model-based geostatistics. *Journal of the Royal Statistical Society, Series* C(Applied Statistics), 47, 299–350.
- Feghali, M.N. Abebe, K.Z. Comer, D.M. Caritis, S. Catov, J.M. Scifres, C.M. (2018). Pregnancy outcomes in women with an early diagnosis of gestational diabetes

mellitus. *Diabetes Res Clin Pract*. 2018;138:177-186. doi:10.1016/j.diabres.2018.02.004

- Ganz, M.L. Wintfeld, N. Li, Q. Alas, V. Langer, J. Hammer, M. (2014). The association of body mass index with the risk of type 2 diabetes: a case-control study nested in an electronic health records system in the United States. *Diabetol Metab Syndr*. 2014;6(1):50. Published 2014 Apr 3. doi:10.1186/1758-5996-6-50
- Gebregiorgis, I. & Hedlin, D. (2013). On the Performance of Chi Square Tests Adjusted for Complex Survey Design.
- Giday, T.K. Aseffa, H. Kidanemariam, A. (2014) Assessment of Risk Factors Associated with Type 2 Diabetes Mellitus in Central Zone of Tigray, North Ethiopia. *IJPBSF International Journal of Pharmaceutical and Biological Sciences Fundamentals*, Vol. 07, Issue 01, October 2014 WWW.IJPBSF.COM ISSN: 2278-3997, IJPBSF www.ijpbsf.com
- Goedecke, J.H. Jennings, C.L. Lambert. Obesity in South Africa; Chronic Diseases of Lifestyle in South Africa since 1995 – 2005: Chapter 7
- Gray, A. Vawda, Y. Jack, C. (2013). Health Policy and Legislation. In: Padarath A, English R, editors. South African Health Review 2012/13. Durban: Health Systems Trust; 2013. Available from: http://www.hst.org.za/publications/south-african-health-review-2012/13 , [accessed 09 July 2018].
- Gray, N. Picone, G. Sloan, F. Yashkin, A. (2015). Relation between BMI and diabetes mellitus and its complications among US older adults. *South Med J. 2015* Jan;108(1):29-36. doi: 10.14423/SMJ.0000000000000214. PMID: 25580754; PMCID: PMC4457375.
- Hall, V. Thomsen, R. W. Henriksen O. & Lohse N. (2011). Diabetes in Sub Saharan Africa 1999-2011: epidemiology and public health implications. A systematic review. *BMC public health 11* 564. https://doi.org/10.1186/1471-2458-11-564
- Hardin J.W. and Hilbe J.M. (2012). Generalized Linear Models and Extensions. Third edition Stata press
- Idemyor V. (2010). Diabetes in sub-Saharan Africa: health care perspectives challenges and the economic burden of disease. *Journal of the National Medical Association 102*(7) 650–653. https://doi.org/10.1016/s0027-9684(15)30643-x
- Igbojiaku O. J. Harbor O. C. & Ross A. (2013). Compliance with diabetes guidelines at a regional hospital in KwaZulu-Natal South Africa. *African Journal of Primary Health Care & Family Medicine 5*(1) 447. https://doi.org/10.4102/phcfm.v5i1.447 [accessed 07 July 2018].

- International Diabetes Federation. *IDF Diabetes Atlas 9th edn.* Brussels Belgium: 2019. Available from: https://www.diabetesatlas.org [accessed 08 February 2021].
- International Diabetes Federation. *IDF Diabetes Atlas 8th edn.* Brussels Belgium: IDF Brussels; 2019. Available from: https://www.idf.org/e-library/epidemiology-research/diabetes-atlas/14-diabetes-atlas-eighth-edition.html [accessed 08 July 2018].
- International Diabetes Federation. *IDF Diabetes Atlas 7th edn.* Brussels Belgium: IDF Brussels; 2015. Available from: https://www.idf.org/e-library/epidemiology-research/diabetes-atlas/13-diabetes-atlas-seventh-edition.html [accessed 06 July 2018].
- Jang Woncheol & Lim Johan. (2005). PQL Estimation Biases in Generalized Linear Mixed Models.
- Jongen V. W. Lalla-Edward S. T. Vos A. G. Godijk N. G. Tempelman H. Grobbee D. E. Devillé W. & Klipstein-Grobusch K. (2019). Hypertension in a rural community in South Africa: what they know what they think they know and what they recommend. *BMC public health 19*(1) 341. https://doi.org/10.1186/s12889-019-6642-3
- Kackar R. N. and D. A. Harville. 1984. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association* 79: 853–862.
- Katchunga P.B. Cikomola J. Tshongo C. *et al 2016.* Obesity and diabetes mellitus association in rural community of Katana South Kivu in Eastern Democratic Republic of Congo: Bukavu Observ Cohort Study Results. *BMC Endocr Disord* 16 60 (2016). https://doi.org/10.1186/s12902-016-0143-5
- Kim S. J. & Kim D. J. (2012). Alcoholism and diabetes mellitus. *Diabetes & metabolism journal 36*(2) 108–115. https://doi.org/10.4093/dmj.2012.36.2.108
- Kruger A. Wissing M.P. Towers G.W Doak C.M. (2012). Sex Differences Independent of Other Psychosociodemographic Factors as a Predictor of Body Mass Index in Black South African Adults. *Journal of health population and nutrition* 30: 56. PMID: 22524120
- Matsha T. E. Hassan M. S. Kidd M. & Erasmus R. T. (2012). The 30-year cardiovascular risk profile of South Africans with diagnosed diabetes undiagnosed diabetes pre-diabetes or normoglycaemia: the Bellville South Africa pilot study. *Cardiovascular journal of Africa 23*(1) 5–11. https://doi.org/10.5830/CVJA-2010-087

- Mbanya J.C. Motala A.A. Sobngwi E. Assah F.K. Enoru S.T. (2010). Diabetes in sub-Saharan Africa. *Lancet* 2010; 375: 2254–66
- Mbuya F. E. Fredrick F. & Kundi B. (2014). Knowledge of diabetes and hypertension among members of teaching staff of higher learning institutions in Dar es Salaam Tanzania. *Tanzania journal of health research 16*(2) 98–103. https://doi.org/10.4314/thrb.v16i2.5
- Mengesha A.Y. (2007). Hypertension and related risk factors in type 2 diabetes mellitus (DM) patients in Gaborone City Council (GCC) clinics Gaborone Botswana. 2007. *African Health Sciences* 2007; 7(1):244-245
- Micklesfield L.K. Lambert E.V. Hume D.J. Chantler S. Pienaar P.R. Dickie K et al. (2013) Socio-cultural environmental and behavioural determinants of obesity in black South African women. *Cardiovascular journal of Africa* 24: 1–7. PMID: 23396408
- Mohamed S. F. Mwangi M. Mutua M. K. Kibachio J. Hussein A. Ndegwa Z. Owondo S. Asiki G. & Kyobutungi C. (2018). Prevalence and factors associated with pre-diabetes and diabetes mellitus in Kenya: results from a national survey. *BMC public health 18*(Suppl 3) 1215. https://doi.org/10.1186/s12889-018-6053-x
- Morojele N. (2010). Alcohol consumption nutrition cardiovascular disease and iron status in South Africa: Implications for South Africa's drinking guidelines. *South African Journal of Clinical Nutrition 22*(5) S2-S3. accessed from http://www.sajcn.co.za/index.php/SAJCN/article/view/480
- National Department of Health (NDoH) Statistics South Africa (Stats SA) South African Medical Research Council (SAMRC) and ICF. 2018. *South Africa Demographic and Health Survey 2016 Key Findings*. Pretoria South Africa and Rockville Maryland USA: NDoH Stats SA SAMRC and ICF.
- Nouh Faiza & Omar Mariam & Younis Manal. (2017). Prevalence of Hypertension among Diabetic Patients in Benghazi: A Study of Associated Factors. *Asian Journal of Medicine and Health*. 6. 1-11. 10.9734/AJMAH/2017/35830.
- Ntontolo P.N. Lukanu P.N. Ogunbanjo G.A. Fina J.P.L. Kintaudi L.N.M. (2017). Knowledge of type 2 diabetic patients about their condition in Kimpese Hospital diabetic clinic Democratic Republic of the Congo. *Afr J Prm Health Care Fam Med*. 2017;9(1) a1385. https://doi. org/10.4102/phcfm. v9i1.1385
- Peer N. Steyn K. Lombard C. Lambert E.V. Vythilingum B et al. (2012) Rising Diabetes Prevalence among Urban-Dwelling Black South Africans. *PLoS ONE 7(9):* e43336. doi:10.1371/journal.pone.0043336
- Pfeffermann D. (1993) The role of sampling weights when modeling survey data. *Int. Statist. Rev.* 61 317–337.

- Pfeffermann D. Skinner C. J. Holmes D. J. Goldstein H. and Rasbash J. (1998) Weighting for unequal Selection Probabilities in Multilevel Models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* vol. 60 no. 1 1998 pp. 23–40. *JSTOR* www.jstor.org/stable/2985969. Accessed 19 June 2020.
- PfeffermannD. (1996) The use of sampling weights for survey data analysis. *Statist. Meth. Med. Res.* 5 239–261.
- Ponesai N. Anderson C. Mufuta T. Gombe N. Lucia T. and Donewell B. (2011). Risk Factors for Diabetic Complications among Diabetic Patients Chirumanzu District Zimbabwe. Austin *J Public Health Epidemiol. 2015*; 2(2): 1021.
- Puoane T. Fourie J. Shapiro M. Rosling L. Tshaka N. Oelefse A. (2005) \'Big is beautiful\'-an exploration with urban black community health workers in a South African township. *South African Journal of Clinical Nutrition 18*. accessed from http://www.sajcn.co.za/index.php/SAJCN/article/view/89
- Puoane T. Steyn K. Bradshaw D. Laubscher R. Fourie J. Lambert V et al. 2002. Obesity in South Africa: the South African demographic and health survey. *Obesity Research* 10: 1038–1048. PMID: 12376585
- Rabe-Hesketh S. and Skrondal A. (2012). Multilevel and Longitudinal Modeling using Stata Vol1 and Vol 2. Stata press.'
- Rao J.N.K & ScottA.J.(1981) .The analysis of categorical data from complex sample surveys. *Chi* Squared Tests for goodness of fit and independence in two-way Tables *Journal of the American Statistical Association* Vol. 76 No. 374 pp. 221-230.
- Rao J.N.K & Thomas D.R. (2003).Analysis of categorical Response data from complex surveys: an appraisal and update John Wiley & Sons.
- Rao J.N.K & Thomas D.R.(1989) chi-squared tests for contingency tables in Analysis of Complex Surveys Eds by Skinner C.J D.Holt & Smith T.M.F. Wiley pp 89-114.
- Reddy S.P. Resnicow K. James S. Funani I.N. Kambaran N.S. Omardien R.G. Masuka P. Sewpaul R. Vaughan R.D. Mbewu A. (2012). Rapid Increases in Overweight and Obesity Among South African Adolescents: Comparison of Data From the South African National Youth Risk Behaviour Survey in 2002 and 2008. *American Journal of Public Health*. 102(2):262-268; 2012.
- Rozi S. Mahmud S. Lancaster G. Hadden W. Pappas G. (2017). Multilevel modeling of binary outcomes with three-level complex health survey data.. *Open Journal of Epidemiology* 7(1) 27-43. Available at: http://ecommons.aku.edu/pakistan_fhs_mc_chs_chs/280
- Sartorius B Veerman LJ Manyema M Chola L Hofman K (2015) Determinants of Obesity and Associated Population Attributability South Africa: Empirical Evidence

from a National Panel Survey 2008-2012. *PLoS ONE* 10(6): e0130218. doi:10.1371/journal.pone.0130218

- Shisana O. Labadarios D. Rehle T. Simbayi L. Zuma K. Dhansay A. et al. (2013). The South African National Health and Nutrition Examination Survey: SANHANES-1. HSRC Press 2013 Aug 15: 1–4. [accessed 06 July 2018].
- Shun Z. and McCullagh P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society Series B Methodological* 57:749–760.
- Silverman R.A. Thakker U. Ellman T. Wong I. Smith K. Graff K. (2011). Hemoglobin A1c as a screen for previously undiagnosed prediabetes and diabetes in an acute-care setting. *Diabetes Care*34(9) 1908-1912.
- South Africa Demographic and Health Survey (SADHS). (2016). *Key Indicators Report*. Pretoria: National Department of Health; 2016. Stats SA website: www.statssa.gov.za
- South Africa Demographic and Health Survey (SADHS). (2016). Selected health indicators among the elderly: Findings from the SADHS 2016 GHS 2016 and MACoD 2016 Report 03-00-13
- Statistics South Africa. (2011). Use of health facilities and levels of selected health conditions in South Africa: Findings from the General Household Survey. Statistics South Africa. Pretoria: Statistics South Africa; 2013 Available from: http://www.statssa.gov.za/publications/Report-03-00-05/Report-03-00-052011.pdf [accessed 09 July 2018].
- Statistics South Africa. (2016). General Household Survey. Statistical release P0318. Pretoria: Statistics South Africa; 2016
- Statistics South Africa. (2016). Mortality and causes of death in South Africa 2016. Findings from death notifications. Statistical release P03093. Pretoria: StatsSA 2018. Available from: http://www.statssa.gov.za/publications/P03093/P030932016.pdf [accessed 06 July 2018].
- Surgeon General Report. (2014). Smoking And Diabetes. Available from: https://www.cdc.gov/tobacco/data_statistics/sgr/50th-anniversary/pdfs/fs_smoking_diabetes_508.pdf [accessed 16 Jan 2019].
- Tavsanli Nurgül & Köken Seçil & Akıncıoğlu Selin & Artuvan Havva & Nerse Zeynep. (2018). Results of scanning for diabetes obesity and blood pressure: a sample from manisa in western turkey. *Journal of Diabetes Metabolic Disorders & Control*. 5. 20-24. 10.15406/jdmdc.2017.05.00133.

- Thomas D.R. and Rao J.N.K. (1987).Small-sample comparisons of level and power for simple goodness of fit statistics under cluster sampling *Journal of the American Statistical Association* Vol.82No.398. pp. 630-636.
- Thomas R. L. Distiller L. Luzio S. D. Chowdhury S. R. Melville V. J. Kramer B. & Owens D. R. (2013). Ethnic differences in the prevalence of diabetic retinopathy in persons with diabetes when first presenting at a diabetes clinic in South Africa. *Diabetes care 36*(2) 336–341. https://doi.org/10.2337/dc12-0683
- Tyrovolas S. Koyanagi A. Garin N et al. 2015. Diabetes mellitus and its association with central obesity and disability among older adults: a global perspective. *Exp Gerontol*. 2015;64:70-77. doi:10.1016/j.exger.2015.02.010
- Tyrovolas S. Koyanagi A. Garin N. Olaya B. Ayuso-Mateos J. L. Miret M. Chatterji S. Tobiasz-Adamczyk B. Koskinen S. Leonardi M. & Haro J. M. (2015). Diabetes mellitus and its association with central obesity and disability among older adults: a global perspective. *Experimental gerontology 64* 70–77. https://doi.org/10.1016/j.exger.2015.02.010
- University of Witwatersrand (WITS). 2017. *SA has highest blood pressure in Southern Africa*. 15 June 2017 from: https://www.wits.ac.za/news/latest-news/research-news/2017/2017-06/sa-has-highest-blood-pressure-in-southern-africa.html [accessed 20 July 2018].
- Vellios N.G. & Van Walbeek C.P. (2017). Self-reported alcohol use and binge drinking in South Africa: Evidence from the National Income Dynamics Study 2014 - 2015. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde 108*(1) 33–39. https://doi.org/10.7196/SAMJ.2017.v108i1.12615
- Warren G.W. Alberg A.J. Kraft A.S. Cummings K.M. The 2014 Surgeon General's report: "The health consequences of smoking--50 years of progress": a paradigm shift in cancer care. *Cancer. 2014* Jul 1;120(13):1914-6. doi: 10.1002/cncr.28695. Epub 2014 Mar 28. PMID: 24687615; PMCID: PMC5928784.
- Werfalli M. Kassanjee R. Kalula S. Kowal P. Phaswana-Mafuya N. & Levitt N.S. (2018). *Diabetes in South African older adults: prevalence and impact on quality of life and functional disability*. Global Health Action Volume 11.
- Whiting D.R. Guariguata L. Weil C. Shaw J. IDF (2011). IDF Diabetes Atlas: Global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Atlas. International Diabetes Federation Brussels Belgium .Baker IDI Heart and Diabetes Institute Australia

- Wild S. Roglic G. Green A. Sicree R. & King H. (2004). Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes care 27*(5) 1047–1053. https://doi.org/10.2337/diacare.27.5.1047
- Wolter K. M. 2007. Introduction to Variance Estimation. 2nd ed. New York: Springer.
- Woolard Ingrid & Leibbrandt Murray & Villiers L. (2010). The South African National Income Dynamics Study: Design and methodological issues. *Journal for Studies in Economics and Econometrics*. 34. 7-24.
- World Health Organization (WHO). (2016). Global report on diabetes. Available from: apps.who.int/iris/bitstream/10665/204871/1/9789241565257_eng.pdf [accessed 20 July 2018].
- World Health Organization (WHO). (2018). Diabetes key facts sheet (online) Available from: http://www.who.int/news-room/fact-sheets/detail/diabetes [accessed 07 July 2018].
- World Health Organization (WHO). (2019). The dual epidemic of TB and Diabetes (online). 2019; Available from: https://www.who.int/tb/publications/diabetes_tb.pdf [accessed 16 Jan 2019].
- World Health Organization (WHO). 2018. Sustainable Development Goals (SDGs) Global Action Plan
- Zhang N. Yang X. Zhu X. Zhao B. Huang T. Ji Q. (2017).Type 2 diabetes mellitus unawareness prevalence trends and risk factors: National Health and Nutrition Examination Survey (NHANES) 1999-2010. *J Int Med Res*. 2017;45(2):594-609. doi:10.1177/0300060517693178

APPENDICES

APPENDIX I: Plagiarism Report

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I MASEDI MENYATSOE (Student number: 0706913V) am a student registered for the degree of MSC BIO Stats in the academic year 2020

I hereby declare the following:

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

Signature: [signature] Date: 14/09/2020

APPENDIX II: Ethics approval

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R14/49 Ms Masedi Menyatsoe

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M190323

NAME: (Principal Investigator)	Ms Masedi Menyatsoe
DEPARTMENT:	Public Health
PROJECT TITLE:	Risk factors associated with diabetes in the South African people aged 15 years and older for the period 2011-2012
DATE CONSIDERED:	29/03/2019
DECISION:	Approved
CONDITIONS:	Permission from HSRC
SUPERVISOR:	Prof Jonathan Levin
APPROVED BY:	Doctor CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	05/04/2019

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

DECLARATION OF INVESTIGATORS

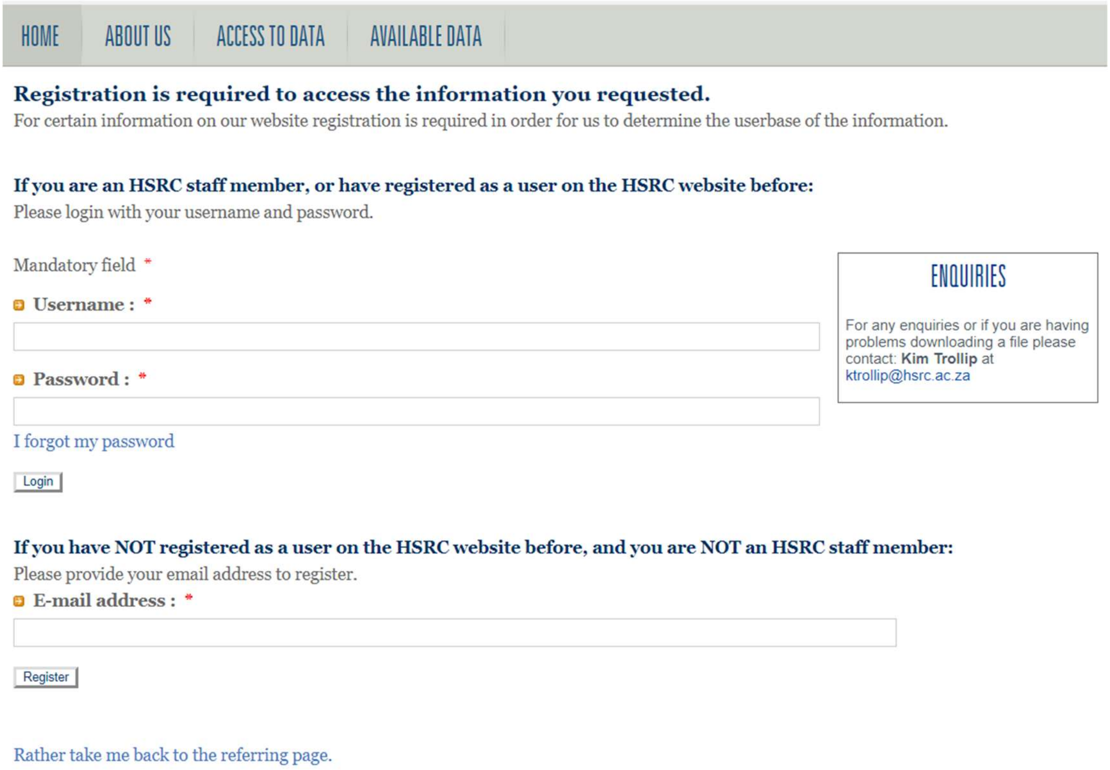
To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the Third Floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are 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**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in March and will therefore be due in the month of March each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

11/02/2021
Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

APPENDIX III: HSRC website link http://curation.hsrc.ac.za/Datasets-XKAHAA.phtml).
Public access to the website with username and password.



HOME | ABOUT US | ACCESS TO DATA | AVAILABLE DATA

Registration is required to access the information you requested.

For certain information on our website registration is required in order for us to determine the userbase of the information.

If you are an HSRC staff member, or have registered as a user on the HSRC website before:

Please login with your username and password.

Mandatory field *

Username : *

Password : *

I forgot my password

Login

ENQUIRIES

For any enquiries or if you are having problems downloading a file please contact: **Kim Trollip** at ktrollip@hsrc.ac.za

If you have NOT registered as a user on the HSRC website before, and you are NOT an HSRC staff member:

Please provide your email address to register.

E-mail address : *

Register

Rather take me back to the referring page.

APPENDIX IV: After logging in using username and password

Logout of Research Data Service | CONTACT US

HSRC Human Sciences Research Council

Research Data Service

Welcome, you are currently logged in with email address menyatsoe@gmail.com

HOME | ABOUT US | ACCESS TO DATA | AVAILABLE DATA

You are now logged in. Please click again the resource that you require.

Data sets

South African National Health and Nutrition Examination Survey (SANHANES-1) 2011-12

The primary objectives of the SANHANES-1 were to assess defined aspects of the health and nutritional status of South Africans with respect to the prevalence of NCDs (specifically cardiovascular disease, diabetes and hypertension) and their risk factors (diet, physical activity and tobacco use):

- The knowledge, attitudes and behaviour of South Africans with respect to NCDs and tuberculosis;
- The nutritional status of South Africans as it relates to food security, dietary intake/ behaviour including alcohol consumption, body image and weight management;
- The perceptions of general and mental health (stress and trauma) and the utilisation of healthcare services;
- The health status of children under the age of five years;
- The health status of children aged 2-9 years with respect to physical and/or mental disabilities
- The behavioural (smoking, diet, physical inactivity) and social determinants of health and nutrition (demographic, socio-economic status and locality) and relate these to the health and nutritional status of the population.

- South African National Health and Nutrition Examination Survey (SANHANES-1) 2011-12: Adult - All provinces

The data set for dissemination contains 728 variables and 17 926 cases of respondents aged 15 years and older who participated in the SANHANES-1 Adult Questionnaire.

The questionnaire covers the following sections: geographic information, biographic details of the respondent, non-communicable diseases, tuberculosis, nutrition, perceptions of respondent's general and mental health, as well as health care utilisation.

- South African National Health and Nutrition Examination Survey (SANHANES-1) 2011-12: Adult Physical Examination - All provinces

APPENDIX V: After logging access is granted with two datasets available. One for adult examination and the adult questionnaire.

HOME | ABOUT US | ACCESS TO DATA | AVAILABLE DATA

Data files

Data files related to South African National Health and Nutrition Examination Survey (SANHANES-1) 2011-12: Adult Physical Examination - All provinces

All data sets | Data set details | Documentation | Data files | Outputs | Access conditions | Contact

Data files

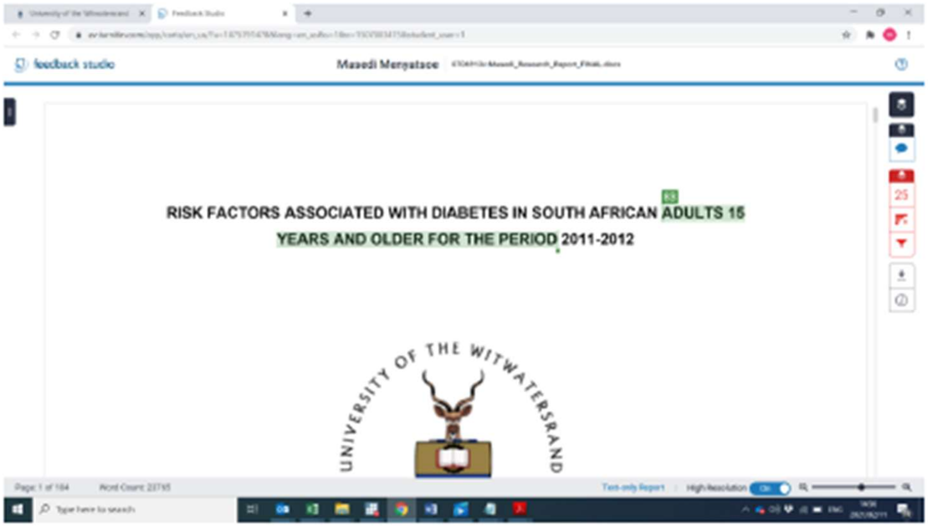
It is advisable to study the introductory information before using the data or related documents as it provides a systematic exposition of what the collection entails and how it should be used.
Note: Old versions of web browsers such as Internet Explorer 6 could cause problems when downloading files.

Download	Access	File	Description
ASCII FIXED FORMAT	✓	SANHANES2011_12_Adult_Exam.dat	Save file to disk and use as input file for the different programs.
COMMA DELIMITED	✓	SANHANES2011_12_Adult_Exam.csv	Save file to disk and open in any text editor or spreadsheet. Refer to the User guide for additional information.
SAS DATA SET	✓	SANHANES2011_12_Adult_Exam.sas7bdat	Save file to disk. See user guide on how to reference the data set and formats in a SAS program.
SAS FORMATS	✓	SANHANES2011_12_Adult_Exam.sas7bcat	Save file to disk. See user guide on how to use the formats in a SAS program.
SAS PROGRAM	✓	SANHANES2011_12_Adult_Exam.sas	Save file to disk and open in SAS.
SPSS DATA SET	✓	SANHANES2011_12_Adult_Exam.sav	Save file to disk and open in SPSS.
SPSS PROGRAM	✓	SANHANES2011_12_Adult_Exam.sps	Save file to disk and open in SPSS.
STATA DATA SET	✓	SANHANES2011_12_Adult_Exam.dta	Save file to disk and open in STATA.
STATA FORMATS	✓	SANHANES2011_12_Adult_Exam.dct	Save file to disk in folder as indicated in .DO file. Refer to User guide.

APPENDIX VI: Turnitin Report

Studen name: Masedi Menyatsoe

Student no: 0706913v



Supervisor 1

Name: Professor Jonathan Levin

Signature: 

Date: 20 February 2021

Supervisor 2

Name: Professor Khangelani Zuma

Signature:______________

Date:____|___