

https://dspace.org/

Risk factors associated with pancreatic cancer at two Johannesburg Academic Hospitals between 2013 and 2015

Kagodora, Shingirai Brenda

2017
https://hdl.handle.net/10539/24868

Downloaded from a DSpace repository. https://dspace.org/

0405483M

Risk factors associated with pancreatic cancer at two Johannesburg Academic Hospitals between 2013 and 2015

Shingirai Brenda Kagodora

Student number: 0405483M



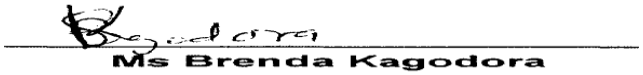
A research report submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirement for the degree of Master of Science in Epidemiology and Biostatistics.

November 2017

0405483M

Declaration

I, Shingirai Brenda Kagodora declare that this research report is my own work, compiled under the supervision of Professor CS Chasela and Dr M Brand. The report is being submitted to the University of the Witwatersrand in partial fulfilment of a degree of Master of Epidemiology in the field of Epidemiology and Biostatistics. The material contained in this research report has not been submitted for any other degree or examination in this university or any other university.



Signature

Shingirai Brenda Kagodora

Date: 3 November 2017

I certify that this study has the approval of the Human Research Ethics Committee of the University of the Witwatersrand, Johannesburg, South Africa. The study ethics number is M160247

Signature 

Shingirai Brenda Kagodora

Date: 3 November 2017

DEDICATION

To all people that are dear and matter most in my life. May you be inspired to achieve your greatest dreams and thank you for standing by my side!

ACKNOWLEDGEMENTS

I would like to thank God almighty for His enabling love and power for my existence. Christinah, Rumbidzai and Munashe I am forever indebted to your kindness and your belief and trust in me. I love and cherish you. I would also like to express my sincere gratitude to my supervisors, Professor Charles Chasela and Dr Martin Brand for their guidance and insight throughout the compilation of this research report. Thank you for your inputs and mentorship. To the research team and patients who contributed the information, thank you for allowing me the opportunity to use this data.

I also thank my family and friends for their support and cheers when the going got tough. To the entire team in the School of Public Health, you were great and you trained and moved me to a level of knowledge that no one can take away from me! I also thank Prof Jonathan Levin, Dr Marietha Nel and Prof Geoffrey Candy. Lastly, I thank the Department of Surgery for recognizing my potential to move to another hierarchy in the academic world.

ABSTRACT

Introduction: Pancreatic ductal adenocarcinoma (PDAC) is a devastating diagnosis for anyone as it is associated with a global mortality rate of about 4%, and has few therapeutic interventions that prolong survival as compared to other cancers. Frequent epidemiological reports on PDAC are available in the developed countries, but in South Africa, there is a paucity of epidemiological data on this aggressive cancer. Understanding risk factors will help to assess and develop relevant interventions for asymptomatic high-risk patient populations.

Aim: To investigate and explore how various risk factors were associated with PDAC at two public academic hospitals in Johannesburg between 2013 and 2015.

Method: This was a secondary unmatched case-control study to assess risk factors for developing PDAC at two public academic hospitals, namely the Chris Hani Baragwanath Academic Hospital (CHBAH) and the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH). All cases of PDAC were histologically and/or cytologically confirmed. All participants were >18 years of age, including 139 cases and 139 controls. Data exported from REDCap database included patient demographics and social and medical histories. Proportions used the chi-square test and bivariate logistic regression estimated ORs between individual variables and PDAC. Multivariate logistic regression analysis investigated all possible confounders present in the data. The likelihood ratio test with a p-value of <0.20 was accepted to assimilate data fitting into the model.

Results: Eighty two percent of the study population was black. The 50-59 age group accounted for 37% of the cases. Multiple logistic regressions showed the following odds ratios 95% CI and p-values for ages (i) 20-29 [0.11(0.11-1.00) p=0.05] and (ii) 50-59 [2.63(1.03-6.70) p=0.04]. As for diet, the following odds were observed (i) high white meat [0.18(0.04-0.86) p=0.03], (ii) low fish intake [2.17(1.06-4.45) p=0.03], (iii) low consumption of fried food [0.48(0.23-1.00) p=0.05] and (iv) high consumption of vegetables [0.17(0.05-0.61) p=0.007]. In terms of occupation, general workers had the following likelihood [1.79(0.93-3.45) p=0.08] of developing PDAC.

Conclusion: Being 50-69 years of age and employed for longer periods than the general norm, was positively associated with PDAC. Additionally, increased consumption of vegetables and white meat was protective against PDAC, whilst a low intake of fish increased PDAC risk.

Keywords: Pancreatic cancer, risk factors, epidemiology and case-control.

TABLE OF CONTENTS

1. Introduction 1
1.1 Background Information 1
1.2 Problem Statement 1
1.3 Justification 2
1.4 Literature Review 2
1.4.1 Incidence of PDAC 2
1.4.2 Risks Factors of PDAC 3
1.4.2.3 Lifestyle 4
1.5 Research Question 6
1.6 Aim 6
1.7 Objectives 6
2 Methods 8
2.1 Primary study 8
2.2 Study site 8
2.3 Study population 8
2.4 Sample size and power 9
2.5 Data collection 9
2.6 Data management and analysis 17
2.6.1 Data management 17
2.6.2 Statistical analysis 17
2.7 Ethical considerations 18
2.8 Study Budget 18
3. Results 19
3.1 Introduction 19
3.1.1 Demographic characteristics of study population 23

3.1.2 Social characteristics of study population....23
3.1.3 Chronic Medical characteristics of study population....23
3.1.4 Dietary characteristics of study population....24
3.1.5 Hospital characteristics for the study population....24
3.2.1 Adjusted and unadjusted demographic factors associated with PDAC....24
3.2.3 Adjusted and unadjusted dietary factors associated with PDAC....24
3.2.3 Adjusted and unadjusted social factors associated with PDAC....24
3.2.5 Final model....25
4. Discussion....28
4.1 Introduction....28
4.2 Risk factors for patients with PDAC....31
4.2.1 Age and Gender....31
4.2.2 Alcohol....31
4.2.3 Diet....32
4.2.4 Employment....32
4.3 Study limitations....33
5. Conclusions and recommendations....34
5.1 Introduction....34
5.2 Conclusions....34
5.3 Recommendations....34
5.4 Future studies....35
6 References....36
7. Appendix....41

List of abbreviations

AIDS - Acquired Immune Deficiency Syndrome

AOR - Adjusted Odd Ratio

BMI - Basal metabolic index

CAGE - Cut Annoyed Guilty Eye-opener

CHBAH - Chris Hani Baragwanath Academic Hospital

CMJAH - Charlotte Maxeke Johannesburg Academic Hospital

COPD - Chronic Obstructive Pulmonary Disease

CP - Chronic Pancreatitis

CT scan - Computerized (or computed) tomography scans

DM - Diabetes mellitus

H Pylori - Helicobacter pylori

HIV- Human Immune Deficiency Virus

IBD - Inflammatory bowel disease

GIT - Gastro-intestinal-tract

LC-PUFA - Long-Chain Polyunsaturated Fatty Acids

NCD - Non-Communicable Disease

NRF - National Research Foundation

OR - Odds Ratio

PDAC - Pancreatic Ductal Adenocarcinoma

REDCap- Research Electronic Data Capture

SD - Standard Deviation

SEER - Surveillance, Epidemiology, and End Results

Stata - Statistical analysis programme

TB - Tuberculosis

UOR - Unadjusted Odds Ratio

WHO - World Health Organization

List of figures

Figure 1. Flow chart of proceedure embarked on for data collection.....10
Figure 2 Letter from the gatekeeper 41
Figure 3 Letter to the gatekeeper 42
Figure 4 Ethic clearance.....44

List of tables

Table 1. Exposure variables and measurements 11
Table 2. Objectives of the study and corresponding analysis undertaken 17
Table 3 Baseline characteristics 20
Table 4 Diet 21
Table 5 Social characteristics 22
Table 6 Unadjusted and Adjusted OR for PDAC 26
Table 7 Demographic characteristics of study population 45
Table 8Social characteristics of study population 47
Table 9 Chronic Medical characteristics of the study population 49
Table 10Dietary characteristics of study population 51
Table 11Hospital characteristics for the study population 53
Table 12Factors associated with PDAC 54
Table 13Chronic medical factors associated with PDAC 56
Table 14Dietary factors associated with PDAC 58
Table 15Social factors associated with PDAC 60
Table 16Hospital factors associated with PDAC 62
Table 17Final model 63

1. Introduction

This chapter discusses the background information regarding pancreatic ductal adenocarcinoma (PDAC), explains the results reported in the literature, describe the problem statement of the study and the justification for this research. The chapter ends with the research question statement and a description of the aim and objectives of the study.

1.1 Background Information

Cancer is one of the most important health problems worldwide due to its high mortality rate (Ferlay, Soerjomataram et al. 2012), primarily as a result of its ability to spread from a primary site to distant organs through a process known as metastasis (Khan and Mukhtar, 2010). Pancreatic cancer commonly known as pancreatic ductal adenocarcinoma (PDAC) is globally ranked thirteenth among all cancers in terms of incidence, but ranked seventh in terms of mortality (Kuzmickiene, Everatt et al. 2013). PDAC is the most common epithelial pancreatic malignancy (Hariharan, Saied et al. 2008) and accounts for the highest percentage (greater than 80%) of all malignant pancreatic tumours (Alexakis, Halloran et al. 2004). In both genders PDAC has an almost equal number of incidences and annual mortality rate (Are, Chowdhury et al. 2016). There is limited data on PDAC in Africa probably due to lack of infrastructure and resource constraints as evidenced by poorly developed primary healthcare systems with its resulting negative impact on referral patterns, as well as the scarcity of well-developed tertiary care and multidisciplinary centres (Shrikhande, Barreto et al. 2012). Furthermore, there is no effective screening test for PDAC and therefore, it is often diagnosed at an advanced stage, contributing to a 5 year survival rate of less than 5% (Fest, Ruiter et al. 2017).

The anatomical location of the occurrence of PDAC contributes to the reason for late presentation of patients. The early stages are asymptomatic thereby making diagnosis difficult (Miroslaw, Sekula et al. 2012), thus in most patients a high index of clinical suspicion is required with the non-specific symptoms of painless obstructive jaundice, weight loss and back pain (Alexakis, Halloran et al. 2004). Other associated signs and symptoms such as late-onset diabetes mellitus may suggest PDAC as well as persistent non-orthopaedic associated back pain, marked rapid weight loss, an epigastric abdominal mass, ascites and supraclavicular lymphadenopathy may all indicate a potentially advanced tumour (Alexakis, Halloran et al. 2004).

1.2 Problem Statement

PDAC is a devastating diagnosis for any patient, as it is associated with a mortality rate of 4 per 100 000 population and few therapeutic interventions that prolong survival compared to

other cancers. The epidemiological data for PDAC have mostly been reported in developed countries. In South Africa however, there is paucity of epidemiological data on this aggressive cancer (Soliman , Zhang et al. 2006). The most recent National Cancer Registry reports a PDAC incidence for the year 2012 and there has been no update since then (Herbst and Joubert 2017). This may give the misconception that the incidence of PDAC is rare in South Africa. In the past decade the epidemics of HIV and TB re-directed resources from NCDs, which resulted in low-cost passive surveillance activities with a 10 year backlog of incidence reports (Singh, Ruff et al. 2015). Besides the backlog the pathology-based data may be clinically underestimated as this data is not linked to mortality data (Singh, Ruff et al. 2015). The other challenges of cancer registries in Africa include: financial constraints for implementation of registries according to the WHO guidelines and lack of trained personnel for long-term sustainability (Singh, Ruff et al. 2015). Since PDAC is not one of the top five causes of death attributed to cancer, not much resource allocation is given. Updated epidemiological data concerning this disease in South Africa will improve the understanding of it and instigate primary health care initiatives that may address issues unique to South Africa.

1.3 Justification

Understanding risk factors associated with the disease will help to assess and develop relevant interventions, as well as allocate primary health care resources for patients at risk. A well-managed research programme, mitigation of the disease impacts and implementation of a contingency plan especially for the less privileged communities in South Africa, would be of value. Diagnosis of PDAC at an advanced stage and knowing the potential factors associated with it will provide information for better screening and management of potential future patients. Early diagnosis may increase the survival of PDAC patients and this study may increase awareness of the disease in the South African population. This study provides an opportunity for the monitoring of risk for developing PDAC within the population and age groups at risk.

1.4 Literature Review

In this section, a review of the incidence rate and risk factors for PDAC will address a gap in the knowledge of the disease in the specific context of South Africa.

1.4.1 Incidence of PDAC

PDAC is a lethal malignancy that accounts for approximately 4% of cancer deaths worldwide (Ferlay, Soerjomataram et al. 2012, Are, Chowdhury et al. 2016). In the United States of America (US), the incidence rate is 2.7% of all new cancer diagnosis. The global

incidence rate of PDAC for all age groups of both sexes is 4.2 per 100 000 with a mortality rate of 4.0 per 100 000 population. The six different WHO regions (Africa, America, Eastern Mediterranean, Europe, South East Asia and Western Pacific) have varying incidence rates. The incidence rates for all age groups of both sexes are as follows: Africa 1.8 per 100 000 population, America 5.9 per 100 000 population, Eastern Mediterranean 1.9 per 100 000 population, Europe 6.5 per 100 000 population, South East Asia 1.5 per 100 000 population and Western Pacific 4.4 per 100 000 population (Ferlay, Soerjomataram et al. 2012, Are, Chowdhury et al. 2016). The incidence rate of PDAC in relation to socio-economic development for all age groups of both sexes varies from 1.2 per 100 000 population in the low socio-economic development category to as high as 7.2 per 100 000 population in the very high socio-economic development category (Ferlay, Soerjomataram et al. 2012, Are, Chowdhury et al. 2016). In 2012 PDAC accounted for 0.52% in all males and 0.39%% in all females as a percentage of all cancers in South Africa (Herbst and Joubert 2017). In 2014, the WHO reported the death rate of PDAC as 4.34 per 100 000 population for South Africa and South Africa was thus ranked 67th for PDAC incidence in the world (Le Duc 2017). The PDAC incidence rate for South Africa in 2012 for all age groups of both sexes, according to Globocan, was 4.7 per 100 000 population with a mortality rate of 4.6 per 100 000 population (Ferlay, Soerjomataram et al. 2012).

1.4.2 Risks Factors of PDAC

PDAC is caused by both genetic coding (such as inherited mutations) and environmental/acquired factors (such as tobacco, diet, radiation, and infectious organisms) (Anand, Kunnumakara et al. 2008, Kuzmickiene, Everatt et al. 2013). Established risk factors for PDAC include smoking (Iodice, Gandini et al. 2008, Lynch, Vrieling et al. 2009, Bosetti, Lucenteforte et al. 2012), alcohol consumption (Michaud, Vrieling et al. 2010, Lucenteforte, La Vecchia et al. 2012), diabetes mellitus (Chari, Leibson et al. 2005, Huxley, Ansary-Moghaddam et al. 2005, Chari, Leibson et al. 2008), obesity (Tang, Wei et al. 2014) and diseases associated with chronic inflammation (Michaud, Vrieling et al. 2010). Chronic inflammation is a recognised factor in the initiation of carcinogenesis (Sobhani, Amiot et al. 2013) and as such is evident in the development of various diseases such as: inflammatory bowel disease (IBD), colon cancer, *Helicobacter pylori* induced gastritis, gastric cancer (Correa, Haenszel et al. 1990), chronic pancreatitis and PDAC (Duell, Lucenteforte et al. 2012). Other risk factors include age, occupation, gender and race (Muniraj, Jamidar et al. 2013).

1.4.2.1 Age

The risk of PDAC increases with age, with most cases from 45 years upwards (Ries, Reichman et al. 2003). Indeed, advancing age is the main risk factor for developing PDAC, a factor not uncommon to most cancers (Yeo and Lowenfels 2012). Data from the Surveillance, Epidemiology, and End Results (SEER) registries show the median age of PDAC diagnosis to be 71 years, where less than 3% of cases were diagnosed before age 44 and 54% of cases between 65 and 84 years of age (Yeo and Lowenfels 2012). The less than 0.5% of cases in the age range of 20 years and below, may be attributed to family history or to genetic factors in those at risk of developing PDAC (Yeo and Lowenfels 2012).

1.4.2.2 Gender and Race

The incidence rate of PDAC is higher in men than in women. This difference may be due to higher tobacco use in men than in women (Yeo and Lowenfels 2012). In addition, men tend to consume more fried, grilled, or barbecued meat than women and this increase the risk of PDAC (Pericleous, Rossi et al. 2014). In the US, the incidence of PDAC is higher in African Americans compared with Caucasians. The differences in incidences reported in gender and in race may be due to higher rates of the associated risk factors such as diabetes and obesity in women, and smoking in men (Yeo and Lowenfels 2012).

1.4.2.3 **Lifestyle** – smoking, alcohol, weight and diet

Smoking is the strongest risk factor for PDAC (Duell, Lucenteforte et al. 2012). It is believed to cause 30% of all cases (Konner and O'Reilly 2002). A meta-analysis by Iodice *et al.*, (2008) indicated that current cigarette smokers, compared with non-smokers, have approximately an 1.7 risk ratio of PDAC and this increases as the number of cigarettes smoked and the number of years of smoking increases (Iodice, Gandini et al. 2008). Tobacco smoking has the greatest risk. The other risk linked to tobacco include smokeless tobacco (chewing tobacco) and environmental tobacco smoke, although it appears discordant or negative (Maisonneuve and Lowenfels 2015).

Heavy and moderate alcohol consumption has been found to have an effect on the risk of PDAC development through the activation of inflammatory pathways in chronic pancreatitis (CP) (Duell, Lucenteforte et al. 2012). Although CP has a low prevalence, alcohol consumption leads to progressive and irreversible tissue destruction following inflammation (Michaud, Vrieling et al. 2010, Lucenteforte, La Vecchia et al. 2012). The incidence of alcohol induced CP resulting in PDAC is 3-5% of all PDAC cases (Konner and O'Reilly 2002). A heavy drinking pattern (>80g alcohol/day, or more than 5–6 drinks/day.)

comprising of wine, beer and other liquor increases the risk for developing PDAC (Ruiz and Hernández 2014)

Being overweight or obese increases the risk of PDAC development due to an increased production of hormones, inflammatory markers and growth factors (Li, Xie et al. 2004, Ruiz and Hernández 2014). Increases in circulating insulin and C-peptide, hyperglycaemia, insulin resistance and diabetes account for a possible development of PDAC. Extra weight around the waistline especially in women, may also be a risk factor of PDAC (Li, Xie et al. 2004). Generally, there is an increasing risk of PDAC with decreasing physical inactivity (Behrens, Jochem et al. 2015).

It has been suggested that diet plays a crucial role in predisposing an individual to PDAC (Al-Majeda, El-Basmib et al. 2013). However, there is conflicting evidence for the effect of diet on PDAC risk (Neale, Clarka et al. 2014). In a review, high red meat (beef, lamb, goat, venison, etc.) consumption increased the risk of developing PDAC, particularly in men. On the other hand, multiple studies have observed that consuming foods of high vegetable, fruit and whole grain content confer protection against PDAC (Neale, Clarka et al. 2014). A healthy eating pattern which includes a high content of fruits, vegetables, poultry, fish, whole grain and low daily intake of fat has a protective effect from PDAC development (Ruiz and Hernández 2014). Processed and smoked foods such as ham, sausages, bacon, burgers, foods high in fat and refined sugars increase the risk of developing PDAC (Jansen, Robinson et al. 2014, Ruiz and Hernández 2014).

1.4.2.4 Some medical conditions

1.4.2.4.1 Chronic pancreatitis (CP)

Chronic pancreatitis (CP) increases the risk of PDAC. Indeed, a review article shows the progression of CP over a period of 20 years until diagnosis of PDAC (Pinho, Chantrill et al. 2014). The K-ras mutation observed in CP is implicated in the activation and the progression to PDAC (Yeo and Lowenfels 2012). Furthermore, the digestive enzyme-secreting acinar cells undergo ductal metaplasia in the inflammatory environment of pancreatitis, and this metaplastic change is recognized as a precursor of PDAC (Pinho, Chantrill et al. 2014).

1.4.2.4.2 Diabetes Mellitus (DM)

Strong evidence exists that DM is associated with PDAC. Type 2 DM is probably second (just after cigarette smoking) on the list of the top five modifiable risk factors for PDAC. Epidemiological investigations found that long-term Type 2 DM is associated with a 1.5-fold to 2.0-fold increase in the risk for developing PDAC (Li 2012). Long-term diabetes was

observed to account for more than 50% of all PDAC sufferers (Maisonneuve and Lowenfels 2015).

1.4.2.5 Infections

1.4.2.5.1 Hepatitis B

There is strong evidence that hepatitis B is associated with PDAC, since hepatitis B positive carriers showed a relative risk of between 1.2-3.8 for developing PDAC (Maisonneuve and Lowenfels 2015). The virus tends to infect the pancreas and liver of chronic hepatitis B carriers (Yeo and Lowenfels 2012). However, both hepatitis B and hepatitis C may be involved in a process of oncogenesis through development of local inflammation in the pancreas. The pancreas serves as a reservoir for the replication of the virus which in turn causes necro-inflammation in the pancreas (Fiorino, Cuppini et al. 2013).

1.4.2.5.2 *Helicobacter pylori* (*H pylori*) and blood group

Helicobacter pylori infection increases the risk of PDAC development (Fiorino, Cuppini et al. 2013). *Helicobacter pylori* and peptic ulcers activate N-nitrosamines which may cause DNA damage resulting in progression to PDAC (Yeo and Lowenfels 2012). Interestingly, in the presence of duodenal ulcers and *H Pylori*, all blood groups except the O blood group, show a risk of developing PDAC (Yeo and Lowenfels 2012). Indeed, blood group A individuals has the worst survival rates of all PDAC sufferers (Kos, Civelek et al. 2012).

1.4.2.6 Occupational exposures

Exposure to chlorinated hydrocarbons and related solvent compounds are among the major occupational risk factors for development of PDAC. Additionally, nickel plating and formaldehyde exposure cause moderate risk of PDAC (Maisonneuve and Lowenfels 2015). Furthermore, Yeo *et al.* (2012) observed that exposure to asbestos, pesticides, herbicides, residential radon, coal products, welding products and radiation are all associated with PDAC (Yeo and Lowenfels 2012).

1.5 Research Question

What are the risk factors associated with PDAC in patients admitted at two public academic hospitals in Johannesburg between June 2013 and December 2015?

1.6 Aim

The aim of this study was to investigate and explore how various risk factors were associated with PDAC at two public academic hospitals in Johannesburg between 2013 and 2015.

1.7 Objectives

1. To describe demographic, social, hospital and dietary characteristics of patients with PDAC in two public academic hospitals in Johannesburg.

2. To assess the risk factors associated with patients diagnosed with PDAC in two public academic hospitals in Johannesburg.

2 Methods

This chapter describes the methods used to collect and manage data, perform statistical analysis and the ethical considerations. First, the chapter describes the study design, the setting, and how the sample size was calculated. It further explains the study population.

The study used an unmatched case-control design to assess detailed information on risk factors for PDAC between June 2013 and December 2015. Data used was from a primary study, which was a hospital based, case-control study that is on-going. Participant's demographics, social and medical history from study questionnaires captured in the REDCap (Research Electronic Data Capture) programme (Harrisa, Taylorb et al. 2009) was imported into Stata 13 and Excel. For this study, data was analysed as explained by **Figure 1**.

2.1 Primary study

The primary study focused on the genetic and environmental factors that influence susceptibility to PDAC. This included evaluation of the possible associations between the inter-individual genetic variation and the risk of developing PDAC, disease progression and the survival of the patients as well as their response to the treatment. More specifically the study objectives were:

- To identify new genetic risk factors for PDAC, in addition to those identified to date.
- To describe genetic factors which influence the outcome of treatment of PDAC patients.
- To assess the genetic factors which influence the survival of PDAC patients.

2.2 Study site

The study sites were two public academic hospitals in the Johannesburg area, namely: the Chris Hani Baragwanath Academic Hospital (CHBAH) in Soweto and the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) in Parktown. These are referral hospitals in Johannesburg and are accessible to the public. Both hospitals have a specialised gastro intestinal tract (GIT) clinic frequented by all patients with GIT diseases, including all PDAC patients, for diagnostic and therapeutic services.

2.3 Study population

The principal study included participants older than 18 years. This secondary study used a selection of the principal study population but with its own set of inclusion and exclusion criteria as follows:

Inclusion criteria

Patients >18 years, admitted at two public Johannesburg academic hospitals, with abdominal CT scans demonstrating PDAC and/or patients with cytological and/or histological diagnoses of PDAC from June 2013-December 2015.

Exclusion criteria

Patients who have not had a CT-scan.

Controls

A case-cohort sampling method allowed the selection of cases and controls independently but at same time.

Inclusion criteria

Patients >18 years, admitted at two public Johannesburg academic hospitals with abdominal CT scans demonstrating a normal pancreas. The controls included trauma patients, abdominal aortic aneurysm patients, and patients with acute abdomens from causes other than any cancer. The control group is a reflection of the study population with the exception of a normal pancreas. Controls were generally easily identified and they were cooperative. Controls resembled the PDAC cases with respect to their tendency to give complete and accurate information, thus reducing potential differences between cases and controls in the quality of their recall of past exposures.

Exclusion criteria

Patients who have not had a CT-scan.

2.4 Sample size and power

The power of the study at 80% using expected proportions of the smoking risk factor in the case group as 0.35, and control group as 0.20, at a 95% confidence level with an OR of 1.7 yielded a sample size of approximately 278 (139 cases and 139 controls) for the period of June 2013- December 2015. Smoking was used as a risk factor for sample size calculation based on the odds ratio of 1.7 between smokers and non-smokers.

2.5 Data collection

Records of participant's demographics, social and medical history was captured in REDCap and imported to Stata 13 and Excel as summarized in **Figure 1**.

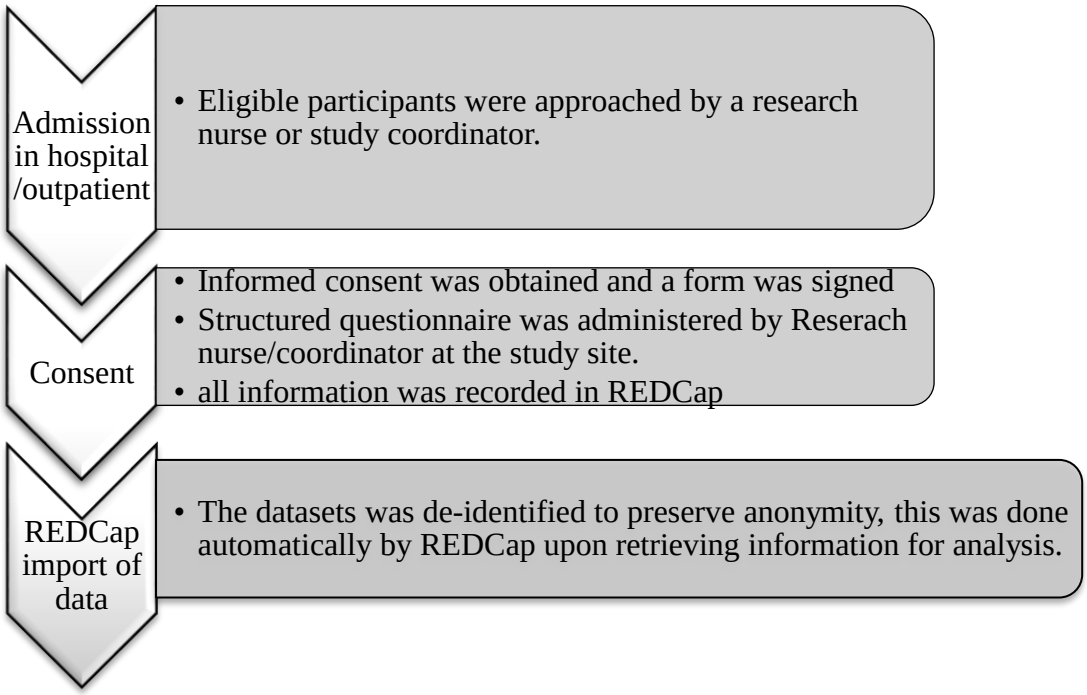


Figure 1. Flow chart of proceedure embarked on for data collection

The following variables were used:

Table 1. Exposure variables and measurements

Exposure variable	Field label	Measurement
Race	Patient race	Black White Indian Coloured Other
Gender	Patient gender	Female Male
Age years	Age at registration	Mean SD
Age	Age category	20-29 30-39 40-49 50-59 60-69 70-79 80+
Schooling	Highest level of education	Never attended school Primary school Secondary school Further Education
Birth place	Province of birth	Gauteng Mpumalanga Limpopo North West Free State KwaZulu-Natal Eastern Cape Northern Cape Western Cape

		Other Country
Employment status	Current Employment	No Yes
Longest job employed	Primary type of employment	Construction & chemical Mining Manufacturing & Factory Engineering Drivers General worker Office & other
Smoked	Have you ever smoked cigarettes?	No Yes
Smoke now	Do you smoke now?	No Yes
Smoke pack years	Number of pack years	Mean SD
Smoke pack category	Pack years category	No (0) Very rare (1-10) Frequent(11-30) Very frequent(31-max)
Smoke home	Is there someone who smokes pipe/cigarettes at home?	No Yes
Smoke work	Is there someone who smokes pipe/cigarettes at work?	No Yes
Snuff	Do you use snuff?	No Yes
Frequency of snuff	How often do you use snuff?	No Frequently-daily Infrequently-2/3 times weekly Rarely-2/3 times monthly
Chew tobacco	Do you chew tobacco?	No

		Yes
Tobacco frequency	How often do you chew tobacco	No Frequently-daily Infrequently-2/3 times weekly Rarely-2/3 times monthly
Alcohol	Have you ever drunk alcohol?	No Yes
Bought beer	Do you drink bought beer?	No Yes
Home brewed	Do you drink home brewed beer?	No Yes
Bought spirits	Do you drink bought spirits?	No Yes
Home-made spirits	Do you drink home-made spirits?	No Yes
Wine	Do you drink wine?	No Yes
Audit score	Sum of drinks containing alcohol, number of drinks per day and occasions of 6 or more drinks.	No(0 drinks) Low (1-6 drinks) High (7-12drinks)
Chronic illness	Do you have any chronic illness?	No Yes
Diabetes	Do you have diabetes?	No Yes
Cancer	Do you have cancer?	No Yes
Hypertension	Do you have hypertension?	No Yes
Vascular	Do you have vascular disease?	No Yes
Asthma	Do you have asthma?	No Yes

COPD	Do you have COPD?	No Yes
HIV/AIDS	Do you have HIV/AIDS?	No Yes
TB	Do you have TB?	No Yes
Acute Pancreatitis	Do you have Acute pancreatitis?	No Yes
Chronic Pancreatitis	Do you have chronic pancreatitis?	No Yes
Hepatitis	Do you have hepatitis?	No Yes
Rheumatoid Arthritis	Do you have rheumatoid arthritis?	No Yes
Other	Do you have other chronic illnesses?	No Yes
Red meat	Do you eat red meat?	No yes
Red meat week	How many times a week do you eat red meat?	No (0) Low (1-4) High (5-7)
White meat	Do you eat white meat?	No Yes
White meat week	How many times a week do you eat white meat?	No (0) Low (1-4) High (5-7)
Vegetables	Do you eat vegetables meat?	No Yes
Vegetable week	How many times a week do you eat vegetables?	No (0) Low (1-4) High (5-7)
Fish	Do you eat fish meat?	No Yes

Fish week	How many times a week do you eat fish?	No (0) Low (1-4) High (5-7)
Fried Fish	What type of fish do you eat?	No Yes
Canned Fish	What type of fish do you eat?	No Yes
Fresh Fish	What type of fish do you eat?	No Yes
Sea food	What type of fish do you eat?	No Yes
Cured food	Do you eat cured food?	No yes
Cured food week	How many times a week do you eat cured food?	No (0) Low (1-4) High (5-7)
Fried food week	How many times a week do you eat fried food?	No (0) Low (1-4) High (5-7)
Maize meal	Do you eat maize meal?	No Yes
Maize meal week	How many times a week do you eat maize meal?	No (0) Low (1-4) High (5-7) Other
Functionality	Karnofsky functional status score.	100%-normal function 90%-capable of normal 80%-normal activity 70%-cares for self 60%-requires help 50%-often requires help 40%-disabled

Symptoms level	What is patient's ECOG score?	0-asymptomattic 1-symptomatic but ambulatory 2-symptomatic, < 50% 3-symptomatic, >50% 4-bedbound
BMI	Body Mass Index	Mean SD
BMI category	Body Mass Index category	Missing (.) Underweight (min-18.5) Healthy(18.6-24.9) Overweight (25-29.9) Obese (30-35.9) Very obese (36-max)

2.6 Data management and analysis

2.6.1 Data management

A de-identified excel spread sheet was imported from the existing REDCap database (Harrisa, Taylorb et al. 2009). Using Stata software version 13, various commands assisted in cleaning the data. Data cleaning processes included checking for duplicates, missing values, recoding and categorising variables. This was a quality assurance process.

2.6.2 Statistical analysis

Table 2. Objectives of the study and corresponding analysis undertaken

Objectives	**Analysis**
1. To describe demographic, social, hospital and dietary characteristics of patients with PDAC in two public academic hospitals in Johannesburg between June 2013 and December2015.	Proportions and percentages gave summary of categorical variables. Differences were analysed using chi-squared test. Summary of continuous variables were by the mean and the standard deviation, or median with interquartile range.
2. To assess the risk factors associated with patients diagnosed with PDAC in two public academic hospitals in Johannesburg between June 2013 and December 2015.	Odds ratios calculated using multiple unconditional logistic regression. The model applied backward analysis. Likelihood ratio test (LRT) tested for confounder or effect modifier or interaction assumptions through adjustment.

From **Table 2** all statistical test analysis were two sided z-test and p values <0.25 were statistically significant. A p value <0.20 meant we rejected the null hypothesis and that the variables were the same in both cases and controls (Hosmer and Lemeshow 2000, Vittinghof, Glidden et al. 2004).

2.7 Ethical considerations

The Human Research Ethics Committee of the University of the Witwatersrand approved the primary study and participants signed an informed consent. The original ethical approval was in 2013 (Clearance Certificates M1305 50 & 51). The Wits Human Research Ethics Committee approved additional ethics for secondary data analysis (Clearance Certificate M160247) **Figure 4**. There was no personal identification in the dataset provided and a unique ID number from REDCap identified each patient. For this secondary data analysis, the principal investigator granted access to the data in the form of an approval letter and this is in **Figure 2** and **Figure 3**.

2.8 Study Budget

The National Research Foundation (NRF) funded the initiation of the principal study. The grant number was 91508 with reference number CSUR13091741850 NRF South Africa. The South African Medical Research Council through a grant awarded to the Wits Common Epithelial Cancer Research Centre provided further funding for continuation of the principal study. Both grants allowed the development of the REDCap database that facilitated this research.

3. Results

3.1 Introduction

The study explored risk factors associated with PDAC at two Johannesburg academic hospitals between 2013 and 2015. This chapter shows the results obtained from the data analysis for both cases and controls.

Summary of the results:

Eighty two percent of all participants were black. The age groups comprising the most cases were the 50-59 and the 60-69 year olds. Most of the participants' had a secondary level school education, and the longest jobs held were that of general workers

Table *3*. Cases had a higher percentage of smoke pack years per category as well as higher alcohol consumption compared to controls **Table 5**. However, smoking and alcohol had no statistical significance in our study. The functional status of the participants was in the 80-90% normal function category. More than 27% of the Body Mass Index (BMI) information was missing. After unconditional multiple-logistic regression, the younger age group was less likely to be at risk of PDAC. White meat and vegetables per week showed a protective effect against PDAC while red meat consumption showed no statistical significance for the study population **Table *6***.

Description of study population:

Table 3 Baseline characteristics

	Overall (n=278)	Control (n=139)	Case (n=139)
Age years			
Mean	54.93	51.94	57.93
SD	14.31	16.07	11.59
Age			
20-29	16(6%)	15(11%)	1(1%)
30-39	36(13%)	27(19%)	9(6%)
40-49	34(12%)	17(12%)	17(12%)
50-59	84(30%)	32(23%)	52(37%)
60-69	69(25%)	28(20%)	41(30%)
70-79	27(10%)	14(10%)	13(9%)
80+	12(4%)	6(4%)	6(4%)
Race			
Black	228(82%)	114(82%)	114(82%)
White	21(8%)	9(6%)	12(8%)
Indian	9(3%)	5(4%)	4(3%)
Coloured	18(8%)	10(7%)	8(6%)
Other	2(1%)	1(1%)	1(1%)
Gender			
Female	141(51%)	80(58%)	61(44%)
Male	137(49%)	59(42%)	78(56%)
Employed			
No	207(74%)	106(76%)	101(73%)
Yes	71(26%)	33(24%)	38(27%)
Longest job employed			
Construction & chemical	16(6%)	6(4%)	10(7%)
Mining	6(2%)	2(1%)	4(3%)
Manufacturing & Factory	26(9%)	17(12%)	9(7%)
Engineering	12(4%)	8(6%)	4(3%)
Drivers	20(7%)	7(5%)	13(9%)
General worker	83(30%)	33(24%)	50(36%)
Office & other	115(41%)	66(48%)	49(35%)
Chronic illness			
No	115(41%)	57(41%)	58(42%)
Yes	163(59%)	82(59%)	81(58%)

Table 4 Diet

	Overall (n=278)	Control (n=139)	Case (n=139)
Red meat week			
No	37(13%)	20(14%)	17(12%)
Low	214(77%)	109(79%)	105(76%)
High	27(10%)	10(7%)	17(12%)
White meat week			
No	15(5%)	6(4%)	9(7%)
Low	208(75%)	98(71%)	110(79%)
High	55(20%)	35(25%)	20(14%)
Vegetable week			
No	22(8%)	6(4%)	16(11%)
Low	138(50%)	63(45%)	75(54%)
High	118(42%)	70(51%)	48(35%)
Fish week			
No	61(22%)	37(27%)	24(17%)
Low	211(76%)	97(70%)	114(82%)
High	6(2%)	5(3%)	1(1%)
Cured food week			
No	130(47%)	64(46%)	66(47%)
Low	143(51%)	70(50%)	73(53%)
High	5(1%)	5(4%)	0
Fried food week			
No	62(22%)	23(17%)	39(28%)
Low	197(71%)	105(76%)	92(66%)
High	19(7%)	11(7%)	8(6%)

Table 5 Social characteristics

Audit score			
No	193(69%)	103(74%)	90(65%)
Low	67(24%)	31(22%)	36(26%)
High	18(7%)	5(4%)	13(9%)
Smoked			
No	125(45%)	70(50%)	84(60%)
yes	153(55%)	69(50%)	55(40%)
Smoke now			
No	211(76%)	108(78%)	103(74%)
Yes	67(24%)	31(22%)	36(26%)
Smoke pack years			
Mean	9.09	7.27	10.9
SD	16.31	14.52	17.79
Smoke pack category			
No	136(50%)	78(56%)	58(41%)
Very rare	72(26%)	34(24%)	38(27%)
Frequent	43(15%)	17(12%)	26(19%)
Very frequent	27(9%)	10(8%)	17(12%)

3.1.1 Demographic characteristics of study population

Table 7 in the appendix explains the entire demographic characteristics. The study population included 278 participants. The youngest study participant (cases plus controls) was 23.8 years old, the oldest participant was 92.8 years old, and the mean age was 54.9 years with a SD of 14.31. The mean age for controls was 51.94 years with a SD of 16.07 while for the cases the mean age was 57.93 years with a SD 11.59. Participants belonged to the following races: black, coloured, Indian, white, and others. The majority of participants fell within the 50-59 age group and represented 30% of the total study population. In this age group, there were 32 controls and 52 cases. The majority of participants represented by race were black making up 82% of the total study population, thus in both case and control groups the number of black participants was 114 of the total. The majority of participants were born in the following provinces: Gauteng (48%), North West (10%), and Free State (9%). The level of education was highest for secondary schooling 138 (50%). The majority of participants were unemployed making up 74% of the total study population and of those who had worked the longest employment was that of the general worker and office/other category each contributing about 30% and 41% respectively of the total study population.

3.1.2 Social characteristics of study population

Table *8* in the appendix explains the social habits of the study population. Fifty five percent of the total study population smoked with the study cases having a mean smoke pack years of 10.90 and a SD of 17.79. This table shows that 64 control participants had drunk alcohol while on the other hand 88 case participants had drunk alcohol. Consumption of bought beer (i.e. factory-brewed beer) was the highest in both cases and controls with 57 and 45 participants respectively. Of the total study population, 18% had high alcohol consumption and within groups, the consumption was 13% for cases and 5% for controls.

3.1.3 Chronic Medical characteristics of study population

Table *9* in the appendix shows the additional chronic medical conditions the participants had besides PDAC. The study population had 14% of participants with diabetes, 35% with hypertension and 17% with HIV/AIDS. Of these three chronic medical conditions, diabetes was present in 16% of cases compared to the 12% for controls.

3.1.4 Dietary characteristics of study population

Table *10* in the appendix shows the entire dietary characteristics of the study population. Low consumption of red meat and white meat made up 214 (77%) and 208 (75%) of the entire study population respectively. Within the cases and controls, the low consumption of red and white meat group had an almost equal percentage above 70%. **Error! Reference source not found.** below shows high consumption of vegetables and maize meal recorded as 42% and 54% respectively for the entire study population.

3.1.5 Hospital characteristics for the study population

Table *11* in the appendix shows that 50 control group participants were capable of 90% normal activity while 47 of the case group had 80% normal activity. The ambulatory symptom level category had the majority of participants at 75 controls and 85 cases respectively, making up 58% of the total study population.

3.2.1 Adjusted and unadjusted demographic factors associated with PDAC

Table *12* in the appendix shows that the odds ratio of PDAC before adjustment in males was 1.73 but after adjustment, it was insignificant. The odds ratio of PDAC in the 20-39 age group was below one showing protection in the young age group while the odds ratio was approximately three in the age group 50-69. The general workers odds ratio for PDAC was approximately two before and after adjustment.

3.2.3 Adjusted and unadjusted dietary factors associated with PDAC

Table *14* in the appendix shows the following unadjusted factors with weak significance and the corresponding odds ratios as follows: white meat 0.38 in the high consumption category, vegetables 0.26 in the high intake category, fish 1.81 in the low intake category, fried foods 0.52 in the low intake category and maize meal 0.22 in the high intake category. After adjustment, the following factors were significant: high consumption white meat per week 0.14, low fish intake per week 2.15, low fried foods consumption per week 0.47, and high vegetable intake per week 0.13.

3.2.3 Adjusted and unadjusted social factors associated with PDAC

Table *15* in the appendix shows the odds ratio of PDAC before adjustment for smoke pack category to be 2.06 in frequent smoking, and 2.98 for high number of alcohol drinks (audit score) but after adjusting, both these factors were insignificant.

3.2.4 Adjusted and unadjusted hospital factors associated with PDAC

Table *16* in the appendix shows the hospital characteristics of the study population where the odds ratio of PDAC in the 90% and 80% normal activity was 1.19 and 3.44 respectively.

3.2.5 Final model

Table *17* shows the final model for my study population. The odds ratios of PDAC in the 20-29 age group was 0.11, in the 50-59 age group was 2.63 and in the 60-69 age group, it was 2.85. The odds ratio of PDAC for general workers was 1.79. High white meat consumption had an odds ratio of 0.18 for PDAC. Low fish consumption per week shows an odds ratio of 2.17. The odds ratio for low fried foods consumption per week was 0.48 and high consumption of vegetables per week had an odds ratio of 0.17 for PDAC.

Table 6 Unadjusted and Adjusted OR for PDAC

Characteristic	UOR(95%CI)	p-value	AOR(95%CI)	p-value
Gender		**0.0224**		0.6992
Female	1 (base)		1(base)	
Male	1.73 (1.07-2.79)		1.49(0.73-3.06)	
Age years	1.03(1.01-1.05)	**0.0004**		
Age		**0**		0.0006
20-29	0.07 (0.008-0.56)		0.11(0.01-1.06)	
30-39	0.33(0.12-0.92)		0.60(0.19-1.90)	
40-49	1(base)		1(base)	
50-59	1.63(0.72-3.63)		2.61(1.00-6.78)	
60-69	1.46(0.64-3.35)		3.04(1.10-8.34)	
70-79	0.93(0.34-2.55)		1.58(0.49-5.11)	
80+	1(0.27-3.73)		1.87(0.42-8.31)	
Long Employment		**0.037**		0.0419
Construction & chemic	2.24(0.76-6.59)		2.08(0.59-7.37)	
Mining	2.69(0.47-15.30)		1.70 (0.24-12.16)	
Manufacturing & Factory	0.71(0.29-1.73)		0.49(0.17-1.40)	
Engineering	0.67(0.19-2.36)		0.30(0.07-1.28)	
Drivers	2.50(0.93-6.73)		1.95(0.60-6.35)	
General worker	2.04(1.15-3.62)		1.83(0.93-3.63)	
Office & other	1 (base)		1 (base)	
Red meat week		0.3406		0.9043
No	1(base)		1(base)	
Low	1.13(0.56-2.28)		0.99(0.42-2.31)	
High	2.00(0.73-5.51)		1.13(0.33-3.92)	
White meat week		**0.0659**		0.0502
No	1(base)		1(base)	
Low	0.75(0.26-2.18)		0.28(0.06-1.31)	
High	0.38(0.12-1.23)		0.14(0.03-0.73)	
Vegetable week		**0.0071**		0.0037
No	1(base)		1(base)	
Low	0.45(0.16-1.21)		0.25(0.07-0.89)	
High	0.26(0.09-0.70)		0.13(0.04-0.49)	
Fish week		**0.0291**		0.0229
No	1(base)		1(base)	
Low	1.81(1.01-3.24)		2.15(1.04-4.45)	
High	0.31(0.03-2.80)		0.17(0.01-2.78)	
Fried food week		**0.0636**		0.07
No	1(base)		1(base)	
Low	0.52(0.29-0.93)		0.47(0.22-0.96)	

High	0.43(0.15-1.22)		0.39(0.11-1.35)	
Smoke pack category		**0.0749**		0.7667
No	1(base)		1(base)	
Very rare	1.50(0.85-2.67)		1.43(0.66-3.08)	
Frequent	2.06(1.02-4.14)		0.91(0.36-2.28)	
Very frequent	2.29(0.98-5.36)		1.10(0.38-3.19)	
Audit score		**0.0849**		0.2135
No	1 (base)		1(base)	
Low	1.33(0.76-2.32)		1.62(0.78-3.35)	
High	2.98(1.02-8.67)		3.17(0.79-12.69)	

4. Discussion

4.1 Introduction

Secondary analysis of data collected over 18 months (June 2013-December 2015) was done for a case-control study aimed at exploring the various risk factors associated with PDAC at two public academic hospitals in Johannesburg. The study population totalled 278, with 50% participants for both cases and controls. The main objectives of this study were, to describe the demographic characteristics of patients with PDAC and to assess the risk factors associated with patients diagnosed with PDAC in two public academic hospitals in Johannesburg between June 2013 and December2015.

Mean and standard deviations for continuous data and proportions and percentages for categorical data was used to describe the demographic characteristics of patients with PDAC. Univariate and multivariate analyses were done using unconditional multiple logistic regression to assess the relationships and associations. Continuous variables were analysed inferentially as categorical to avoid clinically distorted OR at 95% CI and a coefficient slope of zero (Hosmer and Lemeshow 2000). Fitting a logistic model predicted risk of binary outcomes given a set of risk factors. The models isolated single predictors by incorporating other predictors to give clear interpretations of variable effects on the outcome (Vittinghof, Glidden et al. 2004). The models assisted in identifying patients likely to have or not have limitations for future patient management costs. After performing univariate analysis, the variables with a p-value of <0.25 (Hosmer and Lemeshow 2000) were added into the full model for multivariable selection. This selection assisted in having the best model within the scientific context of the PDAC problem. Clinical and intuitive variables were selected regardless of their statistical significance (p-value <0.25 and >0.05) and included so as to control for confounding factors and to avoid losing important variables (Hosmer and Lemeshow 2000). Multiple odds ratios were estimated during the development of the predictive model. For the final model the p-value for the LR test was 0.20 (Hosmer and Lemeshow 2000). All the analysis was done using STATA 13.

With this study, we were able to identify the different exposures for rare PDAC disease. As this was a hospital-based population, only a small catchment area was covered and since patients' details are de-identified, duplication or multiple counting can occur if patients were in different hospitals, as there may be no linking system. The population is standardized thus it is difficult to compare findings to a different setting. Overall, the hospital-based population was good for the case-control study design.

In conclusion, the higher percentage/proportions was being male, consumed alcohol, smoked and being between the ages of 50-69. Being 50-69 years of age and employed for longer periods as the norm for a general worker was positively associated with PDAC. Additionally, increased consumption of vegetables and white meat was protective for PDAC, whilst a low intake of fish increased PDAC risk.

A detailed discussion of the socio-demographic characteristics and risk factors are to follow under the following sub-headings:

4.1.1 Demographic characteristics of patients with PDAC

The study showed the most PDAC cases in the age category 50-59 years, while compared to other described studies, the age 60+ or even 70+ categories had the highest prevalence of PDAC cases. This indicates that this study's PDAC patients were younger compared to other studies (Babb 2011, Yeo 2015). This may be due to the lower life expectancy of approximately 60 years in South Africa compared to in the developed countries. In the 2011 cancer registry, the highest PDAC percentage was in the 60-69 age group followed by the 50-59 age group (Babb 2011). However, when comparing quantitative analysis performed for PDAC, the age range of 50-69 is when most incidences occur (Meza, Jeon et al. 2008), which is in agreement with this study results. Most of this study population attained secondary education and were born in the Gauteng province. Gauteng province is the economic hub of South Africa. Most people reside in Gauteng for employment opportunities and the two academic referral hospitals used in this study are in this province. Long employed general workers, office and other workers account for more than 30% of our study population. The general workers included domestic workers, housewives, cleaners and gardeners. Furthermore, in this study, the longest employed general workers were affected mostly, which is also different from other studies in which industrial plant workers and chemical/pesticide workers were more at risk of developing PDAC (Yeo 2015). The general workers have a higher risk of developing PDAC since they are likely to be of low economic status, which may deprive them of quick access to health services or they may not have sufficient education to be cautious about their health and lifestyle. This is in stark contrast to most published studies which show that the occupations and factors which include: dry cleaning, chemical plant work, sawmills, electrical equipment, Ashkenazi Jewish heritage, manufacturing workers, miners and metal workers are more at risk of developing PDAC (Yeo and Lowenfels 2012, Yeo 2015).

4.1.2 Social characteristics of study population

The smoke pack category was not significant after adjusting and this may be because approximately 25% of the study population smoked a lot (frequently and very frequently), while the rest of the study population (75%) did not smoke. A lower number of participants indicated a high consumption of alcohol (in the audit score) and therefore, the adjustment made the alcohol variable statistically insignificant. Different chronic illnesses were mostly not significant. This insignificance may be because less than 15% of participants had other chronic conditions.

In most studies the PDAC cases smoked more than the controls (Talamini, Polesel et al. 2009), which corresponds with our study in which more cases smoked and additional smoke exposure was from someone else who smoked at work. Sixty-three percent of the cases consumed alcohol, which indicates just how alcohol is an independent risk factor for PDAC if the individual is a smoker (Talamini, Polesel et al. 2009). In this study, smoking and alcohol consumption before adjusting were statistically significant when looking at the variable and outcome relationship. Adjusting for the potential confounding variables stripped away the effects of these factors in the relationship of the main variable and outcome.

4.1.3 Medical/hospital characteristics

More than 50% of our cases had other medical conditions of which among the top of the list was hypertension, diabetes and HIV/AIDS. Most published studies show diabetes, acute pancreatitis, HIV/AIDS, chronic pancreatitis and hepatitis B as additional medical conditions (Yeo and Lowenfels 2012, Ilic and Ilic 2016). In the three above-mentioned chronic medical conditions recorded in the study, the application of logistic regression did not yield any statistical significance values, showing that case and control medical conditions do not have an effect on the likelihood of developing PDAC. The study results may just be by chance or a larger sample size is required. Due to the HIV/AIDS stigma in South Africa, participants may have provided biased information that could have affected our results (Singh, Ruff et al. 2015).

4.1.4 Dietary pattern of study population

This study shows the proportion of the low consumption vs. the high consumption in the following variables: red meat, cured foods, fried foods and fish per week. The high consumption proportion was less than 10% in these variables. The following variables had close to or just above 20% in the high consumption per week: white meat, maize meal and

vegetables. The high consumption in the above-mentioned variables except for maize meal complements most literature (Lu, Shu et al. 2017). High consumption of white meat and vegetables are foods recommended for prevention of most cancers and lifestyle diseases at large. White meat per week, particularly in cases of high consumption shows a protective effect and this may be due to having less saturated fats compared to red meat. Low consumption of fish per week mostly seems to be a risk factor for PDAC. This finding may be due to the preparation method of the fish or it may simply be by chance. Low consumption of fried foods per week was protective. This may be due to the type of oil used and by not frying the foods at extremely high temperatures. In agreement with published studies, vegetables had a protective effect. Vegetables are high in antioxidants and cancer preventative nutrients. Therefore, high consumption of vegetables per week will have a protective effect against PDAC.

4.2 Risk factors for patients with PDAC

4.2.1 Age and Gender

The age category 20-29 had a 0.10 odds ratio of PDAC with a p-value of <0.05, suggesting that people in this age group are less likely to get PDAC. On the other hand, for age categories 50-59 and 60-69 the odds ratios of PDAC are 2.49 and 2.96 respectively with p-values of less than 0.1 indicating a greater risk of developing PDAC. These observations correlate with published literature stating how an increase in age, increase the risk of PDAC development by two to three times. Due to the improved life expectancy of populations globally, the number of aged individuals is on the rise. This rise indirectly increases the incidences of PDAC in these age groups (Fest, Ruiter et al. 2016). Furthermore, in a study using the Bayesian model for the top 20 risk factors, the age of 60+ ranked in the top 10 risk factors on the list (Zhao and Weng 2011), for developing PDAC. Gender was statistically insignificant after adjusting and this was different from other studies in which the male race was shown to be associated with PDAC compared to woman. This may be due to the increase in the number of women who are now smoking in South Africa, unlike in the past where smoking was mostly associated with men.

4.2.2 Alcohol

Alcohol consumption in high quantities was a top risk factor in the Bayesian model and this also agrees with our unadjusted results (Zhao and Weng 2011). Review done around 2005 showed no association of high alcohol consumption and PDAC (Lowenfels and Maisonneuve 2006), and this was confirmed by our adjusted results.

4.2.3 Diet

4.2.3.1 White meat and red meat

High consumption of white meat per week shows a protective effect with an odds ratio of 0.18. This concur with the benefit of regulated healthy dietary requirements compared to red meat consumption per week (Larsson, Hakanson et al. 2006, Lu, Shu et al. 2017). White meat may have a protective effect due to less saturated fat content and upon cooking, it may release minimal nitrosamines compared to red meat. The number of cases and controls consuming red meat was approximately the same and this might be why this category showed no significance during the logistic regression analysis. In a review paper by Zhao et al. (2016) the effects of red meat and processed food consumption was inconclusive (Zhao, Yin et al. 2016).

4.2.3.2 Fish

While low fish consumption per week has an odds ratio risk of 2.17, the observed risk of PDAC with fish consumption could be due to the way the fish is prepared, such as deep-fried, which reduces the amount of LC-PUFA in fish and generates several chemicals that may contribute to carcinogenesis (Pericleous, Rossi et al. 2014).

4.2.3.3 Fried foods

Low fried food consumption per week has a protective effect of 0.48 which may be due to the use of saturated or monounsaturated fatty acids (Nkondjock, Krewski et al. 2005) as seen in the study population.

4.2.3.4 Vegetables

General consumption of vegetables has a protective effect, but this protective effect was enhanced with increased vegetable consumption showing an odds ratio of 0.17. This finding support the other studies on the protection of vegetables in PDAC diagnosis, and it may show that specific vegetables consumed by study participants brought about this effect (Nöthlings, Wilkens et al. 2007).

4.2.4 Employment

In the long employment category, the general workers had a 1.79 times odds ratio for PDAC with a resultant p-value of 0.083. In the 1970s occupational factors such as chemicals and pesticides for example, were suspected to be risk factors for PDAC, but no proof was available, however in around 2005 occupation was not regarded a major risk factor for developing PDAC (Lowenfels and Maisonneuve 2006).

4.3 Study limitations

- The primary study focused on the genetic and environmental factors that influence susceptibility to PDAC, which in itself did not answer the questions of this particular study.
- This study used secondary data. Records did not have all the relevant information required for our study, as most values such as BMI were missing.
- Due to the study being a case-control study, there were difficulties in overcoming potential bias due to the recall or interviewer bias.
- Adjustment of confounders occurred during the analysis stage and not at the design stage or by matching.
- The successful selection of both cases and controls representative of respective populations was difficult, since in our study was hospital based.
- Inference to causality and inadequate data on the chronology of disease and exposure was a problem, considering that this was a retrospective directional study.
- The study was not population based, therefore, it is impossible to calculate incidence of disease since no total population statistics were included.
- Even though there were some limitations, this study design is good for rare diseases like PDAC, and multiple exposures for one outcome allow for checking multiple associations.
- A major disadvantage of the control group selected from diseased individuals is that some of their illnesses may have shared risk factors with the cases, meaning that they may have a higher, or lower, exposure prevalence compared to the population from which the cases arose.
- The ORs measured exposure to the disease and not the disease occurrence.
- In some variables, there was a wide confidence interval due to the sample size being very low in those variables. This is an indication of a larger sample size required.
- Dietary assessment could have been biased as the assessment could have been done while the subject was following a diet prescribed by physician. The norm for recalling patterns should be a healthy lifestyle.
- The hospital based patient data was collected in the hospital, but the catchment population was not defined and the data collection was mostly for administrative purposes.

5. Conclusions and recommendations

5.1 Introduction

This final chapter gives the conclusions, recommendations and future studies that can add valuable information to the epidemiology of PDAC.

5.2 Conclusions

This research showed how many people had PDAC in the period of June 2013 to December 2015 in the two academic hospitals. Age of diagnosis for most cases (50-59) and the birth province where most cases were born (Gauteng) was identified. The black population had the most cases and the male to female ratio was 1.3: 1. In this study, the factors that posed risks for development of PDAC were older age, low consumption of fish and long employment as a general worker. The following weekly diet showed protection against PDAC, high consumption of white meat and vegetables. With the information from this study, identification of whether a particular dietary component influences the risk of developing PDAC was established. Identification of the answers in a public setting gives motivation to check reproducibility of results in the private sector. The results can further assist an individual's daily decisions such as to stop smoking and to eat more vegetables, as this affect their health over a lifetime. As the principal study continues, we may be able to analyse historical trends and current data to project future public health resource needs for PDAC management.

5.3 Recommendations

The evidence presented in this study suggests that a greater emphasis on economic policies focusing on assisting the poor and marginalised communities is needed in the discourse of PDAC control. This is due to the public hospitals having a patient booking system that may take as long as six months or more before the doctor can see a patient. This long period can also be causative of advanced stage diagnosis of PDAC. This study confirms that age is associated with a significant increase in the risk of developing PDAC. There is a need to strengthen the implementation of a better and well-managed lifestyle as an individual age. These include physical activity, improving nutrition and hygiene, as well as improving access to health services. Active periodical case finding among the symptomatic and close contact of the index cases is recommended. Further research on risk factors for PDAC to address the limitations of this study will describe the burden in a large community for public health priorities. PDAC prevention strategies should focus on interventions that reduce or limit the impact of its risk. With a well-functioning population-based registry, good

monitoring and assessment of effectiveness of cancer control activities is possible, opposed to the current hospital-based information only.

5.4 Future studies

Collaboration with both public and private hospitals for a complete report of incidence and prevalence of PDAC across South Africa can add value to these study results. Knowing that South Africa is a poor resourced country, the cost of illness due to PDAC may assist in the resource allocation during the national budget allocation process. Further studies regarding the cost of illness may be determined by the period from admission to diagnosis and all the tests and equipment used for diagnosis of PDAC. The years of life lost to ill health will give an overall picture of the significance of this disease to the society beyond the immediate cost of treatment. Further studies on behaviour related to health and well-being, for example whether exercise has an effect on PDAC would be of great value in future.

6 References

Al-Majeda, H., et al. (March 2013). Pancreatic cancer: Incidence, clinical profile, and frequency of associated factors in Kuwait. *Alexandria Journal of Medicine* **49**(1): 75–80.

Alexakis, N., et al. (2004). Current standards of surgery for pancreatic cancer. *British Journal of Surgery* **91**(11): 1410-1427.

Anand, P., et al. (2008). Cancer is a Preventable Disease that Requires Major Lifestyle Changes. *Pharmaceutical Research* **25**(9): 2097-2116.

Are, C., et al. (2016). Predictive Global Trends in the Incidence and Mortality of Pancreatic Cancer Based on Geographic Location, Socio-Economic Status, and Demographic Shift. *Journal of Surgical Oncology* **114**: 736-742.

Babb, C. (2011). Summary statistics of cancer diagnosed histologically in South Africa in 2011. Retrieved 2017, from http://www.cansa.org.za/files/2016/08/NCR-2011-cancer-tables.pdf.

Behrens, G., et al. (2015). "Physical activity and risk of pancreatic cancer: a systematic review and meta-analysis." European Journal Of Epidemiology **2015**(30): 279–298.

Bosetti, C., et al. (2012). Cigarette smoking and pancreatic cancer: an analysis from the International Pancreatic Cancer Case-Control Consortium (Panc4). *Annals of Oncology* **23**(7): 1880-1888.

Chari, S. T., et al. (2005). Probability of pancreatic cancer following diabetes: a population-based study. *Gastroenterology* **129**(2): 504-511.

Chari, S. T., et al. (2008). Pancreatic cancer-associated diabetes mellitus: prevalence and temporal association with diagnosis of cancer. *Gastroenterology* **134**(1): 95-101.

Correa, P., et al. (1990). Gastric precancerous process in a high risk population: cohort follow-up. 50 **15**(4737-4740).

Duell, E. J., et al. (2012). Pancreatitis and pancreatic cancer risk: a pooled analysis in the International Pancreatic Cancer Case-Control Consortium (PanC4). *Annals Oncology* **23**(11): 2964-2970.

Ferlay, J., et al. (2012). Cancer Incidence and Mortality Worldwide. *Cancer Base No. 11*. Retrieved 30 September 2017, 2017, from http://globocan.iarc.fr.

Fest , J., et al. (2017). Underestimation of pancreatic cancer in the national cancer registry Reconsidering the incidence and survival rates. *European Journal of Cancer* **72**(2017): 186-191.

Fest J., et al. (2016). Underestimation of pancreatic cancer in the national cancer registry e Reconsidering the incidence and survival rates. *European Journal of Cancer* **72**(2017): 186-191.

Fiorino, S., et al. (2013). HBV- and HCV-Related Infections and Risk of Pancreatic Cancer. *Journal of the Pancreas* **14**(6): 603-609.

Hariharan, D., et al. (2008). Analysis of mortality rates for pancreatic cancer across the world. *HPB* (Oxford) **10**(1): 58-62.

Harrisa, P. A., et al. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* **42**(2): 377-381.

Herbst, M. C. and Joubert E. (2017). Fact-Sheet-Pancreatic-Cancer-NCR-2012. Retrieved 25 September 2017, from http://www.cansa.org.za/files/2017/07/Fact-Sheet-Pancreatic-Cancer-NCR-2012-web-July-2017.pdf.

Hosmer, D. W. and S. Lemeshow (2000). *Applied Logistic regression.* New York, Wiley Interscience

Huxley, R., et al. (2005). Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *British Journal of Cancer* **92**(11): 2076-2083.

Ilic, M. and I. Ilic (2016). Epidemiology of pancreatic cancer. *World Journal of Gastroenterology* **22**(44): 9694-9705.

Iodice, S., et al. (2008). Tobacco and the risk of pancreatic cancer: a review and meta-analysis. *Langenbecks Arch Surg* **393**: 535–545.

Jansen, R., J., et al. (2014). Fatty acids found in dairy, protein and unsaturated fatty acids are associated with risk of pancreatic cancer in a case–control study. *International Journal of Cancer* **134**: 1935-1946.

Khan, N. and H. Mukhtar (2010). Cancer and metastasis: prevention and treatment by green tea. *Cancer Metastasis Rev* **29**(3): 435–445.

Konner, J. and E. O'Reilly (2002). Pancreatic Cancer: Epidemiology, Genetics, and Aproaches to screening.

Kos, F. T., et al. (2012). Is There an Association between Blood Group and Survival in Pancreatic Cancer? *Asian Pacific Journal of Cancer Prevention* **13**: 6151-6153.

Kuzmickiene, I., et al. (2013). Smoking and other risk factors for pancreatic cancer: A cohort study in men in Lithuania. *Cancer Epidemiology* **37**(2013): 133–139.

Larsson, S. C., et al. (2006). Meat, fish,poultry and egg consumption in relation to risk of pancreatic cancer: a prospective study. *International Journal of Cancer* (118): 2866-2870.

LeDuc, T. (2017) from http://www.worldlifeexpectancy.com/cause-of-death/pancreas-cancer/by-country/.

Li, D. (2012). Diabetes and Pancreatic Cancer. *Molecular Carcinogenesis* **51**(2012): 64-74.

Li, D., et al. (2004). Pancreatic cancer. *The Lancet* **363**(9414): 1049-1057.

Lowenfels A.B. and Maisonneuve P. (2006). Epidemiology and risk factors for pancreatic cancer. *Best Practice & Research Clinical Gastroenterology* **20**(2): 197–209.

Lu, P., et al. (2017). "Dietary Patterns and Pancreatic Cancer Risk: A Meta-Analysis." Nutrients **9**(39): 18.

Lucenteforte, E., et al. (2012). Alcohol consumption and pancreatic cancer: a pooled analysis in the International Pancreatic Cancer Case-Control Consortium (PanC4). *Annals of Oncology* **23**(2): 374-382.

Lynch, S. M., et al. (2009). Cigarette smoking and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. *American Journal of Epidemiology* **170**(4): 403-413.

Maisonneuve, P. and A. B. Lowenfels (2015). Risk factors for pancraetic cancer: A summary review of meta-analytical studies. *International Journal of Epidemiology* **44**(1): 186-198.

Meza, R., et al. (2008). Age-specific incidence of cancer: Phases, transitions, and biological implications. *PNAS* **105**(42): 16284-16289.

Michaud, D. S., et al. (2010). Alcohol intake and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium (PanScan). *Cancer Causes & Control* **21**(8): 1213-1225.

Miroslaw, J., et al. (2012). Influence of diet and tobacco smoking on pancreatic cancer incidence in Poland in 1960-2008. *Gastroenterology research and practice* **2012**(2): 1-9.

Muniraj, T., et al. (2013). Pancreatic cancer: Acomprehensivereview and update. *Disease-a-Month* **59**(2013): 368–402.

Neale, R. E., et al. (2014). Association between hypermethylation of DNA repetitive elements in white blood cell DNA and pancreatic cancer. *Cancer Epidemiology* **38**(5): 576-582.

Nkondjock, A., et al. (2005). Specific fatty acid intake and the risk of pancreatic cancer in Canada. *British Journal for Cancer* **92**: 971-977.

Nöthlings, U., et al. (2007). Vegetable intake and pancreatic cancer risk: The multiethnic cohort study. *American Journal of Epidemiology* **165**(2): 138-147.

Pericleous M., et al. (2014). Nutrition and Pancreatic Cancer. *Anticancer Research* **34**: 9-22.

Pinho A.V., et al. (2014). Chronic pancreatitis: A path to pancreatic cancer. *Cancer Letters* **345**(2014): 203-209.

Ries, L. A. G., et al. (2003). Cancer survival and incidence from Surveillance, Epidemiology, and End Results (SEER) Program. *The Oncologist* **8**: 541-552.

Ruiz R.B. and Hernández P.S. (2014). Diet and cancer: Risk factors and epidemiological evidence. *Maturitas* **77**(2014): 202-208.

Shrikhande, S. V., et al. (2012). Evolution of pancreatoduodenectomy in a tertiary cancer center in India: Improved results from service reconfiguration. *Pancreatology* **13**(2013): 63-71.

Singh, E., et al. (2015). Establishment of a cancer surveillance programme: the South African experience. *Lancet Oncology* **16**(8): e414-e421.

Sobhani, I., et al. (2013). Microbial dysbiosis and colon carcinogenesis: could colon cancer be considered a bacteria-related disease? *Therapeutic advances in gastroenterology* **6**(3): 215-229.

Soliman , A. S., et al. (2006). Pancreatic cancer mortality in Egypt. *Cancer detection and prevention* **30**(2): 473-479.

Talamini, R., et al. (2009). Tobacco smoking, alcohol consumption and pancreatic cancer risk: A case-control study in Italy. *European Journal of Cancer* **46**(2010): 370 – 376.

Tang, H., et al. (2014). Genes-Environment Interactions in Obesity-and Diabetes-Associated Pancreatic Cancer: A GWAS Data Analysis. *Cancer Epidemiology Biomarkers & Prevention* **23**(1): 98-106.

Vittinghof, E., et al. (2004). Regression methods in Biostatistics. United States of America, Springer Science+Business Media, Inc.

Yeo, T. P. (2015). Demographics, Epidemiology,and Inheritance of Pancreatic Ductal Adenocarcinoma. *Seminar Oncology* **42**: 8-18.

Yeo, T. P. and A. B. Lowenfels (2012). Demographics and Epidemiology of Pancreatic Cancer. *The Cancer Journal* **18**(6): 477-484.

Zhao D. and Weng C. (2011). Combining PubMed knowledge and EHR data to develop a weighted bayesian network for pancreatic cancer prediction. *Journal of Biomedical Informatics* **44**(2011): 859–868.

Zhao, Z., et al. (2016). Association Between Consumption of Red and Processed Meat and Pancreatic Cancer Risk: A Systematic Review and Meta-analysis. *Clinical Gastroenterology and Hepatology* 1-18.

7. Appendix

Figure 2 Letter from the gatekeeper

29 January 2016

To: WITS HREC

Re: Ms Brenda Kagodora's application

Dear ethics committee,

I give permission for Brenda to use the REDCap database that we have developed for pancreas cancer patients and their control group. The original ethics approval is M130551: *Genetic and environmental factors that influence the susceptibility to pancreas cancer*.

I am co-supervising her MSc and thus will be able to keep a close eye on the use of the database and ensure that it is only used for what is described in her application and protocol.

With regards,



Martin Brand

Figure 3 Letter to the gatekeeper

Department of Surgery **University of the Witwatersrand, Johannesburg**
7 York Road, Parktown 2193, Johannesburg. Tel: (011) 717-2478 Fax: (011)484-2717

January 26, 2016

Dr Martin Brand
Department of Surgery
Faculty of Health Sciences

Dear Dr Brand

Re: Letter of request for Shingirai Brenda Kagodora to use Data from REDCap
Addendum to studies undertaken by Dr Brand under Clearance Certificates M1305 50 & 51

I hereby request permission to use data collected for research that you are undertaking under Clearance Certificates M1305 50 & 51. The data will be used as secondary analysis for a protocol titled: Risk factors associated with pancreatic cancer at two Johannesburg Academic Hospitals between 2013 and 2015.

The following will be used:

Demographics	Family History	Diet history	Social history	Admission History
Age 1. 20-30 2. 31-40 3. 41-50 4. 51-60 5. 61-70 6. 71-80 7. +81	Medical history of parents and grandparents	Type of protein 1. red meat, 2. white meat, 3. fish	1.Tobacco smoking status, 2 number of pack years, 3 snuff, 4. chews tobacco)	BMI
Race 1. Black 2. White 3. Indian 4. Coloured 5. not known),		Vegetables	Alcohol 1. Alcohol status, 2. Type of alcohol consumed 3.alcohol dependence (CAGE score).	Any other medical condition

Education 0. never attended school 1.Primary 2. High school 3. Tertiary		Food preparation 1. cured 2. smoked food 3. fried food		
Employment (status of employment, category of type of employment),				
Province (place where patient, parents and grandparents where born),				
Sex 1. male 2. female 3. unknown				

Your help will be greatly appreciated, looking forward to receiving from you.

Yours faithfully

Ms Brenda Kagodora

Shingirai Brenda Kagodora

Figure 4 Ethic clearance

R14/49 Miss Shingirai Brenda Kagodora

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M160247

NAME: (Principal Investigator)	Miss Shingirai Brenda Kagodora
DEPARTMENT:	School of Public Health Chris Hani Baragwanath Academic Hospital Charlotte Maxeke Johannesburg Academic Hospital
PROJECT TITLE:	Risk Factors Associated with Pancreatic Cancer at Two Johannesburg Academic Hospitals between 2013 and 2015
DATE CONSIDERED:	26/02/2016
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Dr Martin Brand and Dr Charles Chasela
APPROVED BY:	Professor P Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	29/02/2016

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 in Room 10004, 10th floor, Senate House/2nd Floor, Phillip Tobias Building, Parktown, 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 February and will therefore be due in the month of February each year.

Principal Investigator Signature

Date 01/03/16

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Table 7 Demographic characteristics of study population

Characteristics	Overall (n=278)	Control (n=139)	Case (n=139)
Race			
Black	228(82%)	114(82%)	114(82%)
White	21(8%)	9(6%)	12(8%)
Indian	9(3%)	5(4%)	4(3%)
Coloured	18(8%)	10(7%)	8(6%)
Other	2(1%)	1(1%)	1(1%)
Gender			
Female	141(51%)	80(58%)	61(44%)
Male	137(49%)	59(42%)	78(56%)
Age years			
Mean	54.93	51.94	57.93
SD	14.31	16.07	11.59
Age			
20-29	16(6%)	15(11%)	1(1%)
30-39	36(13%)	27(19%)	9(6%)
40-49	34(12%)	17(12%)	17(12%)
50-59	84(30%)	32(23%)	52(37%)
60-69	69(25%)	28(20%)	41(30%)
70-79	27(10%)	14(10%)	13(9%)
80+	12(4%)	6(4%)	6(4%)
Schooling			
Never attended school	27(10%)	12(9%)	15(11%)
Primary school	97(35%)	40(29%)	57(41%)
Secondary school	138(50%)	77(55%)	61(44%)
Further Education	16(5%	10(7%)	6(4%)
Birth place			
Gauteng	134(48%)	66(47%)	68(49%)
Mpumalanga	15(5%)	7(5%)	8(6%)
Limpopo	14(5%)	5(4%)	9(6%)
North West	27(10%)	12(9%)	15(10%)
Free State	25(9%)	11(8%)	14(10%)

KwaZulu-Natal	21(8%)	10(7%)	11(8%)
Eastern Cape	14(5%)	9(6%)	5(4%)
Northern Cape	4(1%)	4(3%)	0
Western Cape	2(1%)	1(1%)	1(1%)
Other Country	22(8%)	14(10%)	8(6%)
Employed			
No	207(74%)	106(76%)	101(73%)
Yes	71(26%)	33(24%)	38(27%)
Longest job employed			
Construction & chemical	16(6%)	6(4%)	10(7%)
Mining	6(2%)	2(1%)	4(3%)
Manufacturing & Factory	26(9%)	17(12%)	9(7%)
Engineering	12(4%)	8(6%)	4(3%)
Drivers	20(7%)	7(5%)	13(9%)
General worker	83(30%)	33(24%)	50(36%)
Office & other	115(41%)	66(48%)	49(35%)

Table 8Social characteristics of study population

Characteristic	Overall (n=278)	Control (n=139)	Case (n=139)
Smoked			
No	125(45%)	70(50%)	84(60%)
yes	153(55%)	69(50%)	55(40%)
Smoke now			
No	211(76%)	108(78%)	103(74%)
Yes	67(24%)	31(22%)	36(26%)
Smoke pack years			
Mean	9.09	7.27	10.90
SD	16.31	14.52	17.79
Smoke pack category			
No	136(50%)	78(56%)	58(41%)
Very rare	72(26%)	34(24%)	38(27%)
Frequent	43(15%)	17(12%)	26(19%)
Very frequent	27(9%)	10(8%)	17(12%)
Exposed to smoke at home			
No	183(66%)	88(63%)	95(68%)
yes	95(34%)	51(37%)	44(32%)
Exposed to smoke at work			
No	125(45%)	65(47%)	60(43%)
Yes	153(55%)	74(53%)	79(57%)
Snuff			
No	255(92%)	129(93%)	127(91%)
Yes	23(8%)	10(7%)	12(9%)
Frequency of snuff			
No	256(92%)	129(93%)	127(91%)
Frequently-daily	13(5%)	4(3%)	9(7%)
Infrequently-2/3 times weekly	5(2%)	3(2%)	2(1%)
Rarely-2/3 times monthly	4(1%)	3(2%)	1(1%)
Chew tobacco			
No	270(97%)	136(98%)	134(96%)
yes	8((3%)	3(2%)	5(4%)

Tobacco frequency			
No	270(97%)	136(98%)	134(96%)
Frequently-daily	3(1%)	1(1%)	2(1.5%)
Infrequently-2/3 times weekly	4(1%)	2(1%)	2(1.5%)
Rarely-2/3 times monthly	1(1%)	0	1(1%)
Alcohol			
No	126(45%)	75(54%)	51(37%)
Yes	152(55%)	64(46%)	88(63%)
Bought beer			
No	176(63%)	94(68%)	82(59%)
Yes	102(37%)	45(32%)	57(41%)
Home brewed			
No	246(88%)	122(88%)	124(89%)
Yes	32(12%)	17(12%)	15(11%)
Bought spirits			
No	233(84%)	119(86%)	114(82%)
Yes	45(16%)	20(14%)	25(18%)
Home-made spirits			
No	271(97%)	134(96%)	137(99%)
Yes	7(3%)	5(4%)	2(1%)
Wine			
No	241(87%)	118(85%)	123(88%)
Yes	37(13%)	21(15%)	16(12%)
Audit score			
No	193(69%)	103(74%)	90(65%)
Low	67(24%)	31(22%)	36(26%)
High	18(7%)	5(4%)	13(9%)

Table 9 Chronic Medical characteristics of the study population

Characteristics	Overall (n=278)	Control (n=139)	Case (n=139)
Chronic illness			
No	115(41%)	57(41%)	58(42%)
Yes	163(59%)	82(59%)	81(58%)
Diabetes			
No	239(86%)	122(88%)	117(84%)
Yes	39(14%)	17(12%)	22(16%)
Cancer			
No	271(97%)	138(99%)	133(96%)
Yes	7(3%)	1(1%)	6(4%)
Hypertension			
No	181(65%)	86(62%)	95(68%)
Yes	97(35%)	53(38%)	44(32%)
Vascular			
No	267(96%)	131(94%)	136(98%)
Yes	11(4%)	8(6%)	3(2%)
Asthma			
No	272(98%)	137(99%)	135(97%)
Yes	6(2%)	1(1%)	4(3%)
COPD			
No	276(99%)	138(99%)	138(99%)
Yes	2(1%)	1(1%)	1(1%)
HIV/AIDS			
No	230(83%)	115(83%)	115(83%)
Yes	48(17%)	24(17%)	24(17%)
TB			
No	273(98%)	137(99%)	136(98%)
Yes	5(2%)	2(1%)	3(2%)
Acute Pancreatitis			
No	278(100%)	139(100%)	139(100%)
Yes	0	0	0
Chronic Pancreatitis			

No Yes	277(100%)	139(100%) 0	138(99%) 1(1%)
Hepatitis No Yes	278(100%)	139(100%) 0	139(100%) 0
Rheumatoid Arthritis No Yes	273(98%) 5(2%)	135(97%) 4(3%)	138(99%) 1(1%)
Other No Yes	264(95%) 14(5%)	135(97%) 4(3%)	129(93%) 10(7%)

Table 10Dietary characteristics of study population

Characteristic	Overall (n=278)	Control (n=139)	Case (n=139)
Red meat			
No	26(9%)	11(8%)	15(11%)
yes	252(91%)	128(92%)	124(89%)
Red meat week			
No	37(13%)	20(14%)	17(12%)
Low	214(77%)	109(79%)	105(76%)
High	27(10%)	10(7%)	17(12%)
White meat			
No	10(4%)	1(1%)	9(6%)
Yes	268(96%)	138(99%)	130(94%)
White meat week			
No	15(5%)	6(4%)	9(7%)
Low	208(75%)	98(71%)	110(79%)
High	55(20%)	35(25%)	20(14%)
Vegetables			
No	16(6%)	1(1%)	15(11%)
Yes	262(94%)	138(99%)	124(89%)
Vegetable week			
No	22(8%)	6(4%)	16(11%)
Low	138(50%)	63(45%)	75(54%)
High	118(42%)	70(51%)	48(35%)
Fish			
No	27(10%)	9(6%)	18(13%)
Yes	251(90%)	130(94%)	121(87%)
Fish week			
No	61(22%)	37(27%)	24(17%)
Low	211(76%)	97(70%)	114(82%)
High	6(2%)	5(3%)	1(1%)
Fried Fish			
No	195(70%)	90(65%)	105(76%)
Yes	83(30%)	49(35%)	34(24%)

Canned Fish			
No	145(52%)	64(46%)	81(58%)
Yes	133(48%)	75(54%)	58(42%)
Fresh Fish			
No	100(36%)	37(27%)	63(45%)
Yes	178(64%)	102(73%))	76(55%)
Sea food			
No	269(97%)	131(94%)	138(99%)
Yes	9(3%)	8(6%)	1(1%)
Cured food			
No	105(38%)	44(32%)	61(44%)
Yes	173(62%)	95(68%)	78(56%)
Cured food week			
No	130(47%)	64(46%)	66(47%)
Low	143(51%)	70(50%)	73(53%)
High	5(1%)	5(4%)	0
Fried food week			
No	62(22%)	23(17%)	39(28%)
Low	197(71%)	105(76%)	92(66%)
High	19(7%)	11(7%)	8(6%)
Maize meal			
No	21(8%)	4(3%)	17(12%)
Yes	257(92%)	135(97%)	122(88%)
Maize meal week			
No	21(7%)	4(3%)	17(12%)
Low	106(38%)	57(41%)	49(35%)
High	150(54%)	78(56%)	72(52%)
Other	1(1%)	0	1(1%)

Table 11Hospital characteristics for the study population

Characteristic	Overall (n=278)	Control (n=139)	Case (n=139)
Functionality			
100%-normal function	60(22%)	37(27%)	23(17%)
90%-capable of normal	87(31%)	50(36%)	37(27%)
80%-normal activity	69(25%)	22(16%)	47(34%)
70%-cares for self	39(14%)	15(11%)	24(24%)
60%-requires help	10(4%)	6(4%)	4(3%)
50%-often requires help	9(3%)	6(4%)	3(2%)
40%-disabled	4(1%)	3(2%)	1(1%)
Symptoms level			
0-asymptomattic	60(22%)	37(27%)	23(17%)
1-symptomatic but ambulatory	160(58%)	75(54%)	85(61%)
2-symptomatic, < 50%	44(16%)	19(14%)	25(18%)
3-symptomatic, >50%	11(4%)	6(4%)	5(4%)
4-bedbound	3(1%)	2(1%)	1(1%)
BMI	N=185	N=61	N=124
Mean	24.28	25.61	23.63
SD	7.59	7.80	7.43
BMI category			
Missing	93(33%)	78(56%)	15(11%)
Underweight	41(15%)	11(8%)	30(22%)
Healthy	76(27%)	22(16%)	54(39%)
Overweight	29(10%)	10(7%)	19(14%)
Obese	23(8%)	9(6%)	14(10%)
Very obese	16(6%)	9(6%)	7(5%)

Table 12Factors associated with PDAC

Characteristic	UOR(95%CI)	p-value	AOR(95%CI)	p-value
Race				
Black	1 (base)			
White	1.33(0.54-3.29)	0.9432	†	†
Indian	0.8(0.21-3.06)			
Coloured	0.8(0.30-2.01)			
Other	No values			
Gender				
Female	1 (base)	**0.0224**	1(base)	0.6992
Male	1.73 (1.07-2.79)		1.49(0.73-3.06)	
Age years	1.03(1.01-1.05)	**0.0004**		
Age				
20-29	0.07 (0.008-0.56)		0.11(0.01-1.06)	
30-39	0.33(0.12-0.92)		0.60(0.19-1.90)	
40-49	1(base)	**0.0000**	1(base)	0.0006
50-59	1.63(0.72-3.63)		2.61(1.00-6.78)	
60-69	1.46(0.64-3.35)		3.04(1.10-8.34)	
70-79	0.93(0.34-2.55)		1.58(0.49-5.11)	
80+	1(0.27-3.73)		1.87(0.42-8.31)	
Schooling				
Never attended school	1 (base)			
Primary school	1.14(0.48-2.69)	0.1023	†	†
Secondary school	0.63(0.28-1.45)			
Further Education	0.48(0.14-1.70)			
Birth place				
Gauteng	1 (base)			
Mpumalanga	1.11(0.38-3.23)			
Limpopo	1.75(0.56-5.49)	0.7833	†	†
North West	1.21(0.53-2.79)			
Free State	1.24(0.52-2.92)			
KwaZulu-Natal	1.07(0.43-2.68)			
Eastern Cape	0.54(0.17-1.69)			

Northern Cape Western Cape Other Country	empty 0.98(0.06-15.84) 0.55(0.22-1.41)			
Employed No Yes	 1 (base) 1.21(0.70-2.07)	0.4915	†	†
Long Employment Construction & chemic Mining Manufacturing & Factory Engineering Drivers General worker Office & other	 2.24(0.76-6.59) 2.69(0.47-15.30) 0.71(0.29-1.73) 0.67(0.19-2.36) 2.50(0.93-6.73) 2.04(1.15-3.62) 1 (base)	**0.0370**	2.08(0.59-7.37) 1.70 (0.24-12.16) 0.49(0.17-1.40) 0.30(0.07-1.28) 1.95(0.60-6.35) 1.83(0.93-3.63) 1 (base)	0.0419

† No values since the UOR was statistically insignificant

Table 13Chronic medical factors associated with PDAC

Factor	UOR(95%CI)	p-value	AOR(95%CI)	p-value
Chronic illness			†	†
No	1 (base)	0.9031		
Yes	0.97(0.60-1.56)			
Diabetes			†	†
No	1 (base)	0.3873		
Yes	1.35(0.68-2.67)			
Cancer			∞	∞
No	1 (base)	**0.0440**		
Yes	6.23(0.74-52.41)			
Hypertension			†	†
No	1 (base)	0.2572		
Yes	0.75(0.46-1.23)			
Vascular			†	†
No	1 (base)	0.1174		
Yes	0.36(0.09-1.39)			
Asthma			†	†
No	1 (base)	0.4047		
Yes	2.03(0.37-11.26)			
COPD			†	†
No	1 (base)	1.0000		
Yes	1.00(0.06-16.15)			
HIV/AIDS			†	†
No	1 (base)	1.0000		
Yes	1.00(0.54-1.86)			
TB			†	†
No	1 (base)	0.6507		
Yes	1.51(0.25-9.19)			
Acute Pancreatitis			†	†
No	No values			
Yes				
Chronic Pancreatitis	No values		†	†

No Yes				
Hepatitis No Yes	No values		†	†
Rheumatoid Arthritis No Yes	1 (base) 0.24(0.03-2.22)	0.1615	†	†
Other No Yes	1 (base) 2.62(0.80-8.55)	0.0947	‡	‡

† No values since the UOR was statistically insignificant.

∞The variable not added in the model as the cancer was in exclusion criteria.

‡ AOR was statistically insignificant.

Table 14Dietary factors associated with PDAC

Factor	UOR(95%CI)	p-value	AOR(95%CI)	p-value
Red meat			†	†
No	1 (base)	0.4091		
Yes	0.71(0.31-1.61)			
Red meat week				
No	1(base)	0.3406	1(base)	0.9043
Low	1.13(0.56-2.28)		0.99(0.42-2.31)	
High	2.00(0.73-5.51)		1.13(0.33-3.92)	
White meat				
No	1 (base)	**0.0058**	*	*
Yes	0.10(0.01-0.84)			
White meat week				
No	1(base)	**0.0659**	1(base)	0.0502
Low	0.75(0.26-2.18)		0.28(0.06-1.31)	
High	0.38(0.12-1.23)		0.14(0.03-0.73)	
Vegetables				
No	1 (base)	**0.0001**	*	*
Yes	0.06(0.01-0.46)			
Vegetable week				
No	1(base)	**0.0071**	1(base)	0.0037
Low	0.45(0.16-1.21)		0.25(0.07-0.89)	
High	0.26(0.09-0.70)		0.13(0.04-0.49)	
Fish				
No	1 (base)	**0.0660**	*	*
Yes	0.47(0.20-1.08)			
Fish week				
No	1(base)	**0.0291**	1(base)	0.0229
Low	1.81(1.01-3.24)		2.15(1.04-4.45)	
High	0.31(0.03-2.80)		0.17(0.01-2.78)	
Fried Fish				
No	1 (base)	**0.0011**	*	*
Yes	0.44(0.26-0.72)			

Canned Fish No Yes	 1 (base) 0.61(0.38-0.98)	**0.0410**	*	*
Fresh Fish No Yes	 1 (base) 0.59(0.35-1.00)	**0.0488**	*	*
Sea food No Yes	 1 (base) 0.12(0.01-0.96)	**0.0115**	*	*
Cured food No Yes	 1 (base) 0.59(0.36-0.97)	**0.0352**	*	*
Cured food week No Low High	 1(base) 1.01(0.63-1.63) Empty	0.9632	†	†
Fried food week No Low High	 1(base) 0.52(0.29-0.93) 0.43(0.15-1.22)	**0.0636**	1(base) 0.47(0.22-0.96) 0.39(0.11-1.35)	0.0700
Maize meal No Yes	 1 (base) 0.21(0.07-0.65)	**0.0023**	*	*
Maize meal week No Low High Other	 1(base) 0.20(0.06-0.64) 0.22(0.07-0.68) Empty	**0.0086**	‡	‡

† No values since the UOR was statistically insignificant.

‡ AOR was statistically insignificant.

* The multiple categories in the model and not the binary (yes/no) category.

Table 15Social factors associated with PDAC

Factor	UOR(95%CI)	p-value	AOR(95%CI)	p-value
Smoked			*	*
No	1 (base)	**0.0702**		
yes	1.55(0.96-2.49)			
Smoke now			*†	*†
No	1 (base)	0.4830		
Yes	1.21(0.70-2.11)			
Smoke pack years	1.01(1.00-1.03)	**0.0583**	∞	∞
Smoke pack category				
No	1(base)		1(base)	
Very rare	1.50(0.85-2.67)	**0.0749**	1.43(0.66-3.08)	0.7667
Frequent	2.06(1.02-4.14)		0.91(0.36-2.28)	
Very frequent	2.29(0.98-5.36)		1.10(0.38-3.19)	
Exposed to smoke at home			*†	*†
No		0.3759		
Yes	1 (base)			
	0.80(0.49-1.31)			
Exposed to smoke at work			*†	*†
No		0.5466		
Yes	1 (base)			
	1.16(0.72-1.86)			
Snuff			*†	*†
No	1 (base)	0.8276		
Yes	1.10(0.47-2.58)			
Frequency of snuff			†	†
No	1(base)			
Frequently-daily	2.29(0.69-7.61)	0.3565		
Infrequently-2/3 times weekly	0.68(0.11-4.12)			
Rarely-2/3 times monthly	0.34(0.03-3.30)			

Chew tobacco			*†	*†
No	1 (base)	0.4708		
Yes	1.69(0.40-7.22)			
Tobacco frequency		0.8390	†	†
No	1(base)			
Frequently-daily	2.03(0.18-22.65)			
Infrequently-2/3 times weekly	1.01(0.14-7.31)			
Rarely-2/3 times monthly	Empty			
Alcohol		**0.0037**	*	*
No	1 (base)			
Yes	2.02(1.25-3.27)			
Bought beer		0.1350	†	†
No	1 (base)			
Yes	1.45(0.89-2.37)			
Home brewed		0.7070	†	†
No	1 (base)			
Yes	0.87(0.42-1.82)			
Bought spirits		0.4151	†	†
No	1 (base)			
Yes	1.30(0.69-2.48)			
Wine		0.3767	†	†
No	1 (base)			
Yes	0.73(0.36-1.47)			
Audit score		**0.0849**		0.2135
No	1 (base)		1(base)	
Low	1.33(0.76-2.32)		1.62(0.78-3.35)	
High	2.98(1.02-8.67)		3.17(0.79-12.69)	

*† No values since the UOR was statistically insignificant and the multiple category was added in the model and not the binary (yes/no) category.

‡ AOR was statistically insignificant.

∞ Categorical variable used in the model.

Table 16Hospital factors associated with PDAC

Factor	UOR(95%CI)	p-value	AOR(95%CI)	p-value
Functionality				
100%-normal function	1 (base)			
90%-capable of normal	1.19(0.61-2.33)			
80%-normal activity	3.44(1.66-7.10)	**0.0040**	~~A~~	~~A~~
70%-cares for self	2.57(1.12-5.90)			
60%-requires help	1.07(0.27-4.21)			
50%-often requires help	0.80(0.18-3.53)			
40%-disabled	0.54(0.05-5.47)			
Symptoms level				
0-asymptomattic	1 (base)			
1-symptomatic but ambulatory	0.75(0.20-2.73)	0.1846	†	†
2-symptomatic, < 50%	1.36(0.40- 4.64)			
3-symptomatic, >50%	1.58(0.42-5.96)			
4-bedbound	No values			
BMI	0.97(0.93-1.01)	**0.0991**	≠	≠
BMI category				
Missing	0.08(0.04-0.16)			
Underweight	1.11(0.47-2.60)			
Healthy	1 (base)	**0.0000**	≠	≠
Overweight	0.77(0.31-1.93)			
Obese	0.63(0.24-1.68)			
Very obese	0.32(0.10-0.96)			

† No values since the UOR was statistically insignificant.

~~A~~ Variable not added in the multivariate analysis as this could been a symptoms of illness.

≠ Variable not added due to more than 10% missing data.

Table 17Final model

Characteristics	AOR(95%CI)	p-value
Age Category		
20-29	0.11(0.11-1.00)	0.05
30-39	0.52(0.17-1.61)	0.262
40-49	1(base)	
50-59	2.63(1.03-6.70)	0.042
60-69	2.85(1.08-7.50)	0.034
70-79	1.59(0.51-5.03)	0.426
80+	1.51(0.35-6.51)	0.577
Long employment		
Construction & chemic	2.23(0.66-7.51)	0.196
Mining	2.61(0.38-17.96)	0.329
Manufacturing & Factory	0.55(0.20-1.52)	0.248
Engineering	0.42(0.11-1.62)	0.208
Drivers	2.19(0.70-6.86)	0.178
General worker	1.79(0.93-3.45)	0.083
Office & other	1(base)	
White meat week		
No	1(base)	
Low	0.38(0.09-1.61)	0.190
High	0.18(0.04-0.86)	0.032
Fish week		
No	1(base)	
Low	2.17(1.06-4.45)	0.034
High	0.30(0.02-3.70)	0.346
Fried week		
No	1(base)	
Low	0.48(0.23-1.00)	0.050
High	0.59(0.16-2.15)	0.425
Vegetable week		
No	1(base)	
Low	0.33(0.09-1.17)	0.085

High	0.17(0.05-0.61)	0.007